High-resolution global map of smallholder and industrial closed-canopy oil palm plantations

Adrià Descals[1], Serge Wich[2,3], Erik Meijaard[4,5,6], David L.A. Gaveau[7,8], Stephen Peedell[9], and Zoltan Szantoi[9,10,*]

[1]CREAF, Cerdanyola del Vallès, 08193 Barcelona, Spain; a.descals@creaf.uab.cat
[2]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
[3]Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Science Park, 904, 1098 XH, Amsterdam, The Netherlands
[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, Queensland 4072, Australia
[7]Center for International Forestry Research, P.O. Box 0113 BOCBD, Bogor, Indonesia; d.gaveau@cgiar.org
[8]TheTreeMap, Bagadou Bas 46600 Martel, France
[9]European Commission, Joint Research Centre, 20127 Ispra, Italy; stephen.peedell@ec.europa.eu
[10]Stellenbosch University, Stellenbosch 7602, South Africa

Correspondence to: Zoltan Szantoi (zoltan.szantoi@remote-sensing-biodiversity.org)

Abstract. Oil seed crops, especially oil palm, are among the most rapidly expanding agricultural land uses, and their expansion is known to cause significant environmental damage. Accordingly, these crops often feature in public and policy debates, which are hampered or biased by a lack of accurate information on environmental impacts. In particular, the lack of accurate global crop maps remains a concern. Recent advances in deep learning and remotely-sensed data access make it possible to address this gap. We present a map of closed-canopy oil palm (*Elaeis guineensis*) plantations by typology (industrial vs. smallholder plantations) at the global scale and with unprecedented detail (10-meter resolution) for the year 2019. The DeepLabv3+ model, a convolutional neural network (CNN) for semantic segmentation, was trained to classify Sentinel-1 and Sentinel-2 images into an oil palm land cover map. The characteristic backscatter response of closed-canopy oil palm stands in Sentinel-1 and the ability of CNN to learn spatial patterns, such as the harvest road networks, allowed the distinction between industrial and smallholder plantations globally (overall accuracy = 98.52 ± 0.20%), outperforming the accuracy of existing regional oil palm datasets that used conventional machine learning algorithms. The user's accuracy, reflecting commission error, in industrial and smallholders was 88.22 ± 2.73% and 76.56 ± 4.53%, and the producer's, reflecting omission error, accuracy was 75.78 ± 3.55% and 86.92 ± 5.12% respectively. The global oil palm layer reveals that closed-canopy oil palm plantations are found in 49 countries, covering a mapped area of 19.60 Mha; area estimate was 21.00 ± 0.42 Mha (72.7% industrial and 27.3% smallholder plantations). Southeast Asia ranks as the main producing region with an oil palm area estimate of 18.69 ± 0.33 Mha, or 89% of global closed-canopy plantations. Our analysis confirms significant regional variation in the ratio of industrial versus smallholder growers, but also that, from a typical land development perspective, large areas of legally defined smallholder oil palm resemble industrial-scale plantings. Since our

Deleted: zoltan.szantoi@ec.europa.eu
Deleted: machine
Deleted: an up-to-date
Deleted: an
Deleted: Sentinel-1 and Sentinel-2 data were used to train a
Deleted: the
Deleted: overall accuracy
Deleted: 97
Deleted: 5
Deleted: and kappa = 84.9%
Deleted: 73.8
Deleted: 89.4
Deleted: 85
Deleted: 6
Deleted: 78.8
Deleted: 47
Deleted: tropical
Formatted: Not Highlight
Formatted: Not Highlight
Deleted: 17.47 x
Formatted: Not Highlight
Deleted: 10^6
Deleted: 90
Formatted: Not Highlight
Deleted:

study identified only closed-canopy oil palm stands, our area estimate was lower than the harvested area reported by FAO, particularly in Western Africa, due to the omission of young and sparse oil palm stands, oil palm in non-homogeneous settings, and semi-wild oil palm plantations. An accurate global map of planted oil palm can help to shape the ongoing debate about the environmental impacts of oil seed crop expansion, especially if other crops can be mapped to the same level of accuracy. As our model can be regularly rerun as new images become available, it can be used to monitor the expansion of the crop in monocultural settings. The global oil palm layer for the second half of the year 2019 at a spatial resolution of 10 meters can be found at https://doi.org/10.5281/zenodo.4473715 (Descals et al., 2021).

Deleted: The overall oil palm surface per country is similar to the harvested area reported by FAO, except for countries in Western Africa, where our estimates are lower due to the

Deleted: of

Deleted: feral

Deleted: In Indonesia, the world's largest producer, our planted area estimate is higher because FAO does not report unregistered landholdings. Our model identifies primarily closed-canopy oil palm stands and misses young or sparsely planted oil palm stands.

Deleted: ry

Deleted: is published

Deleted: reliably to

Deleted:

Deleted: a

Deleted: https://doi.org/10.5281/zenodo.3884602 (Descals et al., 2020)

1 Introduction

Crops that produce vegetable oils, such as soy, rapeseed, oil palm, and sunflower, take up ca. 6% of all agricultural land and ca. 2.3% of the total global land area, and are among the world's most rapidly expanding crop types (OECD, 2018). Demand for vegetable oils is increasing, with one estimate foreseeing an increase from 205 Mt in 2019 (OECD, 2018) to 310 Mt in 2050 (Byerlee et al., 2017). This has created a need to optimize land use for vegetable oil production in order to minimize environmental impacts and maximize socio-economic benefits. One of the requirements for this is accurate global maps for all oil-producing crops. The most comprehensive maps available (IFPRI, 2019) map these crops by disaggregating crop statistics identified at national and sub-national units for the year 2005 to 5 arc-minute grid cells, which is a relatively coarse spatial resolution. Direct identification of crops from satellite imagery is likely to result in more accurate maps that delineate where different crops have been planted. One of the most extensively mapped crops is oil palm (Elaeis guineensis) because of societal concerns about the associated environmental impacts on tropical forests and social disruption. However, only the global extent of industrial plantations is reasonably well known, while the more heterogeneous plantings at smallholder scales remain largely unmapped (Meijaard et al., 2018).

Deleted: (Meijaard et al., 2018)

A global map of oil palm at each production scale provides critical insights into the current debate about the social and environmental sustainability of the crop (Meijaard et al., 2018; Meijaard et al., 2020). It would allow for a more accurate determination of the environmental impacts from oil palm expansion, for example, by assessing the deforestation that preceded oil palm development, the related carbon emissions as well as the impacts on species' distributions, key biodiversity areas, and socio-economic impacts. As total and local production volumes of palm oil are reasonably well known, a comparison to the total planted area would allow more accurate average yield estimates and regional variations in yield. Similarly, accurate maps of planted oil palm can determine the extent to which oil palm development has displaced other food crops, an important element in the policy debate in the European Union regarding the use of palm oil in biofuels (Meijaard & Sheil, 2019). Such information is important for comparing oil palm to other vegetable oil crops, such as soy, rapeseed, sunflower, groundnut, and coconut, once global maps for these crops become available. The challenge is thus to develop a method to accurately map large industrial plantations as well as smallholder oil palm areas.

Deleted: Meijaard, Brooks, et al., 2020;

Deleted: (Meijaard et al., 2018)

Deleted: (Erik Meijaard & Sheil, 2019)

Previous studies have demonstrated the usefulness of radar imagery for the detection of closed-canopy oil palm stands. Palm-like trees have a characteristic backscatter response which consists of a ~~low vertical transmit and vertical receive (VV)~~ and ~~high vertical~~ transmit and ~~horizontal~~ receive ~~(VH) in Sentinel-1, or a high horizontal transmit and vertical receive (HV)~~ and low horizontal transmit and horizontal receive (HH) in PALSAR imagery (Miettinen & Liew, 2011). This characteristic backscatter response is a consequence of the canopy structure of palm-like trees and allows the detection of closed-canopy palm plantations, particularly oil palm. Several studies have taken advantage of this characteristic backscatter response for mapping oil palm at the local and the regional scale (Koh et al., 2011; Lee et al., 2016; Nomura et al., 2019; Oon et al., 2019), and similarly using supervised classification models ~~(Descals et al., 2019; Shaharum et al., 2020; Xu et al., 2020)~~.

The mapping of oil palm plantations by typology (smallholder versus industrial) with remotely-sensed data presents a more challenging classification problem than the detection of only closed-canopy oil palm. In addition to the backscatter response of radar data, texture analysis also offers a complementary method to distinguish between smallholders and industrial-scale plantations (Descals et al., 2019). Contextual information, such as the presence and shape of harvesting road network and drainage structures, can be included as predictive variables for the classification of industrial and smallholder plantations.

Deep learning, in particular semantic segmentation, is a subfield of machine learning with characteristics suitable for the distinction of smallholder and industrial oil palm plantations. Deep learning employs a series of models for computer vision that excel in very complex classification scenarios (LeCun et al., 2015) and, in particular, convolutional neural networks (CNN) have recently been embraced by the remote sensing community due to the ability to recognize intricate patterns in the images (Ma et al., 2019). To date, there are no studies that consider CNN for the land use classification of oil palm plantations at regional or global scales. One study used deep learning for object detection, focusing on the identification of single palm trees (Li et al., 2017).

The aim of this study is i) present an up-to-date map of oil palm plantations by typology (industrial vs. smallholder plantations) at the global scale and with unprecedented detail (10-meter resolution) for the year 2019 and ii) show the suitability of deep learning in remote sensing for complex classification scenarios in which contextual information may be useful.

Deleted: high

Deleted: lower

Deleted: horizontal

Deleted: vertical

Deleted: HV

Formatted: English (United Kingdom)

Formatted: English (United Kingdom)

Deleted: complimentary

Deleted: identify distinctions

Deleted: observes spatial patterns such

Deleted: in plantations, drainage structure, and association with palm oil mills, which

Deleted: help to distinguish

Deleted: between

Deleted: Only a few studies

Deleted: , and have only

Deleted: achieved

2 Methods

2.1 Overview

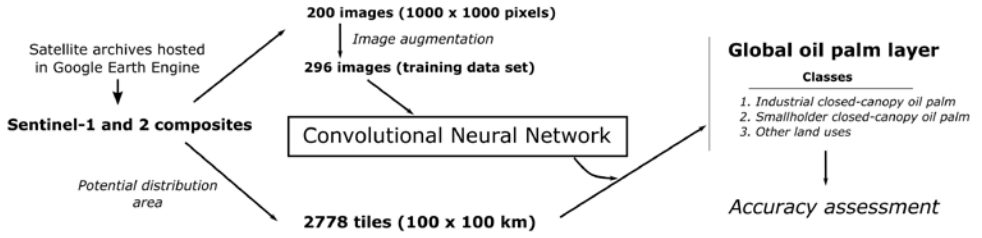
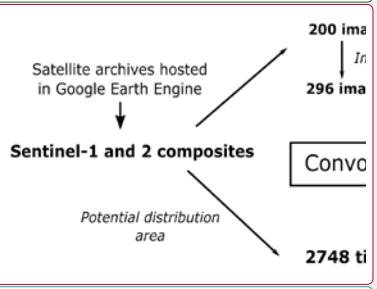


Figure 1: Diagram of the algorithm used to generate the global oil palm layer. The input images, Sentinel-1 and Sentinel-2 half-year composites, were obtained from Google Earth Engine in a grid of 100 x 100 km. The Sentinel-1 and Sentinel-2 tiles were classified with a convolutional neural network (CNN). The CNN model was trained with labeled images with constant size (1000 x 1000 pixels). The output classification layer was validated with 13,495 points that were randomly distributed.

The classification model for oil palm plantations used the Sentinel-1 and Sentinel-2 half-yearly composites as input images (Figure 1). The maps presented in this study correspond to the second half-year of 2019. We used a deep learning model that was trained with 296 images of 1000 x 1000 pixels distributed throughout the main oil palm producing regions and applied over Sentinel-1 and Sentinel-2 composites in the potential area (Figure 2) where oil palm can grow. Table 1 shows the geospatial data used in the study. The links to the datasets appear in section *6. Data availability*. The processing steps depicted in figure 1 were implemented in different computing environments (Supplementary figure 1) depending on the convenience of the processing. The annual compositing of Sentinel-1 and Sentinel-2 images was done in Google Earth Engine (GEE) (Gorelick et al., 2017) since a cloud-processing platform was suited for this task considering the high amount of satellite data required in the compositing. The visual interpretation of training and validation data was also done in GEE. The training of the CNN and the classification of images, however, was performed with a local computer using Matlab 2019a since the implementation of the CNN model was less feasible in GEE. The CNN model can be also trained and used for the prediction of images with Python (Code accessible through section 5. *Code availability*). The Sentinel-1 and Sentinel-2 images taken in 2019 are the only data necessary to reproduce the results of the global oil palm map. The rest is auxiliary data used for the identification of the oil palm distribution, the visual interpretation of oil palm plantation, and the comparison with other oil palm maps.

Moved (insertion) [3]



Deleted:

Formatted: Font color: Custom Color(RGB(34,34,34)), Highlight

Deleted: ¶

Deleted: 252

Formatted: Font color: Auto

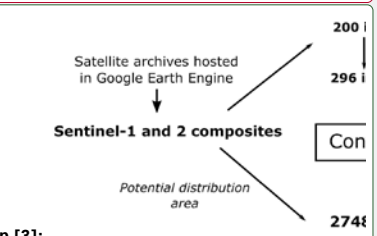
Formatted: Normal

Deleted: uses

Deleted: The deep learning model that we used is the DeepLabv3+ (Chen et al., 2017, 2018) based on MobileNetV2 (Sandler et al., 2018) as the CNN backbone. The model was

Formatted: Font: Italic

Formatted: Font: Italic



Moved up [3]:

¶

¶

Figure 1: Diagram of the algorithm used to generate the global oil palm layer. The input images, Sentinel-1 and Sentinel-2 half-year composites, were obtained from Google Earth Engine in a grid of 100 x 100 km. The Sentinel-1 and Sentinel-2 tiles were classified with a convolutional neural network (CNN). The CNN model was trained with labeled images with constant size (1000 x 1000 pixels).

Formatted: Font color: Custom Color(RGB(34,34,34)), Highlight

Formatted: Normal

Formatted: Font color: Auto

Table 1. Data sources used in the study.

Source	Band/Input	Spatial resolution	Usage*	Reference
Sentinel-1 GRD	VV and VH	10 m	1, 3	(Torres et al., 2012)
Sentinel-2 Level-2A	B4	10 m	1, 3	(Drusch et al., 2012)
IUCN industrial layer	Land Cover (Oil palm map)	30 m	2, 4	(Meijaard et al., 2018)
Oil palm layer Sumatra	Land Cover (Oil palm map)	10 m	2	(Descals et al., 2019)
Oil palm layer Indonesia	Land Cover (Oil palm map)	30 m	2	(Gaveau et al., 2021)
Oil palm layer in SE Asia	Land Cover (Oil palm map)	100 m	2	(Xu et al., 2020)
WorldClim V1 Bioclim	19 bioclimatic variables	30 arcsec	4	(Hijmans et al., 2005)
DigitalGlobe imagery	RGB orthoimages	< 1 m	3	(Google Earth Engine, 2020)
FAOSTAT	Oil palm harvested area	country-level statistics	2	(FAO, 2020)

* The column describes how the data was used in the study: (1) input of the convolutional neural network (CNN), (2) used for comparison with the results of the CNN, (3) base layers for the visual interpretation of oil palm plantations, and (4) used for the identification of the potential distribution of oil palm.

2.2 Potential distribution of oil palm

The classification of oil palm plantations was restricted to those areas where the climatic conditions were favorable for oil palm growth. In order to delimit the potential distribution of oil palm, we used climate data and an existing global oil palm dataset. The climate dataset was obtained from WorldClim V1 Bioclim (Hijmans et al., 2005), which provides nineteen gridded variables at a spatial resolution of 30 arc seconds that are generated from monthly temperature and precipitation. This study's existing oil palm layer was obtained from the IUCN (Meijaard et al., 2018), and shows the industrial oil palm plantations at the global scale (link to the IUCN layer is available in Section 6. *Data Availability*). This map was derived from a compilation of all published spatial data on oil palm combined with manual digitizing of characteristics spatial signatures of industrial-scale oil palm using cloud-free Landsat mosaics acquired in 2017 and created in GEE.

The potential area where oil palm can grow was estimated with the climate variable range in the IUCN layer. We estimated the histogram of the nineteen bioclimatic variables in the areas that were classified as industrial oil palm plantations in the IUCN layer. Supplementary Table 1 shows the minimum and maximum of each bioclimatic variable for the industrial plantations. A pixel in the WorldClim dataset was considered favorable for oil palm growth when at least seventeen out of the nineteen bioclimatic variables fell within the climate range observed in the IUCN layer (Supplementary figure 2). The resulting potential oil palm distribution map encompasses similar areas as used in previous studies (Pirker et al., 2016; Strona et al., 2018; Wich et al., 2014). The classification of oil palm plantations was processed in a grid of 100 x 100 kilometers that covers the area with favorable conditions for oil palm growth (Figure 2).

Formatted ...
Formatted ...
Formatted Table ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Deleted: (Descals et al., 2019)(E Meijaard et al., 2018)
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Deleted: (Gaveau et al., 2021)
Formatted ...
Formatted ...
Deleted: (Descals et al., 2019)
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...
Formatted ...

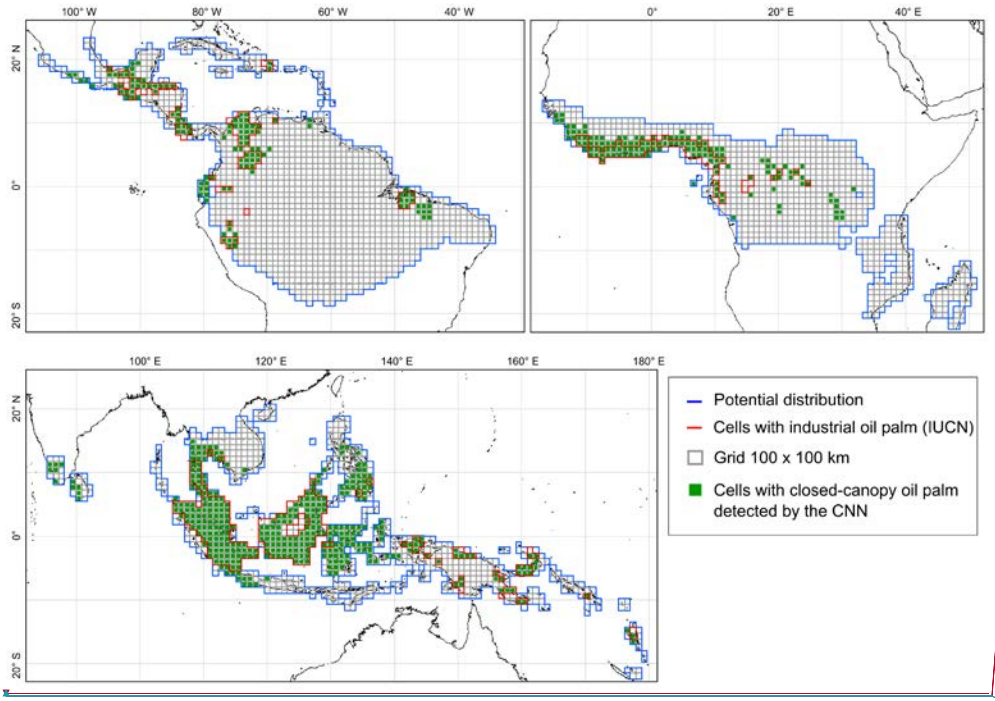


Figure 2: Localization map of the grid cells where the convolutional neural network (CNN) was applied for the classification of industrial and smallholder plantations. The grid cells cover a potential distribution area (blue line) over 7 tropical regions of the world where oil palm can grow: Central and South America, Central and Western Africa, South and Southeast Asia, and the Pacific. Cells in red depict the areas where there is presence of industrial oil palm plantations in the IUCN layer. Cells filled with green signify areas where closed-canopy oil palm was detected by the CNN.

2.3 Sentinel-1 and Sentinel-2 pre-processing

The CNN classifies radar and optical images collected by Sentinel-1 (C-band) (Torres et al., 2012) and Sentinel-2 (multispectral) (Drusch et al., 2012) satellites, respectively, both missions launched by the European Space Agency and part of the Copernicus Programme (www.copernicus.eu). The images were pre-processed and downloaded from GEE (code is available in section 5. *Code availability*). We used the Sentinel-1 SAR Ground Range Detected (GRD), which has a temporal resolution of 12 days, in both ascending and descending orbits. We used the Interferometric Wide Swath images processed at a spatial resolution of 10 meters. The scenes were processed with the local incident angle (LIA) correction and

Deleted:
Formatted: Font color: Custom Color(RGB(36,41,46))
Deleted: Convolutional
Deleted: Neural
Deleted: Network
Deleted: 6
Deleted: Southeast Asia,
Deleted: a
Deleted: signifiy
Deleted: oil palm was detected with the convolutional neural network
Deleted: spectral
Formatted: Font: Italic

then the median value was computed over the second half-year of 2019 for the ascending and descending scenes separately. The final composite is the average of the two orbit composites.

We also used Band 4 (red band; central wavelength = 665 nm) of Sentinel-2 Level 2A (surface reflectance). Different feature selection algorithms highlighted the relevance of Band 4 for predicting industrial and oil palm plantations in a previous study (Descals et al., 2019). Band 4 is the 10-meter resolution band that best shows the roads in industrial plantations because of the high contrast in terms of reflectance between the road and the surrounding oil palm. The high light scattering of vegetation in the near-infrared spectrum makes the recognition of roads less feasible in the 10-meter near-infrared band (Band 8). The Sentinel-2 images were masked with the quality flag provided in Level 2A, which is produced by the ATCOR algorithm and provides information about the clouds, cloud shadows, and other non-valid observations (Drusch et al., 2012). The images were aggregated for the second half-year of 2019 using the normalized difference vegetation index as the quality mosaic. The 5-day revisit time of Sentinel-2 allowed the generation of cloud-free composites over the study area.

2.4 Image labeling

Semantic segmentation models require input images with a constant size for both training and prediction. The size of the input images in this study was set to 1000 x 1000 pixels, which corresponds to an area of 10 x 10 km in a 10-meter-resolution image. We set an input size of 10 km because it captures the contextual spatial information necessary for identifying smallholders and industrial plantations (e.g., harvesting road network). Consequently, the model was trained with Sentinel-1 and Sentinel-2 half-yearly annual composites of 10 x 10 km. The oil palm plantations that were present within the Sentinel composites were labeled by visual interpretation. We digitized the oil palm plantations also by interpreting the very-high-resolution DigitalGlobe images that are displayed as the base layer in GEE. The DigitalGlobe images have a sub-meter spatial resolution and are displayed as true-color composites in GEE. These images are updated regularly and the date depends on the location, but usually, the images are taken during the past one to two years. The DigitalGlobe images were used as complementary data to the Sentinel-1 and 2 composites in the visual interpretation. We used the geometry editing tool in GEE for labeling smallholder and industrial plantations. Once the training areas were labelled, we downloaded the truth images from GEE along with the Sentinel-1 and Sentinel-2 composites for the second half-year of 2019. The image labeling was carried out in 84 different regions of the world where oil palm is cultivated (Supplementary figure 3) and resulted in 200 training images.

Deep learning algorithms require large amounts of data to ensure good performance and data augmentation is a technique used to improve the performance of the models when the size of the training data is small (Shorten & Khoshgoftaar, 2019). Data augmentation aims to generate a more diverse training dataset with certain affine transformations applied to the original training data. Data augmentation techniques have been used in remote sensing studies (Yu et al., 2017), in which affine transformations such as flips, translations, and rotations have improved the accuracy results of deep learning models. We

Deleted: (more information in section 5. Code availability)

Deleted: top-of-atmosphere reflectances

Deleted: the

Deleted: (more information in section 5. Code availability)

Deleted:

Deleted: I

Deleted: roads, mills

Deleted: the size of

Deleted: that were labeled by visual interpretation of the

Deleted: provided by DigitalGlobe and used

Deleted: a

Deleted: Google Earth

Deleted: and for downloading the labeled images

Deleted: 2

Deleted: image

Deleted: that is often

Deleted:

used the rotation of images (90 degrees clockwise) as the data augmentation technique for this study (Supplementary figure 4). The rotation was applied only to the training images that presented more than 10% of the pixels labeled as smallholders in order to reduce the class imbalance between industrial and smallholder plantations. We also clipped the central area of 4x4 blocks of labeled images and rotated them with an angle of 45°. This process resulted in 96 additional images that were added to the 200 original training images.

2.5 Definition of industrial and smallholder plantations

Definitions of smallholders and industrial plantations differ per country, and many variations within each of these classes exist (Bronkhorst et al., 2017; Glenday & Gary, 2015; Meijaard & Sheil, 2019). For the current study, we used the following generalized classifications. An *industrial oil palm* plantation typically covers several thousand hectares of land and is very well structured and homogeneous in tree age. It consists of an area bounded by long linear, sometimes rectangular boundaries. It has a dense trail and a road/canal network. Roads in industrial plantations are developed at the start of plantation development and, therefore, equidistantly placed for optimal harvesting. 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). A *smallholder oil palm plantation* must be typically smaller than 25 ha to be recognized as 'small' by the Indonesian government. These definitions vary by country with Malaysia using a 4 ha cut-off, while in Cameroon this varies from 8 to 40 ha (for an overview, see (Meijaard et al., 2018), Table 2). Compared to an industrial plantation, smallholder plantation tends to be less structured in shape and more heterogeneous in tree age. Smallholder plantations tend to form a landscape mosaic, composed of small plantations of varying shape and size, mixed with other types of land cover (e.g., idle land or other plantation types) (Figure 3d). When smallholder plantations form a large homogenous cluster, this cluster has a less dense trail network than industrial plantations (Figure 3a, c).

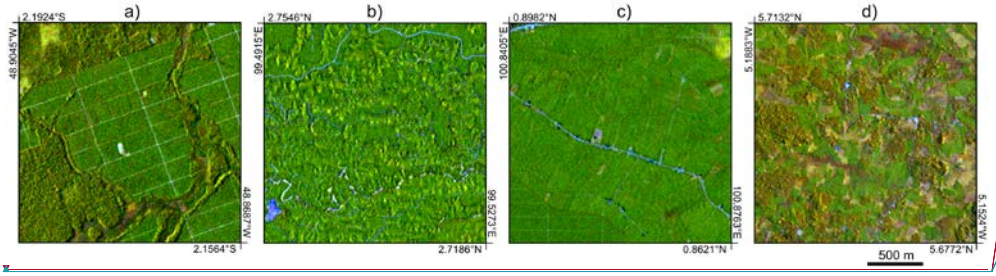


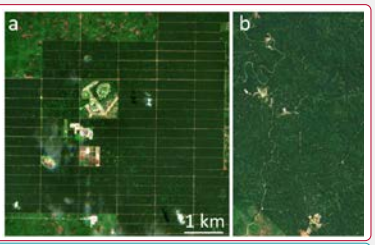
Figure 3. Examples of industrial and smallholder oil palm plantations seen by a 10-meter resolution Sentinel-1 and Sentinel-2 composite (R: VV, G: VH, and B: Band 4). The VV and VH bands were transformed and stretched so that the closed-canopy oil palm appears in green. a) An industrial plantation on a flat surface in Brazil, with harvesting trails built in straight lines and thus

Deleted: 3

Deleted: l

Deleted: the purposes of

Deleted: Meijaard et al. 2018



Deleted:

Formatted: Font color: Custom Color(RGB(36,41,46))

Deleted: imagery

Deleted: 10 m spatial resolution

Deleted: displayed in true colors

Deleted: Here

Deleted: ,

Deleted: dark

Deleted: , open-canopy oil palm has different shades of green

Deleted: Nigeria

forming rectilinear grids. b) An industrial plantation on hilly terrain in Indonesia, with curvy harvesting trails. c) Smallholder plantations forming a large homogeneous cluster in Indonesia. d) Smallholder plantations of varying shape, size, and tree age in Côte d'Ivoire.

2.6 Semantic segmentation

Image segmentation is the subfield of deep learning that aims to link each pixel of an image to a class label. Thus, semantic segmentation is the analog of the standard pixel-wise machine learning algorithms that are used in remote sensing for image classification (Ma et al., 2019). The difference is that semantic segmentation, as any model based on CNN, automatically learns and exploits the spatial patterns within the image by tuning the parameters of different convolutional operations.

This study employed the classification model DeepLabv3+ (Chen et al., 2017, 2018) with the MobileNetV2 (Sandler et al., 2018) as a backbone network. DeepLab has a series of versions for semantic segmentation. DeepLabv3+ is the latest version (link to the code in section 5. *Code availability*). The model uses an encoder-decoder architecture, in which the image is downsampled with max-pooling layers during the encoder part and spatial information is retrieved during the decoder part. A characteristic of DeepLabv3+ is that the CNN uses atrous convolutions, which enhances the field of view of filters to incorporate a larger spatial and informational context. The second-last layer of the CNN shows the probability that a pixel belongs to a certain class and the last operation of the CNN assigns the class with the maximum value in the probability layers, resulting in the final classification layer.

2.7 Validation

The accuracy of the global oil palm classification layer was evaluated with 10,816 reference points; 544 points were industrial plantations, 305 were smallholders, and 9,967 were other types of land uses. The points were randomly distributed using a simple random sampling, which means that each pixel in the map had an equal chance of being selected, and were distributed in the 100 x 100 km cells where the IUCN oil palm layer showed the presence of industrial plantations (cells outlined in red in Figure 2). This sample method led to a high imbalance between the points labeled as 'Other land uses' and the points labeled as oil palm, both industrial and smallholder, since oil palm plantations present a rare occurrence in the study area. The rare occurrence of oil palm implied that the probability of randomly selecting an oil palm plantation was also low. This low representation of oil palm plantations in the simple random sampling resulted in a high uncertainty in the oil palm area estimates at the regional and country level. For this reason, we included 2,679 points that were distributed with a stratified random sampling in order to achieve a minimum sample size in the industrial and smallholder oil palm classes. The size of each strata was 977 points in the class industrial oil palm, 802 in the class smallholder oil palm, and 900 were other types of land uses. The 2,679 stratified points were merged with the 10,816 simple random points, making a total of 13,495 points that were used to calculate the oil palm area estimates.

Deleted: s
Deleted: Southern Thailand
Deleted: analog
Deleted: For instance, CNN can automatically learn contextual information such as the road network in industrial plantations or the heterogeneity in landscape mosaics of smallholders.
Deleted: emloyed
Formatted: Font: Italic
Deleted: the image convolutions
Deleted: validation
Deleted: performed
Deleted: 13,252
Formatted: Not Highlight
Deleted: 865
Deleted: were
Deleted: 355
Deleted: 12,058
Deleted: (class 'Other')
Deleted:

Since the study aims to classify closed-canopy oil palm against other land uses, we included young oil palm and plantations that have not reached the full canopy coverage in the class "Other land uses". The points were visually interpreted using the Sentinel-1 and 2 annual composites of the year 2019 (See Subsection *2.3. Sentinel-1 and Sentinel-2 pre-processing*) and the DigitalGlobe orthoimages (< 1-meter spatial resolution) that are displayed as the base layer in the GEE code editor.

The accuracy metrics that we reported were the Overall Accuracy (OA), the user's accuracy (UA), and the producer's accuracy (PA) (Olofsson et al., 2014). The OA is the proportion of reference points that have been correctly classified and is calculated by summing the number of correctly classified points and dividing by the total number of points. The OA represents the probability that a randomly sampled pixel is correctly classified. The PA results from dividing the number of correctly classified points in each class by the number of visually interpreted points for each class. The PA is the complement of the omission error; PA = 100% - omission error. Thus, the PA for the classes "industrial" and "smallholder" is a relevant accuracy metric that shows the rate at which the oil palm plantations were missed in the classification image. On the other hand, the UA results from dividing the number of correctly classified points in each class by the number of points classified in each class. The UA is complement of the commission error; UA = 100% - commission error). The UA for the classes "industrial" and "smallholder" indicates the rate at which land uses have been incorrectly classified as oil palm plantations.

The accuracy metrics were evaluated following the good practices for estimating area and assessing accuracy reported by Olofsson et al., 2014. The practices explain the post-stratified estimation of the OA, PA, and UA, with a confidence interval. Olofsson et al., 2014 also describes the formulation for the area estimation for the classes that are present in the land cover map. The area estimates are also calculated with a confidence interval (here, we used a 95% confidence interval for both accuracy metrics and area estimates). Here, we used the term *area mapped* for the total area classified as a given class, and the term *area estimate* for the estimation of the actual area and the associated uncertainty following the practices in Olofsson et al., 2014. The area mapped is subject to the good accuracy of the classification; for instance, a high omission rate in the class 'industrial closed-canopy oil palm' would potentially lead to a low area mapped, which would represent an underestimate of the actual industrial oil palm area. The area estimate and its confidence interval, however, cover the actual area with a given confidence level.

Owing to the high imbalance in the validation dataset, we tested whether the overall accuracy of the CNN was higher than the no-information rate. The no-information rate was computed as the overall accuracy obtained if all pixels were classified as the major class, which is the class "Other land uses" in our study. The hypothesis test evaluates whether the overall accuracy obtained in the CNN classification is significantly higher than the no-information rate, with a 95% confidence level. If the null hypothesis is rejected (OA > no-information rate), we can assure that the CNN did better than predicting indiscriminately all pixels with the class "Other land uses".

Formatted: Space Before: 0 pt, After: 0 pt
Deleted: '
Deleted: Other'
Deleted: The points were randomly distributed
Deleted: and
Formatted: Font: Italic
Deleted: with high-resolution images and with Sentinel-1 and Sentinel-2 composites. 11,570 points were randomly distributed in the 100 x 100 km cells where the IUCN showed the presence of industrial plantations. The amount of grid cells with industrial plantations is 384 (Figure 2), 80 in America, 67 in Africa, and 237 in Southeast Asia and Pacific. The remaining 1,682 points were randomly distributed in a buffering area of 5 km around the industrial plantations of the IUCN layer.
Formatted: Space Before: 12 pt, After: 6 pt
Formatted: Not Highlight
Formatted: Not Highlight
Deleted: (
Deleted:)
Formatted: Font: Italic
Formatted: Font: Italic
Formatted: Highlight
Formatted: Highlight

2.7 Comparison with other oil palm datasets

The accuracy of the CNN classification was compared with existing oil palm maps of Sumatra for the year 2019 (Descals et al., 2019) and Southeast Asia for the year 2016 (Xu et al., 2020). Also, we compared our oil palm area estimates with the oil palm harvested area included in FAOSTATS at the country level, and with the area estimates obtained from an oil palm map developed in (Gaveau et al., 2021) over Indonesia for the year 2019. The oil palm maps in Descals et al., 2019 and Xu et al., 2020 were generated with a Random Forest Classification, while the map developed by (Gaveau et al., 2021) was generated by digitizing the oil palm plantations in Landsat and SPOT6 images.

In order to compare the current results with our previous study in Descals et al., 2019, we reclassified the young oil palm classes in this existing dataset to the class 'Other land uses'. We also kept only the validation points that cover Sumatra; this resulted in 2,463 points out of the 13,495 total points. For the comparison with Xu et al., 2020, we used our CNN model to classify Sentinel-1 and Sentinel-2 composites for the second half-year of 2016. Besides, we reclassified the smallholders and industrial plantations as a single class since the oil palm map in Xu et al. (2020) does not make distinctions between oil palm typology (industrial versus smallholder plantations). We also removed the validation points that were placed in young plantations because the temporal analysis in Xu et al. (2020) aimed to detect young oil palm and the plantations that have been clear-cut in the previous years (Figure 7). Note that Xu's dataset includes a 100-meter multi-year classification for the years 2001-2016 and we only compared the last year (2016) to ensure data availability in Sentinel-1 and 2 over the study area.

3 Results

The global map of industrial and smallholder plantations reveals the importance of high-resolution images (10 meters) for the accurate delimitation of smallholder plantations. Figure 4 shows the degree of detail of the classification image obtained with Sentinel-1 and Sentinel-2 composites. The figure also exemplifies the classification of industrial plantations, with the characteristic road network and the surrounding smallholder plantations. Supplementary Figure 4 shows examples of landscape types of oil palm plantations that were successfully detected and others that were omitted.

Formatted: Font color: Auto
Formatted: Heading 2
Moved (insertion) [1]
Deleted: We evaluated t
Deleted: results
Deleted: CNN
Deleted: with 13,252 randomly distributed points and
Deleted: them
Deleted: previous
Deleted: data sets in
Deleted: and in
Deleted: (D. Gaveau et al., 2021)(D. Gaveau et al., 2021)
Deleted: .
Deleted: (D. Gaveau et al., 2021)
Formatted: Font color: Auto, Not Highlight
Deleted: the
Deleted: (
Deleted:)
Deleted: .
Deleted: T
Deleted: in
Deleted:
Deleted: ¶
Deleted: Similarly to the image labeling, we used Sentinel-1 and 2 composites and the high-resolution images in Google Earth for the visual interpretation of the points.¶

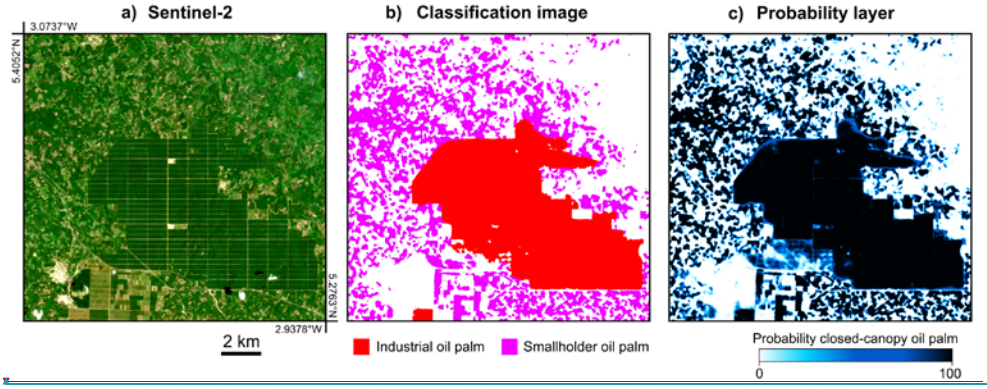
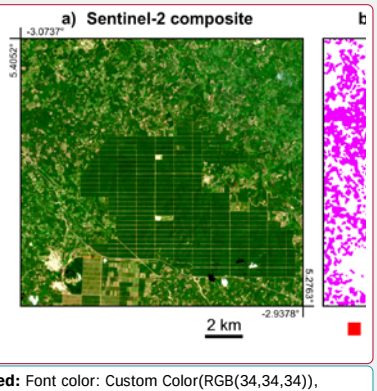


Figure 4. Example of the global oil palm layer in Côte d'Ivoire. Panel a) shows a Sentinel-2 true color image. Panel b) shows the resulting classification image obtained with a convolutional neural network (CNN). The classification image depicts an industrial plantation (red) surrounded by smallholder plantations (purple). The CNN learns contextual information such as the rectilinear road network in the industrial plantation, which is noticeable in the Sentinel-2 composite. Panel c) shows the probability of closed-canopy oil palm. The probability layer was generated from the second last layer of the CNN, which reflects the probability of each class.

We estimated the global area of planted closed-canopy oil palm at 21.00 ± 0.42 Mha, of which 15.26 ± 0.40 Mha (72.7%) was industrial plantations and 5.72 ± 0.22 Mha (27.3%) was smallholders. The map confirms that Southeast Asia is the highest producing region in the world (Figure 5) with a total surface area of 18.69 ± 0.33 Mha. It is followed by South America (0.91 ± 0.06 Mha), Western Africa (0.79 ± 0.11 Mha), Central America (0.52 ± 0.04 Mha), Central Africa (0.21 ± 0.6 Mha), and the Pacific (0.14 ± 0.00 Mha). Oil palm plantations were found in 49 tropical countries (See Supplementary Table 2). However, the estimated oil palm area varies greatly among countries, with Indonesia and Malaysia representing the bulk of the total surface area, while most other countries have a plantation area below 2 Mha (Figure 6).



Deleted:

Formatted: Font color: Custom Color(RGB(34,34,34)), Highlight

Deleted: Cote …ôte d'Ivoire. Panel a) shows a Sentinel-2 true color composite …

Formatted …

Deleted: 19.50 x 10^6 ha

Formatted …

Deleted: 13.10 x 10^6 ha…(67.2 …

Formatted: Not Highlight

Deleted: is

Formatted …

Deleted: 6.40 x 10^6 ha …32.8 …

Formatted: Not Highlight

Deleted: is

Formatted: Not Highlight

Deleted: 17.47 x 10^6 ha

Formatted …

Deleted: 0.82 x 10^6 ha

Formatted …

Deleted: 0.49 x 10^6 ha…, Western Africa (0.42 x 10^6 ha), …

Formatted: Not Highlight

Deleted: 0.16 x 10^6 ha

Formatted: Not Highlight

Deleted: 0.14 x 10^6 ha…. . …il palm plantations were found in 47 49 tropical countries (See Supplementary table …able 2). However, the estimated oil palm area varies greatly among countries, with Indonesia and Malaysia representing the bulk of the total surface area, while most other and with most of the …ountries with …ave a plantation area below 2 Mha 2 ha x 10^6 …

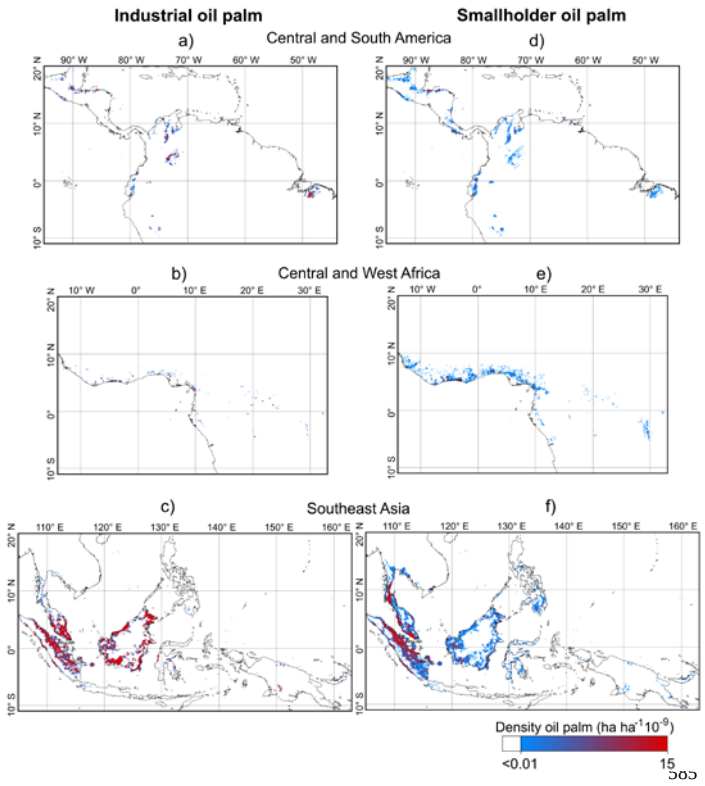


Figure 5. Density maps generated with the global oil palm layer. Panels a), b), and c) show the density maps of industrial oil palm plantations and panels d), e), and f) show the density maps for smallholder plantations. The maps have a spatial resolution of 10 km and represent the surface of closed-canopy oil palm, in hectares, in an area of 10^9 hectares. The values in the map were obtained by dividing the area of the oil palm within the 10-kilometer pixel by the total area covered in the pixel.

The region with the highest percentage of smallholder oil palm was West Africa (68.7 % of total plantings) (Supplementary figure 6). Elsewhere, the percentage of smallholders varied from 14.5% in Central Africa to 26.8% in the Pacific. As Figure 6 illustrates, however, countries in the same region might show different proportions of smallholders and industrial plantations. For instance, Thailand showed the highest proportion of smallholders (71.5%) which differed from the low ratio in neighboring Malaysia (15.4%). Countries in Southeast Asia also showed the highest oil palm surface per total land area,

Deleted: ¶
¶
¶
¶
¶
¶
¶
¶
¶
¶
¶
¶
¶
¶
¶
¶
¶
¶
...
Deleted: [6]
Formatted: Not Highlight
Deleted: is
Formatted: Not Highlight
Deleted: 1
Deleted: 5
Deleted: s
Formatted: Not Highlight
Deleted: 22
Deleted: 3
Formatted: Not Highlight
Deleted: in South America
Deleted: 36.5
Deleted: Central America
Deleted: As Figure
Deleted: s
Formatted: Not Highlight
Deleted: (95.1%)
Formatted: Not Highlight
Formatted: Pattern: Clear (Yellow), Not Highlight
Formatted: Not Highlight
Formatted: Not Highlight
Deleted: s
Formatted: Not Highlight
Deleted: neighboring
Deleted: 21
Deleted: 7
Formatted: Not Highlight

followed by smaller countries that allocate the majority of their cropland for oil palm production (Guatemala, Honduras, Costa Rica, and São Tomé and Príncipe).

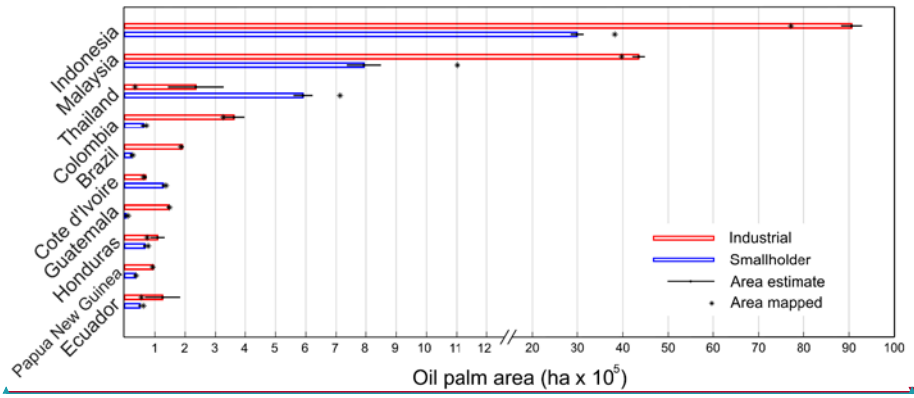


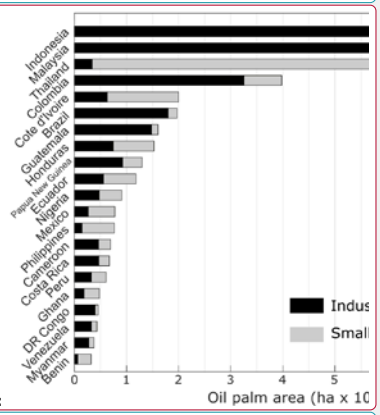
Figure 6. Oil palm plantation area per typology (industrial versus smallholder) for the second half-year of 2019 in the ten first countries with the largest oil palm area. The figure reflects the area mapped (asterisk mark), which resulted from the classification of Sentinel-1 and Sentinel-2, and the area estimate with a confidence level of 95%.

The accuracy metrics obtained with the 10,816 points show an OA of 98.52 ± 0.20% (Table 2) for the global oil palm map (Supplementary Table 3 shows the confusion matrix). This OA is significantly higher than the no-information rate (92.00 ± 0.51%) and, thus, we can assure that the CNN classification did better than assigning the major class to all the validation points. The UA and PA were lower in industrial and smallholder plantations than the same accuracies obtained in the class 'Other land uses'. Smallholder plantations showed the lowest UA (76.56 ± 4.53%), while the industrial plantations showed the lowest PA (75.78 ± 3.55%). The UA and PA accuracies were lower when evaluated only in Sumatra (smallholder UA = 63.27 ± 7.82% and industrial PA = 69.15 ± 4.62%). However, these accuracies were considerably lower in Descals et al., 2019, which presented a UA = 45.85 ± 6.84% for smallholders and PA = 54.26 ± 4.42% for industrial plantations. The state-of-the-art methodology using CNN also showed a higher overall accuracy than the Random Forest classification for the case study in Sumatra (91.31 ± 0.97% compared to the 94.02 ± 0.89% in the current study).

Table 2. Accuracy assessment of the global oil palm layer for the second half of 2019 and comparison of the global layer with the results of a previous study (Descals et al., 2019), which used a Random Forest in Sumatra for the same year. The accuracy metrics

Deleted: (Supplementary figure 6), …ollowed by smaller countries which …

Formatted: Font: 9 pt, Bold, Font color: Custom Color(RGB(34,34,34)), Highlight



Deleted:

Formatted: Font color: Custom Color(RGB(34,34,34)), Highlight

Deleted: surface …rea per country and …ypology (industrial versus smallholder) for the second half-year of 2019 in the ten first countries with the largest oil palm area. The area was estimated using the global oil palm layer that resulted from the classification of Sentinel-1 and Sentinel-2 composite at 10-meter resolution.…The figure reflects the area mapped (asterisk …

Moved up [1]: We evaluated the results of the CNN with 13,252 distributed points and compared them with previous data sets in Sumatra for the year 2019 (Descals et al., 2019) and in Southeast Asia for the year 2016 (Xu et al., 2020). In order to compare the current results with the previous study (Descals et al., 2019), we kept

Deleted: ¶

Table 1. Accuracy assessment of the global oil palm layer for the second half of 2019 and comparison of the global layer with the results of a previous study (Descals et al., 2019) in Sumatra for the same year. The accuracy metrics of the global layer were …

Formatted Table

Deleted: 13,252…0,816 points show an OA of 987…5…2 ± 0.20% and kappa of 84.9% …Table 1…) for the global oil palm map (Supplementary Table 3 shows the confusion matrix). This OA is significantly higher than the no-information rate (92.00 ± 0.51%) and, thus, we can assure that the CNN classification did better tha …

Formatted: Font: 10 pt

Deleted: 3… and smallholder …ndustrial PA = 69.15 ± 4.6278.8…). However, these accuracies are …ere considerably lower in Descals et al., 2019, which presents …resented a PA …A = 45.859…± 6.84% % …or smallholders and UA …A = 58…4.26 ± 4.428… for industrial plantations. Overall, t…he state-of-the-art …

of the global layer were estimated with 10,816 points randomly distributed in the main oil palm producing areas in the world, while the comparison used only the validation points that were located in Sumatra (2,463 points). The reported metrics are the Overall Accuracy (OA), the user's accuracy (UA), and the producer's accuracy (PA). The accuracy metrics are reported with a confidence interval (95% confidence level).

		Global OP	Global OP (Sumatra)	Descals et al., 2019 (Sumatra)
OA (%)		98.52 (99.42, 99.61)	94.02 (93.13, 94.91)	91.31 (90.34, 92.28)
UA (%)	Other	99.19 (99.01, 99.36)	97.00 (96.27, 97.73)	96.97 (96.24, 97.71)
	Industrial	88.22 (85.49, 90.96)	89.25 (85.10, 93.40)	88.70 (84.04, 93.36)
	Smallholder	86.92 (81.80, 92.04)	63.27 (55.47, 71.06)	45.85 (39.03, 52.67)
PA (%)	Other	99.52 (99.42, 99.61)	97.99 (97.41, 98.57)	96.59 (95.86, 97.31)
	Industrial	75.78 (72.23, 79.33)	69.15 (64.54, 73.77)	54.26 (49.83, 58.68)
	Smallholder	84.94 (81.36, 88.51)	81.44 (75.26, 87.63)	83.30 (77.47, 89.13)

Our results (OA = 96.59 ± 0.50%) performed better than Xu's classification image (OA = 91.35 ± 0.69%) for 2016 (Supplementary Table 4). The producer's accuracy for industrial plantations in Xu's results (PA = 76.41 ± 3.08%) is higher than our results (PA = 73.65 ± 2.94%), although this difference is not significant for a confidence level of 95%. The main difference between the data sets, however, was found in the user's accuracy for smallholders, in which our results excelled (UA = 96.60 ± 0.51% compared to 57.36 ± 3.76% in Xu's data set). The comparison with Xu's data, however, only reflects the accuracies for closed-canopy oil palm plantations; the multi-annual analysis in Xu et al., 2020 also included the detection of disturbances in the time series to classify young plantations. Similar to Xu's data, the dataset produced in (Gaveau et al., 2021) also mapped young oil palm and areas that were clear-cut for oil palm plantation in Indonesia. For this reason, our closed-canopy oil palm area estimate was 12.05 ± 0.23 Mha in Indonesia (area mapped was 11.54 Mha, with 7.71 Mha (66.8%) industrial and 3.83 Mha (33.2%) smallholder) but, by comparison, Gaveau found a higher oil palm area for Indonesia for the same year: 16.26 Mha. Despite this difference, Gaveau found a similar ratio between industrial and smallholder plantation extent: 10.33 Mha industrial (64%) and 5.93 Mha smallholder (36%).

Deleted: For the comparison with Xu et al. (2020), we reclassified the smallholders and industrial plantations as a single class. We also removed the validation points that were placed in young plantations, and classified as 'Other', because the temporal analysis in Xu et al. (2020) also aims to detect young oil palm and the plantations that have been clear-cut in the previous years (Figure 7). Note that Xu's dataset includes a 100-meter multi-year classification and we only compared the last year (2016) to ensure data availability in Sentinel-1 over the study area.

Deleted: 6

Deleted: and kappa = 77.4%

Deleted: 92.2

Deleted: and kappa = 61.2%

Deleted: omission

Deleted: error

Deleted: 3

Deleted: are

Deleted: lower

Deleted: 67.1

Deleted: However, t

Deleted: 4

Deleted: ¶

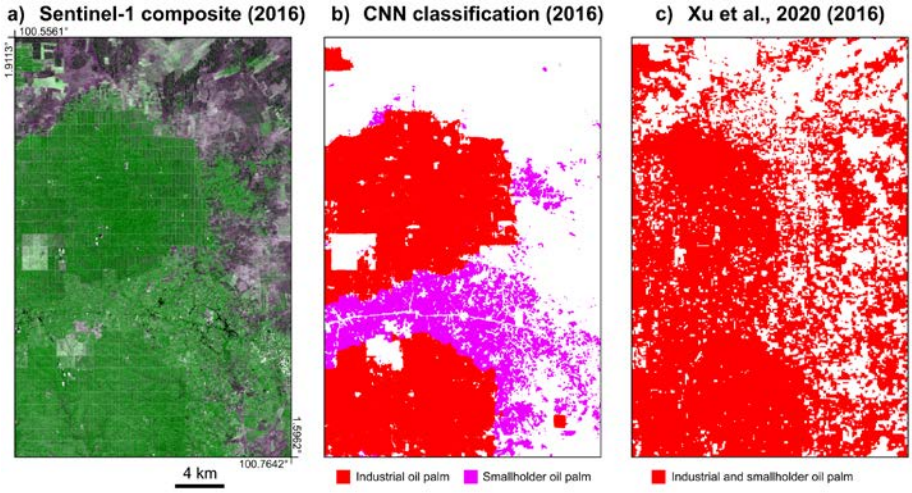


Figure 7. Comparison of the classification image obtained with the convolutional neural network (CNN) and the last year of the multiannual analysis presented in Xu et al., 2020. Panel a) shows a Sentinel-1 composite (VV-VH-VV) for the second half-year of 2016 in Riau province (Indonesia). The VV and VH bands were transformed and stretched so that closed-canopy oil palm appears in green. Panel b) shows the classification image that results from the CNN using the Sentinel-1 and Sentinel-2 composites for 2016. Panel c) shows the oil palm layer presented in Xu et al., 2020 for the year 2016.

The comparison with inventories from FAOSTAT also evidences a large omission of oil palm plantations in West Africa (Supplementary figure 7). The total surface reported as harvested area in FAOSTAT is 4.16 Mha in Western Africa, while our oil palm area estimate was 0.79 ± 0.11 Mha and the area mapped 0.42 Mha. The country with the highest difference is Nigeria, with an area estimate of 3.02 Mha reported by FAOSTAT that contrasts with the 0.01 Mha classified by the CNN and the 0.25 ± 0.07 Mha total closed-canopy oil palm area estimate.

4 Discussion

The results confirm previous findings on the suitability of radar satellite data for mapping closed-canopy oil palm plantations at the regional scale (Miettinen & Liew, 2011) and the improved accuracies obtained with the combined use of radar and optical data for mapping smallholder and industrial oil palm plantations (Descals et al., 2019). Our study further shows that these plantations can be mapped globally and by typology at high spatial resolution (10 meters). The results obtained with the CNN outperformed previous studies and provide evidence that deep learning is more suitable than standard machine learning algorithms, such as Random Forests, when contextual information is required for class prediction. Overall, the

Deleted: Comparison with Xu's data requires a degree of caution because it only reflects the accuracies for closed-canopy oil palm plantations, while the multi-annual analysis in Xu et al., 2020 aimed to detect disturbances in the time series in order to classify young plantations. This is also evidenced if we compare the global oil palm layer for 2019 with the data set in (Gaveau et al., n.d.), which maps oil palm and areas that have been clear-cut for oil palm plantation in Indonesia. We estimate that 11.54 x 10^6 ha of closed-canopy oil palm plantations existed in Indonesia in 2019, with 7.71 x 10^6 ha (66.8%) Industrial and 3.83 x 10^6 ha (33.2%) smallholder. By comparison, Gaveau et al. (forthcoming) found a higher oil palm area for Indonesia for the same year: 16.26 x 10^6 ha, but a similar ratio between industrial and smallholder plantation extent: 10.33 x 10^6 ha industrial (64%) and 5.93 x 10^6 ha smallholder (36%). ¶

Deleted: 7

Deleted: x 10^6

Deleted:

Deleted: the total area estimated with the global

Deleted: layer

Deleted: is

Deleted: 0.42 x 10^6 ha

Deleted: x 10^6

Deleted:

Deleted: x 10^6

Deleted: Our results also report more plantation area in Indonesia (11.54 x 10^6) than the reported by FAOSTAT (6.78 x 10^6). Overall, there is high agreement between the CNN data and FAOSTAT data for other countries and the difference in oil palm area can be attributed to the different years used in the comparison; we used the last year of the FAOSTAT data (2018) while the global oil palm layer was generated for 2019.

Deleted: .

results show a high accuracy (OA = 97.46 ± 0.26%). The accuracy assessment and the comparison with other products also evidences uncertainties associated to the oil palm definition: young plantations, plantations that have open canopies, and plantations mixed with non-palm tree species, such as semi-wild oil plantations in Africa.

We compared our findings with three studies, Descals et al. (2019), Xu et al. (2020), and the dataset developed by (Gaveau et al., 2021). Our CNN model applied to Sentinel-1 and Sentinel-2 classified closed-canopy oil palm stands with higher detail (10-meter spatial resolution) than existing datasets. Although, at coarser resolution (100-meter spatial resolution), the temporal analysis used in Xu et al., 2020 aimed to detect disturbances (e.g., land clearing and burning) that may or may not result in the development of oil palm plantations. Thus, Xu et al. classified opened-canopy plantations that remained undetected in our classification. Accordingly, the omission error for oil palm was lower in Xu's case, although this difference was not significant. However, Xu et al. detected much more than oil palm plantation, including scrubs and grasslands and, therefore, commission error for oil palm was significantly higher in Xu's map (UA = 57.36 ± 3.76 % compared to 96.55 ± 1.92 % in our study).

Gaveau et al., 2021 did not directly measure planted areas, but instead, they identified areas that were "cleared to develop plantations". An area may have been cleared for oil palm, left idle because of several constraints, or the area may have been planted but the plantation may have failed. The comparison with Xu's and Gaveau's maps evidences an important shortcoming with our method; the classification of oil palm with radar data can only detect closed-canopy oil palm stands and, thus, excludes areas cleared for oil palm that has been left idle or where oil palm trees died. Besides, oil palm must be at least 3-years-old (Descals et al. 2019) to reach the full canopy closure. Therefore, it is likely that our maps missed young oil palm plantations developed after 2016. The dataset developed in Gaveau et al., 2021 is more suited to verify the impacts of the oil palm industry on forests, while our method is more suited to map the productive planted area, i.e. closed-canopy oil palm stands > 3 years old. In contrast, Gaveau's dataset was produced mostly by visual interpretation and manual delimitation of oil palm development while our method consisted of a supervised learning algorithm; our trained CNN can automatically classify remotely-sensed data into oil palm maps for future land cover monitoring.

Since our method classified only closed-canopy oil palm, it also struggled to detect oil palm in non-homogeneous settings (e.g., oil palm mixed with other crops), plantation with low canopy coverage, and naturally occurring and semi-wild oil palm trees, known as feral oil palm, that are present in Africa. These semi-wild oil palm plantations explain the large difference with the harvested area reported by FAOSTAT in West Africa. This means that our global estimate of total planted oil palm areas (21.00 ± 0.42 Mha) is an underestimate, which considers only closed-canopy oil palm plantations. It is difficult to say by how much we underestimate the total planted area if considering young, non-homogeneous settings, and sparse oil palm plantations, but assuming constant planting rates and an average age of 25 years of palms until replanting, we could miss 3/25 = 12% of the total planted area.

Deleted: Overall Accuracy
Deleted: 97.5
Deleted: and kappa = 84.9%
Deleted: shortcomings primarily in the classification of smallholder plantations,
Deleted: Open-source deep learning libraries, such as TensorFlow or PyTorch, are democratizing deep learning to the broader scientific community. The model is planned to be used for follow-up monitoring once a year to generate global oil palm maps. The ...
Moved down [4]: In this study, only 296 images of 1000 x 1000
Moved (insertion) [4]
Deleted: two
Formatted: Not Highlight
Deleted: Gaveau et al. (forthcoming)
Formatted: Not Highlight
Formatted: Highlight
Formatted: Highlight
Deleted: 100m
Deleted: which
Deleted: (PA = 76.3% vs PA = 67.1% in our estimate for 2016)
Deleted: , UA = 57.4% compared to 96.6% in our classification
Deleted: ¶
Deleted: Gaveau
Deleted: et al. (forthcoming) used a combination of Landsat an ...
Deleted: have been
Deleted: a number of
Deleted: This
Deleted: have
Deleted: In addition
Deleted: closed-canopy
Deleted: smallholder
Deleted:
Deleted: Gaveau et al.
Deleted: Contextual to these insights, we note several caveats ...
Deleted: struggles
Deleted: . This means that our global estimate of total planted o ...
Moved (insertion) [5]
Deleted: that
Deleted: . Additionally, there are palms in agroforestry type ...
Moved up [5]: Additionally, there are palms in agroforestry type
Formatted: Not Highlight

Despite the caveats regarding the total areas of planted oil palm, our findings of the ratio between smallholders and industrial-type plantings are still relevant, as both types of oil palm are similarly affected by omission in our analysis (except the oil palm we miss in agroforestry type settings, which is mostly smallholders). Evidence of this is that the ratios reported by (Gaveau et al., 2021) ~~are similar to our estimates in Indonesia. Globally, our data indicate that 72.7% of planted oil palm is under industrial-scale management and 27.3% managed by smallholders. These percentages diverge from the commonly~~ stated claim that 40% of the palm oil produced globally is from smallholders (Meijaard et al., 2018). Not only is the land managed in smallholder-type settings less than 40%, but there is also a significant yield gap between industrial-scale and smallholder-scale operators. Smallholder yields are often 40% or lower than yields in industrially-managed plantations (Woittiez et al., 2017), which suggests that the overall contribution of smallholders to global palm oil production is about 18% rather than 40%. Industrial-scale operators thus appear to produce about 82% of the global palm oil. We note that this excludes the locally produced palm oil in agroforestry-type settings in the African tropics, where oil palm is traditionally produced for local consumption.

Our findings on the ratio between smallholder and industrial-scale oil palm are different from those reported by various governments. For example, the government of Indonesia estimates that 40.8% of the country's oil palm planted area is developed under smallholder licenses, whereas our analysis of the typical characteristics of planted crops indicate that this ratio is 66.8% industrial and 33.2% smallholder for the country. To qualify as a 'smallholder farmer' in Indonesia, according to the government, farms must be less than 25 ha. Those that cultivate less than 25 ha of oil palm are required to apply for a Plantation Registration Certificate (STD-B), while those producers cultivating more than 25 ha require a Plantation Business License (IUP-B) (Jelsma et al., 2017). The latter involves more complex procedures and regulatory requirements such as an environmental impact assessment (Paoli et al., 2013). Those with an STD-B are exempted from most of these requirements (Jelsma et al. 2017). This creates an incentive for producers to classify their plantations as non-industrial scale because the paperwork and licensing involve fewer hurdles. This mismatch between land occupancy (de facto) and legal allocation (de jure) was also noted by (Gaveau et al., 2017) in Sumatra who noted unregistered medium-sized landowners operating like companies in terms of their approach to oil palm development, but without formal company status. Missing young plantations cannot explain the large difference we noticed in Indonesia between our planted area estimate and FAO harvested area because expansion has gone down in recent years (Gaveau et al., 2019). In important producing regions (Riau, Sumatra), only 15% of all agricultural land parcels have a national-level registration and 26% of all oil palm plantations were only registered at the village level (Meijaard et al., 2018). Unregistered plantations explain why we found more plantations than FAOSTAT. Discrepancies between this study's findings and those of various governments on the ratio between smallholder and industrial-scale oil palm could result from underestimations by authorities as identified by Oon and colleagues (2019). As with Indonesia, it indicates how difficult it is to accurately map smallholder oil palm because of the heterogeneous characteristics of this land use, the lack of legal registration of smallholder lands, and potentially vested interests in running large-scale operations under smallholder type licenses (Supplementary figure 8).

Deleted: about

Deleted: An e

Deleted: Gaveau et al. (forthcoming)

Deleted: 67.2

Deleted: 32.8

Deleted: s

Deleted: operate

Deleted: O

Deleted: (Meijaard et al., 2018)

Deleted: FAO

Deleted: ¶
In Malaysia, we also found a discrepancy between what is reported by the government and our data. Malaysia reports that smallholder farms covered about 17% of the total area of oil palm cultivation in this country in 2018 (Senawi et al., 2019) as opposed to 21.7% identified in our dataset. The difference may relate to types of smallholders, with "organized smallholders" (as opposed to independent) operating in schemes that look like industrial-scale operation on imagery.

Deleted: industrial

Deleted: of

Deleted: -

The CNN model trained for the year 2019 is planned to be used for follow-up monitoring once a year to generate global oil palm maps. The shortcomings of deep learning include the high computational cost for training the models and the high cost for gathering labeled data compared to the standard machine learning algorithms commonly used in remote sensing, such as Random Forest. In this study, 296 images of 1000 x 1000 pixels were used as a training data set, consisting of 200 labeled images and 96 augmented images, and the computing time for training a pre-trained DeepLabv3+ was nearly 8 days with an office computer. Despite this, the computing time and the size of our training dataset were considerably lower than state-of-the-art deep learning studies in computer vision (i.e. more than 200,000 labeled images in the Common Objects in Context (COCO) dataset) in which the number of classes and complexity of the classification problem surpasses the current study.

Formatted: Font: Not Italic, Font color: Custom Color(RGB(34,34,34))

5 Code availability

The code that generates the Sentinel-1 and Sentinel-2 composites can be found at:
https://github.com/adriadescals/oil_palm_global

Deleted: :

Deleted: https://github.com/adriadescals/oil_palm_global

Formatted: No underline, Font color: Blue, Not Highlight

The original code of the semantic segmentation model DeepLabv3+ can be found at:
https://github.com/tensorflow/models/tree/master/research/deeplab

6 Data availability

Deleted: ¶

The dataset presented in this study is freely available for download at https://doi.org/10.5281/zenodo.4473715 (Descals et al., 2021). The dataset contains 634 100x100 km tiles, covering areas where oil palm plantations were detected. The file '*grid.shp*' contains the grid that covers the potential distribution of oil palm. The file '*grid_withOP.shp*' shows the 100x100 grid squares with presence of oil palm plantations. The classified images ('*oil_palm_map*' folder, in GeoTIFF format) are the output of the convolutional neural network based on Sentinel-1 and Sentinel-2 half-year composites. The images have a spatial resolution of 10 meters and contain three classes: [1] Industrial closed-canopy oil palm plantations, [2] Smallholder closed-canopy oil palm plantations, and [3] other land covers/uses that are not closed canopy oil palm. The file '*Validation_points_GlobalOilPalmLayer_2019.shp*' includes the 13,495 points that were used to validate the product. Each point includes the attribute 'Class', which is the labelled class assigned by visual interpretation, and the attribute 'predClass, which reflects the predicted class by the convolutional neural network. The 'Class' and 'predClass' values are the same as the raster files: [1] Industrial closed-canopy oil palm plantations, [2] Smallholder closed-canopy oil palm plantations, and [3] other land covers/uses that are not closed canopy oil palm.

Deleted: https://doi.org/10.5281/zenodo.3884602 (Descals et al., 2020)

Deleted: 622

Deleted: geotiff

Deleted: ,252

Deleted: The values of the 'Class' attribute are [0] other land covers/uses that are not closed canopy oil palm, [1] Industrial closed-canopy oil palm plantations, [2] Smallholder closed-canopy oil palm plantations. The 'predClass' values are identical to the 'Class' values.

The data can be visualized online at the BIOPAMA application portal: https://apps.biopama.org/oilpalm/. The BIOPAMA application portal also includes the probability layer, which shows the probability (from 0 to 100) that a pixel is closed-canopy oil palm plantation. The data can also be visualized at the Google Earth Engine (GEE) experimental app:

Deleted: https://apps.biopama.org/oilpalm/

Deleted:

https://adriadescals.users.earthengine.app/view/global-oil-palm-map-2019. The global oil palm map is hosted in GEE as an *Image Collection*: https://code.earthengine.google.com/?asset=users/adriadescals/shared/OP/global_oil_palm_map_v1.

The Sentinel-1 SAR GRD and Sentinel-2 Level-2A used in this study (scenes taken in the second half year of 2019 in the tropics, and second half year of 2016 in Sumatra) are available at https://scihub.copernicus.eu/ and can be retrieved in GEE. When using GEE, the Sentinel-1 and 2 data are hosted and accessed in the Earth Engine data catalog (the links to the data are https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S1_GRD and https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR respectively). Despite the data are hosted by GEE, these satellite data are the same as accessed via the official portal (Copernicus Open Access Hub: https://scihub.copernicus.eu/); Data ingested and hosted in GEE are always maintained in their original projection, resolution, and bit depth (Gorelick et al., 2017).

The WorldClim bioclimatic variables (WorldClim V1 Bioclim) (Hijmans et al., 2005) were also accessed through GEE: https://developers.google.com/earth-engine/datasets/catalog/WORLDCLIM_V1_BIO. The data can be accessed in the official portal at https://www.worldclim.org/data/v1.4/worldclim14.html.

The IUCN industrial oil palm layer (Meijaard et al., 2018) can be found at https://doi.org/10.5061/dryad.ghx3ffbn9. The oil palm layer of Indonesia and Malaysia for the year 2016 (Xu et al., 2020) can be found at https://doi.org/10.5281/zenodo.3467071. The oil palm layer of Sumatra for the year 2019, developed with the same methodology as in Descals et al., 2019, for Riau province (Indonesia), is hosted as a GEE asset at https://code.earthengine.google.com/?asset=users/adriadescals/shared/Sumatra_oilPalm_L2_Descals_et_al_2019.

Very-high-resolution images (spatial resolution < 1 meter) from DigitalGlobe can be visualized in the GEE code editor (i.e. https://code.earthengine.google.com/7b35908829cdc62617138a9759eb370e) or Google Maps (i.e. https://www.google.com/maps/@-3.969372,105.048514,782m/data=!3m1!1e3)

The country-wide harvested area of oil palm was extracted from the FAOSTAT database (accessed on 10 Jun 2020): http://www.fao.org/faostat/en/.

7 Conclusions

This study presents the first global map of oil palm plantations, for the year 2019, derived from remotely-sensed data with a spatial resolution of 10 meters. We classified Sentinel-1 and Sentinel-2 data into a map that discriminates between smallholders and industrial oil palm plantations. We obtained high accuracies, with user's and consumer's accuracy generally above 80%, thanks to the use of cutting-edge deep learning algorithms. The method is deployable and can generate yearly maps for oil palm monitoring in a cloud processing environment and based on freely available satellite imagery.

Formatted: Font: Italic
Formatted: Font: Italic
Deleted:
Formatted: Not Highlight
Deleted:
Deleted: Google Earth Engine (see Code availability section)
Deleted: ADDIN ZOTERO_ITEM CSL_CITATION {"citationID":"XTNc3pRR","properties":{"formattedCitation":"(Xu et al., 2020)","plainCitation":"(Xu et al., 2020)","noteIndex":0},"citationItems":[{"id":"J0Qyq45D/Pkahj8jL","uris":["http://zotero.org/users/3744214/items/XGWPFK8L"],"uri":["http://zotero.org/users/3744214/items/XGWPFK8L"],"itemData":{"id":"J0Qyq45D/Pkahj8jL","type":"article-journal","container-title":"Earth System Science Data","issue":"2","page":"847–847","title":"Annual oil palm plantation maps in Malaysia and Indonesia from 2001 to 2016.","volume":"12","author":[{"family":"Xu","given":"Yidi"},{"family":"Yu","given":"Le"},{"family":"Li","given":"Wei"},{"family":"Ciais","given":"Philippe"},{"family":"Cheng","given":"Yuqi"},{"family":"Gong","given":"Peng"}],"issued":{"date-parts":[["2020"]]}}}],"schema":"https://github.com/citation-style-language/schema/raw/master/csl-citation.json"}
Formatted: Font: (Default) +Headings (Times New Roman)
Formatted: Font: (Default) +Headings (Times New Roman), 10 pt
Formatted: Font: (Default) +Headings (Times New Roman), 10 pt
Formatted: Font: (Default) +Headings (Times New Roman)
Deleted: (
Deleted:)
Moved (insertion) [2]
Deleted: (FAO, 2020)¶
Formatted: Font: (Default) +Headings (Times New Roman)
Moved up [2]: (FAO, 2020)
Deleted: (FAO, n.d.)¶

Our global oil palm map makes an important contribution to the palm oil debate. It will be useful to solve or at least clarify a range of social and environmental debates. We know that oil palm plantations are a major cause of deforestation in Indonesia and Malaysia (Austin et al., 2019; Gaveau et al., 2019), but the share of oil ~~palm-~~driven deforestation to global tropical forest loss is not known. This map will help to inform the debate on oil palm driven deforestation globally. Forest clearing for oil palm is associated with negative socio-economic impacts on forest-dependent communities (Santika et al., 2019). This map validates a novel approach to mapping where oil palm is grown by smallholders who generate direct income or consumption from their own plantations, as opposed to industrial-scale oil palm (or industrial-scale plantings disguised as smallholder) where plantations provide labo~~r opportunity but profits primarily attribute to company owners and the government~~ (through taxes). The data can thus guide better planning for maximizing socio-economic benefits from oil palm. The global oil palm layer also assists in the discussion about environmental impacts of oil palm, including on biodiversity (Fitzherbert et al., 2008; Meijaard et al., 2018) and regional climate (McAlpine et al., 2018). These negative impacts are real but need to be considered in the light of meeting the global demand for vegetable oil through optimal allocation of land not just to oil palm, but to all major oil-producing crops. This requires high-resolution spatial data for all oil seed crops (Meijaard, Abrams, et al., 2020~~) so that informed decisions can be made about land use based on yield differences, past environmental and social~~ impacts of different crops and the different characteristics of oils from different crops and their particular end uses. Finally, and relevant to the current COVID-19 pandemic, our global map can help localize areas where zoonotic diseases can originate from, especially in areas where oil palm expansion was associated with recent deforestation. Such insights are essential for the health of people and the economy (Wardeh et al., 2020).

Supplement. Supplementary material related to this article is available online at: https://doi.org/10.5194/essd-XX-XXXX-XXXX-supplement.

Author contributions. The concept for this work originated by ZS during scientific project discussions with SW, EM and SP for the goal of providing the scientific and policy-making community with up-to-date and repeatable monitoring capabilities of oil palm plantations. AD and ZS designed the study, AD collected the training data, and AD and SW collected the validation points. AD implemented the data processing workflow. AD, ZS, SW, EM, DG and SP were involved in the analysis, writing and the revision of the manuscript, and AD generated the figures and tables.

Competing interests. Adrià Descals performed remote-sensing consulting work for the ICT & GIS department at PT Austindo Nusantara Jaya Tbk. Serge Wich received research funding from PT Austindo Nusantara Jaya Tbk and is a member of the IUCN Oil Palm Task Force and he has done oil palm related work for this task force. Erik Meijaard chairs and has received funding from the IUCN Oil Palm Task Force and he has done work paid by palm oil companies and the

Deleted: palm

Deleted: u

Deleted: ,

Deleted: a

Deleted: policy

Deleted:

Deleted:

Deleted: and the

Deleted: The authors declare that they have no conflict of interest

Roundtable on Sustainable Palm Oil. David Gaveau is a member of the IUCN Oil Palm Task Force, a group tasked by the IUCN to investigate the sustainability of palm oil and he has done oil palm related work for this task force, and Greenpeace.

Deleted: .

Disclaimer. All features and data are provided "as is" with no warranties of any kind. The views expressed are purely those of the writers and may not in any circumstance be regarded as stating an official position of the European Commission.

Acknowledgements. The authors would like to thank the BIOPAMA (European Commission, Joint Research Centre) technical developer team members' effort with the online visualization tool (L. Battistella, M. Boni, J. Davy). The authors also thank Mr. A. McKinnon (Copernicus Emergency Management Service and Global Land Service – Communication Manager) for proofreading the paper.

Financial support. This study was supported by the Biodiversity and Protected Areas Management (BIOPAMA) programme, an initiative of the African, Caribbean, and Pacific Group of States financed by the 10th and 11th European Development Funds of the European Union and co-managed by the European Commission Joint Research Centre and The International Union for Conservation of Nature.

Review statement. This paper was edited by Dr. David Carlson and reviewed by XX anonymous referees.

References

Austin, K. G., Schwantes, A., Gu, Y., & Kasibhatla, P. S. (2019). What causes deforestation in Indonesia? *Environmental Research Letters*, *14*(2), 024007.

Bronkhorst, E., Cavallo, E., van Dorth tot Medler, M., Klinghammer, S., Smit, H. H., Gijsenbergh, A., & van der Laan, C. (2017). *Current practices and innovations in smallholder palm oil finance in Indonesia and Malaysia: Long-term financing solutions to promote sustainable supply chains* (Vol. 177). CIFOR.

Byerlee, D., Falcon, W. P., & Naylor, R. (2017). *The tropical oil crop revolution: Food, feed, fuel, and forests*. Oxford University Press.

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2017). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(4), 834–848.

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. *Proceedings of the European Conference on Computer Vision (ECCV)*, 801–818.

Descals, A., Szantoi, Z., Meijaard, E., Sutikno, H., Rindanata, G., & Wich, S. (2019). Oil Palm (Elaeis guineensis) Mapping with Details: Smallholder versus Industrial Plantations and their Extent in Riau, Sumatra. *Remote Sensing*, *11*(21), 2590.

Descals, A., Wich, S., Meijaard, E., Gaveau, D., Peedell, S., & Szantoi, Z. (2021). *High resolution global industrial and smallholder oil palm map for 2019* (Version v1) [Map]. Zenodo. https://doi.org/10.5281/zenodo.4473715

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., & others. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, *120*, 25–36.

FAO. (2020). *FAOSTAT statistical database*. http://www.fao.org/faostat/en/

Fitzherbert, E. B., Struebig, M. J., Morel, A., Danielsen, F., Brühl, C. A., Donald, P. F., & Phalan, B. (2008). How will oil palm expansion affect biodiversity? *Trends in Ecology & Evolution*, *23*(10), 538–545.

Gaveau, D., Locatelli, B., Salim, M. A., Yaen, H., Pacheco, P., & Sheil, D. (2019). Rise and fall of forest loss and industrial plantations in Borneo (2000–2017). *Conservation Letters*, *12*(3), e12622.

Gaveau, D., Locatelli, B., Salim, M., Husnayaen, H., Manurung, T., Descals, A., Angelsen, A., Meijaard, E., & Sheil, D. (2021). Slowing deforestation in Indonesia follows declining oil palm expansion and lower oil prices. *In Review*. https://doi.org/10.21203/rs.3.rs-143515/v1

Gaveau, D., Pirard, R., Salim, M. A., Tonoto, P., Yaen, H., Parks, S. A., & Carmenta, R. (2017). Overlapping land claims limit the use of satellites to monitor no-deforestation commitments and no-burning compliance. *Conservation Letters*, *10*(2), 257–264.

Glenday, S., & Gary, P. (2015). Indonesian oil palm smallholder farmers: A typology of organizational models, needs, and investment opportunities. *Bogor (ID): Daemeter Consulting*.

Google Earth Engine. (2020). *DigitalGlobe imagery* [Map]. https://code.earthengine.google.com

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *25*(15), 1965–1978.

IFPRI. (2019). *Global Spatially-Disaggregated Crop Production Statistics Data for 2010, Version 1.0*. Harvard Dataverse.

Jelsma, I., Schoneveld, G. C., Zoomers, A., & Van Westen, A. (2017). Unpacking Indonesia's independent oil palm smallholders: An actor-disaggregated approach to identifying environmental and social performance challenges. *Land Use Policy*, *69*, 281–297.

Koh, L. P., Miettinen, J., Liew, S. C., & Ghazoul, J. (2011). Remotely sensed evidence of tropical peatland conversion to oil palm. *Proceedings of the National Academy of Sciences*, *108*(12), 5127–5132. https://doi.org/10.1073/pnas.1018776108

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.

Lee, J. S. H., Wich, S., Widayati, A., & Koh, L. P. (2016). Detecting industrial oil palm plantations on Landsat images with Google Earth Engine. *Remote Sensing Applications: Society and Environment*, *4*, 219–224.

Li, W., Fu, H., Yu, L., & Cracknell, A. (2017). Deep learning based oil palm tree detection and counting for high-resolution remote sensing images. *Remote Sensing*, *9*(1), 22.

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *152*, 166–177.

McAlpine, C. A., Johnson, A., Salazar, A., Syktus, J., Wilson, K., Meijaard, E., Seabrook, L., Dargusch, P., Nordin, H., & Sheil, D. (2018). Forest loss and Borneo's climate. *Environmental Research Letters*, *13*(4), 044009.

Meijaard, E., Abrams, J., Juffe-Bignoli, D., Voigt, M., & Sheil, D. (2020). Coconut oil, conservation and the conscientious consumer. *CURRENT-BIOLOGY-D-20-00608*.

Meijaard, E., Brooks, T. M., Carlson, K. M., Slade, E. M., Garcia-Ulloa, J., Gaveau, D. L., Lee, J. S. H., Santika, T., Juffe-Bignoli, D., Struebig, M. J., & others. (2020). The environmental impacts of palm oil in context. *Nature Plants*, *6*(12), 1418–1426.

Formatted: Spanish (Spain)

Formatted: Spanish (Spain)

Meijaard, E., Garcia-Ulloa, J., Sheil, D., Wich, S., Carlson, K. M., Juffe-Bignoli, D., & Brooks, T. M. (2018). *Oil palm and biodiversity: A situation analysis by the IUCN Oil Palm Task Force*.

Meijaard, E., & Sheil, D. (2019). The moral minefield of ethical oil palm and sustainable development. *Frontiers in Forests and Global Change*, *2*, 22.

Miettinen, J., & Liew, S. C. (2011). Separability of insular Southeast Asian woody plantation species in the 50 m resolution ALOS PALSAR mosaic product. *Remote Sensing Letters*, *2*(4), 299–307.

Nomura, K., Mitchard, E. T., Patenaude, G., Bastide, J., Oswald, P., & Nwe, T. (2019). Oil palm concessions in southern Myanmar consist mostly of unconverted forest. *Scientific Reports*, *9*(1), 1–9.

OECD. (2018). *OECD-FAO agricultural outlook 2018-2027*. OECD Publishing.

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., & Wulder, M. A. (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, *148*, 42–57.

Oon, A., Ngo, K. D., Azhar, R., Ashton-Butt, A., Lechner, A. M., & Azhar, B. (2019). 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 Sensing Applications: Society and Environment*, *13*, 183–190.

Paoli, G. D., Gillespie, P., Wells, P. L., Hovani, L., Sileuw, A., Franklin, N., & Schweithelm, J. (2013). Oil palm in Indonesia: Governance, decision making and implications for sustainable development. *The Nature Conservancy, Jakarta, Indonesia*.

Pirker, J., Mosnier, A., Kraxner, F., Havlík, P., & Obersteiner, M. (2016). What are the limits to oil palm expansion? *Global Environmental Change*, *40*, 73–81.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4510–4520.

Santika, T., Wilson, K. A., Budiharta, S., Kusworo, A., Meijaard, E., Law, E. A., Friedman, R., Hutabarat, J. A., Indrawan, T. P., St. John, F. A., & others. (2019). Heterogeneous impacts of community forestry on forest conservation and poverty alleviation: Evidence from Indonesia. *People and Nature*, *1*(2), 204–219.

Shaharum, N. S. N., Shafri, H. Z. M., Ghani, W. A. W. A. K., Samsatli, S., Al-Habshi, M. M. A., & Yusuf, B. (2020). Oil palm mapping over Peninsular Malaysia using Google Earth Engine and machine learning algorithms. *Remote Sensing Applications: Society and Environment*, 100287.

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, *6*(1), 60.

Strona, G., Stringer, S. D., Vieilledent, G., Szantoi, Z., Garcia-Ulloa, J., & Wich, S. A. (2018). Small room for compromise between oil palm cultivation and primate conservation in Africa. *Proceedings of the National Academy of Sciences*, *115*(35), 8811–8816.

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Floury, N., Brown, M., & others. (2012). GMES Sentinel-1 mission. *Remote Sensing of Environment*, *120*, 9–24.

Wardeh, M., Sharkey, K. J., & Baylis, M. (2020). Integration of shared-pathogen networks and machine learning reveals the key aspects of zoonoses and predicts mammalian reservoirs. *Proceedings of the Royal Society B*, *287*(1920), 20192882.

Wich, S. A., Garcia-Ulloa, J., Kühl, H. S., Humle, T., Lee, J. S., & Koh, L. P. (2014). Will oil palm's homecoming spell doom for Africa's great apes? *Current Biology*, *24*(14), 1659–1663.

Woittiez, L. S., van Wijk, M. T., Slingerland, M., van Noordwijk, M., & Giller, K. E. (2017). Yield gaps in oil palm: A quantitative review of contributing factors. *European Journal of Agronomy*, *83*, 57–77.

Xu, Y., Yu, L., Li, W., Ciais, P., Cheng, Y., & Gong, P. (2020). Annual oil palm plantation maps in Malaysia and Indonesia from 2001 to 2016. *Earth System Science Data*, *12*(2), 847–847.

Yu, X., Wu, X., Luo, C., & Ren, P. (2017). Deep learning in remote sensing scene classification: A data augmentation enhanced convolutional neural network framework. *GIScience & Remote Sensing*, *54*(5), 741–758.

Deleted: ¶
¶