ELSEVIER

Monitoring direct drivers of small-scale tropical forest disturbance in near real-time with Sentinel-1 and -2 data

Bart Slagter [a,*], Johannes Reiche [a], Diego Marcos [b,a], Adugna Mullissa [a], Etse Lossou [a], Marielos Peña-Claros [c], Martin Herold [a,d]

[a] Laboratory of Geo-Information Science and Remote Sensing, Wageningen University & Research, Wageningen, the Netherlands
[b] Inria, University of Montpellier, Montpellier, France
[c] Forest Ecology and Forest Management Group, Wageningen University & Research, Wageningen, the Netherlands
[d] Helmholtz GFZ German Research Centre for Geosciences, Remote Sensing and Geoinformatics, Potsdam, Germany

ARTICLE INFO

ABSTRACT

Advancements in satellite-based forest monitoring increasingly enable the near real-time detection of small-scale tropical forest disturbances. However, there is an urgent need to enhance such monitoring methods with automated direct driver attributions to detected disturbances. This would provide important additional information to make forest disturbance alerts more actionable and useful for uptake by different stakeholders. In this study, we demonstrate spatially explicit and near real-time methods to monitor direct drivers of small-scale tropical forest disturbance across a range of tropical forest conditions in Suriname, the Republic of the Congo and the Democratic Republic of the Congo. We trained a convolutional neural network with Sentinel-1 and Sentinel-2 data to continuously classify newly detected RAdar for Detecting Deforestation (RADD) alerts as *smallholder agriculture, road development, selective logging, mining* or *other*. Different monitoring scenarios were evaluated based on varying sensor combinations, post-disturbance time periods and confidence levels. In general, the use of Sentinel-2 data was found to be most accurate for driver classifications, especially with data composited over a period of 4 to 6 months after the disturbance detection. Sentinel-1 data showed to be valuable for more rapid classifications of specific drivers, especially in areas with persistent cloud cover. Throughout all monitoring scenarios, *smallholder agriculture* was classified most accurately, while *road development, selective logging* and *mining* were more challenging to distinguish. An accuracy assessment throughout the full extent of our study regions revealed a Macro-F1 score of 0.861 and an Overall Accuracy of 0.897 for the best performing model, based on the use of 6-month post-disturbance Sentinel-2 composites. Finally, we addressed three specific monitoring use cases that relate to rapid law enforcement against illegal activities, ecological impact assessments and timely carbon emission reporting, by optimizing the trade-off in classification timeliness and confidence to reach required accuracies. Our findings demonstrate the strong capacities of high spatiotemporal resolution satellite data for monitoring direct drivers of small-scale forest disturbance, considering different user interests. The produced forest disturbance driver maps can be accessed via: https://bartslagter94.users.earthengine.app/view/forest-disturbance-drivers.

1. Introduction

Tropical forests play a crucial role in global climate stability, biodiversity conservation and human livelihoods, but remain threatened by unsustainable human activities (Barlow et al., 2016; Harris et al., 2021), leading to deforestation and forest degradation. Small-scale forest disturbances, such as road development, selective logging and mining, comprise a large portion of the total disturbance in the tropics (Kalamandeen et al., 2018; Tyukavina et al., 2018). They can have negative effects, such as severe carbon emissions or long-lasting ecological impacts and are often considered unsustainable or illegal (Alvarez-Berrios and Mitchell Aide, 2015; Kleinschroth and Healey, 2017; Pearson et al., 2017; Piponiot et al., 2019; Umunay et al., 2019).

Satellite-based forest disturbance alerting systems are of importance to increase transparency in forest activities and are a primary tool for law enforcement against illegal actions (Finer et al., 2018; Lynch et al.,

2013; Moffette et al., 2021). Still, most of the established operational forest disturbance alerting systems have limited capacities to detect small-scale disturbances, since they rely on medium spatial resolution (30 m or coarser) optical satellite data (e.g. Diniz et al., 2015; Hansen et al., 2016; Vargas et al., 2019). Apart from limited spatial resolutions, optical satellite time series can have long data gaps in the tropics due to cloud cover, which is especially problematic for small-scale disturbance detections. For example, remote sensing-based signals of selective logging disturbances can disappear within weeks due to rapid canopy and understory regrowth (Asner et al., 2004; Souza et al., 2005; Verhegghen et al., 2015), leaving them undetected after long periods of cloud cover.

Recent advancements in satellite-based forest disturbance alerting systems have enabled better detection of small-scale forest disturbances in high spatial detail and near real-time (Bouvet et al., 2018; Doblas et al., 2020; Hoekman et al., 2020; Mermoz et al., 2021; Pickens et al., 2020; Reiche et al., 2021). The Global Land Analysis and Discovery Sentinel-2 system (GLAD-S2) (Pickens et al., 2020), based on 10 m spatial resolution Sentinel-2 data, has largely improved the spatial detail of disturbance detections but still has limitations for timely detections in areas with persistent cloud cover. The recently developed RAdar for Detecting Deforestation system (RADD) (Reiche et al., 2021) employs cloud-penetrating Sentinel-1 radar sensors and consistently provides alerts on a weekly basis at 10 m spatial scale. This has enabled an unprecedented near real-time view on small-scale forest disturbances at the pantropical level.

However, mapping only the location of new disturbances is not always sufficient for targeted intervention and timely reporting of disturbance impacts. A critical advancement for alerting systems is to include a rapid and automated attribution of a forest disturbance's direct driver (Finer et al., 2018; Weisse et al., 2019). A direct driver is commonly defined as the proximate cause or human activity that directly impacted forest cover (Geist and Lambin, 2002). Having this additional information as fast as possible could enable alert prioritization and provide more actionable and enhanced information for alerting system users (Weisse et al., 2019). For example, knowing whether a detected disturbance is related to logging, mining or agriculture is important for determining the legality of the event and the different government agencies that may be involved in the response (Finer et al., 2018). In addition, rapid and spatially explicit information of direct forest disturbance drivers can provide options for more detailed and frequent reporting of, for example, ecological impacts or carbon emissions associated with forest disturbance.

In recent years, several studies have focused on classifying and mapping direct forest disturbance drivers based on satellite data. Large-scale retrospective insights into drivers have been gained in sample-based studies (De Sy et al., 2019; Laso Bayas et al., 2022; Tyukavina et al., 2018) or classifications on a coarse grid cell level (Curtis et al., 2018). More spatially explicit classifications methods have also been studied, with annual driver classifications at the disturbance pixel- or patch-level, in temperate and boreal forests (Hermosilla et al., 2015; Huo et al., 2019; Nguyen et al., 2018; Oeser et al., 2017; Schroeder et al., 2011, 2017; Sebald et al., 2021; Senf and Seidl, 2021; Stewart et al., 2009; Vogeler et al., 2020; Zhang et al., 2022) and in tropical (dry) forests (De Marzo et al., 2022; Shimizu et al., 2019). In order to obtain spatially explicit outputs, these studies relied predominantly on medium spatial resolution (30 m) Landsat data and used a variety of rule-based and machine learning methods.

The combination of machine learning and remote sensing provides a major opportunity for forest disturbance driver classifications (Finer et al., 2018). Recent machine learning-based studies emphasized that an accurate classification of a forest disturbance driver requires a large quantity of complex input features related to a disturbance's geometric features, multispectral reflectance, temporal development and spatial context (Alonso et al., 2022; Hermosilla et al., 2015; Nguyen et al., 2018; Schroeder et al., 2017; Sebald et al., 2021; Senf and Seidl, 2021; Shimizu et al., 2019; Stewart et al., 2009; Vogeler et al., 2020). The use of deep learning methods, especially convolutional neural networks (CNN), offers a great potential for remote sensing-based driver classifications because such complex features can be learned automatically from multi-sensor data without the need for intensive feature engineering (LeCun et al., 2015; Reichstein et al., 2019; Yuan et al., 2020). To date, three studies are known to have applied CNNs to retrospectively classify forest disturbance drivers or follow-up land use in deforested areas (Irvin et al., 2020; Masolele et al., 2021, 2022), presenting accurate classifications in the tropics with the use of optical data.

The established rule-based, machine learning and deep learning methods have enabled accurate forest disturbance driver classifications on an annual level retrospectively, but they cannot be used for near real-time monitoring and have often neglected small-scale disturbances. With Landsat as the predominant data source in past studies, forest disturbance drivers are challenging to classify in near real-time because of the medium spatial resolution (30 m) and data gaps due to cloud cover. This is especially problematic for classifying small-scale disturbance drivers in the tropics.

These limitations can now be partly overcome since the launch of the European Space Agency's Sentinel-1 and Sentinel-2 satellites (in 2014 and 2015 respectively), offering openly available radar and optical imagery in high spatiotemporal detail. The Sentinel-1 satellites are equipped with C-Band synthetic aperture radar sensors that provide imagery at approximately 20×22 m spatial resolution and have a revisit time of 6–12 days in the tropics (Torres et al., 2012). This ensures a consistent stream of dense time series data in high spatial detail, independent from daylight and cloud cover conditions. The Sentinel-2 satellites are equipped with multispectral optical sensors that provide imagery in resolutions of 10-, 20- and 60 m in the visible, near-infrared (NIR) and short-wave infrared (SWIR) portions of the electromagnetic spectrum and have a revisit time of 5 days in the tropics (Drusch et al., 2012). The availability of high spatiotemporal resolution Sentinel-1 and Sentinel-2 time series data now provides the opportunity to move towards more rapid forest disturbance driver classifications in high spatial detail (Finer et al., 2018), including small-scale drivers. The use of Sentinel-1 and Sentinel-2 data for near real-time driver classifications of small-scale tropical forest disturbances has not been studied yet. However, the potential of Sentinel-2 data was recently demonstrated for timely classifications of forest disturbances related to fire and harvesting in temperate and boreal forests (Alonso et al., 2022; Cardille et al., 2022).

There is an urgent need to extend forest disturbance monitoring methods with near real-time and spatially explicit classifications of disturbance drivers. In this research, we address three main aspects related to near real-time monitoring of small-scale forest disturbance drivers in the tropics, making use of high spatiotemporal resolution satellite data. First, we demonstrate the use of a CNN applied to Sentinel-1 and Sentinel-2 data to classify forest disturbance alerts into a set of small-scale driver classes, across a range of tropical forest conditions. Second, we assess the classification methods for different near real-time monitoring scenarios. Third, we evaluate the trade-off between classification timeliness and confidence in the context of specific use cases.

2. Materials & methods

2.1. Study regions

Three study regions were selected for this research: The whole country of Suriname, four provinces in the north of the Republic of the Congo (ROC) and one province in the east of the Democratic Republic of the Congo (DRC) (Fig. 1). All study regions contain mainly humid tropical forests. The areas were selected because they have varying deforestation and forest degradation patterns and are mainly subject to small-scale disturbances (Table 1). In Suriname, forest disturbances are mainly driven by mining, road development and selective logging (Government of Suriname, 2018). In ROC, road development, selective
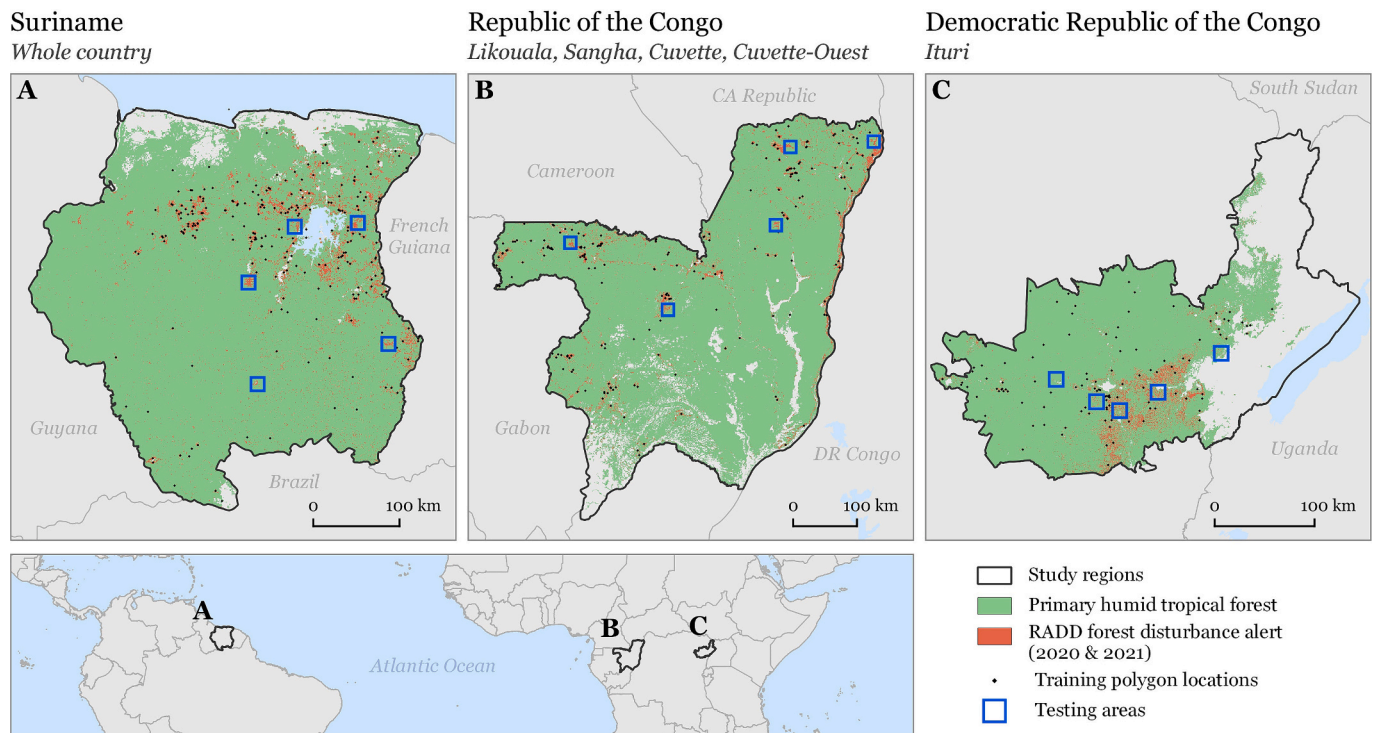
Suriname
Whole country

Republic of the Congo
Likouala, Sangha, Cuvette, Cuvette-Ouest

Democratic Republic of the Congo
Ituri

Fig. 1. The three study regions selected for this research. Displayed are the extent of primary humid tropical forest (Turubanova et al., 2018), the detected RADD alerts in 2020 and 2021, the locations of collected training polygons and the testing areas used for model evaluation in this study (further described in chapter 2.3).

Table 1
Overview of forest disturbance sizes in the study regions based on the detected RADD alert patches (spatially connected groups of alert pixels) in 2020 and 2021. The proportions of small-scale disturbances are shown as percentages of the total number and total area.

	RADD alert numbers (patch-based)			RADD alert area		
	Total number of alerts	Alerts <0.25 ha (%)	Alerts <1 ha (%)	Total area of alerts (kha)	Area of alerts <0.25 ha (%)	Area of alerts <1 ha (%)
Suriname	28,425	40.1	84.9	150	9.1	35.6
ROC (4 provinces)	37,795	41.4	89.2	144	12.2	50.1
DRC (Ituri)	67,109	34.5	78.5	457	6.8	34.4

logging and smallholder agriculture are the major drivers (Kleinschroth et al., 2019; Tyukavina et al., 2018). The study region in DRC experiences severe forest disturbances mostly driven by smallholder agriculture and mining (IPIS, 2016; Tyukavina et al., 2018).

The regions also vary in terms of cloud cover. In general, high cloud coverage limits the number of suitable Sentinel-2 observations, and this may hinder the rapid classification of forest disturbance drivers. Suriname and ROC are especially severely cloud-covered. Based on a 25% cloud-probability cut-off point in the s2cloudless dataset (Zupanc, 2017), Suriname has on average 73.1% of Sentinel-2 pixels obstructed by clouds and the study region in ROC has 71.3% obstructed. The study region in DRC has slightly lower levels of cloud cover, with 64.9% of the pixels obstructed.

2.2. Data and preprocessing

We used the near real-time Sentinel-1-based RADD alerts (Reiche et al., 2021) from 2020 and 2021 as a basis for our forest disturbance driver classifications. In the RADD alerts methods, a new forest disturbance alert is triggered by a single observation from the latest Sentinel-1 image. This is done by determining per pixel the backscatter deviation from expected backscatter values in stable forest, as derived from historical time series metrics. Subsequent Sentinel-1 observations are used to increase the confidence and either reject or confirm the alert. For this study, images from 2018 and 2019 were used as historical time series to derive stable forest backscatter metrics, and images from 2020 and 2021 were used to produce forest disturbance alerts. Only high-confidence alerts were included in the outputs and a minimum mapping unit of 0.1 ha was used, corresponding to 10 Sentinel-1 pixels.

For the driver classifications of the detected RADD alerts, we sourced satellite imagery from Sentinel-1 and Sentinel-2 through Google's cloud-based geospatial processing platform Google Earth Engine (GEE) (Gorelick et al., 2017). For Sentinel-1, we used the Ground Range Detected (GRD) imagery acquired in Interferometric Wide Swath mode. In the study regions, Sentinel-1 acquires dual-polarized imagery (VV and VH) only in descending orbit. The images have a spatial resolution of approximately 20 × 22 m and are provided with 10 m pixel spacing. Sentinel-1 preprocessing prior to GEE ingestion includes the application of the orbit files, thermal noise and GRD border noise removal, radiometric calibration to sigma naught and range-Doppler terrain correction, based on the European Space Agency's Sentinel-1 Toolbox methods (European Space Agency, 2022). We further preprocessed the images by removing remaining border noise and applying radiometric terrain correction (Hoekman and Reiche, 2015; Vollrath et al., 2020), following the GEE-based implementations of Mullissa et al. (2021). Speckle filtering was not applied, as this was expected to be redundant in combination with the use of a CNN.

For Sentinel-2, we used the Level 2A Surface Reflectance imagery and selected the Green, Red, NIR and SWIR bands (i.e. bands 3, 4, 8 and 12). The images provided in GEE are atmospherically corrected with the

Sen2Cor methods (Main-Knorn et al., 2017). We further preprocessed the images by applying cloud- and shadow masking, following the GEE-based methods from s2cloudless (Zupanc, 2017). We produced for each Sentinel-2 image a cloud mask based on a 25% cloud-probability threshold. Additionally, we detected cloud shadows based on low NIR reflectance in proximity of detected clouds and the solar azimuth angle. To remove errors in the cloud- and shadow masks, we applied a 5×5 pixel focal majority filter and expanded the resulting masks with a 50 m buffer.

For reference data collection, we used monthly Planet mosaics (Planet Team, 2017), openly available through the NICFI program. The Planet mosaics were accessed in GEE, where they are provided in 4.8 m spatial resolution and contain Blue, Green, Red and NIR bands.

2.3. Class definitions and reference data

Direct forest disturbance drivers in this study are defined as the proximate causes that directly affect forest cover (Geist and Lambin, 2002). Direct drivers refer to the human activity or land use that is responsible for a disturbance and can commonly be observed from the interpretation of high-resolution satellite or in-situ data (De Sy et al., 2019). Our study did not consider underlying drivers or reasons for forest change related to fundamental socio-economic processes, such as demographic pressures, economic markets or national policies (Geist and Lambin, 2002).

We defined five classes of direct forest disturbance drivers. These were *smallholder agriculture, road development, selective logging, mining* and *other* (Fig. 2, Table 2). Given that the selected study regions experience mainly small-scale forest disturbances, larger-scale disturbance drivers in the tropics, such as industrial agriculture or wildfire, were not included in the classification scheme.

The class *other* was added to enable the model to recognize small-scale disturbances that do not fit into one of the four classes, such as natural disturbances (Table 2). This class also included isolated small disturbances, for which no clear human-induced cause was visible.

We included skid trails as part of the *selective logging* class (Table 2). However, wider skid trails can cause narrow linear canopy openings which can appear visually similar to smaller logging roads in satellite imagery, making it in some cases ambiguous to distinguish what is part of *selective logging* or part of *road development*. Still, when skidding causes linear disturbance visibility in satellite imagery, this signal usually disappears quickly due to canopy recovery (Asner et al., 2004), while logging roads stay visible for a longer time. Consequently, for linearly

shaped logging-related disturbances, we reduced ambiguity of what was considered to be a skid trail or a logging road by using a threshold for visibility duration. Linearly shaped disturbances disappearing within three months due to canopy regrowth were considered to be skid trails and part of *selective logging,* while linearly shaped disturbances with a longer visibility duration were considered to be logging roads and part of *road development*. Finally, we did not differentiate between logging roads and roads connecting settlements, which were both considered part of *road development.*

Two interpreters collected training data for the defined driver classes by labelling RADD alert patches based on visual interpretation of monthly Planet mosaics of up to one year following the disturbance. One interpreter did the initial labelling, and the second interpreter did a final verification of all labels. Disturbance patches were labelled by manually drawing polygons around groups of forest disturbance patches that fit in one of the defined classes (Fig. 1, point locations). With this method, large numbers of forest disturbance patches (12,016 in total, Table 3) could be labelled in an efficient way for model training. Based on acquired knowledge from the initial model training attempts, we focused our training data collection mainly on edge cases, i.e. training samples that were expected to be particularly challenging to classify correctly (e. g. small *road development* and *mining* samples in landscapes dominated by *smallholder agriculture,* or *selective logging* and *road development* samples in areas with a complex mix of disturbances related to tree felling, skidding and logging road construction).

Additionally, we collected a held-out testing dataset, which was exclusively used for model evaluation. For the testing dataset, we selected five rectangular 10×10 km areas per study region, located at specifically selected forest disturbance hotspots with a variety of disturbance patterns and drivers (Fig. 1, areas in blue). Within these rectangles, we labelled all detected forest disturbance patches with their observed direct driver. The total number of training and testing samples collected per study region and driver class can be found in Table 3.

It was challenging to confidently label all forest disturbance patches in the testing areas based on visual interpretation of Planet imagery. For this reason, 5.5% of the testing samples had to be left out of the testing dataset (reported as *undefined* in Table 3).

The representation of the separate driver classes differed widely among the individual study regions (Table 3). For this reason, we did not have a solid basis for accuracy comparisons among the three study regions. All collected reference data was aggregated throughout the three regions to ensure a regionally varied representation of all classes.
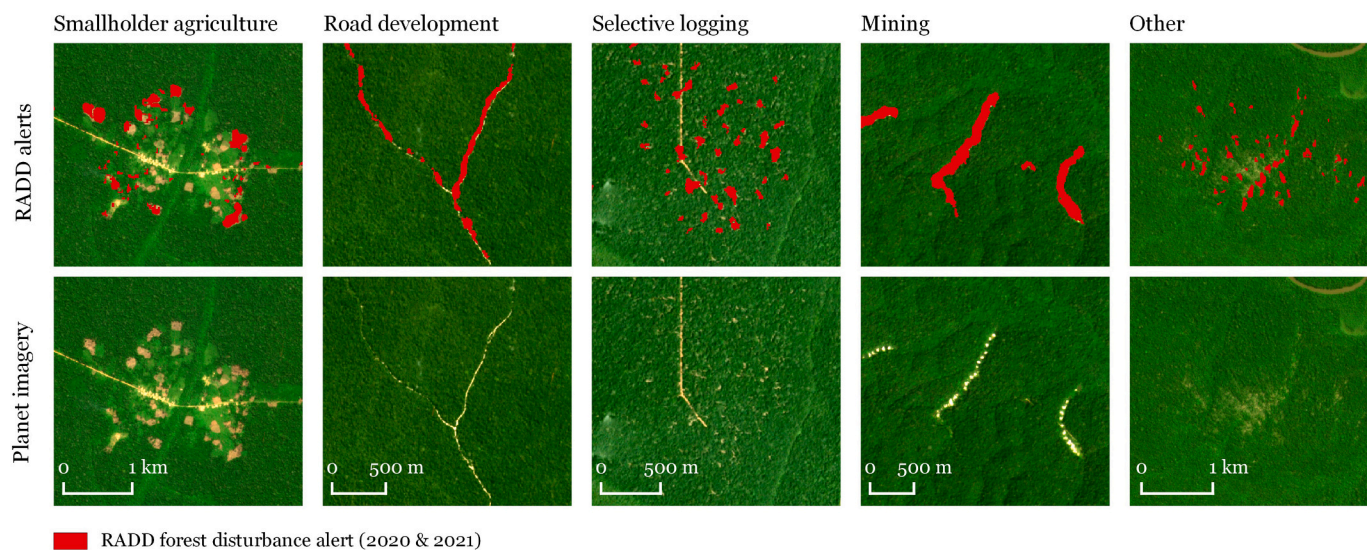

Fig. 2. Examples of forest disturbances related to a direct driver within the study regions, detected in the RADD alerts and visualized with post-disturbance 4.8 m spatial resolution Planet imagery.

Table 2

The defined direct forest disturbance driver classes, their class descriptions and their distinct visible features as they can be derived from 4.8 m spatial resolution Planet imagery.

Class & description	Visible features (based on Planet imagery)
Smallholder agriculture: Forest clearing mainly driven by subsistence and/or shifting agriculture, but occasionally mixed with other smallholder forest activities, such as charcoal production and fuelwood collection.	- Clearing of forest patches with appearance of bare soil. - Disturbance is followed by vegetation regrowth within one year. - Located in vicinity of settlements and/or roads. - Located at forest edges (commonly). - Disturbance patterns are seasonal and rotational (commonly). - Multiple smallholder agriculture disturbances are co-located (commonly).
Road development: Forest clearing mainly driven by road infrastructure development, commonly related to the construction of logging roads to facilitate log hauling.	- Clearing in linear shapes, with appearance of bare soil. - Disturbance is composed of spatially continuous canopy openings, with little interruption. - Disturbance stays visible for at least three months before canopy recovers. - In direct or indirect connection with settlements or existing roads. - Co-located with selective logging (commonly).
Selective logging: Small-scale forest disturbances related to tree felling and skidding.	- Canopy disturbance at small scale. - Visibility of bare soil is restricted to only several small patches of pixels. - Canopy recovers and bare soil visibility mostly disappears within three months. - Located in vicinity of (logging) roads. - Multiple selective logging disturbances are co-located (commonly).
Mining: Forest clearing to facilitate open-pit mineral extraction, commonly related to alluvial gold mining.	- Clearing with appearance of bright bare soil and/or water ponds. - Disturbance stays clearly visible for at least one year. - Located in vicinity of rivers and/or wetland areas. - Located in vicinity of roads or other infrastructure. - Multiple mining disturbances are co-located (commonly). - Clearing in wide linear shapes (commonly).
Other: Any forest disturbance not fitting in the abovementioned classes, commonly related to natural disturbances, such as flooding or windthrows.	- Disturbance has no visible human-induced cause. - Co-located disturbances have mostly similar disturbance dates. - Disturbance is followed by vegetation regrowth within three months (commonly). - Disturbance is isolated from other disturbances or infrastructure (commonly).

2.4. Driver classification and monitoring

Our patch-based forest disturbance driver classification approach focused on the use of standard CNN models tailored for multi-class image classifications. The CNN was designed to accept multi-channel image data as an input and assign five class probability scores as an output. For each detected RADD forest disturbance patch, a 2 × 2 km bounding box was defined and used as extent to delineate small Sentinel-1 and Sentinel-2 composites, based on images acquired in a specified time period directly after the disturbance date of the patch. We trained the model with input bands from Sentinel-1, Sentinel-2 and with the combination of both sensors and evaluated their performance in the held-out testing areas. We also tested the effects of using different lengths for the post-disturbance time period to assess a practical trade-off between accurate and rapid classifications of drivers. Furthermore, we assessed whether the class probability scores could be used to eradicate classification errors. The best performing model based on the testing dataset was selected to produce wall-to-wall forest disturbance driver maps for the three study regions and validated in an independent accuracy assessment.

2.4.1. Forest disturbance patch definition

We simulated near real-time monitoring scenarios by using RADD forest disturbance alerts in our study regions for 2020 and 2021 and spatiotemporally delineated individual forest disturbance patches through time. During monitoring, the aim was to continuously produce driver classifications of the latest detected alerts, defined as all forest disturbances mapped in the latest 6 months. These latest detected alert pixels were grouped into patches of connected pixels, to create single alert objects to be assigned a forest disturbance driver. The patches were defined with a connection tolerance of 3 pixels (i.e. 30 m) and a minimum size of 10 pixels (i.e. 0.1 ha).

2.4.2. Input images

RADD alert patches were generated on a monthly basis to produce the training dataset. The labelled training data polygons (as described in chapter 2.3) were used to target the specific forest disturbance patches to be included in the training dataset. This method caused most patches to be sourced multiple times for training because a progressing forest disturbance could occur several times throughout the time series of monthly forest disturbance alert maps. This effect was intentionally introduced to increase the input data volume for model training. With this approach, 82,974 labelled forest disturbance patches were sampled for the training dataset. Note that this number is substantially higher than the numbers presented in Table 3, due to the repeated segmentation of forest disturbance patches on a monthly basis. Eventually, 48.9% of the resulting training samples represented *smallholder agriculture*, 16.5% represented *road development*, 9.4% represented *selective logging*, 17.1% represented *mining* and 8.1% represented *other*.

Each selected forest disturbance patch was transformed into a small input image, by assigning a bounding box of 200 × 200 pixels (2 × 2 km) located around the centroid of the patch. Within the bounding boxes,

Table 3

Numbers of labelled reference forest disturbance patches throughout the study regions for training and testing. Numbers are reported for a spatial segmentation of patches, without segmentation in the temporal dimension. Note that no suitable reference samples were found for *selective logging* in DRC.

	Number of collected reference patches											
	Training dataset						Testing dataset					
	Sm. agric.	Road dev.	Sel. logging	Mining	Other	Total	Sm. agric.	Road dev.	Sel. logging	Mining	Undefined	Total
Suriname	780	579	782	1119	492	3752	125	162	218	79	34	618
ROC	2781	841	1017	165	701	5505	268	177	257	64	29	795
DRC	2035	50	0	206	468	2759	1372	33	0	43	100	1548
Total	5596	1470	1799	1490	1661	12,016	1765	372	475	186	163	2961

Table 4

Overview of the 2 × 2 km images sampled per individual forest disturbance patch. All 8 input band values were scaled between 0 and 1. Note that the disturbance dates in band 2 required a special form of scaling to values related to alert recency.

Image band	Data / Sensor	Pixel spacing	Metric	Scaling
1	RADD	10 m	Disturbance patch	0 = non-patch
				1 = patch
2		10 m	Disturbance dates	0 = undisturbed
				0.1 = disturbed >1 year ago
				0.1–1 = disturbed 1 year ago until now
3	Sentinel-1	10 m	VV backscatter	0–1 = −25 to 0 dB
4		10 m	VH backscatter	0–1 = −30 to −5 dB
5	Sentinel-2	20 m	Short-wave infrared	0–1 = 0%–30% reflectance
6		10 m	Near-infrared	0–1 = 0%–60% reflectance
7		10 m	Green	0–1 = 0%–30% reflectance
8		10 m	Red	0–1 = 0%–30% reflectance

preprocessed Sentinel-1 and Sentinel-2 imagery was sourced for a specified post-disturbance time period (Table 4). The post-disturbance images were compressed to a median composite for Sentinel-1 and a quality composite for Sentinel-2 based on the lowest cloud probability per pixel (Zupanc, 2017). With a 2 × 2 km bounding box around each disturbance patch, multispectral and backscatter information was captured not only for the disturbance itself, but also for its direct surrounding. Spatial context in the direct surrounding of a disturbance can be an important predictor for a driver classification (Sebald et al., 2021; Senf and Seidl, 2021). In the tropics, for example, selective logging is commonly in proximity of roads and vice versa. Smallholder agriculture is commonly in proximity of settlements and located at forest edges in shifting agriculture landscapes. Mining is often located in wetland areas or along rivers.

Along with the satellite composites, two layers from the RADD alerts were added (Table 4): the mask of the disturbance patch itself and the detected disturbance dates from all the disturbances within the bounding box up to the moment of classification. This second RADD alert layer was included to add spatiotemporal information of the detected forest disturbances that occurred in the immediate surrounding of the patch. The disturbance dates in this layer were converted to continuous values related to the alert pixel's recency to ensure generalizability of the values through time.

2.4.3. Convolutional neural network

We designed a CNN based on a simplified VGG-16 architecture (Simonyan and Zisserman, 2015) (Fig. 3). The model was developed and trained with TensorFlow 2.1. We used a batch size of 128, Adam

optimization (Kingma and Ba, 2015) with a learning rate of 0.0001, and trained for 8 epochs with a learning rate decay of factor 3 after every second epoch. We used class-weights proportionate to the number of training samples per class to account for the effects of class-imbalance. During model training, we held out 10% of the training data as validation samples to assess the model learning process. To avoid spatial overlap in this 90/10% split, the partition was made at the level of the collected reference data polygons (Fig. 1, point locations).

2.4.4. Monitoring scenarios

We trained separate models with Sentinel data composited over several post-disturbance time periods ranging from 1 to 6 months. We also trained a model for an immediate classification scenario, with only a single Sentinel-1 image from directly after the disturbance. The trade-off between classification timeliness and accuracy is of importance for different monitoring applications. Short post-disturbance time periods enable rapid classifications, but these can be prone to errors due to cloudy Sentinel-2 data, speckle in Sentinel-1 data, or because it is too early to determine a disturbance driver. Long post-disturbance time periods can lead to more accurate but less timely classifications. Since our data did not extend into 2022, and the longest post-disturbance time period we tested was 6 months, we trained and applied our models only on disturbance patches detected in 2020 and the first half of 2021, to maintain consistency throughout all tested post-disturbance time periods.

We trained models with input bands from Sentinel-1 only, Sentinel-2 only and both sensors combined. In the immediate classification scenario based on a single image, we used only Sentinel-1 because single
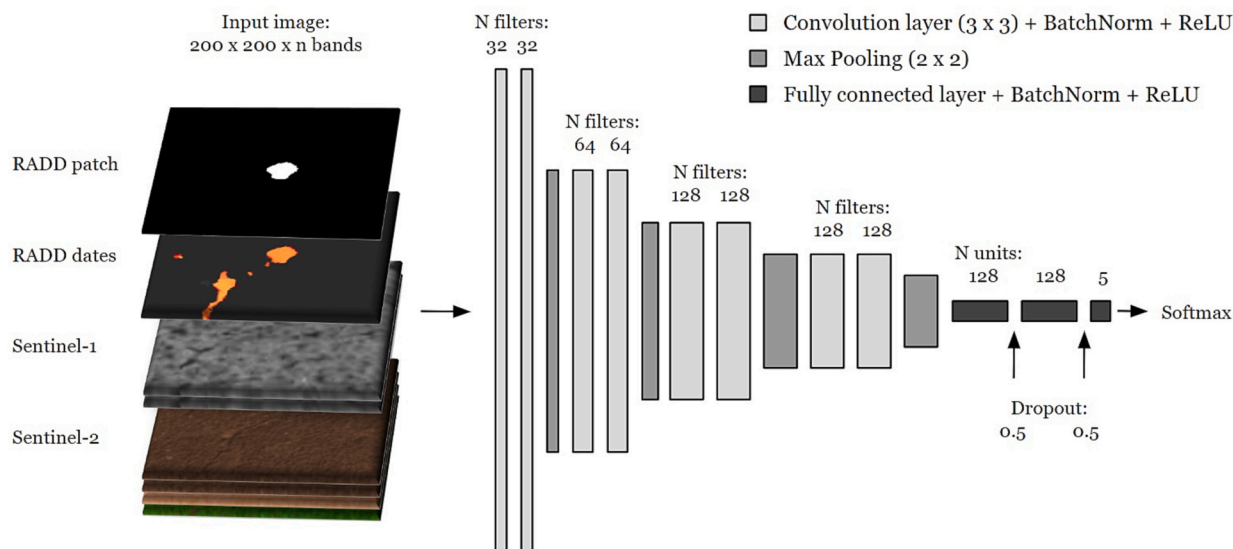

Fig. 3. The convolutional neural network architecture for multi-class image classification.

Sentinel-2 images are predominantly cloud-covered in the tropics and usually not a valuable basis for classification. Additionally, A model was trained based on just the RADD alert data, with no Sentinel imagery added. This was done to produce a baseline classification scenario based on only the alert's geometric features, such as shape, size and temporal development.

The different monitoring scenarios were evaluated by applying the classifications in the 15 held-out testing areas (Fig. 1, areas in blue). For all forest disturbances in the testing areas, the driver classifications were produced on a quarterly basis for 2020 and the first half of 2021. To evaluate model performance, the classifications were compared against the true driver labels as were visually interpreted from Planet imagery. Based on the True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN) classifications, we calculated the User's Accuracy (UA), Producer's Accuracy (PA) and F1 score per class:

$$UA = \frac{TP}{TP + FP} \tag{1}$$

$$PA = \frac{TP}{TP + FN} \tag{2}$$

$$F1\ score = 2 \times \frac{UA \times PA}{UA + PA} \tag{3}$$

The Macro-F1 score was used as a final evaluation metric (the average of all class-specific F1 scores), because we aimed to weigh each driver class equally in the testing areas. It should be noted that the reported accuracies refer to classifications at the patch level and do not reflect pixel-based map accuracies.

In the testing areas, we evaluated the classification of the four main human-induced driver classes *smallholder agriculture, road development, selective logging* and *mining*. The class *other* was not sufficiently represented in the selected testing areas and also not of main interest for the evaluation of the different scenarios. Therefore, the class *other* was only evaluated for the misclassification of the four main driver classes as *other* and reflected in their UA scores.

Filtering out uncertain driver classifications from the output maps can produce more effective results in practice when an undefined classification is preferred over a false classification. Therefore, we evaluated the effect of applying confidence thresholds to the classification outputs, based on the class probability scores calculated from the Softmax activation function in the CNN. The application of a threshold to the class probability scores always comes at the cost of excluding disturbances

from the classification. To incorporate this exclusion in the assessment, confidence threshold effects were evaluated by comparing the Macro-F1 score increases versus the percentages of classifications that remained included after thresholding.

2.4.5. Forest disturbance driver maps and accuracy assessment

The best performing model from the various monitoring scenarios was selected based on the Macro-F1 scores obtained in the held-out testing areas. We applied this model to produce forest disturbance driver maps throughout the full extent of our three study regions (Fig. 1) for 2020 and the first half of 2021, based on the classification of disturbance patches on a biannual basis. We performed an independent accuracy assessment based on a stratified random sample of classified forest disturbance patches throughout the full extent of the study regions. Before sampling, we excluded all areas that were covered by our training data polygons buffered with 1 km (Fig. 1, point locations). This step was necessary to obtain an unbiased accuracy assessment, unaffected by samples that were also included for model training. Without the exclusion of training areas, a random sample would have overlapped for 28.9% with the training data and thus have led to a considerable overestimation of the accuracies of our methods.

We used three strata for sample stratification, based on disturbance patch sizes. The first stratum included patches with a size smaller than 0.25 ha, the second stratum included patches with a size between 0.25 and 1.5 ha, and the third stratum included patches larger than 1.5 ha. We sampled 150 patches per stratum in each of the three study regions. This resulted in a total of 1350 randomly sampled disturbance patches.

All sampled disturbance patch classifications were independently verified based on the visual interpretation of monthly Planet mosaics. We calculated proportionally adjusted accuracies conforming with the representation of the different strata in the study regions. We calculated the UA, PA and Macro-F1 score, as was similarly done for the model evaluations based on the testing dataset. Additionally, we calculated the Overall Accuracy (OA):

$$OA = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

3. Results

The testing accuracies revealed high variations for the model performances when different satellite sensors and post-disturbance time periods were used (Fig. 4, Appendix 1). Classifications based on
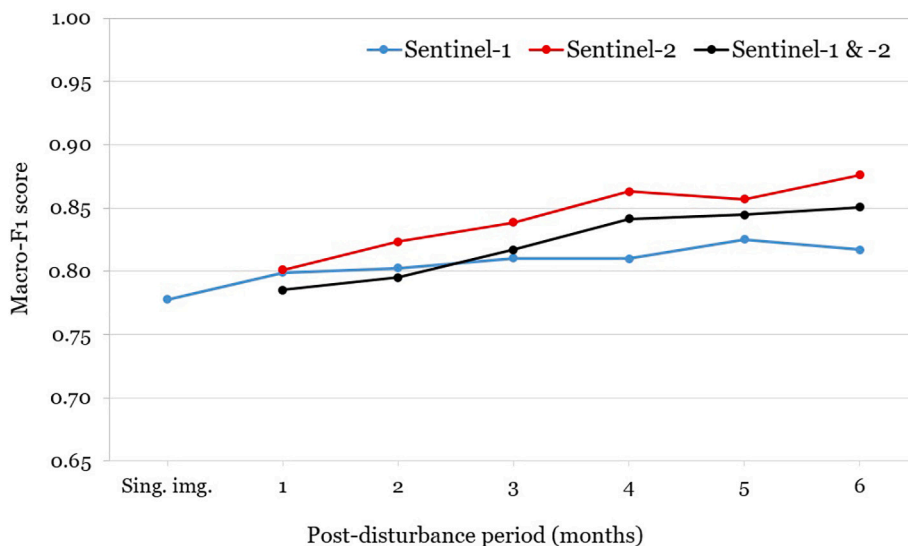

Fig. 4. Testing accuracies of the driver classifications in the different monitoring scenarios. Macro-F1 scores are presented per sensor and post-disturbance time period.
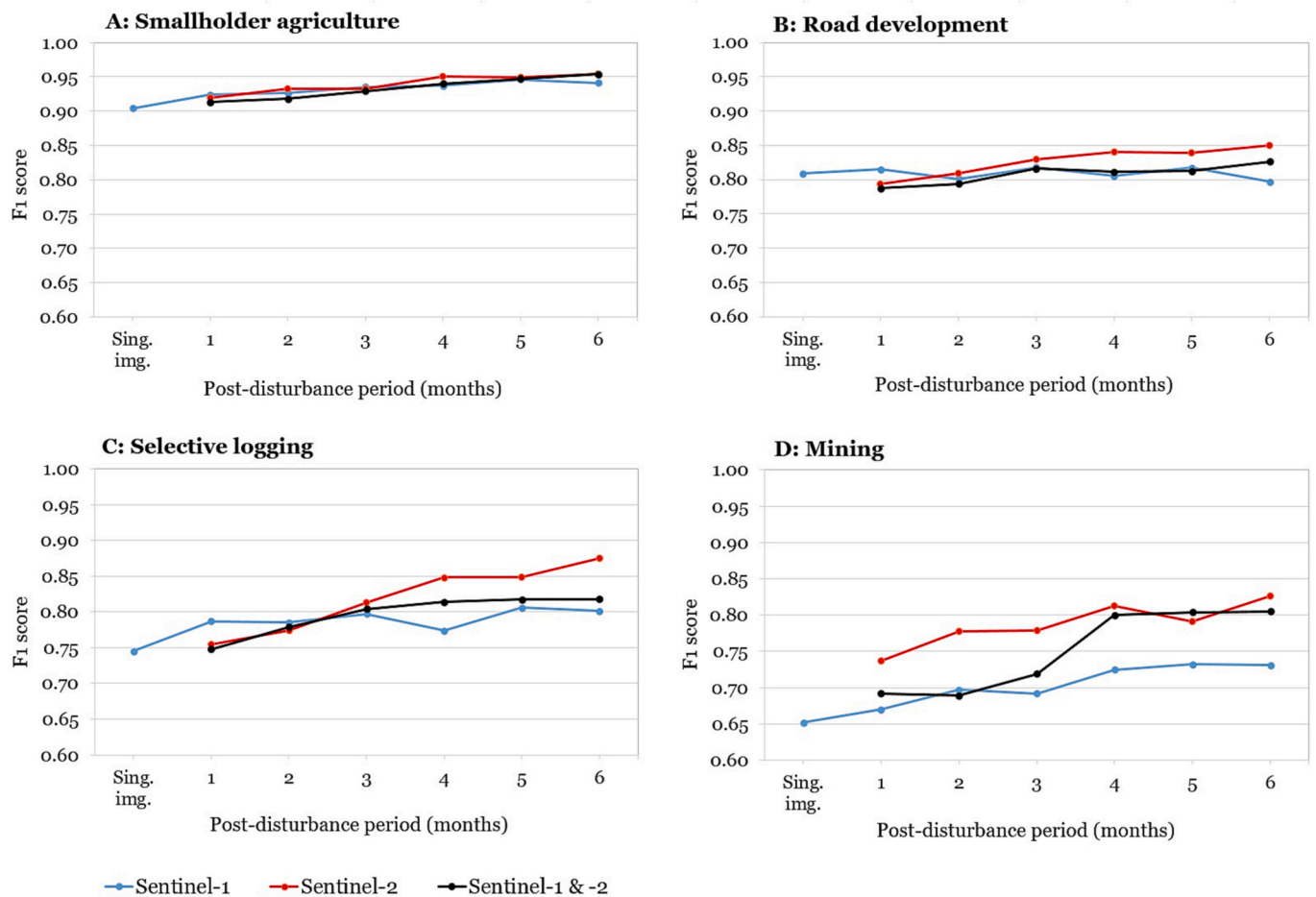
Fig. 5. Testing accuracies of the driver classifications in the different monitoring scenarios for the defined classes *smallholder agriculture* (A), *road development* (B), *selective logging* (C) and *mining* (D). Class-specific F1 scores are presented per sensor and post-disturbance time period.

Sentinel-2 data showed to be more accurate than classifications based on Sentinel-1 data, with Macro-F1 scores ranging up to 0.876 and 0.825 for the sensors respectively in the testing areas. Combining the input bands from Sentinel-1 and Sentinel-2 did not lead to accuracy increases. In general, the best accuracies were obtained with the use of Sentinel-2 data composited over long post-disturbance periods of 4 to 6 months. The sensors obtained comparable Macro-F1 scores in a more rapid monitoring scenario, with the use of 1-month post-disturbance composites. A Macro-F1 score of 0.861 and an OA of 0.897 were obtained in the accuracy assessment of the best performing model, based on the use of 6-month post-disturbance Sentinel-2 composites.

Over all the testing areas, high variations were observed in the class-specific F1 scores (Fig. 5, Appendix 1). *Smallholder agriculture* was consistently classified most accurately, with F1 scores ranging up to 0.954. *Road development* and *selective logging* were occasionally confused (Table 5) and were classified with F1 scores ranging up to 0.850 and 0.875 respectively. *Mining* was classified with an F1 score ranging up to 0.826 and showed most confusion with *road development* and *smallholder agriculture*. Examples of the model classification outputs for the different classes are shown in Fig. 6.

Rapid driver classifications were achieved with the use of short post-disturbance time periods, but this came at the cost of having lower

Table 5
Four normalized confusion matrices based on the testing accuracies, related to the monitoring scenarios with a 1-month and 6-month post-disturbance period per sensor. Note that the class 'other' was only incorporated in classification results and not in the testing dataset. SA = *smallholder agriculture*, RD = *road development*, SL = *selective logging*, M = *mining*, O = *other*.

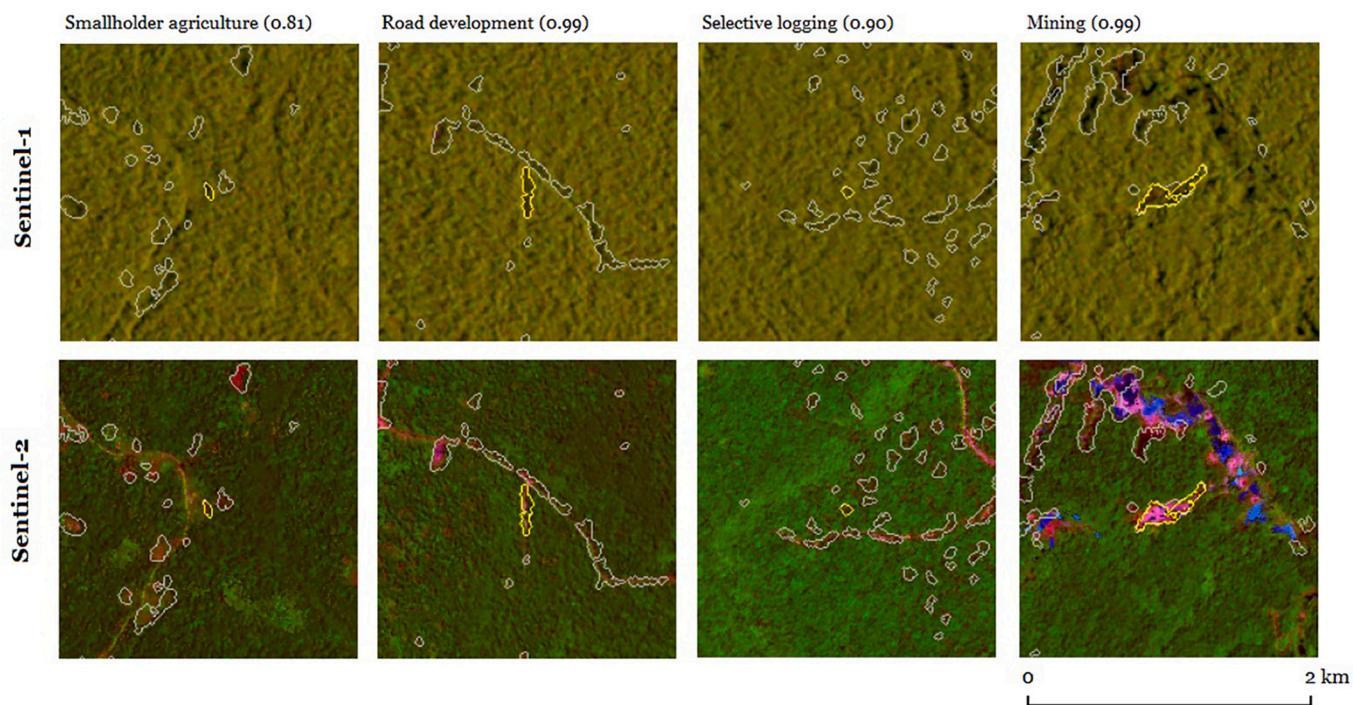|			Classification									
			1 month post-disturbance					6 months post-disturbance				
			SA	RD	SL	M	O	SA	RD	SL	M	O
Reference (testing dataset)	Sentinel-1	SA	0.883	0.011	0.007	0.068	0.031	0.905	0.021	0.008	0.042	0.024
		RD	0.032	0.808	0.069	0.063	0.028	0.019	0.806	0.099	0.062	0.013
		SL	0.015	0.177	0.745	0.012	0.053	0.002	0.160	0.790	0.005	0.043
		M	0.123	0.031	0.005	0.815	0.025	0.104	0.038	0.020	0.826	0.012
	Sentinel-2	SA	0.867	0.022	0.020	0.038	0.053	0.924	0.013	0.004	0.022	0.036
		RD	0.012	0.824	0.073	0.048	0.043	0.018	0.843	0.059	0.060	0.020
		SL	0.001	0.201	0.745	0.000	0.053	0.002	0.103	0.858	0.001	0.036
		M	0.120	0.065	0.005	0.793	0.016	0.063	0.027	0.009	0.899	0.002

Fig. 6. Correct image classification outputs visualized for four example forest disturbance patches in the testing areas, along with their RADD alert patterns (disturbance patch = yellow, disturbance surrounding = grey), post-disturbance Sentinel-1 composite (RGB = VV, VH and VV-VH), and Sentinel-2 composite (RGB = SWIR, NIR, Red), all sampled from the disturbance's corresponding 2 × 2 km bounding boxes. Classifications were produced with Sentinel-1 and Sentinel-2 data, composited over a 6-month post-disturbance period.
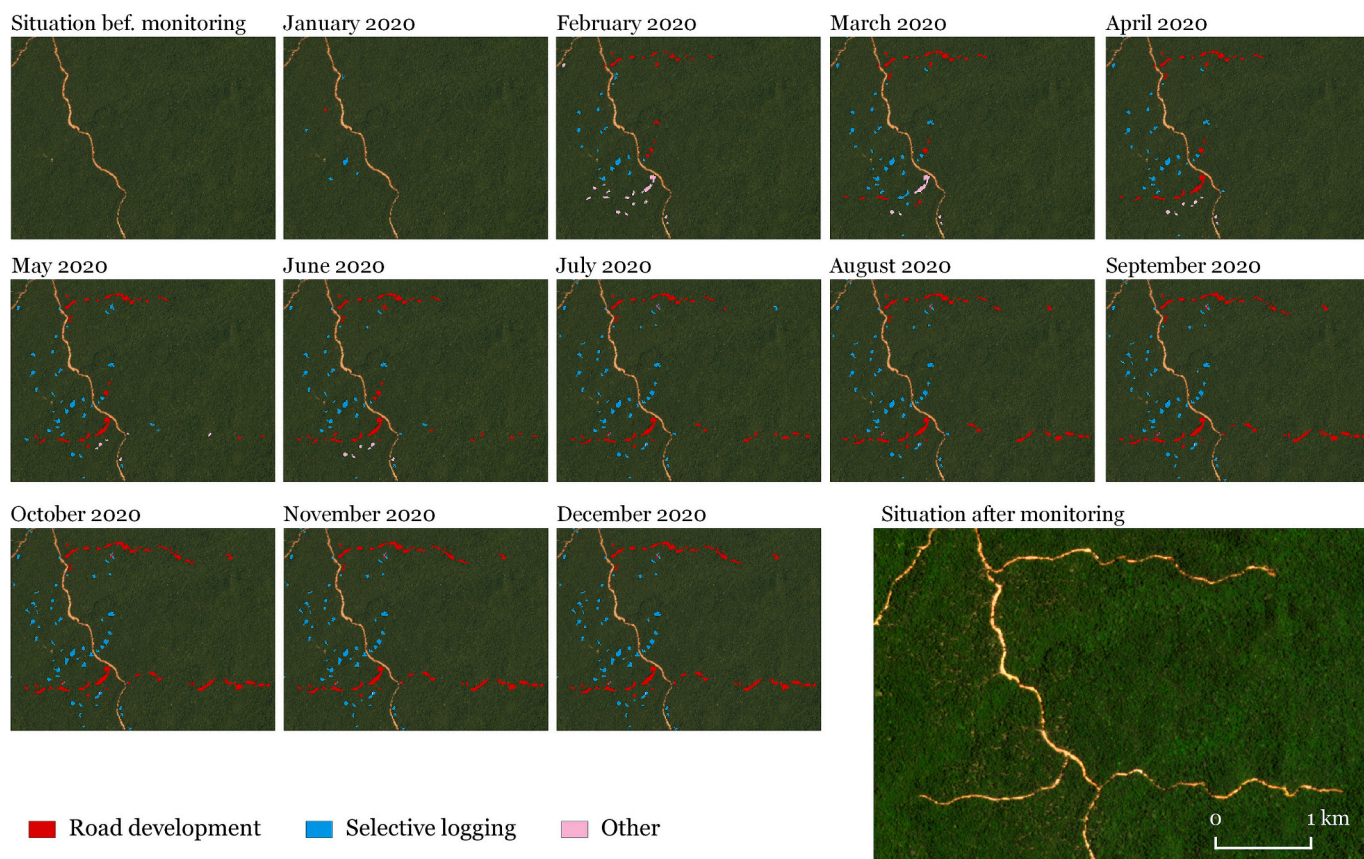

Fig. 7. A simulation of the immediate classification scenario, with driver classifications based on the use of a single post-disturbance Sentinel-1 image. The example shows part of a logging area in Suriname (also presented in Fig. 10D). Monthly outputs were overlaid to show the most recent classification. Planet imagery is shown as a background.

Fig. 8. The relationship between cloud cover levels in the 1-month post-disturbance composites (binned at 5% intervals) and the resulting Macro-F1 scores in the testing areas. 21.8% of the testing data had a cloud cover level over 50%.

classification accuracies (Fig. 4). With a 1-month post-disturbance period, Macro-F1 scores decreased to 0.799 with the use of Sentinel-1 composites and 0.801 with Sentinel-2 composites. Although the Macro F1-scores were comparable for both sensors in this monitoring scenario, Sentinel-1 data showed to be more accurate for classifying *smallholder agriculture* (F1 score: 0.924)*, road development* (F1 score: 0.815) and *selective logging* (F1 score: 0.787), while Sentinel-2 data was more accurate for classifying *mining* (F1 score: 0.737) (Fig. 5). An immediate classification scenario with the use of a single post-disturbance Sentinel-1 image revealed a Macro-F1 score of 0.778 and using no Sentinel imagery at all led to a Macro-F1 score of 0.714. Although rapid classifications lead to lower accuracies, the timeliness of these classifications is an important advantage, as is demonstrated in Fig. 7.

It was observed that the level of cloud cover in the Sentinel-2 composites had a major influence on the resulting accuracies for rapid classifications (Fig. 8). Approximately 21.8% of the 1-month post-disturbance composites were severely cloud-covered, with over 50% of their pixels obstructed by clouds. This subset of data could be classified with a Macro-F1 score of only 0.694 with the use of Sentinel-2, which was substantially lower than what could be achieved with Sentinel-1 over the same subset of cloud-covered data.

The effect of the post-disturbance time period length was more pronounced with the use of Sentinel-2 composites than with Sentinel-1
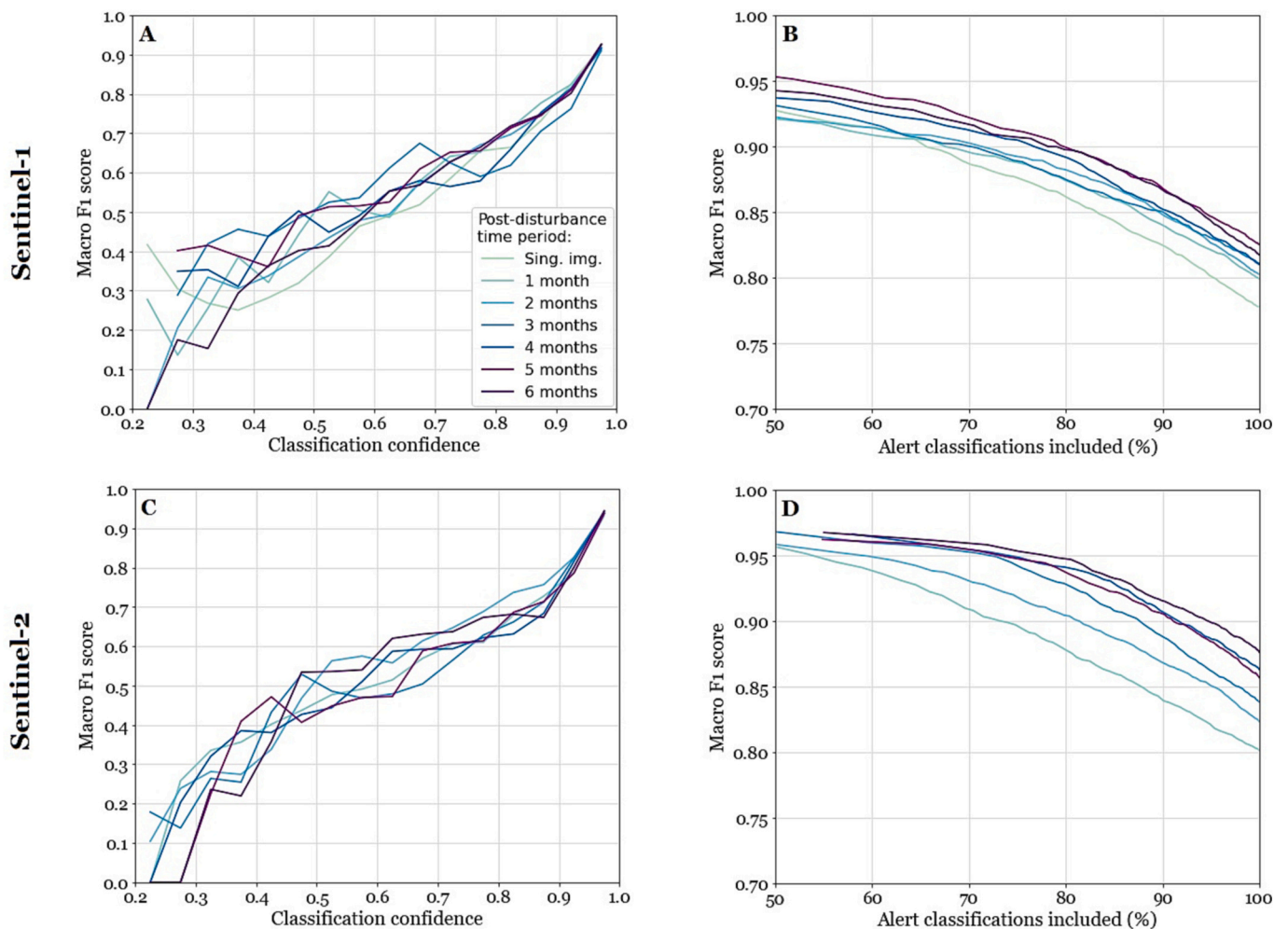

Fig. 9. The relationship between classification confidence based on class probability scores (binned at 5% intervals) and the resulting Macro-F1 scores (A, C) and the effect of applying confidence thresholds to the classification outputs (B, D), based on the use of Sentinel-1 data (A, B) and Sentinel-2 data (C, D) in the testing areas. Applying a confidence threshold increases the classification accuracy (B, D, y-axis) but leads to a lower percentage of alerts included in the classification (B, D, x-axis). Note that the immediate classification scenario based on a single post-disturbance image is only included in the Sentinel-1 data.
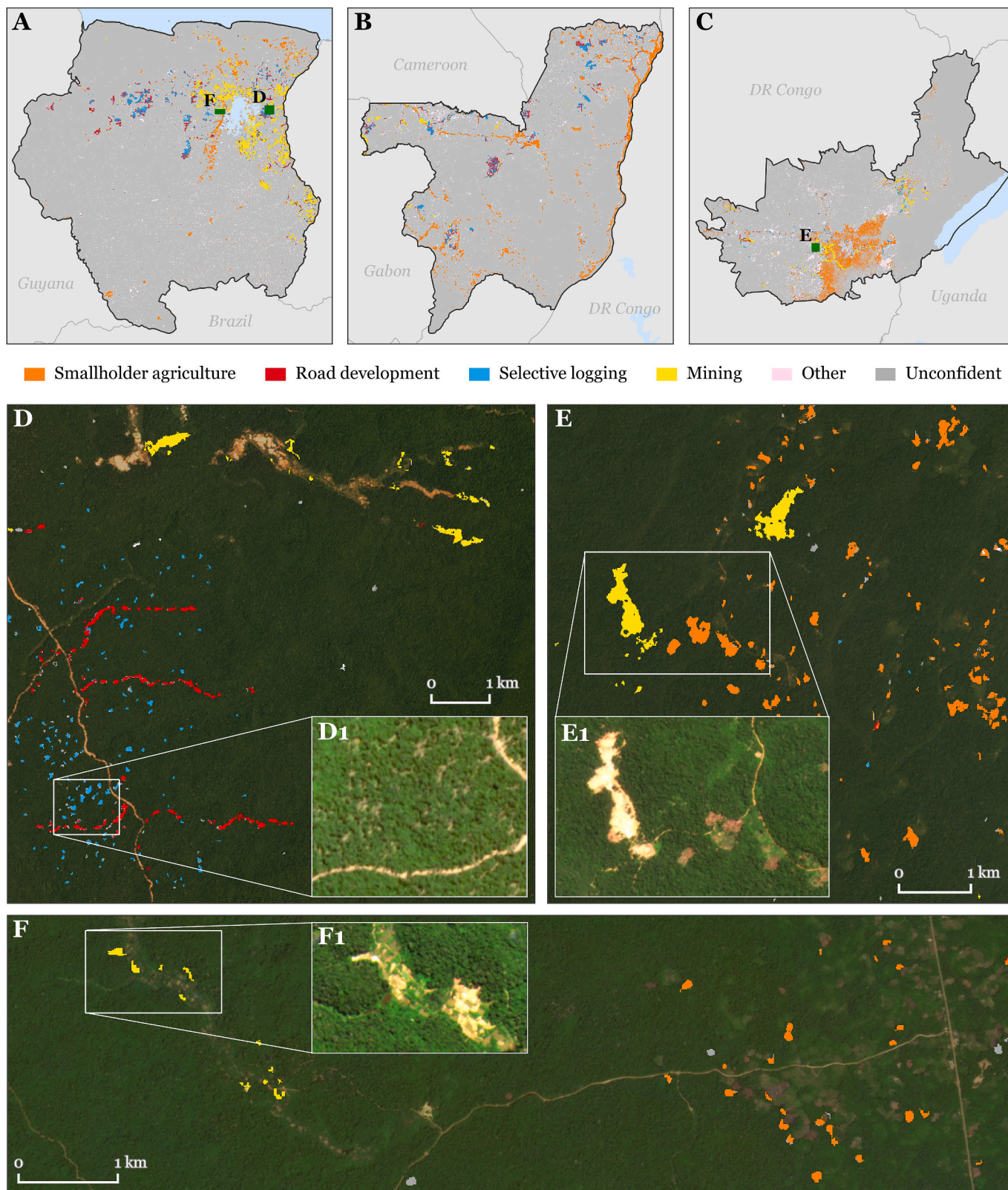
Fig. 10. Output maps of the forest disturbance driver classifications for 2020 and the first half of 2021 (A, B & C), based on biannual classifications with the use of 6-month post-disturbance Sentinel-2 composites. Detailed examples from the held-out testing areas are shown in D, E & F, where the temporal detail was enhanced by producing monthly classification outputs, overlaid with a pixel-based majority vote. Planet imagery is shown as a background with small-scale details added in D1, E1 and F1. A complete overview of the maps can be accessed via: https://bartslagter94.users.earthengine.app/view/forest-disturbance-drivers.

Table 6

The accuracy assessment for the best performing model, based on the use of 6-month post-disturbance Sentinel-2 composites. All accuracies were proportionally adjusted to the strata sizes. The accuracies refer to classifications at the forest disturbance patch level. (UA = User's Accuracy, PA = Producer's Accuracy, OA = Overall Accuracy).

	Sm. agriculture	Road dev.	Sel. logging	Mining	Other	
N samples	744	137	53	231	185	
UA	0.959	0.814	0.803	0.904	0.773	
PA	0.913	0.858	0.786	0.914	0.895	
F1 score	0.935	0.835	0.794	0.909	0.830	
Macro-F1 score						0.861
OA						0.897

composites. The Macro-F1 score difference between the least and most accurate monitoring scenario with the use of composites was 0.026 with Sentinel-1 data and 0.075 with Sentinel-2 data (Fig. 4). With Sentinel-2 composites, longer post-disturbance time periods led to accuracy increases mainly in the range of 1-to-4-month periods, and no substantial accuracy increases were observed for post-disturbance periods longer than 4 months.

Testing accuracy increases were achieved by excluding unconfident classifications, based on the class probability scores as a measure of classification confidence (Fig. 9). For example, keeping only 90% of the classifications in the testing areas with the highest confidence increased the Macro-F1 score to 0.916 with the use of 6-month post-disturbance Sentinel-2 composites. Keeping only the 80% highest confidence classifications increased the Macro-F1 score to 0.947. With a post-disturbance period of 1-month, the Macro-F1 scores increased to 0.839 and 0.878 when keeping only the 90% and 80% highest confidence classifications respectively.

We produced wall-to-wall maps for the three study regions on a biannual basis for alerts detected throughout 2020 and the first half of 2021 (Fig. 10A, B & C). The best performing model, based on the use of 6-month post-disturbance Sentinel-2 composites, was applied for the classifications. For specific regions of interest in our testing areas, we showcased more temporally dense repeated driver classifications on a monthly basis, overlaid with a pixel-based majority vote (Fig. 10D, E & F). Here, a threshold was applied to keep only the 90% highest confidence driver classifications to filter out possible false classifications before the pixel-based majority vote.

The accuracy assessment for the best performing model revealed a Macro-F1 score of 0.861 and an OA of 0.897 throughout the three study regions (Table 6). A similar Macro-F1 score was obtained in testing dataset (0.876). However, the accuracy assessment throughout the entire study regions revealed lower class-specific accuracies for *selective logging,* and higher accuracies for *mining,* compared to what was obtained in the testing areas.

4. Discussion

This study demonstrated the use of Sentinel-1 and Sentinel-2 data for spatially explicit and near real-time monitoring of direct forest disturbance drivers. Our findings confirm the strong capacities of applying a CNN to high spatiotemporal resolution satellite data for rapid classification of newly detected disturbance patches as *smallholder agriculture, road development, selective logging, mining* or *other,* with a Macro-F1 score of 0.861 and an OA of 0.897. Previous studies have applied various approaches for spatially explicit forest disturbance driver classifications, of which some also focused on tropical (dry) forests (De Marzo et al., 2022; Irvin et al., 2020; Masolele et al., 2021, 2022; Shimizu et al., 2019). However, our study is the first to demonstrate near real-time monitoring methods specifically for small-scale disturbance drivers in the tropics.

4.1. Driver classification with Sentinel-1 and Sentinel-2

It was found that the use of Sentinel-2 data led to more accurate results for driver classifications than the use of Sentinel-1 data (Fig. 4). This indicates that the most relevant information to determine a disturbance's driver can be inferred from optical, rather than radar data. Only in the most rapid monitoring scenarios, Sentinel-1 data proved to be more accurate than Sentinel-2, mainly for distinguishing specific classes (Fig. 5), or for classifications in areas with severe cloud cover (Fig. 8).

It was also observed that combining Sentinel-1 and Sentinel-2 data did not lead to increased testing accuracies compared to the use of Sentinel-2 only (Fig. 4), which was contrary to our expectations. The underperformance of the Sentinel-1 and Sentinel-2 combination was not observed during model training (based on the 10% held-out validation samples). Here, the Macro-F1 scores were notably higher than what was achieved for the testing dataset, by 0.081 points on average. This shows that the model based on the Sentinel-1 and Sentinel-2 combination did not transfer as well to the held-out testing areas as the models based on the sensors individually. The models based on the combination of Sentinel-1 and Sentinel-2 have most likely started to slightly overfit on input bands that may not have had much added value for the classifications. Better accuracies may be achieved with multi-sensor input data when more advanced deep learning methods are used, such as multi-modal CNN architectures (Li et al., 2022).

Throughout all monitoring scenarios, *smallholder agriculture* was classified most accurately, with F1 scores ranging up to 0.954 in the testing areas (Fig. 5A) High accuracies for this class may be explained by two reasons. First, almost half of all the training samples represented *smallholder agriculture,* which was the most dominant driver class throughout the study regions. This means that models were trained with a relatively high number of training samples for this class. Second, it is likely that *smallholder agriculture* disturbances are relatively easy to separate from other driver classes based on their distinct features related to geometric patterns, spatial context and post-disturbance backscatter and spectral reflectance. It should be noted that distinguishing *smallholder agriculture* can potentially become more challenging with the use of classification schemes that include additional agriculture-related classes. For example, studies that included classes related to smallholder agriculture as well as classes such as large-scale agriculture, plantations or pasture observed more class confusion (Irvin et al., 2020; Masolele et al., 2021, 2022; Shimizu et al., 2019). Such related classes were not included in our study.

The logging-related driver classes *road development* and *selective logging* showed most confusion in the testing areas (Fig. 5B & C, Table 5). This result was somewhat expected, since these forest disturbance drivers are often co-located, and their distinctive features are challenging to derive from satellite imagery. This is especially true when main logging roads transit into secondary roads or are connected to linear disturbances caused by skidding, where the class boundaries between *road development* and *selective logging* can become somewhat ambiguous. The ambiguity can potentially be solved by combining these

two classes into one overarching *logging* class, which was also done in other driver classification studies in the tropics (De Marzo et al., 2022; Shimizu et al., 2019). In these studies, based on Landsat data, *logging* still proved to be a problematic class to distinguish. In our study, the combined *logging* class could be classified with a high F1 score of 0.940 with Sentinel-2 data in the testing areas (derived from Table 5). This demonstrates the promising capacities of Sentinel-2 data to distinguish logging activities in general from other forest disturbances.

Mining was generally classified with the lowest accuracies and was mostly confused with *road development* and *smallholder agriculture* in the testing areas (Fig. 5D, Table 5). The confusion with *road development* can be explained by the fact that both disturbance drivers are often associated with the post-disturbance appearance of bright bare soil in linear shapes. We observed that wider roads, in particular, were occasionally classified as *mining*. The confusion of *mining* and *smallholder agriculture* was mostly observed in DRC, where these drivers are often co-located, and especially with the use of 1-month post-disturbance composites.

The independent accuracy assessment for the best performing model, based on the use 6-month post-disturbance Sentinel-2 composites, showed that our methods are well generalizable at large-scale, with a Macro-F1 score of 0.861 and an OA of 0.897 (Table 6). The Macro-F1 score obtained for the testing dataset was similar (0.876), but class-specific accuracies differed slightly. Lower accuracies for *selective logging* were obtained in the accuracy assessment compared to what was obtained for the testing dataset, showing that this class is more challenging to distinguish in large-scale generalizations of our methods. The accuracy assessment also showed a substantially higher UA for *mining* than what was obtained for the testing dataset. This shows that false mining classifications were more abundant in our selected testing areas, and accuracies are higher for model applications at large scale. The accuracy assessment also revealed a large number of *smallholder agriculture* disturbances misclassified as *other*.

In operational monitoring set-ups, combining multiple single-moment classification outputs could lead to improved disturbance driver map accuracies. We demonstrated the use of a pixel-based majority vote applied to monthly patch-based classification outputs (Fig. 10D, E & F). However, various advanced methods are possible to produce more confident outputs over time. For example, single-moment class probability scores of the latest detected alerts can be used in sequence to increase or decrease classification confidence during monitoring.

Noteworthy is the moderate classification accuracy that was already reached without any added satellite imagery (Macro-F1 score: 0.714). This classification was solely based on geometric features of the detected RADD alerts, such as shape, size and temporal development. This indicates that some relevant disturbance features can already be inferred from geometrical patterns, as was also demonstrated in earlier studies (Alonso et al., 2022; Hermosilla et al., 2015; Senf and Seidl, 2021). Still, adding only a single Sentinel-1 image as an additional input to the classification increased the Macro-F1 score to 0.778, and adding Sentinel-2 composites increased accuracies even further. This demonstrates the added value of including post-disturbance backscatter and spectral information associated with the different forest disturbance drivers. For example, *mining* disturbances are characterized by the appearance of bright bare soil, which may be distinguished by relatively high Green and Red reflectance and low backscatter, while *smallholder agriculture* disturbances are characterized by the appearance of growing crops, which may be distinguished by relatively high NIR reflectance and high backscatter. Moreover, the added imagery reveals the spatial context of a disturbance, such as the presence of settlements, roads or rivers.

The importance of geometric features for the classifications suggests that an accurate spatial representation of the defined forest disturbance alerts is critical. We observed that the RADD alerts are not always optimal for accurately delineating disturbance shapes, especially for small-scale disturbances related to road development and selective logging. In the RADD alerts, small-scale road developments are occasionally represented as irregularly placed patches instead of continuous linear shapes. In areas with a complex mix of disturbances related to tree felling, skidding and logging road construction, this challenges the classification because *road development* and *selective logging* become undistinguishable based on alert shapes, and the classification has to be entirely inferred from small-scale details in the added post-disturbance composites. Further studies could apply driver classification methods based on underlying forest disturbance alerts with more accurate spatial representations. This could potentially be achieved by complementing radar-based with optical-based disturbance detection methods. For example, initial results from the GLAD-S2 system (Pickens et al., 2020) indicate better representations of road developments as linear shapes at 10 m spatial resolution.

Other improvements can be achieved for near real-time driver classifications by incorporating higher spatiotemporal resolution satellite data. For example, in the optical domain, monthly Planet mosaics (4.8 m spatial resolution) released through the NICFI program now present more spatial detail than Sentinel-2 data. This enables opportunities for even more detailed insights in small-scale forest disturbances and their direct drivers (Csillik et al., 2019; Masolele et al., 2022).

The absence of consistently collected high thematic quality ground-truth data on forest disturbance drivers was a main limitation for this study. We used 4.8 m spatial resolution monthly Planet mosaics to collect reference data, as has been done in several other studies related to forest disturbance mapping (e.g. Reiche et al., 2021; Zhao et al., 2022). Unfortunately, the spatial resolution of Planet imagery does not always suffice to confidently identify or distinguish lower magnitude changes, such as charcoal production, fuelwood collection or skidding. For this reason, we have chosen a rather limited level of thematic detail, with only five generic driver classes that were feasible to distinguish with high confidence based on visual interpretation of Planet imagery. With high-quality ground-truth data, future studies may be able to better assess remote sensing-based forest disturbance driver classification methods.

4.2. Driver monitoring in near real-time

We simulated several near real-time monitoring scenarios based on the use of Sentinel composites derived over different post-disturbance time periods, ranging from 1 to 6 months. The length of the post-disturbance time period was found to have a high influence on the classification accuracies. The use of longer post-disturbance time periods (4 to 6 months) led to the highest classification accuracies (Fig. 4). As was expected, classification accuracies decreased towards more rapid monitoring scenarios.

Rapid driver classifications with the use of a single post-disturbance Sentinel-1 image were already relatively accurate, with a Macro-F1 score of 0.778 (Fig. 4). However, compositing data over a 1-month post-disturbance period increased the accuracy to a Macro-F1 score of 0.801. The best-performing sensor for rapid driver classifications differed per class (Fig. 5). With a 1-month post-disturbance composite, Sentinel-1 data was more accurate for rapidly classifying *smallholder agriculture, road development* and *selective logging,* while Sentinel-2 data was more accurate for classifications of *mining*. It cannot be confidently

stated whether this difference is caused by specific driver features observed in radar and optical imagery. Higher classification accuracies with the use of Sentinel-1 data compared to Sentinel-2 data may also be caused simply by negative effects of cloud cover in Sentinel-2 data with such a short post-disturbance time period. It was observed that Sentinel-2 data composited over a 1-month post-disturbance period had on average 23.5% of its pixels obstructed by clouds, while this was only 0.03% in the 6-month composites. In case Sentinel-2 composites were severely cloud-covered, the CNN was expected to learn classifications based on the best available information, which was in such cases the RADD alert's geometric features. It was found that classifications based on severely cloud-covered Sentinel-2 composites led to substantially lower accuracies (Fig. 8). Sentinel-1 data is unaffected by cloud cover and would be the preferred sensor to rapidly classify drivers in areas with persistent cloud cover.

With the use of Sentinel-2 data, increasing the post-disturbance time period length had the most pronounced effect on the accuracies, especially in the range of 1-to-4-month periods, with minor improvements beyond this length (Fig. 4). This indicates that 4 months after a disturbance, most features related to a direct driver are already well represented in Sentinel-2 composites, as well as that sufficient cloud-free imagery can be sourced to produce a quality composite as a basis for the classification. In contrast to Sentinel-2 data, there were no substantial accuracy increases observed for extending the post-disturbance time period for Sentinel-1 data (Fig. 4). This indicates that any backscatter-related disturbance driver features are already well represented in Sentinel-1 composites produced over a short post-disturbance time period.

Increasing the post-disturbance time period length had a positive effect on accuracies of *selective logging* and *mining* in particular (Fig. 5). For *selective logging*, most class confusion was observed with *road development* in rapid monitoring scenarios, and this confusion decreased with the use longer post-disturbance time periods (Table 5). This may be explained by the fact that canopy and understory regrowth commonly appear within months at *selective logging* disturbances, while *road development* disturbances have a longer visibility duration. This contrast becomes stronger after several post-disturbance months, and this may lead to better distinction of these classes. For *mining*, class confusion with co-located *smallholder agriculture* disturbances was particularly noticeable with a 1-month post-disturbance period, and this confusion decreased with the use of longer periods (Table 5). Based on this observation, we assume that these two classes cannot be accurately separated rapidly after a disturbance; multiple post-disturbance months need to pass before it can be determined whether either *smallholder agriculture* or *mining* develops at a disturbance location.

4.3. Driver monitoring for specific user needs

Throughout all monitoring scenarios, higher Macro-F1 scores were achieved when a threshold was applied to the classification confidence, based on the class probability scores (Fig. 9B & D). By adjusting this confidence threshold, possibly combined with selecting a desired post-disturbance time period, our methods can be adapted to serve specific user needs. For example, when a user requires a certain classification timeliness, minimum accuracy, or has an interest in a specific driver, the methods could be adapted to maximize their usability. Here, we use three example use cases to demonstrate how the monitoring methods can be tuned for targeted applications (Table 7).

The first use case is to support rapid law enforcement activities against illegal logging. Although determining the legality of disturbances is complex, rapidly knowing their direct drivers is already a major advantage (Weisse et al., 2019). In this case, a user would focus on rapidly and confidently distinguishing forest disturbances driven by *road development* and *selective logging*. These two classes are commonly both related to logging activities and their distinction is not needed for rapid law enforcement. For this reason, these classes are combined in this scenario to classify one overarching *logging* class, including disturbances related to tree felling, skidding and logging road construction, with higher accuracy. The classifications need to be rapid and require a high UA to avoid false alarms. We demonstrate two scenarios: An immediate classification scenario based on a single post-disturbance Sentinel-1 image and a scenario based on a 1-month post-disturbance Sentinel-1 composite, both with a confidence threshold optimized to reach a high UA of 0.975. For the immediate classification based on a single image, the required UA is achieved with a minimum class probability score of 0.830 and leads to a PA of 0.735 (Table 7). For the scenario based on a 1-month post-disturbance composite, the required UA is achieved with a minimum class probability score of 0.635 and leads to a PA of 0.813 (Table 7).

The second use case is to support ecological impact assessments of forest disturbances related to mining. Mining activities in tropical forests can lead to severe air, soil and water pollution and impacts can be long-lasting (Alvarez-Berrios and Mitchell Aide, 2015). In this case, a user is interested in planning dedicated field campaigns to assess or measure mining impacts on the ground. To avoid field campaigns targeted at the wrong sites, classifications of *mining* require a high UA in combination with a reasonable classification timeliness. Therefore, a post-disturbance period of 3 months was selected, and a confidence threshold was optimized to reach a UA of 0.850. For this scenario, we propose the use of Sentinel-2 data for the most accurate results. The required UA of 0.850 is achieved with a minimum class probability score of 0.865 and leads to a PA of 0.744 (Table 7).

The third use case is to support timely and frequent carbon emission reporting. Carbon emissions associated with small-scale forest disturbances differ per disturbance type (Houghton, 2012; Umunay et al., 2019). Therefore, quantifications of emissions require accurate accounting of disturbances and their underlying direct drivers (Csillik et al., 2022). In this case, a user is interested in mapping all driver classes as accurate as possible and obtaining a good balance of over- and underestimations of the reported areas. Classification timeliness is less important (commonly, carbon emission reporting is done on an annual basis). Therefore, a post-disturbance period of 6 months was selected,

Table 7
Use cases related to rapid law enforcement, ecological impact assessments and carbon emission reporting. For each monitoring scenario, a desired classification timeliness and User's Accuracy (UA) were defined, and the associated Producer's Accuracy (PA) and required minimum class probability score (as a measure of classification confidence) were computed in the testing areas. Classification omissions due to the application of a confidence threshold are reflected in the PA.

Use case	Classification timeliness	Proposed sensor	Relevant drivers	UA	PA	Applied minimum class probability
Rapid law enforcement	Immediate (single image)	Sentinel-1	Road dev. + Sel. logging	0.975	0.735	0.830
	1 month	Sentinel-1	Road dev. + Sel. logging	0.975	0.813	0.635
Ecological impact assessments	3 months	Sentinel-2	Mining	0.850	0.744	0.865
Carbon emission reporting	6 months	Sentinel-2	Sm. agriculture	0.985	0.924	None
			Road dev.	0.856	0.843	
			Sel. logging	0.893	0.858	
			Mining	0.764	0.899	

and all alert classifications were included regardless of their classification confidence. In this use case, we propose the use of Sentinel-2 data for the most accurate results. The average UA and PA are respectively 0.875 and 0.881 (Table 7).

4.4. Next steps for implementation in operational alerting systems

Our methods have proven to be robust across a range of tropical forest conditions and could potentially be implemented in operational forest disturbance alerting systems. For integration in alerting systems, our driver classification methods may need several further optimizations. First, a comprehensive forest disturbance driver classification scheme needs to be used that is applicable throughout all tropical regions. This could include extra classes beyond the five driver classes presented in this study, such as industrial agriculture, wildfire or riparian changes. Second, models may need additional training with samples that represent a wider variety of tropical regions and forest change conditions. Third, method optimizations need to be explored based on region-specific characteristics. For example, the choice for the optimal sensor or post-disturbance time period may differ per region based on environmental conditions, such as cloud cover levels.

Adding extra classes will pose new challenges for near real-time driver monitoring. For example, it may be ambiguous to define what is considered smallholder agriculture versus industrial agriculture. For this distinction, previous studies adopted definitions mainly based on disturbance size, shape and context (rectangular shapes, linear boundaries, homogeneous vegetation, organized infrastructure, straight roads) (e.g. Descals et al., 2021). Furthermore, disturbances may initially start as smallholder agriculture and be classified as such, but eventually grow to larger-scale industrial agriculture. Pantropical near real-time alerting systems will need to clearly define how such complex disturbance dynamics are handled over time.

It is important to mention the complications of our study related to the application of deep learning methods for the classifications. Deep learning methods are highly suitable to incorporate the complex features related to forest disturbance drivers and their spatial context to obtain accurate classifications. However, there are limitations of deep learning methods, such as limited model interpretability, the need for large amounts of training data and the demand for high computational power.

5. Conclusion

Near real-time monitoring of small-scale tropical forest disturbances and their direct drivers is challenging but has recently become possible with high spatiotemporal resolution satellite data. In this study, we demonstrated the use of Sentinel-1 and Sentinel-2 data to monitor direct drivers of small-scale forest disturbance spatially explicit and in near real-time. We trained a convolutional neural network to classify forest disturbance alerts as *smallholder agriculture, road development, selective logging, mining* or *other* and obtained a Macro-F1 score of 0.861 and an Overall Accuracy of 0.897. It was found that Sentinel-2 data is more suitable than Sentinel-1 data for classifying drivers, except for rapid

classifications of specific classes or in areas with persistent cloud cover. Driver classifications are most accurate with the use of Sentinel-2 data composited over a post-disturbance period of 4 to 6 months. Although more rapid driver classifications lead to lower accuracies, they can still be valuable when timeliness is of high importance.

The value of the different monitoring scenarios was demonstrated for three use cases related to rapid law enforcement against illegal activities, ecological impact assessments and carbon emission reporting. Implementations of our methods in operational forest disturbance alerting systems could contribute to better alert prioritization and provide more actionable and enhanced information for users.

CRediT authorship contribution statement

Bart Slagter: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Visualization, Funding acquisition. **Johannes Reiche:** Conceptualization, Methodology, Writing – review & editing, Supervision, Funding acquisition. **Diego Marcos:** Methodology, Writing – review & editing. **Adugna Mullissa:** Methodology, Writing – review & editing. **Etse Lossou:** Investigation, Writing – review & editing. **Marielos Peña-Claros:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition. **Martin Herold:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix

Appendix 1: Overview of the testing accuracies of the different monitoring scenarios. All class-specific User's Accuracies (UA), Producer's Accuracies (PA), F1 scores and the final Macro-F1 scores are presented per sensor and post-disturbance time period.

		Sentinel-1			Sentinel-2			Sentinel-1 & Sentinel-2		
		UA	PA	F1	UA	PA	F1	UA	PA	F1
Single image	Sm. agriculture	0.975	0.843	0.904	NA	NA	NA	NA	NA	NA
	Road dev.	0.816	0.802	0.809	NA	NA	NA	NA	NA	NA
	Sel. logging	0.713	0.780	0.745	NA	NA	NA	NA	NA	NA
	Mining	0.549	0.804	0.652	NA	NA	NA	NA	NA	NA
	Macro-F1 score			**0.778**			NA			NA
1 month	Sm. agriculture	0.969	0.883	0.924	0.977	0.867	0.919	0.977	0.857	0.913
	Road dev.	0.821	0.808	0.815	0.766	0.824	0.794	0.793	0.781	0.787
	Sel. logging	0.835	0.745	0.787	0.765	0.745	0.754	0.741	0.755	0.748
	Mining	0.569	0.815	0.670	0.689	0.793	0.737	0.588	0.842	0.692
	Macro-F1 score			**0.799**			**0.801**			**0.785**
2 months	Sm. agriculture	0.970	0.887	0.926	0.984	0.885	0.932	0.983	0.860	0.918
	Road dev.	0.808	0.794	0.801	0.800	0.819	0.809	0.783	0.806	0.794
	Sel. logging	0.795	0.776	0.785	0.759	0.789	0.774	0.883	0.697	0.779
	Mining	0.634	0.774	0.697	0.738	0.822	0.778	0.554	0.912	0.689
	Macro-F1 score			**0.802**			**0.823**			**0.795**
3 months	Sm. agriculture	0.969	0.903	0.935	0.983	0.888	0.933	0.984	0.880	0.929
	Road dev.	0.825	0.811	0.818	0.818	0.840	0.829	0.817	0.816	0.816
	Sel. logging	0.862	0.741	0.797	0.837	0.791	0.813	0.897	0.728	0.804
	Mining	0.580	0.857	0.692	0.710	0.862	0.779	0.589	0.921	0.719
	Macro-F1 score			**0.810**			**0.838**			**0.817**
4 months	Sm. agriculture	0.963	0.912	0.937	0.989	0.916	0.951	0.988	0.897	0.940
	Road dev.	0.860	0.756	0.805	0.846	0.835	0.840	0.804	0.819	0.811
	Sel. logging	0.852	0.709	0.774	0.878	0.820	0.848	0.861	0.773	0.814
	Mining	0.654	0.813	0.725	0.727	0.920	0.813	0.701	0.931	0.800
	Macro-F1 score			**0.810**			**0.863**			**0.841**
5 months	Sm. agriculture	0.973	0.919	0.946	0.987	0.914	0.949	0.992	0.905	0.947
	Road dev.	0.841	0.795	0.817	0.847	0.832	0.839	0.797	0.827	0.812
	Sel. logging	0.855	0.763	0.806	0.866	0.832	0.849	0.824	0.810	0.817
	Mining	0.642	0.851	0.732	0.694	0.918	0.791	0.715	0.918	0.804
	Macro-F1 score			**0.825**			**0.857**			**0.845**
6 months	Sm. agriculture	0.978	0.905	0.941	0.985	0.924	0.954	0.983	0.926	0.954
	Road dev.	0.787	0.806	0.797	0.856	0.843	0.850	0.835	0.817	0.826
	Sel. logging	0.812	0.790	0.801	0.893	0.858	0.875	0.879	0.765	0.818
	Mining	0.656	0.826	0.731	0.764	0.899	0.826	0.710	0.930	0.805
	Macro-F1 score			**0.817**			**0.876**			**0.851**

References

Alonso, L., Picos, J., Armesto, J., 2022. Automatic identification of forest disturbance drivers based on their geometric pattern in Atlantic forests. Remote Sens. 14 https://doi.org/10.3390/rs14030697.

Alvarez-Berrios, N.L., Mitchell Aide, T., 2015. Global demand for gold is another threat for tropical forests. Environ. Res. Lett. https://doi.org/10.1088/1748-9326/10/1/014006.

Asner, G.P., Keller, M., Pereira, R., Zweede, J.C., Silva, J.N.M., 2004. Canopy damage and recovery after selective logging in Amazonia: field and satellite studies. Ecol. Appl. 14, 280–298. https://doi.org/10.1890/01-6019.

Barlow, J., Lennox, G.D., Ferreira, J., Berenguer, E., Lees, A.C., Nally, R.Mac, Thomson, J.R., Ferraz, S.F.D.B., Louzada, J., Oliveira, V.H.F., Parry, L., Ribeiro De Castro Solar, R., Vieira, I.C.G., Aragaõ, L.E.O.C., Begotti, R.A., Braga, R.F., Cardoso Jr., T.M., Souza, C.M., Moura, N.G., Nunes, S.S., Siqueira, J.V., Pardini, R., Silveira, J.M., Vaz-De-Mello, F.Z., Veiga, R.C.S., Venturieri, A., Gardner, T.A., R.C.D.O, 2016. Anthropogenic disturbance in tropical forests can double biodiversity loss from deforestation. Nature. https://doi.org/10.1038/nature18326.

Bouvet, A., Mermoz, S., Ballère, M., Koleck, T., Le Toan, T., 2018. Use of the SAR shadowing effect for deforestation detection with Sentinel-1 time series. Remote Sens. 10, 1–19. https://doi.org/10.3390/rs10081250.

Cardille, J.A., Perez, E., Crowley, M.A., Wulder, M.A., White, J.C., Hermosilla, T., 2022. Multi-sensor change detection for within-year capture and labelling of forest disturbance. Remote Sens. Environ. 268, 112741 https://doi.org/10.1016/j.rse.2021.112741.

Csillik, O., Kumar, P., Mascaro, J., O'Shea, T., Asner, G.P., 2019. Monitoring tropical forest carbon stocks and emissions using planet satellite data. Sci. Rep. 9, 1–12. https://doi.org/10.1038/s41598-019-54386-6.

Csillik, O., Reiche, J., De Sy, V., Araza, A., Herold, M., 2022. Rapid remote monitoring reveals spatial and temporal hotspots of carbon loss in Africa's rainforests. Commun. Earth Environ. https://doi.org/10.1038/s43247-022-00383-z.

Curtis, P.G., Slay, C.M., Harris, N.L., Tyukavina, A., Hansen, M.C., 2018. Classifying drivers of global forest loss. Science 1979 (361), 1108–1111. https://doi.org/10.1126/science.aau3445.

De Marzo, T., Gasparri, N.I., Lambin, E.F., Kuemmerle, T., 2022. Agents of forest disturbance in the argentine dry Chaco. Remote Sens. 14 https://doi.org/10.3390/rs14071758.

De Sy, V., Herold, M., Achard, F., Avitabile, V., Baccini, A., Carter, S., Clevers, J.G.P.W., Lindquist, E., Pereira, M., Verchot, L., 2019. Tropical deforestation drivers and associated carbon emission factors derived from remote sensing data. Environ. Res. Lett. 14 https://doi.org/10.1088/1748-9326/ab3dc6.

Descals, A., Wich, S., Meijaard, E., Gaveau, D.L.A., Peedell, S., Szantoi, Z., 2021. High-resolution global map of smallholder and industrial closed-canopy oil palm plantations. Earth Syst. Sci. Data 13, 1211–1231. https://doi.org/10.5194/essd-13-1211-2021.

Diniz, C.G., Souza, A.A.D.A., Santos, D.C., Dias, M.C., Luz, N.C.Da., Moraes, D.R.V.De, Maia, J.S.A., Gomes, A.R., Narvaes, I.D.S., Valeriano, D.M., Maurano, L.E.P., Adami, M., 2015. DETER-B: the new Amazon near real-time deforestation detection system. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. https://doi.org/10.1109/JSTARS.2015.2437075.

Doblas, J., Shimabukuro, Y., Sant'Anna, S., Carneiro, A., Aragão, L., Almeida, C., 2020. Optimizing near real-time detection of deforestation on tropical rainforests using Sentinel-1 data. Remote Sens. (Basel) 12, 3922. https://doi.org/10.3390/rs12233922.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote Sens. Environ. 120, 25–36. https://doi.org/10.1016/j.rse.2011.11.026.

European Space Agency, 2022. SNAP - ESA Sentinel Application Platform.

Finer, B.M., Novoa, S., Weisse, M.J., Petersen, R., Mascaro, J., Souto, T., Stearns, F., Martinez, R.G., 2018. Combating deforestation: from satellite to intervention. Science 1979 (360), 1303–1305. https://doi.org/10.1126/science.aat1203.

Geist, H.J., Lambin, E.F., 2002. Proximate causes and underlying driving forces of tropical deforestation. Bioscience. https://doi.org/10.1641/0006-3568(2002)052[0143:PCAUDF]2.0.CO;2.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031.

Government of Suriname, 2018. Forest Reference Emission Level for Suriname's REDD+ Programme.

Hansen, M.C., Krylov, A., Tyukavina, A., Potapov, P.V., Turubanova, S., Zutta, B., Ifo, S., Margono, B., Stolle, F., Moore, R., 2016. Humid tropical forest disturbance alerts using Landsat data. Environ. Res. Lett. 11 https://doi.org/10.1088/1748-9326/11/3/034008.

Harris, N.L., Gibbs, D.A., Baccini, A., Birdsey, R.A., de Bruin, S., Farina, M., Fatoyinbo, L., Hansen, M.C., Herold, M., Houghton, R.A., Potapov, P.V., Suarez, D.R., Roman-Cuesta, R.M., Saatchi, S.S., Slay, C.M., Turubanova, S.A., Tyukavina, A., 2021. Global maps of twenty-first century forest carbon fluxes. Nat Clim. Chang. 11, 234–240. https://doi.org/10.1038/s41558-020-00976-6.

Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., 2015. Regional detection, characterization, and attribution of annual forest change from 1984 to 2012 using Landsat-derived time-series metrics. Remote Sens. Environ. 170, 121–132. https://doi.org/10.1016/j.rse.2015.09.004.

Hoekman, D., Kooij, B., Quiñones, M., Vellekoop, S., Carolita, I., Budhiman, S., Arief, R., Roswintiarti, O., 2020. Wide-area near-real-time monitoring of tropical forest degradation and deforestation using Sentinel-1. Remote Sens. 12, 1–32. https://doi.org/10.3390/rs12193263.

Hoekman, D.H., Reiche, J., 2015. Multi-model radiometric slope correction of SAR images of complex terrain using a two-stage semi-empirical approach. Remote Sens. Environ. 156 https://doi.org/10.1016/j.rse.2014.08.037.

Houghton, R.A., 2012. Carbon emissions and the drivers of deforestation and forest degradation in the tropics. Curr. Opin. Environ. Sustain. 4 https://doi.org/10.1016/j.cosust.2012.06.006.

Huo, L.Z., Boschetti, L., Sparks, A.M., 2019. Object-based classification of forest disturbance types in the conterminous United States. Remote Sens. 11 https://doi.org/10.3390/rs11050477.

IPIS, 2016. Analysis of the Interactive Map of Artisanal Mining Areas in Eastern DR Congo: 2015 Update.

Irvin, J., Sheng, H., Ramachandran, N., Johnson-Yu, S., Zhou, S., Story, K., Rustowicz, R., Elsworth, C., Austin, K., Ng, A.Y., 2020. ForestNet: Classifying Drivers of Deforestation in Indonesia using Deep Learning on Satellite Imagery.

Kalamandeen, M., Gloor, E., Mitchard, E., Quincey, D., Ziv, G., Spracklen, D., Spracklen, B., Adami, M., Aragaõ, L.E.O.C., Galbraith, D., 2018. Pervasive rise of small-scale deforestation in Amazonia. Sci. Rep. 8 https://doi.org/10.1038/s41598-018-19358-2.

Kingma, D.P., Ba, J., 2015. Adam: A method for stochastic optimization, in: 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings.

Kleinschroth, F., Healey, J.R., 2017. Impacts of logging roads on tropical forests. Biotropica. https://doi.org/10.1111/btp.12462.

Kleinschroth, F., Laporte, N., Laurance, W.F., Goetz, S.J., Ghazoul, J., 2019. Road expansion and persistence in forests of the Congo Basin. Nat. Sustain 2, 628–634. https://doi.org/10.1038/s41893-019-0310-6.

Laso Bayas, J.C., See, L., Georgieva, I., Schepaschenko, D., Danylo, O., Dürauer, M., Bartl, H., Hofhansl, F., Zadorozhniuk, R., Burianchuk, M., Sirbu, F., Magori, B., Blyshchyk, K., Blyshchyk, V., Rabia, A.H., Pawe, C.K., Su, Y.F., Ahmed, M., Panging, K., Melnyk, O., Vasylyshyn, O., Vasylyshyn, R., Bilous, A., Bilous, S., Das, K., Prestele, R., Pérez-Hoyos, A., Bungnamei, K., Lashchenko, A., Lakyda, M., Lakyda, I., Serediuk, O., Domashovets, G., Yurchuk, Y., Koper, M., Fritz, S., 2022. Drivers of tropical forest loss between 2008 and 2019. Sci. Data 9, 1–8. https://doi.org/10.1038/s41597-022-01227-3.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539.

Li, J., Hong, D., Gao, L., Yao, J., Zheng, K., Zhang, B., Chanussot, J., 2022. Deep learning in multimodal remote sensing data fusion: a comprehensive review. Int. J. Appl. Earth Obs. Geoinf. 112 https://doi.org/10.1016/j.jag.2022.102926.

Lynch, J., Maslin, M., Balzter, H., Sweeting, M., 2013. Choose satellites to monitor deforestation. Nature 496, 293–294.

Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., Gascon, F., 2017. Sen2Cor for Sentinel-2. https://doi.org/10.1117/12.2278218.

Masolele, R.N., De Sy, V., Herold, M., Marcos Gonzalez, D., Verbesselt, J., Gieseke, F., Mullissa, A.G., Martius, C., 2021. Spatial and temporal deep learning methods for deriving land-use following deforestation: a pan-tropical case study using Landsat time series. Remote Sens. Environ. 264, 112600 https://doi.org/10.1016/j.rse.2021.112600.

Masolele, R.N., De Sy, V., Marcos, D., Verbesselt, J., Gieseke, F., Mulatu, K.A., Moges, Y., Sebrala, H., Martius, C., Herold, M., 2022. Using high-resolution imagery and deep learning to classify land-use following deforestation: a case study in Ethiopia. GISci. Remote. Sens. 59, 1446–1472. https://doi.org/10.1080/15481603.2022.2115619.

Mermoz, S., Bouvet, A., Koleck, T., Ballère, M., Le Toan, T., 2021. Continuous detection of forest loss in Vietnam, Laos, and Cambodia using Sentinel-1 data. Remote Sens. 13 https://doi.org/10.3390/rs13234877.

Moffette, F., Alix-Garcia, J., Shea, K., Pickens, A.H., 2021. The impact of near-real-time deforestation alerts across the tropics. Nat. Clim. Chang. 11, 172–178. https://doi.org/10.1038/s41558-020-00956-w.

Mullissa, A., Vollrath, A., Odongo-Braun, C., Slagter, B., Balling, J., Gou, Y., Gorelick, N., Reiche, J., 2021. Sentinel-1 SAR backscatter analysis ready data preparation in Google Earth Engine. Remote Sens. 13, 5–11. https://doi.org/10.3390/rs13101954.

Nguyen, T.H., Jones, S.D., Soto-Berelov, M., Haywood, A., Hislop, S., 2018. A spatial and temporal analysis of forest dynamics using Landsat time-series. Remote Sens. Environ. 217, 461–475. https://doi.org/10.1016/j.rse.2018.08.028.

Oeser, J., Pflugmacher, D., Senf, C., Heurich, M., Hostert, P., 2017. Using intra-annual Landsat time series for attributing forest disturbance agents in Central Europe. Forests 8. https://doi.org/10.3390/f8070251.

Pearson, T.R.H., Brown, S., Murray, L., Sidman, G., 2017. Greenhouse gas emissions from tropical forest degradation: an underestimated source. Carbon Balance Manag. 12 https://doi.org/10.1186/s13021-017-0072-2.

Pickens, A.H., Hansen, M.C., Adusei, B., Potapov, P., 2020. Sentinel-2 Forest Loss Alert. Global Land Analysis and Discovery (GLAD) [WWW Document]. URL (accessed 11.18.21). www.globalforestwatch.org.

Piponiot, C., Rödig, E., Putz, F.E., Rutishauser, E., Sist, P., Ascarrunz, N., Blanc, L., Derroire, G., Descroix, L., Guedes, M.C., Coronado, E.H., Huth, A., Kanashiro, M., Licona, J.C., Mazzei, L., D'Oliveira, M.V.N., Peña-Claros, M., Rodney, K., Shenkin, A., De Souza, C.R., Vidal, E., West, T.A.P., Wortel, V., Hérault, B., 2019. Can timber provision from amazonian production forests be sustainable? Environ. Res. Lett. 14 https://doi.org/10.1088/1748-9326/ab195e.

Planet Team, 2017. Planet Application Program Interface: In Space for Life on Earth.

Reiche, J., Mullissa, A., Slagter, B., Gou, Y., Tsendbazar, N., Odongo-Braun, C., Vollrath, A., Weisse, M., Stolle, F., Pickens, A., Donchyts, G., Clinton, N., Gorelick, N., Herold, M., 2021. Forest disturbance alerts for the Congo Basin using Sentinel-1. Environ. Res. Lett. 16, 024005 https://doi.org/10.1088/1748-9326/abd0a8.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat, 2019. Deep learning and process understanding for data-driven Earth system science. Nature 566, 195–204. https://doi.org/10.1038/s41586-019-0912-1.

Schroeder, T.A., Schleeweis, K.G., Moisen, G.G., Toney, C., Cohen, W.B., Freeman, E.A., Yang, Z., Huang, C., 2017. Testing a Landsat-based approach for mapping disturbance causality in U.S. Forests. Remote Sens. Environ. 195, 230–243. https://doi.org/10.1016/j.rse.2017.03.033.

Schroeder, T.A., Wulder, M.A., Healey, S.P., Moisen, G.G., 2011. Mapping wildfire and clearcut harvest disturbances in boreal forests with Landsat time series data. Remote Sens. Environ. 115, 1421–1433. https://doi.org/10.1016/j.rse.2011.01.022.

Sebald, J., Senf, C., Seidl, R., 2021. Human or natural? Landscape context improves the attribution of forest disturbances mapped from Landsat in Central Europe. Remote Sens. Environ. 262, 112502 https://doi.org/10.1016/j.rse.2021.112502.

Senf, C., Seidl, R., 2021. Storm and fire disturbances in Europe: distribution and trends. Glob. Chang. Biol. 27, 3605–3619. https://doi.org/10.1111/gcb.15679.

Shimizu, K., Ota, T., Mizoue, N., Yoshida, S., 2019. A comprehensive evaluation of disturbance agent classification approaches: strengths of ensemble classification, multiple indices, spatio-temporal variables, and direct prediction. ISPRS J. Photogramm. Remote. Sens. 158, 99–112. https://doi.org/10.1016/j.isprsjprs.2019.10.004.

Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition. In: 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, pp. 1–14.

Souza, C.M., Roberts, D.A., Cochrane, M.A., 2005. Combining spectral and spatial information to map canopy damage from selective logging and forest fires. Remote Sens. Environ. 98, 329–343. https://doi.org/10.1016/j.rse.2005.07.013.

Stewart, B.P., Wulder, M.A., McDermid, G.J., Nelson, T., 2009. Disturbance capture and attribution through the integration of Landsat and IRS-1C imagery. Can. J. Remote. Sens. 35, 523–533. https://doi.org/10.5589/m10-006.

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B.Ö., Floury, N., Brown, M., Traver, I.N., Deghaye, P., Duesmann, B., Rosich, B., Miranda, N., Bruno, C., L'Abbate, M., Croci, R., Pietropaolo, A., Huchler, M., Rostan, F., 2012. GMES Sentinel-1 mission. Remote Sens. Environ. 120 https://doi.org/10.1016/j.rse.2011.05.028.

Turubanova, S., Potapov, P.V., Tyukavina, A., Hansen, M.C., 2018. Ongoing primary forest loss in Brazil, Democratic Republic of the Congo, and Indonesia. Environ. Res. Lett. 13 https://doi.org/10.1088/1748-9326/aacd1c.

Tyukavina, A., Hansen, M.C., Potapov, P., Parker, D., Okpa, C., Stehman, S.V., Kommareddy, I., Turubanova, S., 2018. Congo Basin forest loss dominated by increasing smallholder clearing. Sci. Adv. 4 https://doi.org/10.1126/sciadv.aat2993.

Umunay, P.M., Gregoire, T.G., Gopalakrishna, T., Ellis, P.W., Putz, F.E., 2019. Selective logging emissions and potential emission reductions from reduced-impact logging in the Congo Basin. For. Ecol. Manag. 437, 360–371. https://doi.org/10.1016/j.foreco.2019.01.049.

Vargas, C., Montalban, J., Leon, A.A., 2019. Early warning tropical forest loss alerts in Peru using Landsat. Environ. Res. Commun. https://doi.org/10.1088/2515-7620/ab4ec3.

Verhegghen, A., Eva, H., Achard, F., 2015. Assessing forest degradation from selective logging using time series of fine spatial resolution imagery in Republic of Congo. In: International Geoscience and Remote Sensing Symposium (IGARSS). https://doi.org/10.1109/IGARSS.2015.7326202.

Vogeler, J.C., Slesak, R.A., Fekety, P.A., Falkowski, M.J., 2020. Characterizing over four decades of forest disturbance in Minnesota, USA. Forests 11, 1–19. https://doi.org/10.3390/F11030362.

Vollrath, A., Mullissa, A., Reiche, J., 2020. Angular-based radiometric slope correction for Sentinel-1 on Google Earth Engine. Remote Sens. https://doi.org/10.3390/rs12111867.

Weisse, M.J., Noguerón, R., Eduardo, R., Vicencio, V., Arturo, D., Soto, C., 2019. Use of Near-Real-Time Deforestation Alerts: a Case Study From Peru.

Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., Xu, H., Tan, W., Yang, Q., Wang, J., Gao, J., Zhang, L., 2020. Deep learning in environmental remote sensing: achievements and challenges. Remote Sens. Environ. 241, 111716 https://doi.org/10.1016/j.rse.2020.111716.

Zhang, Y., Woodcock, C.E., Chen, S., Wang, J.A., Sulla-Menashe, D., Zuo, Z., Olofsson, P., Wang, Y., Friedl, M.A., 2022. Mapping causal agents of disturbance in boreal and arctic ecosystems of North America using time series of Landsat data. Remote Sens. Environ. 272, 112935 https://doi.org/10.1016/j.rse.2022.112935.

Zhao, F., Sun, R., Zhong, L., Meng, R., Huang, C., Zeng, X., Wang, M., Li, Y., Wang, Z., 2022. Monthly mapping of forest harvesting using dense time series Sentinel-1 SAR imagery and deep learning. Remote Sens. Environ. 269 https://doi.org/10.1016/j.rse.2021.112822.

Zupanc, A., 2017. Improving Cloud Detection with Machine Learning [WWW Document]. https://medium.com/sentinel-hub/improving-cloud-detection-with-machine-learning-c09dc5d7cf13.