Remote Sensing of Environment 331 (2025) 115037

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Novel unsupervised Bayesian method for Near Real-Time forest loss detection using Sentinel-1 SAR time series: Assessment over sampled deforestation events in Amazonia and the Cerrado

Marta Bottani [a,b,c,d,*], Laurent Ferro-Famil [b,d], Juan Doblas Prieto [e], Stéphane Mermoz [e], Alexandre Bouvet [d], Thierry Koleck [c,d], Thuy Le Toan [e]

[a] TéSA, 7 boulevard de la Gare, Toulouse, 31500, France
[b] ISAE-Supaero, 10 Avenue Marc Pélegrin, Toulouse, 31400, France
[c] Centre National d'Etudes Spatiales (CNES), 18 Avenue Edouard Belin, Toulouse, 31400, France
[d] CESBIO, University of Toulouse, CNES/CNRS/INRAE/IRD/UT3, 18 Avenue Edouard Belin, Toulouse, 31400, France
[e] GlobEO, 18 Avenue Edouard Belin, Toulouse, 31400, France

ARTICLE INFO

Edited by Marie Weiss

Keywords:
Forest loss
Change detection
Bayesian inference
Sentinel-1
Time series
Rainforest
Savanna

ABSTRACT

Over the past four decades, forests have experienced major disturbances, highlighting the need for Near Real-Time (NRT) monitoring. Traditional optical-based detection is cloud-sensitive, whereas Synthetic Aperture Radar (SAR)-based frameworks enable all-weather observation. Yet, SAR monitoring has mainly focused on humid tropical forests, with reduced performance in regions showing strong seasonal backscatter variation, such as tropical savannas. Detecting small-scale forest loss also remains difficult due to the spatial resolution loss from speckle filtering. This paper presents an unsupervised SAR-based disturbance detection method with NRT capabilities, using Bayesian inference. Building on an existing methodology, the approach processes single-polarization Sentinel-1 SAR time series through Bayesian conjugate analysis. Forest disturbance is framed as a changepoint detection problem, where each new observation updates the probability of forest loss using prior information and a data model. The algorithm uses a hidden Markov chain to adapt recursively to seasonal variation and bypasses spatial filtering, preserving native data resolution and enhancing small-scale forest loss detection. Additionally, a methodology accounts for proximity to past disturbances. The method is tested on two 2020 reference datasets from the Brazilian Amazon and Cerrado savanna. The first covers small validation polygons (0.1–1 ha, excluding selective logging), totaling 2,650 ha in the Amazon and 450 ha in the Cerrado. The second includes larger clearings totaling 11,200 ha in the Amazon, and 12,700 ha in the Cerrado. A further comparison is conducted with operational NRT forest loss monitoring approaches. Results show substantial gains in detecting small-scale disturbances with reduced false alarms. In the Amazon, the method achieves an F1-score of 97.3% versus 93.1% for the current leading NRT approach. In the Cerrado, it reaches an F1-score of 97.4%, far exceeding the 33.3% of the optical-based method. For larger clearings, performance matches existing SAR approaches in the Amazon. While combined optical-SAR monitoring increases true positives, it also raises false alarm rates. In the Cerrado, the proposed method clearly outperforms optical monitoring, and in both regions it improves timeliness relative to individual operational approaches.

1. Introduction

Over the past decades, about 17% of moist tropical forests have been lost (Vancutsem et al., 2021). From 2001 to 2019, forests emitted an average of 8.1 billion metric tonnes of CO_2 due to deforestation and degradation, while absorbing 16 billion through photosynthesis (Harris et al., 2021). Acting as natural carbon sinks, forests remove roughly one-third of annual CO_2 emissions (Forzieri et al., 2022). Their loss not only accelerates global warming by reducing carbon uptake and releasing stored carbon, but also contributes to soil erosion, water cycle disruption, and biodiversity decline through habitat destruction (Hoang and Kanemoto, 2021). Effective tools for timely disturbance detection are therefore essential to enable rapid intervention and support forest conservation.

* Corresponding author at: ISAE-Supaero, 10 Avenue Marc Pélegrin, Toulouse, 31400, France.
E-mail address: marta.bottani@isae-supaero.fr (M. Bottani).

https://doi.org/10.1016/j.rse.2025.115037
Received 29 November 2024; Received in revised form 12 September 2025; Accepted 16 September 2025
Available online 24 September 2025

Earth Observation (EO) data provide effective monitoring of vast and often inaccessible forest areas. In recent years, the number, quality, and accessibility of EO satellites and imagery have greatly increased (Finer et al., 2018). Multiple high-resolution satellite datasets for global forest loss monitoring are now freely available, enabling the development of many operational Near Real-Time (NRT) disturbance detection methods. A pioneer solution in this context is the Global Land Analysis and Discovery system (GLAD-L, Hansen et al., 2016), which was originally based on Landsat imagery and has recently been extended to Sentinel-2 data (GLAD-S2). However, a major limitation of optical-based methods is their sensitivity to cloud cover, which significantly reduces the availability of usable images (Verbesselt et al., 2012). Unlike optical imagery, Synthetic Aperture Radar (SAR) data are insensitive to cloud cover, enabling consistent and regular time series acquisition in tropical regions. An early approach utilizing L-band SAR data (wavelength, $\lambda \approx 24$ cm) from ALOS-2/PALSAR-2 is JJ-FAST (Watanabe et al., 2018, 2021). This method leverages HH and HV polarizations, as well as their ratio, to detect different stages of deforestation based on their polarimetric scattering characteristics. However, the NRT capabilities of JJ-FAST are constrained by the 14-day revisit interval of ALOS-2/PALSAR-2.

Since the launch of the European Space Agency's (ESA) C-band Sentinel-1 mission (wavelength $\lambda \approx 5.6$ cm) in 2014, several SAR-based frameworks with enhanced NRT capabilities have been developed, some of which are now operational across the tropics. A recent deforestation monitoring approach using Sentinel-1 data applies the CuSum (Cumulative Sum) algorithm to dual-polarization VV-VH observations to detect changes in vegetation cover (Ygorra et al., 2021). While computationally efficient, its performance depends heavily on parameter tuning tailored to specific regional conditions, which can limit its generalizability. Another example of a method leveraging Sentinel-1 data is DETER-R (Doblas et al., 2020, 2022), which is currently operational in the Brazilian Amazon. This approach employs the Adaptive Linear Threshold (ALT) algorithm to identify pixels exhibiting backscatter values below a dynamically set threshold, indicative of potential deforestation. Additionally, TropiSCO (Mermoz et al., 2021; Ballère et al., 2021) is a fully Sentinel-1-based method that detects forest loss by identifying shadow patterns formed at the boundaries between intact forest and cleared areas (Bouvet et al., 2018).

In the tropics, deforestation — defined as the permanent conversion of forest to other land uses — results from a variety of drivers with distinct characteristics, scales, and management practices. The primary driver is agricultural expansion, which ranges from small-scale subsistence farming to large-scale commercial operations. Other contributing factors include infrastructure development and urban growth, while activities such as timber logging and fires are typically classified as forest degradation rather than permanent deforestation (Bourgoin et al., 2024). Despite the complexity and diversity of deforestation practices, most SAR-based forest loss detection methods rely on thresholds to identify abrupt changes (Bouvet et al., 2018; Mermoz et al., 2021; Ballère et al., 2021; Doblas et al., 2022; Carstairs et al., 2022). However, this approach can fail to accurately detect forest loss in scenarios where SAR backscatter varies significantly before and after disturbance due to factors such as seasonality, weather conditions, and post-logging management activities (Zhao et al., 2019). Threshold-based methods generally perform well in dense rainforests where SAR backscatter remains relatively stable. However, in regions with pronounced seasonal dynamics, their limited ability to adapt to changing conditions often results in higher rates of missed detections and false alarms. An exception is a deep learning–based method, adapted to seasonal variation, that utilizes Sentinel-1 data and weakly supervised learning to automatically generate training labels from existing global forest loss datasets, reducing the need for manual annotations (Mullissa et al., 2023). By leveraging a deep neural network and incorporating a Bayesian updating approach for probabilistic disturbance detection, the method demonstrates improved detection accuracy and earlier alert generation compared to optical-based approaches in tropical dry forests. However, deep learning approaches are known to introduce high computational costs and require region-specific training, limiting their generalizability across different ecosystems without retraining. Another methodology adapted to seasonal variation is used in the LUCA dataset (Mullissa et al., 2024), which provides global information on forest land-use changes. This approach employs feature engineering to characterize the statistical properties of stable forest areas, facilitating the distinction between stable forests and areas undergoing change. Reference labels and training samples are collected across diverse forest biomes to train a machine learning model that estimates probabilities of forest presence and land-use changes. These probabilities are subsequently refined within a probabilistic framework to improve accuracy.

To better accommodate the complexity and variability inherent in forest loss detection, it is possible to rely on Bayesian approaches, which interpret probability as a degree of belief in an event given the available evidence. In the context of NRT forest disturbance monitoring, the RAdar for Detecting Deforestation (RADD) approach (Reiche et al., 2021), exploits Sentinel-1 data and uses an algorithm based on the Bayesian update principle (Reiche et al., 2018b). Specifically, the algorithm learns the statistical properties of the Forest (F) and Non-Forest (NF) classes in an unsupervised manner, using a two-year time series of Sentinel-1 data acquired prior to the monitoring period, under the assumption of undisturbed forest. These likelihoods are then processed to compute the posterior probability of NF for each Sentinel-1 acquisition. The time series of NF posteriors is then used to evaluate the conditional Probability Density Function (PDF) of deforestation, and forest loss is decided if this probability exceeds a user-defined threshold for a given number of consecutive acquisitions.

All currently available SAR-based monitoring methods, rely on pre-processing steps, such as temporal or spatial filtering, and post-processing techniques like Minimum Mapping Unit (MMU) filtering to manage false alarm rates. Several approaches also incorporate seasonality compensation or removal to address backscatter variability (Reiche et al., 2018a,b; Watanabe et al., 2021). While these steps improve overall detection performance, spatial filtering in particular can substantially degrade spatial resolution, hindering the detection of small-scale forest loss (Carstairs et al., 2022).

This work introduces an unsupervised forest loss monitoring technique based on Bayesian inference with NRT capabilities. Specifically, the approach builds upon the Bayesian Online Changepoint Detection algorithm first developed in Adams and MacKay (2007), adapting it to work with unfiltered Sentinel-1 Radiometrically Terrain Corrected (RTC) data and extending it to incorporate spatial context from neighboring pixels when detecting forest loss events. The method is evaluated using sample datasets from two Brazilian biomes, covering analyzed areas of approximately 13,800 hectares in the Amazon rainforest and 13,200 hectares in the Cerrado savanna. These areas were selected due to the availability of reliable reference data provided by MapBiomas Alerta (MapBiomas, 2024), and their contrasting backscatter patterns, with the Cerrado exhibiting notably stronger seasonal variations than dense rainforests. Specifically, validation is performed using two reference datasets per biome: one including small-scale disturbances (0.1–1 hectares, excluding selective logging), and another comprising randomly selected larger clearings obtained through stratified sampling. The results are compared with forest loss alerts from GLAD-L (Hansen et al., 2016), RADD (Reiche et al., 2021), and GFW (Global Forest Watch) (GFW, 2024).

2. Research area and data inputs

2.1. Study area

The experimental part of this study focuses on two areas in Brazil: the Brazilian Amazon and the Cerrado woodland-savanna. The Brazilian Amazon (also referred to as Amazonia), covering about 40% of the country's landmass, contains the largest continuous tropical rainforest in the world and is a deforestation hotspot largely monitored by existing operational NRT methods. On its part, the Cerrado is the world's most biodiverse woodland-savanna, and the largest one in South America. Spanning over 20% of Brazil's territory, the Cerrado plays a crucial role as a carbon sink, making its preservation critical (Rodrigues et al., 2022). Over the last two decades, the Cerrado has experienced agricultural expansion, including the establishment of soy farms and cattle ranching, resulting in the disappearance of its vegetation at twice the rate of the Amazon rainforest (Miranda et al., 2019). The native vegetation of the Cerrado spans a variety of distinct physiognomies (vegetation communities with different structures, heights, spacing, and dominant species), which can be grouped into broad vegetation formation categories, including grasslands, savannas, and forests. The precipitation patterns and vegetation phenology of the Cerrado (de Freitas Bussinguer et al., 2024) cause strong seasonality in the C-band SAR backscatter, which represents a challenge for existing NRT forest disturbance monitoring methods. As a result, this biome remains under-monitored in NRT.

The choice of the study area is further motivated by the availability of reliable reference data covering both Amazonia and the Cerrado, as well as by Brazil's significant contribution to global deforestation. In the MapBiomas Annual Deforestation Report of 2023 (RAD2023, 2024), it is documented that over 6.5 million hectares were deforested in Brazil between 2019 and 2022, indicating a 22.3% increase in deforested areas in 2022 compared with 2021. The largest forest loss increases occurred in Amazonia and in the Cerrado, accounting for approximately 90% of Brazilian deforestation in 2022. The report's findings highlight agriculture expansion and mining as primary drivers of deforestation, notably occurring within protected areas.

While large-scale deforestation accounts for most forest loss, small-scale disturbances are becoming increasingly frequent and damaging, posing growing challenges for monitoring efforts (Slagter et al., 2023). Activities such as artisanal mining (Asner et al., 2013), illegal logging (Tacconi et al., 2019), and subsistence agriculture (Kalamandeen et al., 2018) collectively contribute to increased forest fragmentation (Haddad et al., 2015), widespread degradation, biodiversity loss, carbon emissions, and ultimately, a loss of ecosystem resilience. In the Amazon, a notable shift in deforestation patterns has been observed, with large forest clearings (>50 ha) declining over time, while the number of small clearings (<1 ha) increased by 34% between 2001–2007 and 2008–2014 (Kalamandeen et al., 2018). The study found that 94.6% of deforested patches were smaller than 6.25 ha — the minimum threshold detected by PRODES, Brazil's official deforestation monitoring program (Shimabukuro et al., 2016) — with the vast majority (81.1%) being under 1 ha. The increasing prevalence of small deforestation parcels in Brazil partly reflects an adaptive strategy by larger landowners seeking to evade detection, particularly following the implementation of strict monitoring programs (Assunção et al., 2019).

2.2. Reference data

For validation, a subset of the MapBiomas Alerta dataset is used (MapBiomas, 2024). MapBiomas is a multi-institutional initiative that focuses on mapping land cover in Brazil. It aggregates alerts from various operational detection systems and validates them with high-resolution satellite imagery (RAD2023, 2024). Specifically, it combines data from sources such as DETER, SAD, GLAD, PRODES, and SIRAD-X, using imagery with resolutions ranging from 10 to 60 m. Alerts are first filtered and then verified through visual assessments using PlanetScope imagery (3.7-m resolution). The validated deforestation areas are refined to improve spatial precision using a supervised classification algorithm (Random Forest) and cross-referenced with public databases, such as Indigenous Lands and Conservation Units. Each alert undergoes an audit to ensure accuracy before being published weekly on the MapBiomas Alerta platform (https://alerta.mapbiomas.org/), with the possibility of post-publication revisions as new data becomes available to maintain ongoing reliability.

In this study, the baseline for forested areas is derived from MapBiomas Alerta, which marks deforestation on previously vegetated parcels and provides a temporal window for the deforestation events. Both forest and savanna formations are considered, classified according to MapBiomas Alerta's criteria by biome and based on vegetation type and canopy height:

- *Amazonia*. Forests are characterized by high-density, continuous canopies with a height >7 m, including evergreen, semi-deciduous, and deciduous non-floodable vegetation.
- *Cerrado*. Forests have a continuous canopy with a height >7 m, including riparian, gallery, dry forests, and forested savannas. Savannas are defined by a tree and shrub-herbaceous stratum with a canopy height between 3 and 6 m.

Forest loss, as defined by MapBiomas Alerta (RAD2023, 2024), refers to the complete suppression of native vegetation, excluding selective logging or degradation that does not result in full vegetation loss. The term "deforestation" is used broadly to include all forms of native vegetation suppression, including non-forest areas like savannas. Furthermore, the minimum size for a mapped disturbance is set at 0.1 hectares.

For this analysis, MapBiomas alerts (or polygons) from the 2020 operational year are used, as this period corresponds to the final full year of concurrent operation for both Sentinel-1 A and Sentinel-1B (Sentinel-1B was decommissioned in December 2021). This ensures a 6-day revisit interval and provides a complete year of reference polygons for robust testing. The proposed algorithm is applied to an area surrounding each reference polygon, defined by a rectangular buffer that is, on average, 2.6 times larger than the polygon. Forest loss is estimated at the polygon scale. A disturbed polygon is detected when a significant portion of its area (i.e., T_{poly}) is identified as forest loss by the proposed method. True positives (TP) correspond to polygons labeled as forest losses by both the proposed method and the MapBiomas dataset for the 2020 monitoring year. Conversely, a false negative represents a case where a polygon is missed by the proposed algorithm but is present in the MapBiomas dataset for the same monitoring year.

To address potential bias from undetected disturbed areas in the MapBiomas dataset, false positives are estimated using a temporal approach. Specifically, each MapBiomas Alerta polygon includes two key timestamps: one indicating the last time the area was observed as vegetated, and another marking the first time it was detected as deforested. This temporal window ensures that deforestation occurred within a specific timeframe. If a polygon is flagged as deforested in a given year, it means the area will be excluded from analysis in the following year. Hence, MapBiomas Alerta polygons acquired in 2021 are used to identify false positives (FP). A false positive is defined as a forest loss detected by proposed method in 2020, whereas the loss is recorded in the MapBiomas dataset as occurring in 2021. A polygon is assigned to the forest loss class if:

$$\frac{A_{\text{detected}}}{A_{\text{polygon}}} \geq T_{poly}, \tag{1}$$

where:

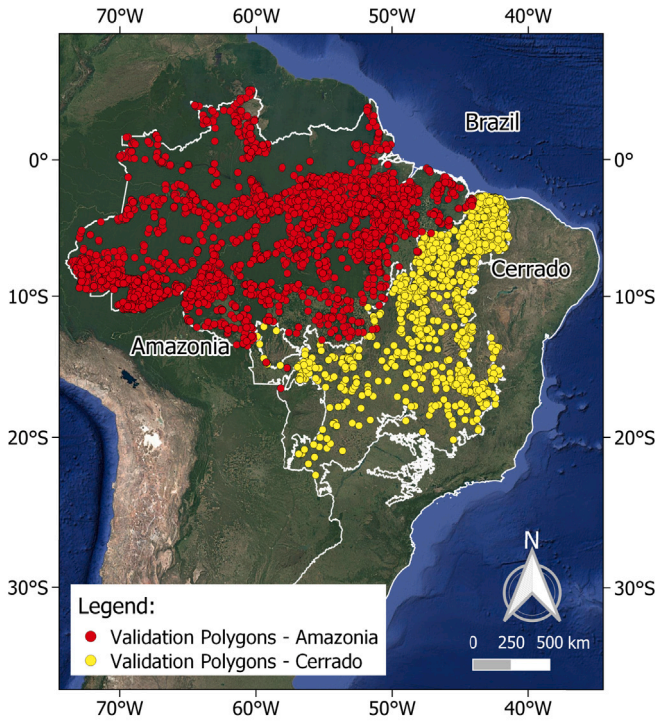


Fig. 1. Study area highlighting the validation polygons from MapBiomas Alerta for Amazonia and the Cerrado. Optical background image from Google Earth (©2024 Google).

- A_{detected} is the area of the polygon where the event is detected;
- A_{polygon} is the total area of the validation polygon;
- T_{poly} is the threshold representing the portion of a detected polygon required to determine forest loss.

The following sections define two tailored reference datasets used for validation: one representing small-scale clearings and the other representing larger clearings, both derived from the MapBiomas Alerta data and illustrated in Fig. 1.

2.2.1. Small-scale clearings reference data

To test the effectiveness of the proposed method in detecting small-scale forest loss, the first reference dataset includes all polygons of 0.1–1 hectare from the MapBiomas Alerta dataset; the number of polygons per biome is provided in Table 1, and histograms of deforested polygon sizes are shown in Fig. 2. These polygons represent 7.65% (2650 hectares) of the total MapBiomas Alerta alerts for the Amazon biome and 5.14% (450 hectares) of alerts for the Cerrado. The vast majority of these clearings have been converted to small-scale subsistence agriculture (88.9%), with smaller shares attributed to mining (2.9%) and urban expansion (0.3%). The difference in polygon counts between the Amazon and Cerrado reference datasets reflects both biome size and distinct deforestation patterns. In the Amazon, logging and smallholder agriculture result in many small deforested areas, whereas in the Cerrado, large-scale agriculture typically produces fewer but larger clearings.

2.2.2. Larger-scale clearings reference data

A stratified sampling approach (Robb and Cochran, 1963) is used to select large-scale clearings based on the total polygon size distribution in the 2020 MapBiomas Alerta dataset (Fig. 3). Deforested polygons are classified into three size categories (1–5 hectares, 5–20 hectares, and ≥20 hectares), with a proportional number selected from each category,

Table 1
Number of polygons in the reference datasets for 2020 and 2021, separated by biome (Amazon and Cerrado) and by polygon size category (small-scale and larger clearings).

2020 reference dataset				
Size	Small-scale dataset	Larger-scale dataset		
	$A_0 < 1$ ha	$1 \le A_1 < 5$ ha	$5 \le A_2 < 20$ ha	$A_3 \ge 20$ ha
Amazonia	3590	528	346	126
Cerrado	629	200	185	115
2021 reference dataset				
Amazonia	1657	220	188	92
Cerrado	196	127	75	48

as reported in Table 1. The polygons represent 2.1% (11,200 hectares) of the total MapBiomas Alerta alerts in the Amazon and 4.0% (12,700 hectares) of the alerts in the Cerrado. The vast majority of deforested areas were converted to agriculture, ranging from 94.0% in the 1–5 ha class up to 99.6% in areas greater than 20 hectares, while mining and urban expansion accounted for the rest.

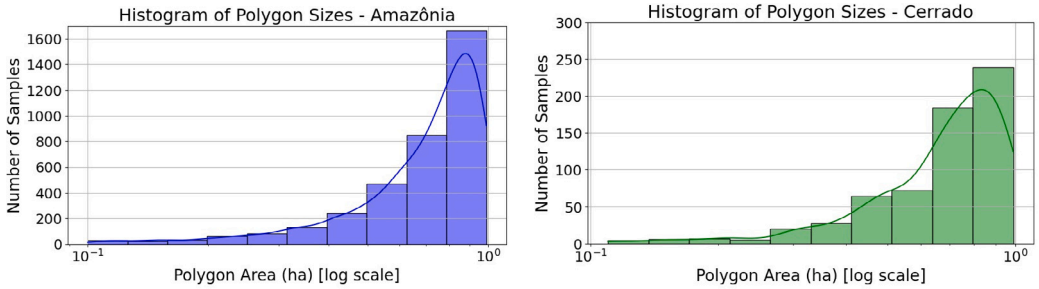
2.3. Operational products for comparison

Forest disturbance results in this study are compared against three operational NRT forest monitoring approaches: GLAD-L (Hansen et al., 2016), RADD (Reiche et al., 2021), and the integrated alerts provided by GFW. GLAD-L produces alerts from Landsat imagery with 30-m spatial resolution, covering the entire tropics from January 2018 onward, and selected countries from 2015. GLAD-S2 offers alerts at 10-m resolution derived from Sentinel-2 optical imagery, limited to primary humid tropical forests and available from January 2019 to the present. RADD provides 10-m pixel-spacing alerts from Sentinel-1 imagery, also restricted to primary humid tropical forests starting in January 2019. Consequently, neither GLAD-S2 nor RADD operate in the Cerrado biome. GFW combines alerts from GLAD-L, GLAD-S2, and RADD, where available, by resampling GLAD-L to a 10-m grid for integration. Each GFW alert is assigned a confidence level: low confidence (detected once by a single method), high confidence (detected multiple times by one method), and highest confidence (detected by multiple methods). For the purpose of comparison in this study, only GFW alerts labeled as high confidence or highest confidence are considered.

2.4. Sentinel-1 input data

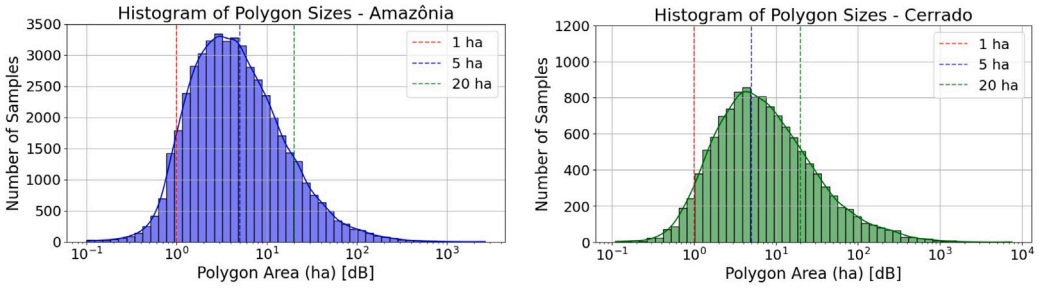
This work uses Sentinel-1 A/B C-band (wavelength, $\lambda \approx 5.6$ cm), Interferometric Wide swath (IW), RTC images processed by Catalyst (accessed on: https://planetarycomputer.microsoft.com/catalog), with a revisit time of 6 to 12 days. The analysis focuses on cross-polarization, as previous studies have identified higher contrast in the cross-polarized channel between forested and deforested areas (Mermoz et al., 2021; Doblas et al., 2020). Additionally, either ascending or descending orbits are selected, with the only constraint that each single-pixel time series is processed using a single relative orbit. The data pre-processing that allows transforming Ground Range Detected (GRD) into RTC products can be summarized as:

- Radiometric calibration: conversion of GRD products to calibrated intensity using gamma values normalized for the incidence angle (Miranda and Meadows, 2015);
- Radiometric terrain correction (Small, 2011): using the 30 m PlanetDEM provided by Planet Observer as the elevation source;
- Orthorectification: to the appropriate UTM projection.



(a) Amazonia (log-scale) (b) Cerrado (log-scale)

Fig. 2. Histograms of MapBiomas Alerta deforested polygon sizes under 1 hectare in Amazonia ($\mu_{size} = 0.715$ ha and $\sigma_{size} = 0.214$) and Cerrado ($\mu_{size} = 0.670$ ha and $\sigma_{size} = 0.206$) for year 2020.



(a) Amazonia (log-scale) (b) Cerrado (log-scale)

Fig. 3. Histograms of MapBiomas Alerta deforested polygon sizes in Amazonia and Cerrado for year 2020.

No additional spatial filtering is applied to the original 10-m pixel spacing multi-looked GRD products. This results in a ground resolution of 20 m in azimuth and 22 m in ground range. This approach is intended to demonstrate the proposed method's ability to retain the native spatial resolution of the data and enhance the detection of small-scale forest loss.

3. Methodology

3.1. Bayesian online changepoint detection

Bayesian approaches estimate the posterior probability of an event using a likelihood function derived from a statistical model of the data, and a prior distribution of the model parameters. In other words, the probability of a hypothesis — in this case, a forest loss occurrence — is updated as more information becomes available and flagged when a significant change in the statistical properties of the distribution is detected. The method developed in this work, subsequently named BOCD, builds upon the Bayesian Online Changepoint Detection algorithm (Adams and MacKay, 2007).

In the following, $\mathbf{x}_{1:t} = (x_1, \dots, x_t)^T \in \mathbb{R}^t$ represents a single-pixel Sentinel-1 RTC cross-polarized backscatter time series, organized in chronological order. It should be noted that the index t is hereafter referred to as the "acquisition time". The starting assumption is that the time series is partitioned by changepoints, and the samples within each partition are independent, identically distributed (i.i.d.) according to some distribution. The algorithm aims at splitting the time series in a set of connected segments, where the changepoints are the initial time instants of each new segment. The number of changepoints and their locations are the parameters to be determined. To do so, the algorithm estimates the run length, denoted as r_t, which represents the number of acquisitions since the last changepoint. The run length is considered as a random variable associated with the hidden state of a Markov model (Fearnhead and Liu, 2007; Barry and Hartigan, 1993), and its processing is conducted in a message-passing manner (Adams and MacKay, 2007) using the observable x_t measured at each date. At each timestamp, the run length value either increments by 1, indicating a growth phase, or resets to 0, signaling a changepoint. Changes are detected by tracking the Maximum A Posteriori (MAP) estimate of r_t.

Fig. 4 illustrates an example of a time series and its corresponding r_t map. In the map, all possible paths of the run length are indicated in light gray. The true path corresponding to the time series on the left is highlighted in black, with changepoints marked by red lines. Fig. 5 depicts the hidden infinite-state Markov model on which BOCD is based, as r_t can take arbitrarily large values.

3.2. Conjugate Bayesian framework and likelihood modeling

The sequence of backscatter observations, $\mathbf{x}_{1:t}$, is considered to be divided by changepoints, as detailed in Section 3.1. The primary goal of the BOCD algorithm is to deduce the run length of the current segment by tracking its posterior distribution, which can be expressed as:

$$p(r_t|\mathbf{x}_{1:t}) = \frac{p(r_t, \mathbf{x}_{1:t})}{\sum_{r_t=0}^{t} p(r_t, \mathbf{x}_{1:t})}. \tag{2}$$

According to the properties of hidden Markov models (Adams and MacKay, 2007), it is possible to express the joint distribution in (2) as:

$$p(r_t, \mathbf{x}_{1:t}) = \sum_{r_{t-1}=0}^{t-1} \underbrace{p(x_t|\mathbf{r}_{t-1:t}, \mathbf{x}_{1:t-1})}_{\text{posterior predictive}} \underbrace{p(r_t|r_{t-1})}_{\text{run length prior}} \underbrace{p(r_{t-1}, \mathbf{x}_{1:t-1})}_{\text{message}}, \tag{3}$$

where $\mathbf{r}_{t-1:t} = (r_{t-1}, r_t)^T$ denotes the considered branch of the lattice in Fig. 4. The recursive expression in (3) enables the evaluation of the joint distribution using a message-passing algorithm, in which $p(r_{t-1}, \mathbf{x}_{1:t-1})$ is passed from the previous recursion to the current one. The other elements of (3) are the conditional prior on the run length, $p(r_t|r_{t-1})$, representing the probability of transition, and the posterior

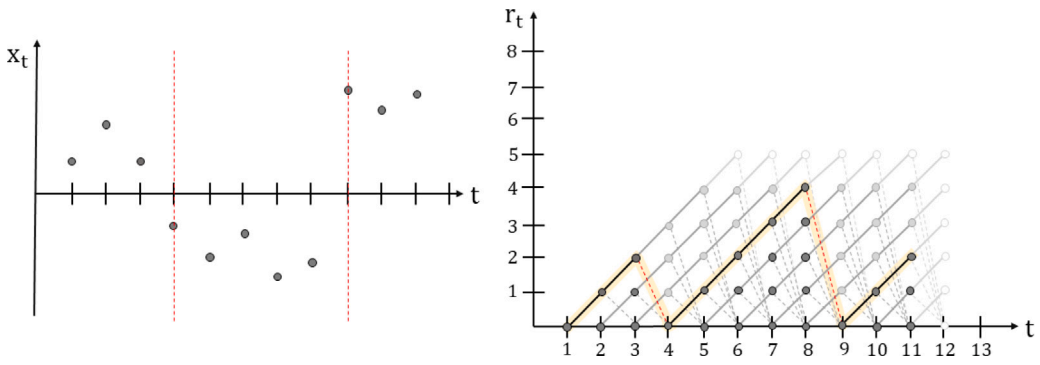


Fig. 4. Example of single-pixel backscatter time series with abrupt changes marked with red dashed lines (left), and corresponding r_t map (right).

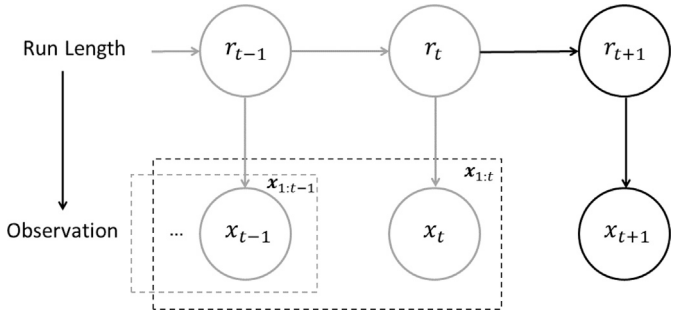


Fig. 5. BOCD represented by a hidden infinite-state Markov model. x_t represents an observation performed at time t, whereas r_t indicates the run length and represents the hidden state.

predictive distribution $p(x_t|\mathbf{r}_{t-1:t}, \mathbf{x}_{1:t-1})$, which is the probability density of the newly observed datum given the data sequence since the last changepoint. For a given run length value r_{t-1}, the posterior predictive is $p(x_t|\mathbf{x}_{t-1}^{(r_{t-1})})$, where $\mathbf{x}_{t-1}^{(r_{t-1})} = (x_{t-1-r_{t-1}}, \dots, x_{t-1})^T$ denotes the subset of measured data for the considered segment.

The transition probability can be selected based on a user-defined detection strategy or prior beliefs about the likelihood of the events under study. Meanwhile, within a data segment assumed to be i.i.d. between two consecutive changepoints, the posterior predictive distribution is modeled using the likelihood function $p(x_t|\boldsymbol{\theta})$ and the parameter prior distribution $\pi_{\boldsymbol{\eta}_0}(\boldsymbol{\theta})$, which depends on the parameter vector $\boldsymbol{\eta}_0$:

$$p\left(x_t\middle|\mathbf{x}_{t-1}^{(r_{t-1})}\right) = \int p(x_t|\boldsymbol{\theta})\,\pi_{\boldsymbol{\eta}_0}\left(\boldsymbol{\theta}\middle|\mathbf{x}_{t-1}^{(r_{t-1})}\right) d\boldsymbol{\theta}, \tag{4}$$

where the posterior distribution of the parameters is given by:

$$\pi_{\boldsymbol{\eta}_0}\left(\boldsymbol{\theta}\middle|\mathbf{x}_{t-1}^{(r_{t-1})}\right) = \frac{p\left(\mathbf{x}_{t-1}^{(r_{t-1})}|\boldsymbol{\theta}\right)\pi_{\boldsymbol{\eta}_0}(\boldsymbol{\theta})}{\int p\left(\mathbf{x}_{t-1}^{(r_{t-1})}|\boldsymbol{\theta}\right)\pi_{\boldsymbol{\eta}_0}(\boldsymbol{\theta})d\boldsymbol{\theta}}. \tag{5}$$

The relationships in Eqs. (2) to (5) illustrate the Bayesian approach, treating unknown model parameters as random variables and incorporating their prior distributions into decision-making. Nevertheless, calculating the parameters posterior generally represents a significant computational burden, due to the evaluation of the integral term in (5). Prior conjugacy represents a solution to this problem, as it reduces the computation to a simple parameter update (Feller, 1968). A prior distribution is conjugate to a given likelihood function, if the posterior and the prior belong to the same family of probability distributions. In particular, a conjugate prior always exists when the likelihood belongs to an exponential family (Adams and MacKay, 2007).

The statistical model of Single-Look Complex (SLC) Sentinel-1 data follows a complex circular normal distribution (Goodman, 1976), while the multi-look intensity, obtained by averaging the squared modulus of multiple SLC realizations, follows a Gamma distribution. The logarithmic reflectivity follows an Exponential-Gamma distribution, with a Gaussian-like shape emerging for a high number of looks, e.g., ≥ 20. Speckle-filtered Sentinel-1 GRD data can safely be assumed to follow a normal distribution (Reiche et al., 2018b), but this assumption typically breaks down for unfiltered data with fewer looks. Interestingly, the statistical analysis of unfiltered RTC Sentinel-1 data, presented in Appendix A, indicates that normality can still be assumed over homogeneous vegetated areas, even with fewer looks.

The normal distribution belongs to the exponential family, and its conjugate prior is, in this case of unknown mean and variance, a normal inverse-gamma probability density, i.e., $\boldsymbol{\theta} \sim \mathcal{NIG}(\boldsymbol{\theta}; \boldsymbol{\eta}_0)$. According to the conjugacy principle, the parameter posterior is also a normal inverse-gamma density function $\boldsymbol{\theta}|\mathbf{x}_m \sim \mathcal{NIG}(\boldsymbol{\theta}; \boldsymbol{\eta}_m)$, with $\mathbf{x}_m = (x_0, \dots, x_m)^T$, and $\boldsymbol{\eta}(m) = (\alpha_m, \beta_m, \mu_m, \kappa_m)$, which may be expressed as:

$$\begin{aligned}
\alpha_m &= \alpha_0 + \frac{m}{2} \\
\beta_m &= \beta_0 + \frac{1}{2}\sum_{i=1}^{m}(x_i - \overline{x})^2 + \frac{m\kappa_0}{\kappa_0 + m}\frac{(\overline{x} - \mu_0)^2}{2} \\
\mu_m &= \frac{\kappa_0\mu_0 + m\overline{x}}{\kappa_0 + m} \\
\kappa_m &= \kappa_0 + m,
\end{aligned} \tag{6}$$

with $\boldsymbol{\eta}(0) = \boldsymbol{\eta}_0$ and $\overline{x} = \sum_{i=0}^{m} x_i/(m+1)$. In (6), α is the number of degrees of freedom, β is the precision, μ corresponds to the location, and κ represents the shape. The posterior predictive can be derived in closed form and corresponds to a t-distribution (Murphy, 2007):

$$p(x|\mathbf{x}_m) = p_{\boldsymbol{\eta}(m)}(x) = t_{2\alpha_m}\left(x; \mu_m, \frac{\beta_m(\kappa_m + 1)}{\alpha_m\kappa_m}\right). \tag{7}$$

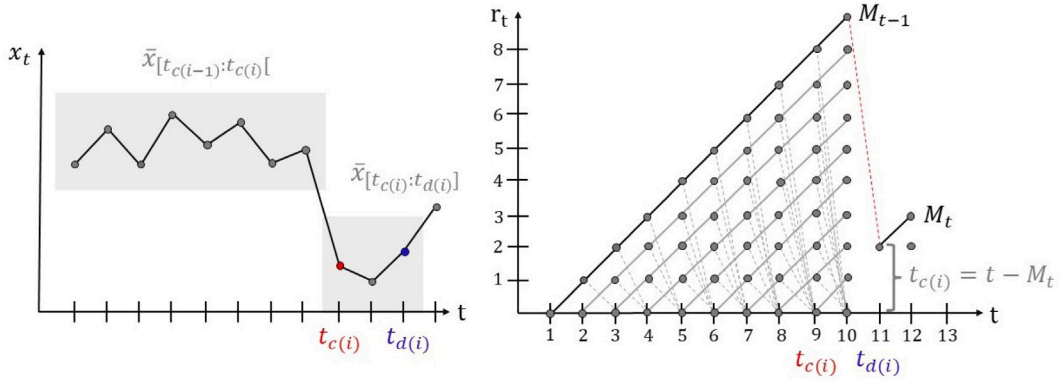


Fig. 6. Posterior run length probability for a study case. $t_{c(i)}$ represents the ith time of change, whereas $t_{d(i)}$ represents the corresponding time of detection.

3.3. Algorithmic design of BOCD

Following the developments presented in Section 3.2, the BOCD algorithm can be summarized as follows:

- Observe a new Sentinel-1 backscatter value x_t at time t.
- Evaluate the posterior predictive distribution $p(x_t|\mathbf{r}_{t-1:t}, \mathbf{x}_{t-1}^{(r_{t-1})})$ for each possible value of $\mathbf{r}_{t-1:t}$.
- Apply the message-passing principle: from the run length posterior probability at $t-1$, compute the probability that r_t is increased by 1, or growth probability as:

$$p(r_t = r_{t-1} + 1, \mathbf{x}_{1:t}) = p(r_{t-1}, \mathbf{x}_{1:t-1})\, p(x_t \mid \mathbf{x}_{t-1}^{(r_{t-1})})\,(1 - p(r_t = 0 \mid r_{t-1})), \tag{8}$$

and the probability that r_t drops to 0, i.e., the probability of a changepoint, for all run length values, $r_{t-1} = 1, 2, \ldots$, as:

$$p(r_t = 0, \mathbf{x}_{1:t}) = \sum_{r_{t-1}} p(r_{t-1}, \mathbf{x}_{1:t-1})\, p(x_t \mid \mathbf{x}_{t-1}^{(r_{t-1})})\, p(r_t = 0 \mid r_{t-1}). \tag{9}$$

The computation of growth and changepoint probabilities follows the evaluation of the transition probability discussed in Section 3.5, and the posterior predictive, detailed in Section 3.2. The updates of the posterior predictive parameters can be easily derived from (6) as:

$$\begin{aligned} \alpha_{m+1} &= \alpha_m + 1/2 \\ \beta_{m+1} &= \beta_m + \frac{\kappa_m (x_i - \mu_m)^2}{2(\kappa_m + 1)} \\ \mu_{m+1} &= \frac{\kappa_m \mu_m + x_i}{\kappa_m + 1} \\ \kappa_{m+1} &= \kappa_m + 1. \end{aligned} \tag{10}$$

- The result is a run length posterior probability vector. The detection of forest loss events from the vector is addressed in Section 3.4.

3.4. Forest loss detection methodology

The output of the BOCD at time t is a vector of posterior run length probabilities, whose elements are specified in (2). Changes are detected by tracking the position of the most probable run length at time t, M_t, defined as:

$$M_t = \arg\max_{r_t} p(r_t|\mathbf{x}_{1:t}). \tag{11}$$

In the absence of change, i.e., when the most probable last changepoint remains unchanged, the maximum position satisfies the following equation:

$$M_t = M_{t-1} + 1. \tag{12}$$

An ideal change, such as the one represented in Fig. 4, corresponds to $M_t = 0$. However, in practice, a change may be detected with some temporal delay, and the minimum value does not always reach this lower bound. This case is handled by defining a change if:

$$M_t < M_{t-1} - \Delta M. \tag{13}$$

The minimal detectable shift, $\Delta M > 0$, is designed to make the process robust against small oscillations in M_t, caused by the inherent variability in the input data. Fig. 6 illustrates the evolution of the posterior run length probability for a study case where a decrease in M_t to a value satisfying the condition in (13) indicates a change that actually occurred at time $t_c = t - M_t$ and is detected at time t. Furthermore, considering $t_{c(i-1)}$ as the time instant at which the last changepoint was observed, a change is classified as forest loss only if the mean backscatter of acquisitions between the last changepoint and the current one is greater than the mean backscatter of acquisitions between the current changepoint and the detection instant $t_d(i)$:

$$\bar{x}_{[t_{c(i-1)}:t_{c(i)}[[v=10pt]} < \bar{x}_{[t_{c(i)}:t_{d(i)}]}. \tag{14}$$

3.5. Incorporating spatial context in change detection

In Adams and MacKay (2007), the conditional prior on the run length $p(r_t|r_{t-1})$, used in (3), is set to a constant value. In the present study, this quantity is modeled using the survival analysis formalism.

Considering T as the time at which a change occurs in a time series, the survival function is defined as the probability that the time of failure exceeds T, $S(t) = p(t > T)$ (Cox, 2001). Such a function can be parameterized as:

$$S(t) = e^{-\int_0^t h(u)du} \text{ with } S(0) = 1,\ S(+\infty) = 0, \tag{15}$$

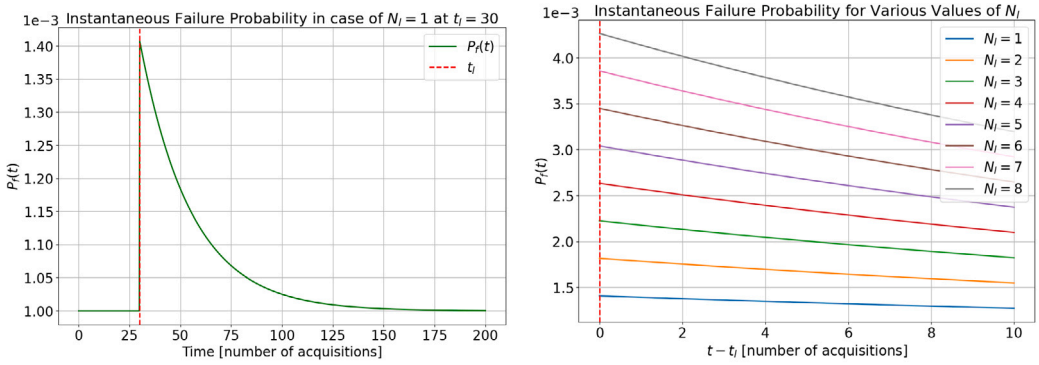
where $h(t)$ is the instantaneous hazard rate, defined as (Cox, 2001):

$$h(t) = \lim_{\Delta t \to 0} \frac{p(t \le T \le t + \Delta t | T \ge t)}{\Delta t} = \frac{S'(t)}{S(t)} \ge 0, \tag{16}$$

with $p(t \le T \le t + \Delta t|T \ge t)$ the conditional probability that the event occurs between time t and $t + \Delta t$, given that the event has not occurred before time t, and Δt a small time interval. In the framework of a time series represented as a hidden infinite-state Markov model (as shown in Fig. 5), the survival probability can be rewritten as follows:

$$S(t) = S(t-1)p(r_t = r_{t-1} + 1|r_{t-1}) = S(t-1)e^{-\int_{t-1}^t h(u)du}. \tag{17}$$

This study further exploits the survival formalism to incorporate contextual information into the decision process, recognizing that in forest loss monitoring, the probability of an event is not directly related to time. Instead, the probability of deforestation increases with spatial proximity to previous deforestation — a pattern demonstrated for



(a) Instantaneous failure probability in the case of a single deforested neighbor at time $t_l = 30$

(b) Instantaneous failure probability for various numbers of deforested neighbors

Fig. 7. Time varying instantaneous failure probability.

the Amazon (Doblas et al., 2022) and also observed in the Cerrado, though with regionally varying intensity (Trigueiro et al., 2020). The instantaneous hazard rate is modeled as a time-varying function that accounts for recent forest loss events in the eight neighboring pixels of the pixel of interest. The selected instantaneous hazard rate is inspired by the Gompertz-Makeham model (Stevenson, 2007) and modified to implicitly modulate the probability of transition according to the spatial context:

$$h(t) = c + N_l a e^{b(t-t_l)}, \tag{18}$$

where N_l represents the number of deforested neighbors, t_l corresponds to the date of the last detected forest loss in the neighborhood, a and b are negative constants, and c is a baseline constant rate used when no forest loss has been observed in the surroundings of the considered location, i.e., $N_l = 0$.

The instantaneous survival probability $P_S(t)$ and the instantaneous failure probability $P_F(t)$ are defined as:

$$P_S(t) = p(r_t = r_{t-1} + 1) \quad P_F(t) = p(r_t = 0) = 1 - P_S(t). \tag{19}$$

Fig. 7 illustrates an example of the instantaneous failure probability in the case of a single deforested neighbor at a specific time (Fig. 7(a)). In a typical scenario, where survival is largely favored, the failure probability is found to be a quasi-linear function of N_l, as:

$$P_F(t, N_l) = 1 - P_S(t, N_l) \approx 1 - P_S^{N_l}(t, 1) \approx 1 - (1 - P_F(t, 1))^{N_l} \approx N_l P_F(t, 1). \tag{20}$$

Following a detailed explanation of the BOCD algorithm for forest loss monitoring, Fig. 8 illustrates its working principle.

3.6. Parameter setting and susceptibility trade-off in BOCD

Susceptibility can be adapted to detect more or fewer changepoints by modifying some key parameters. The ΔM threshold presented in Section 3.4, should be large enough to avoid random oscillations of M_t causing false alarms, yet small enough to allow for delayed detections. Another parameter is the constant term of the instantaneous hazard rate, c, introduced in (18), which represents the knowledge of how rare a forest loss event is. A higher value of c increases the algorithm's responsiveness to detect forest loss events. The other parameters influencing the instantaneous hazard rate are a, which represents the immediate increase in hazard rate due to a forest loss event in the neighborhood, and b, which controls the rate at which the algorithm forgets this event.

Furthermore, susceptibility can be adjusted by modifying the priors, $\boldsymbol{\eta}_0$, which represent the initial parameters of the posterior predictive distribution. In the case of the t-distribution, the parameter α represents the degrees of freedom. A higher α value corresponds to lighter tails in the t distribution, making BOCD more conservative in detecting changes. Conversely, lowering α increases the probability of larger drops in M_t and, consequently, susceptibility to changes. Another tunable prior is κ, which controls the spread of the distribution. A larger value reduces susceptibility to changes, whereas a smaller one accommodates more variability in the data. The relationships in (6) indicate that the influence of the initial values on α and κ remains unchanged over time, as these parameters grow linearly with n. Conversely, with a sufficiently large number of observations, μ and β converge to the sample estimates of the mean and variance of x_t, respectively. Consequently, variations in μ and β have minimal impact on forest loss detection.

To test the method's capabilities, four configurations with varying levels of responsiveness to changes are considered. The configurations share a set of fixed parameters: $\Delta M = 10$, $c = 0.001$, $\mu_0 = \overline{x}_1$, and $\beta_0 = 0.01$, but differ in the values assigned to the most influential priors, α_0 and κ_0:

- Configuration 1 (C1, most conservative): $\alpha_0 = 1$, $\kappa_0 = 0.01$;
- Configuration 2 (C2): $\alpha_0 = 1$, $\kappa_0 = 0.005$;
- Configuration 3 (C3): $\alpha_0 = 0.1$ $\kappa_0 = 0.01$;
- Configuration 4 (C4, most responsive): $\alpha_0 = 0.1$, $\kappa_0 = 0.005$.

3.7. Assessment of detection timeliness

In the context of NRT forest disturbance monitoring, both spatial precision and the timeliness of detections are fundamental. Timely detection of changes is crucial for initiating mitigation efforts, such as actions to prevent further deforestation. Therefore, when evaluating the performance of a forest disturbance detection method with NRT capabilities, both the date of detection and detection delay should be taken into account. In this context, "detection delay" refers to the time interval between the occurrence and detection of a forest loss event. It depends on the revisit time of the satellite constellation providing the input imagery and the confirmation procedure adopted by the different methods. According to Doblas et al. (2023), RADD needs approximately 3.1 Sentinel-1 acquisitions to confirm an alert, while GLAD-L/S2's detection delay is defined as the interval between the first anomaly detection and the fourth cloud-free optical observation showing the same anomaly. However, determining the detection delay of an optical method is challenging, as it strongly depends on cloud coverage. The determination of the detection delay of BOCD for a single event, i.e., a single pixel, is described in Section 3.4. At the polygon level, the mean detection delay is computed by averaging the individual delays of all

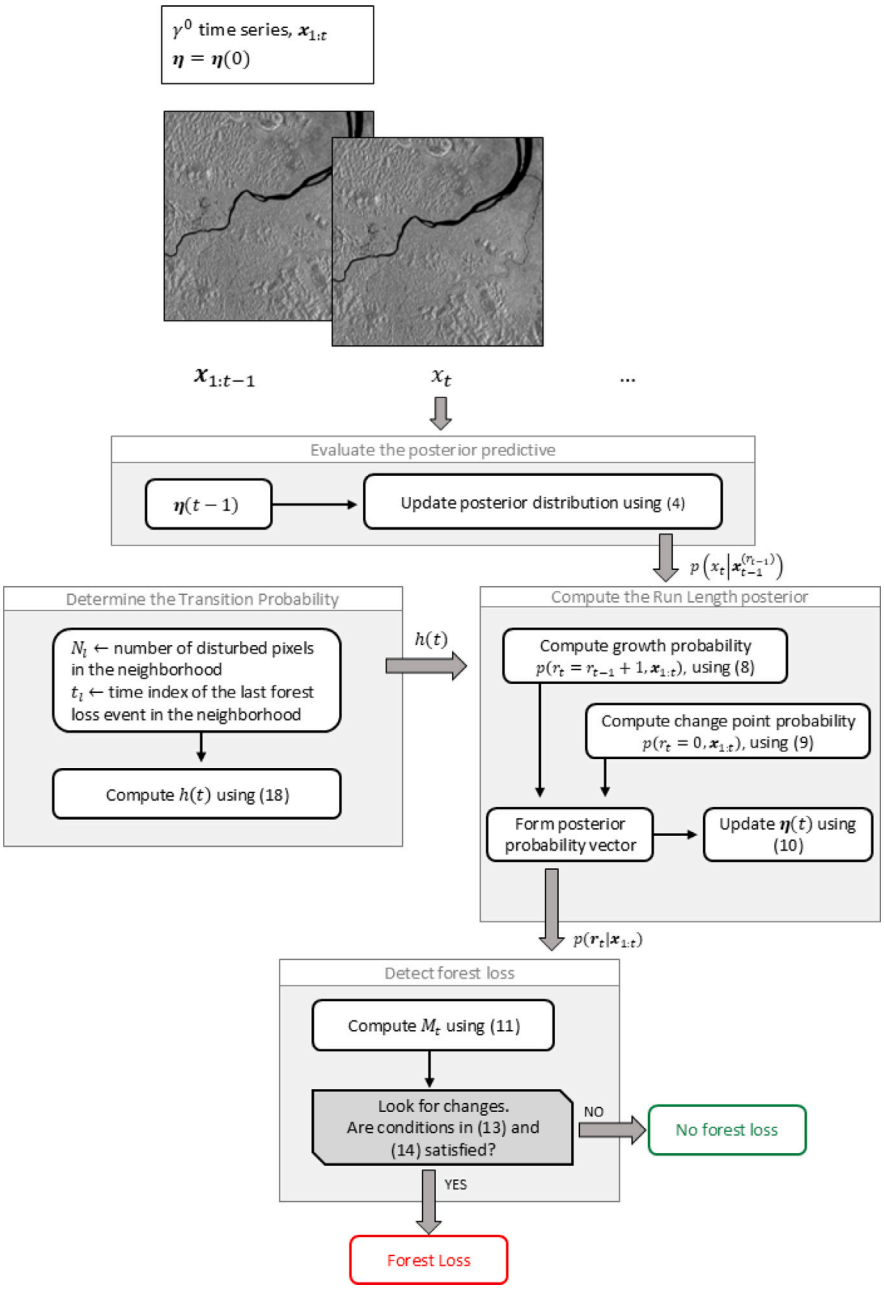


Fig. 8. Synopsis of the BOCD forest loss detection method at time t and for a given image pixel.

pixels within the polygon. At the dataset level, the global detection delay is computed as the mode of all polygon-level mean delays, in order to reduce the influence of outliers.

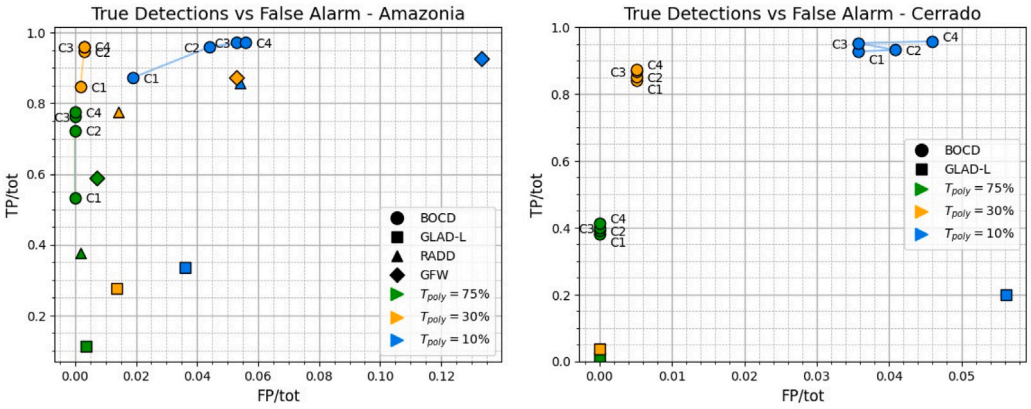
This study compares the detection dates estimated by the considered NRT methods with those estimated by BOCD. The objective is to assess the proposed method's detection timeliness relative to existing approaches, and to identify which method most closely matches the MapBiomas Alerta information, given that the true deforestation date is typically unknown. Each polygon in the MapBiomas Alerta dataset has two timestamps: the last date the parcel was observed as vegetated ("before date") and the first date it was observed as deforested ("after date"). These timestamps define a time range within which deforestation is assumed to have occurred. For BOCD, RADD, GLAD-L, and GFW, the detection date is first checked to determine whether it falls within (i.e., > "before date") or after (i.e., ≥ "after date") the temporal window provided by MapBiomas Alerta. Secondly, the detection dates of existing methods and those estimated by BOCD are compared with the "after date" provided by MapBiomas Alerta, considering only a subset of reference polygons where both approaches detected deforestation in at least 75% or 50% of the polygon area.

4. Results

4.1. Validation of NRT forest loss monitoring methods

4.1.1. Small-scale clearings

In the Amazon (Fig. 9(a)), BOCD demonstrates superior detection performance compared to other methods, particularly in configurations C2 to C4. The only exception occurs when compared to GFW in the most conservative configuration (C1). Nonetheless, GFW suffers from the highest rate of false alarms among considered methods. Notably, BOCD C4 achieves nearly 19% more true detections than GFW, the



(a) Amazon (b) Cerrado

Fig. 9. Normalized true positives vs. false positives for deforested polygons in the 0.1–1 ha range, across the Amazon and Cerrado. Colors indicate different T_{poly} values (1), and labels 'C1'–'C4' refer to different BOCD configurations.

best-performing existing method, at $T_{\text{poly}} = 75\%$. RADD demonstrates significantly better detection performance compared to GLAD-L. As expected, BOCD C4 achieves the highest number of true detections, although with a higher false alarm rate compared to other BOCD configurations, especially at very low T_{poly}. Nevertheless, all BOCD configurations remain robust to false alarms relative to other methods, even at lower T_{poly} values. At $T_{\text{poly}} = 10\%$, BOCD C4 exhibits almost 2.5 times fewer false alarms compared with GFW, while the extreme conservativeness of BOCD C1 results in 7 times fewer false alarms than GFW at the same threshold, although at the cost of lower detections. Appendix B presents additional results, including the confusion matrix (Table B.6) and spatial accuracy metrics — precision, sensitivity, and F1-score (Table B.8) — as well as forest loss detection maps (Figs. B.16 and B.17). To summarize, BOCD C4 achieves an F1-score of 87.4% compared to GFW's 74.1% at $T_{\text{poly}} = 75\%$. At a lower T_{poly} of 10%, BOCD C4 further improves to 97.3%, surpassing GFW's 93.1%. All other NRT methods yield F1-scores lower than that of GFW.

In the Cerrado biome (Fig. 9(b)), the results consistently demonstrate a substantial improvement over GLAD-L — the only available alerts in this biome — in terms of true detections, while maintaining a relatively low false alarm rate. It is noted that, to the authors' knowledge, GLAD-L detects changes exclusively within forested areas. This limits its performance in the Cerrado biome, where much of the disturbance involves the suppression of savanna-like vegetation rather than forested areas. Within the selected MapBiomas Alerta dataset for the Cerrado, forests account for only 37.5% (236 polygons) of the overall dataset, while savannas account for the remaining 62.5% (393 polygons). The confusion matrix (Table B.7) and spatial accuracy metrics (Table B.9) are provided in Appendix B. In the Cerrado, BOCD C4 achieves an F1-score of 58.5%, significantly outperforming GLAD-L at $T_{\text{poly}} = 75\%$. At a low T_{poly} (i.e., 10%), BOCD C4 achieves an F1-score of 97.4%, surpassing GLAD-L's 33.3%.

4.1.2. Larger-scale clearings

Validation over larger-scale clearings (≥1 ha) has been performed only for the BOCD C3 configuration. In the Amazon, BOCD shows performance comparable to RADD at $T_{\text{poly}} = 75\%$, with progressively improved true positive detections as T_{poly} decreases (Table 2). GFW achieves significantly higher true positives than BOCD at high T_{poly}; however, it also exhibits substantially higher false alarm rates, reaching up to +50% at low T_{poly}. Overall, BOCD proves to be the most robust approach against false alarms among the evaluated methods. In the Cerrado, BOCD greatly outperforms GLAD-L regarding true positive detections (Table 3). A trend inversion is observed when analyzing larger deforested parcels, as BOCD performance is approximately twice as good in the Cerrado compared to the Amazon at $T_{\text{poly}} = 75\%$, contrary to what was observed for smaller deforestation areas.

Table 2
Confusion matrix for deforested polygons ≥1 ha over Amazonia. T_{poly} defined in (1); 'small' refers to small-scale polygons (0.1–1 ha) reported for comparison, and 'large' to categories A_1–A_3 combined.

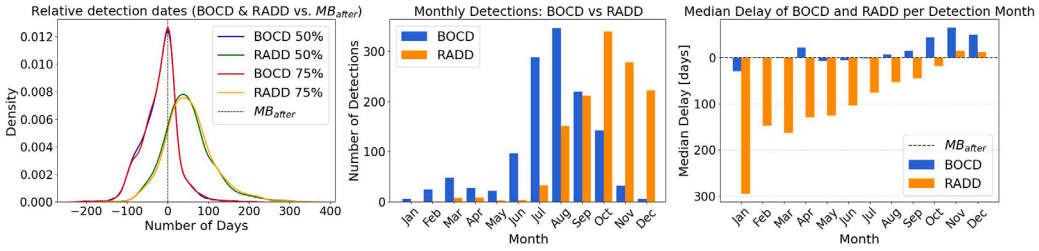
True positive [%]					
T_{poly}	BOCD C3 - small	BOCD C3 - large	RADD	GLAD-L	GFW
75%	76.3	32.2	33.3	11.6	68.3
50%	90.2	65.3	62.2	24.9	86.0
30%	95.9	79.9	76.9	32.7	91.5
10%	97.2	95.9	90.3	40.0	96.1
False positive [%]					
75%	0	0	0	0	0.4
50%	0.1	0	0	0.8	4.0
30%	0.3	0	0.4	1.2	16.0
10%	5.43	0.8	2.8	3.4	52.6

For the interested reader, confusion matrices showing the detection performance of BOCD C3 over large clearings divided by size categories (A_1–A_3) are provided in Appendix B for both the Amazon (Table B.10) and Cerrado (Table B.11) biomes.

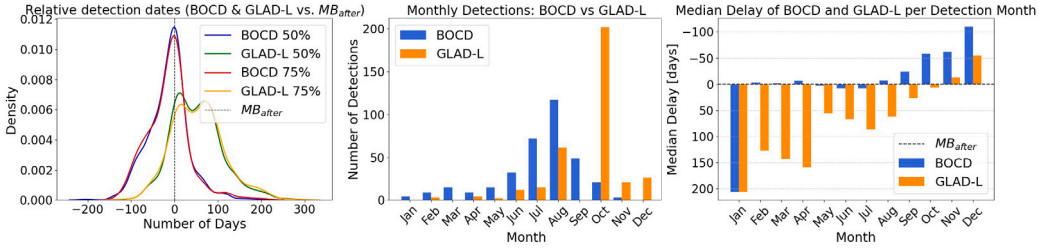
4.2. Comparison of temporal accuracy of forest loss monitoring methods

At the dataset level, and for the BOCD C3 configuration, the mode detection delay consistently equals 3 acquisitions across both biomes and for both polygon sizes.

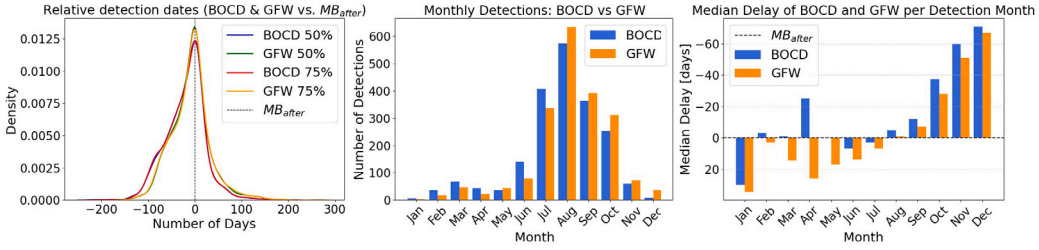
The comparison of detection dates between methods is limited to the Amazon, as GLAD-L detections in the Cerrado are insufficient for a meaningful comparison. The resulting density curves of detection dates for small-scale clearings (0.1–1 ha) are shown in Fig. 10, while results for larger clearings (≥1 ha) are presented in Fig. 11. For larger clearings, GLAD-L is not reported due to an insufficient number of



(a) BOCD vs. RADD; $T_{poly} = 75\%$ (1260 polygons) and $T_{poly} = 50\%$ (2254 polygons).



(b) BOCD vs. GLAD-L; $T_{poly} = 75\%$ (346 polygons) and $T_{poly} = 50\%$ (691 polygons).



(c) BOCD vs. GFW; $T_{poly} = 75\%$ (1998 polygons) and $T_{poly} = 50\%$ (2518 polygons).

Fig. 10. Detection timeliness and delay analysis (0.1–1 ha, Amazonia). Density curves of detection dates relative to MapBiomas Alerta ($t_{MB_{after}}$) for $T_{poly} = 75\%$ and $T_{poly} = 50\%$ (1) (left). Monthly number of detections for $T_{poly} = 75\%$ (middle). Monthly median detection delay relative to MapBiomas Alerta for $T_{poly} = 75\%$ (right).

Table 3
Confusion matrix for deforested polygons ≥1 ha over Cerrado. T_{poly} defined in (1); 'small' refers to small-scale polygons (0.1–1 ha) reported for comparison, and 'large' to categories A_1–A_3 combined.

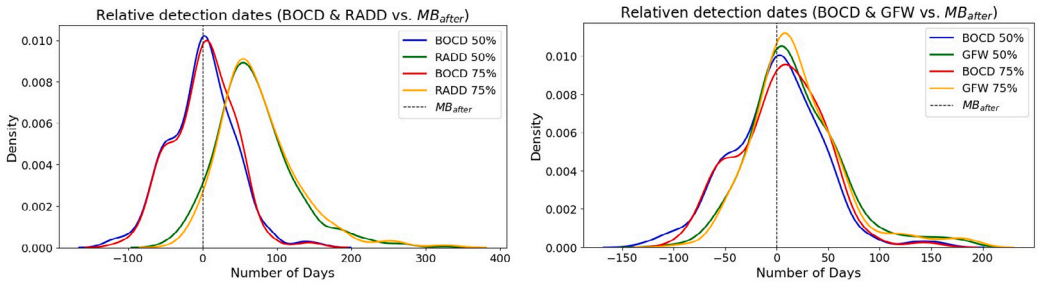
True positive [%]			
T_{poly}	BOCD C3 - small	BOCD C3 - large	GLAD-L
75%	39.9	64.2	1.6
50%	71.1	73.8	6.4
30%	86.7	86.6	12.2
10%	95.2	94.4	30.0
False positive [%]			
75%	0	0	0.4
50%	0	0	0.4
30%	0.5	0	0.4
10%	3.6	2.4	2.4

detected polygons. Similar findings are observed for both clearing sizes, with BOCD detecting disturbances, on average, two months earlier than the individual systems RADD and GLAD-L. When compared with GFW, however, the median difference in detection dates is reduced to approximately three days. Moreover, in most cases, BOCD can identify disturbances before the parcels are officially recorded as deforested in MapBiomas Alerta.

4.3. Spatial and temporal detection patterns across biomes

Detection examples from different methods across several MapBiomas Alerta polygons (0.1–1 ha) in Amazonia reveal that existing methods, especially GLAD-L and GFW, generate detections extending well beyond the boundaries of the reference polygons (Fig. 12). Although, in this work, false alarms are quantified temporally rather than spatially, this over-detection still provides a visual indication of BOCD's finer detection capabilities. Specifically, incorporating spatial context, as discussed in Section 3.5, refines the probability of deforestation without compromising spatial accuracy (Table 4).

The results, presented in Table 5, highlight strong seasonal variability in the Sentinel-1 time series over the Cerrado, with seasonality increasing as canopy closure and vegetation density decrease, i.e., seasonality is higher in savannas than in forests. Examples of BOCD detections over seasonality-affected time series in Fig. 13 show the robustness of BOCD in detecting disturbances despite seasonal effects. Further examples in Fig. 14 illustrate the spatial precision of BOCD detections over savanna formations, even at very low T_{poly}. Lowering T_{poly}, although not a user-tunable parameter, serves to highlight the resilience of BOCD to false alarms while still maintaining acceptable spatial precision in the detections.



(a) BOCD vs. RADD; $T_{poly} = 75\%$ (191 polygons) and $T_{poly} = 50\%$ (530 polygons).

(b) BOCD vs. GFW; $T_{poly} = 75\%$ (196 polygons) and $T_{poly} = 50\%$ (404 polygons).

Fig. 11. Density curves of detection dates (≥1 ha, Amazonia) relative to MapBiomas Alerta ($t_{MB_{after}}$) at two overlap thresholds T_{poly} (1).

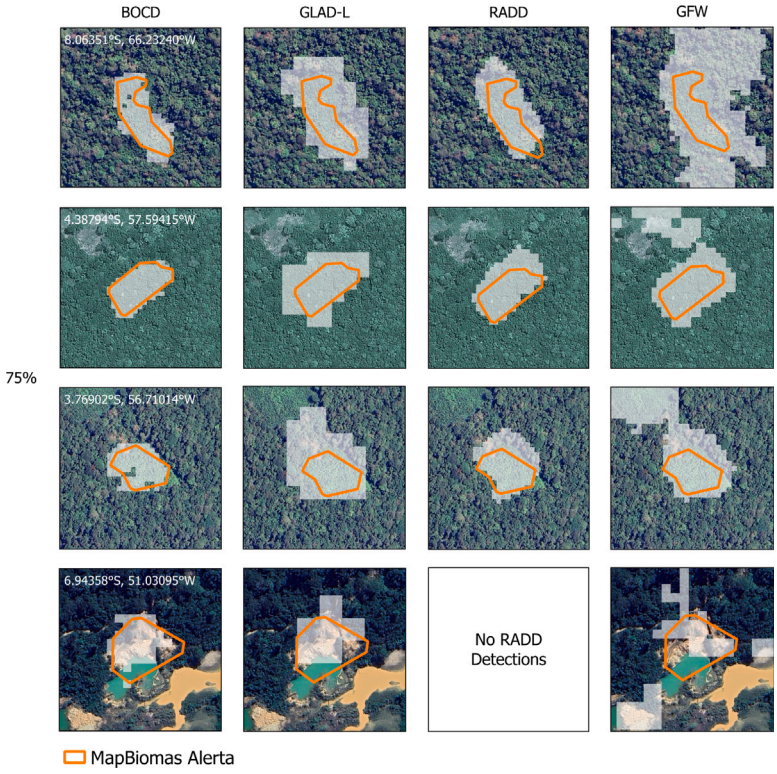


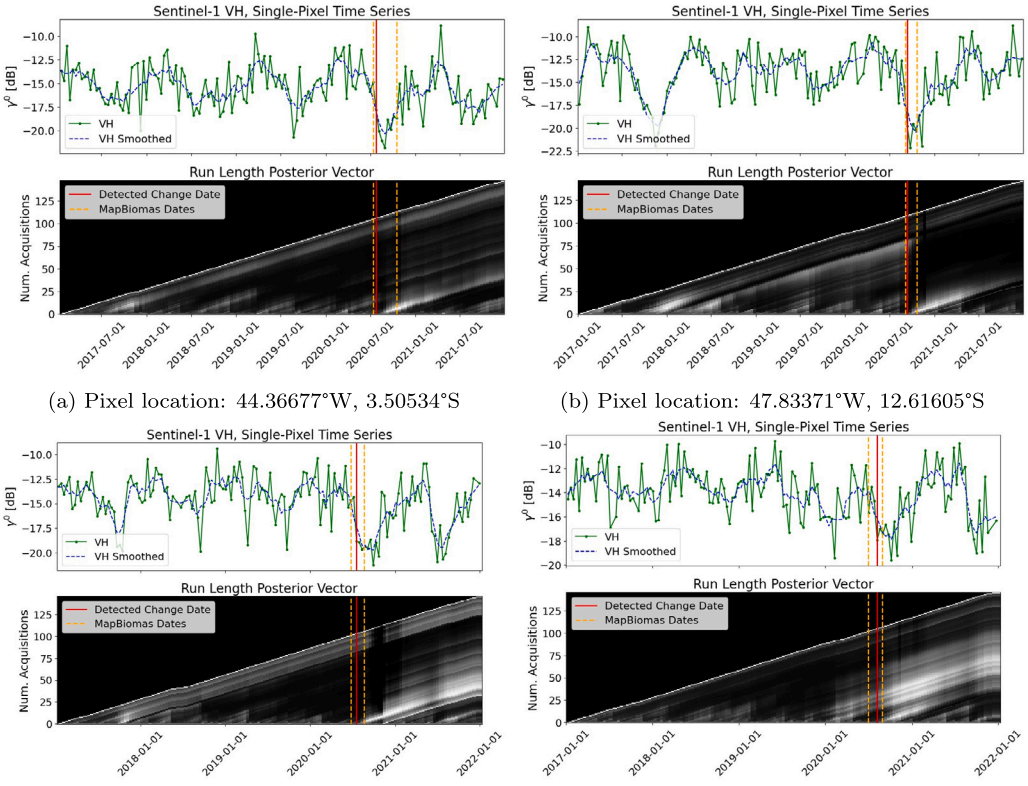
Fig. 12. Examples of BOCD C3, GLAD-L, RADD and GFW's detections (0.1–1 ha, light gray) of MapBiomas Alerta polygons (orange) over Amazonia ($T_{poly} = 75\%$, (1)). Optical background image from ©Google Earth (taken in 2024).

For the interested reader, maps of BOCD C3 detections over the MapBiomas Alerta small-clearings reference dataset (0.1–1 ha) for Amazonia and the Cerrado are included in Appendix B.

5. Discussion

5.1. BOCD as a forest loss detector

The analysis of small clearings (0.1–1 ha, excluding selective logging) in the Amazon biome (see Section 4.1.1) highlights BOCD's superior ability to detect small-scale forest loss compared to existing methods. Moreover, BOCD demonstrates considerably lower false alarm rates than the top-performing NRT method, GFW, even at very low T_{poly} values and less conservative configurations. The choice of priors in BOCD notably influenced detection rates, especially when transitioning from the most conservative configuration (C1) to less conservative ones. In contrast, adjustments to priors among the less conservative configurations had a less pronounced impact. While BOCD C2 and C3 are optimized for intermediate susceptibility, BOCD C4 is more prone to flagging changes. Given the comparable performance of BOCD C3 and C4 in terms of true detections and false alarms, adjusting the degrees of freedom, α, of the t-posterior predictive distribution proved to be the most effective method for enhancing responsiveness to changes. BOCD C2, which features a lower shape parameter, κ, for the t-distribution, exhibited slightly reduced performance in detecting



(a) Pixel location: 44.36677°W, 3.50534°S

(b) Pixel location: 47.83371°W, 12.61605°S

(c) Pixel location: 42.41667°W, 2.74804°S

(d) Pixel location: 44.54985°W, 15.91964°S

Fig. 13. Examples of single-pixel backscatter time series showing deforestation events and posterior run length probabilities (brighter colors indicate higher probabilities). BOCD detection dates are in red; MapBiomas Alerta pre- and post-change dates in orange. Blue lines show smoothed time series for illustration only. For interpretation of the color references in this figure legend, please refer to the web version of the article.

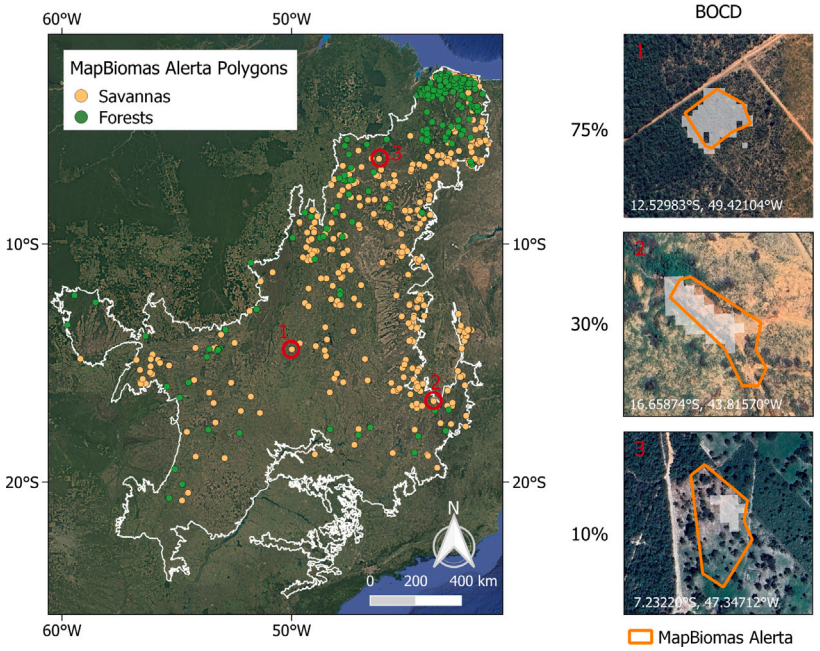


Fig. 14. Spatial distribution of BOCD C3 detections in the Cerrado (0.1–1 ha) highlighting the original land cover, and examples of BOCD detections (light gray) over savanna formations for varying T_{poly} (1). Optical background image from ©Google Earth (taken in 2024).

Table 4
Confusion matrix comparing BOCD C3 with and without spatial context (H) for deforested polygons in the 0.1–1 ha range in Amazonia and Cerrado. T_{poly} defined in (1).

	T_{poly}	True positive [%]		False positive [%]	
		BOCD C3	BOCD C3 - No H	BOCD C3	BOCD C3 - No H
Amazonia	75%	76.30	48.55	0	0
	50%	90.17	64.20	0.12	0.12
	30%	95.93	89.11	0.30	0.36
	10%	97.16	94.29	5.43	5.55
Cerrado	75%	39.90	33.07	0	0
	50%	71.07	66.93	0	0
	30%	86.65	82.51	0.51	1.02
	10%	95.23	90.46	3.57	4.59

Table 5
Seasonal variability metrics by vegetation type in the Cerrado, based on the MapBiomas Alerta 2020 dataset prior to deforestation. Sentinel-1 time series over 4 years were smoothed to reduce speckle and highlight seasonal patterns.

Variability metric	Forests	Savannas
Mean seasonal peak-to-peak amplitude [dB]	7.37	7.42
Mean root-mean-square amplitude [dB]	1.57	1.65

forest loss compared to BOCD C3. However, it was less affected by false alarms, especially at lower T_{poly} values.

In the Cerrado biome, BOCD demonstrates significantly improved performance in detecting small clearings compared to GLAD-L. Unlike in the Amazon biome, adjusting the priors does not have as significant an impact in the Cerrado. At $T_{\text{poly}} = 10\%$ (Fig. 9(b)), an unexpectedly lower false alarm rate is observed for BOCD C3 compared to BOCD C2, contrasting with the findings in the Amazon biome. This observation illustrates that varying the values of BOCD priors can influence susceptibility in different ways. Although the parameter adjustments in Section 3.6 guide optimization of change detection responsiveness, their effects can vary by dataset and should be carefully considered. Ultimately, it is important to note that user preferences play a crucial role in system optimization. For example, some users may prioritize minimizing false alarms, even at the cost of occasional omissions, especially when the method is used to inform deforestation deterrent actions. Others, however, may prioritize detecting all forest disturbances, regardless of the false alarms.

The evaluation of larger clearings demonstrates that BOCD significantly outperforms GLAD-L in the Cerrado. In the Amazon, BOCD delivers performance comparable to RADD but registers fewer true positive detections than GFW, which, however, exhibits a higher false alarm rate. A reversed trend emerges for larger deforested parcels compared to smaller ones, with BOCD performing roughly twice as well in the Cerrado compared to the Amazon at $T_{\text{poly}} = 75\%$. This fact can be explained by the different deforestation practices in the two biomes. In the Cerrado, large clearings are typically associated with the establishment of soybean plantations, which require completely cleared soils (RAD2023, 2024). In contrast, in the Amazon, large clearings are often intended for pasture, and previous studies (Balling et al., 2023) have shown that deforestation of extensive parcels often results in substantial amounts of debris and fallen trees being left on the ground. This debris can remain throughout the dry season and is only burned prior to the onset of the rainy season, due to the greater difficulty of clearing dense forest biomass. As a result, the presence of residual vegetation delays or even prevents the detection of deforestation with C-band SAR data, an effect that is particularly pronounced for larger parcels, thus explaining the observed trends. The performance of GFW highlights the potential of optical data to overcome this problem and more effectively detect larger disturbances.

5.2. BOCD adaptability to seasonality

The performance of BOCD in the Cerrado, which is affected by strong seasonality as shown in Section 4.3, demonstrates the iterative adaptability of the method. At each new observation, BOCD adapts to the current data statistics through the parameter update described in (6). A change is detected when variations fall outside the modeled statistics, even after adaptation, and persist for several consecutive observations, resulting in a significant drop in the most probable run length. This temporal persistence requirement makes BOCD resilient to single outliers, such as those caused by rain cells, which typically affect only one acquisition and do not suffice to trigger a false alarm. As shown in Fig. 13, sudden and short-lived drops in reflectivity do not lead to false detections.

The results obtained in the Cerrado highlight BOCD's superior detection capabilities compared to optical-based monitoring. For completeness, it is noted that no comparison was made with the LUCA dataset because its forest loss products are not publicly available, and its accuracy metrics are reported at the continental rather than biome level (Mullissa et al., 2024).

5.3. Spatial detection accuracy

Section 4.3 demonstrates the tendency of existing NRT forest loss monitoring methods to over-detect small-scale forest loss, particularly in optical-based approaches such as GLAD-L and GFW. In contrast, BOCD exhibits finer detection capabilities, demonstrating its potential in identifying small clearings. This capability is attributed to the use of unfiltered Sentinel-1 RTC data (≈ 20 m). To the authors' knowledge, all existing SAR-based forest loss monitoring methods, including RADD, employ spatial filtering to reduce speckle-related data variability. However, spatial filtering reduces spatial resolution, which may lead to over-detection and omission of finer disturbances.

The incorporation of spatial context into the change detection process, as described in Section 3.5, involves adjusting the probability of deforestation for pixels directly adjacent to a confirmed disturbance. This approach leverages spatial information within a single sensor's data without artificially boosting confidence through cross-sensor validation. Cross-sensor integration using spatial neighborhood rules across multiple satellite datasets can increase false alarms by aggregating errors from each sensor (Reiche et al., 2024), as observed in the integrated GFW alerts. In contrast, the use of spatial context in BOCD reduces false detections rather than amplifying them. The results presented in Section 4.3 show a notable increase in true positive detections within the Amazon — particularly at high T_{poly} values — when spatial context is considered. In the Cerrado, the improvement in true detections is more modest but still significant. Additionally, the inclusion of spatial context leads to a slight reduction in false alarms across both biomes.

5.4. Temporal detection accuracy

The temporal results in Section 4.2 show that, in most cases, BOCD detects disturbances before the parcels are officially recorded as deforested in MapBiomas Alerta. Compared with the individual methods RADD and GLAD-L, BOCD detects both small- and large-scale disturbances, on average, two months earlier. The results shown in Fig. 10 indicate that BOCD's timeliness is not affected by seasonal variations. As a SAR-based method, BOCD generally remains unaffected by rainy season conditions. Fig. 10(c) shows that the median detection delay

of BOCD is comparable to that of GFW, mainly because most clearings occur during the dry season. During this time, GFW likely benefits from reduced cloud cover, enabling faster detection using optical imagery. However, as an integrated approach, GFW's timeliness comes at the cost of increased false alarms.

5.5. Potential sources of omissions and false alarms

The potential sources of omissions and false alarms within the BOCD method are primarily due to limitations in the backscatter signals and environmental factors that influence the radar data. Particularly, omissions can occur when the backscatter signal processed by BOCD does not show sufficient signs of change. Another factor contributing to omissions is heavy rainfall immediately following deforestation. The moisture can saturate the C-band backscatter, masking the effects of forest clearing. Additionally, residual biomass left on the ground after deforestation, especially in the cross-polarized channel, which is more sensitive to volume scattering, can also cause omissions (Picard et al., 2004). This behavior is expected to be consistent across both analyzed biomes.

False alarms, on the other hand, may arise when the algorithm mistakenly attributes changes in backscatter to deforestation. One of the primary causes of false alarms is seasonality (Magagi et al., 2022). Although the algorithm adapts to seasonal fluctuations through a hidden Markov model, strong seasonal changes in backscatter can still trigger false detections. Another major contributor is short-term flooding or variations in the water table (Mermoz et al., 2021). In flood-prone regions, such as parts of the Amazon floodplain, fluctuations in water levels can cause abrupt changes in backscatter, which may be misinterpreted as deforestation. Additionally, forest degradation and canopy disturbances, especially in dense rainforests like the Amazon, can lead to false alarms (Hoekman and Quiñones, 2000). Events such as selective logging, natural treefall, or edge effects can alter backscatter without actual deforestation, and when these changes are coupled with soil visibility, they may be incorrectly classified as forest loss.

5.6. Future work

Building on the findings of this study, several promising directions for future research emerge. First, by avoiding data stack filtering, the proposed method shows the capability to detect small-scale forest loss, highlighting the need to assess its effectiveness in detecting selective logging — particularly in regions with well-established reference data. Second, the adaptability of the Bayesian approach to seasonal trends opens the door to making the method operational and to evaluating deforestation monitoring performance in regions with pronounced seasonality. Third, the stronger performance of GFW compared to BOCD in detecting larger-scale clearings underscores the potential of extending BOCD to a multi-source framework, such as integrating Sentinel-1 and Sentinel-2 data. Exploring these directions would not only enhance the versatility of the proposed method but also offer valuable insights for addressing forest loss across diverse landscapes by leveraging multiple data sources.

6. Conclusions

This paper presented an unsupervised, SAR-based forest disturbance detection method with NRT capabilities, grounded in Bayesian inference. Building on the methodology proposed in Adams and MacKay (2007), the BOCD approach was adapted to process single-polarization Sentinel-1 SAR data through Bayesian conjugate analysis. Furthermore, a strategy was developed to incorporate proximity to previously observed disturbances, using a survival analysis framework.

The method was tested on two subsets of the 2020 MapBiomas Alerta reference dataset: one comprising small-scale clearings (0.1–1 ha) and the other comprising larger-scale clearings. The results demonstrated the method's efficacy in detecting small-scale disturbances, made possible by preserving the spatial resolution of the measurements through the avoidance of spatial filtering. The adaptability of the Bayesian framework enabled the method to account for variations and trends in the data, showing strong potential for monitoring both dense rainforests and seasonally dynamic regions such as the Cerrado. When detecting small-scale disturbances, BOCD outperformed existing operational NRT methods by achieving higher true detection rates in both biomes while maintaining low false alarm rates. In the case of larger clearings, BOCD performed comparably to RADD but achieved lower detection rates than GFW, which integrates three alerting systems. However, GFW's increased detection rate came at the cost of higher false alarms, likely due to the aggregation of false detections from multiple sensors. Furthermore, a timeliness analysis showed that BOCD's performance aligned with that of GFW and significantly surpassed that of GLAD-L and RADD individually.

Overall, the BOCD method showed strong potential as an advancement in forest loss monitoring, offering a versatile and robust approach for NRT detection across diverse ecosystems and supporting more precise and timely forest conservation efforts.

CRediT authorship contribution statement

Marta Bottani: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Laurent Ferro-Famil:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Juan Doblas Prieto:** Validation, Resources. **Stéphane Mermoz:** Validation. **Alexandre Bouvet:** Supervision. **Thierry Koleck:** Resources. **Thuy Le Toan:** Supervision.

Declaration of competing interest

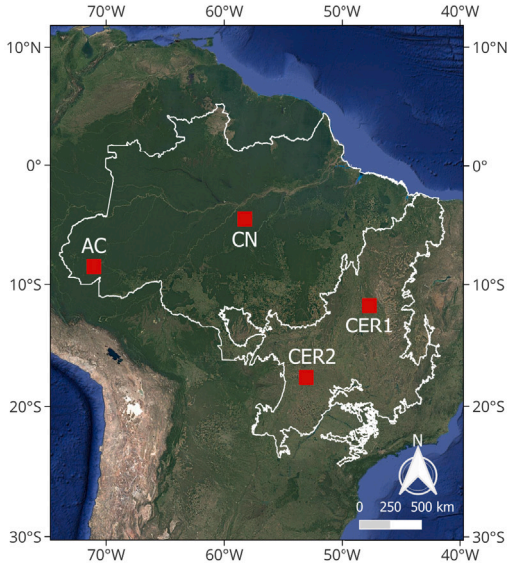
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

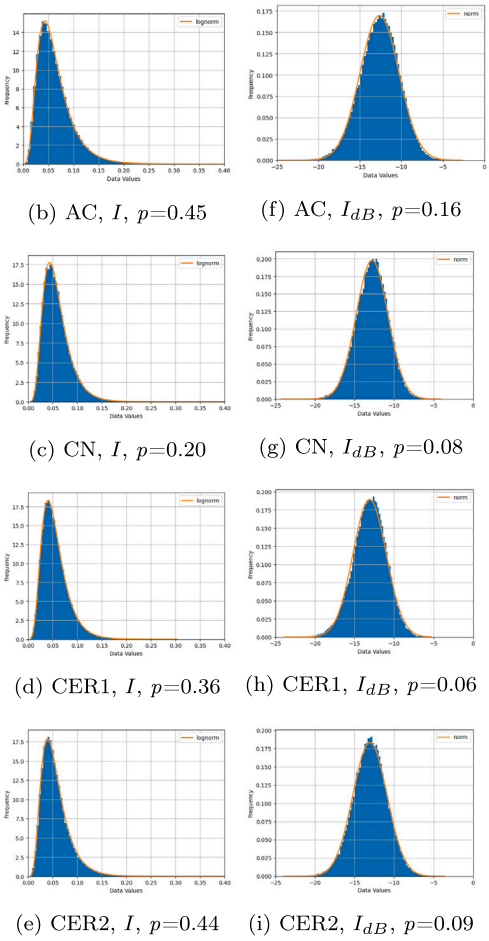
Marta Bottani's research was supported by the Centre National d'Études Spatiales (CNES), France and ISAE Supaero, France, and was hosted at the TéSA and CESBIO laboratories in Toulouse, France. This study also received partial funding through the EUR TESS grant (N° ANR-18-EURE-0018) as part of the Programme des Investissements d'Avenir, along with additional support from ISAE Supaero, France and GdR IASIS mobility, France grants. The authors also extend their gratitude to the Universidade Federal do Pará (UFPA), Brazil for providing Marta Bottani with the opportunity to engage in a valuable and enriching exchange experience on their campus.

Appendix A. Statistical model of Sentinel-1 unfiltered RTC data

To establish the statistical model for the Sentinel-1 data and subsequently determine the posterior probability and posterior predictive distributions required by the BOCD algorithm, various candidate density functions are fitted to both the backscatter intensity, I, and the logarithmic intensity, I_{dB}, data. The fits are performed using non-disturbed vegetated data samples, each consisting of approximately 300 by 300 Sentinel-1 RTC pixels, extracted from several areas within the study region. For illustrative purposes, histograms from two areas in the Amazon region and two areas in the Cerrado region are shown in Fig.



(a) Location of the reference sites: Acre (AC, centered on: 8.41°S and 70.58°W), Calha Norte (CN, centered on: 4.15°S and 58.1°W), Cerrado 1 (CER1, centered on: 12.18°S and 47.42°W), Cerrado 2 (CER2, centered on: 16.75°S and 53.17°W). Optical background image from Google Earth (©2024 Google).



(b) AC, I, p=0.45 (f) AC, I_{dB}, p=0.16

(c) CN, I, p=0.20 (g) CN, I_{dB}, p=0.08

(d) CER1, I, p=0.36 (h) CER1, I_{dB}, p=0.06

(e) CER2, I, p=0.44 (i) CER2, I_{dB}, p=0.09

Fig. A.15. PDF fits of I and I_{dB}, Sentinel-1 RTC data (≈300 × 300 pixels) over different reference sites. p indicates the p-value.

A.15. The analysis results indicate that the I data best fits a log-normal distribution, meaning that the I_{dB} statistics are well represented by a normal distribution.

The quality of the PDF fits is assessed using the Kolmogorov–Smirnov (KS) test (Chakravarti et al., 1967), which compares a dataset's empirical distribution to a theoretical distribution. This test evaluates the statistical significance of the null hypothesis, which asserts that the sample data aligns with the distribution PDF. The KS statistic quantifies the difference between observed and expected cumulative distributions, whereas the p-value indicates the likelihood of this difference occurring if the null hypothesis is true. A p-value below a set significance level (typically 0.05) suggests a notable discrepancy, while a higher value indicates no strong evidence to reject the hypothesis. As presented in Fig. A.15, the p-values corresponding to the KS test for different fits indicate statistical significance, thereby supporting the hypothesis that the data samples are generated from the considered distributions. The choice adopted within this work is to use log-scale Sentinel-1 data and assume a normal likelihood.

Appendix B. Supplementary results

Table B.6
Confusion matrix for deforested polygons in the 0.1–1 ha range over Amazonia. T_{poly} defined in (1). Labels 'C1'–'C4' refer to different BOCD configurations.

True positive [%]							
T_{poly}	BOCD C1	BOCD C2	BOCD C3	BOCD C4	GLAD-L	RADD	GFW
75%	53.3	72.3	76.3	77.6	11.3	37.6	59.0
50%	74.8	88.0	90.2	90.7	20.4	64.9	77.6
30%	84.6	94.6	95.9	96.1	27.7	77.5	87.3
10%	87.2	96.0	97.2	97.3	33.7	85.8	92.6
False positive [%]							
75%	0	0	0	0	0.4	0.2	0.7
50%	0.1	0.1	0.1	0.1	0.7	0.5	1.6
30%	0.2	0.3	0.3	0.3	1.4	1.5	5.4
10%	1.9	4.5	5.4	5.7	3.7	5.6	13.6

Table B.7
Confusion matrix for deforested polygons in the 0.1–1 ha range over Cerrado. T_{poly} defined in (1). Labels 'C1'–'C4' refer to different BOCD configurations.

True positive [%]					
T_{poly}	BOCD C1	BOCD C2	BOCD C3	BOCD C4	GLAD-L
75%	38.2	39.1	39.9	41.3	1.0
50%	67.7	70.3	71.1	72.3	3.7
30%	83.9	85.1	86.7	87.3	9.7
10%	92.7	93.2	95.2	95.7	20.0
False positive [%]					
75%	0	0	0	0	0
50%	0	0	0	0	0
30%	0.5	0.5	0.5	0.5	0
10%	3.6	4.1	3.6	4.6	0.5

Table B.8
Spatial accuracy metrics for deforested polygons in the 0.1–1 ha range over Amazonia. T_{poly} defined in (1). Labels 'C1'–'C4' refer to different BOCD configurations. p: precision, s: sensitivity, F_1: F1-score, TP: true positive, FP: false positive, FN: false negative. Metrics expressed as percentages.

	T_{poly}	BOCD C1	BOCD C2	BOCD C3	BOCD C4	GLAD-L	RADD	GFW
$p = \frac{TP}{TP+FP}$	75%	100	100	100	100	99.4	99.8	98.5
	50%	99.9	99.9	99.9	99.9	98.5	99.7	99.1
	30%	99.9	99.8	99.8	99.8	97.7	99.1	97.2
	10%	99.0	97.9	97.5	97.4	95.3	95.2	93.6
$s = \frac{TP}{TP+FN}$	75%	53.3	72.3	76.3	77.6	11.3	37.6	59.0
	50%	74.8	88.0	90.2	90.7	20.4	65.0	77.6
	30%	84.6	94.6	95.9	96.1	27.7	77.5	87.3
	10%	87.2	96.0	97.2	97.3	33.6	85.8	92.6
$F_1 = \frac{2ps}{p+s}$	75%	69.6	83.9	86.6	87.4	20.2	54.6	74.1
	50%	85.6	93.6	94.8	95.1	33.8	78.6	87.1
	30%	91.6	97.2	97.8	97.9	43.2	87.0	92.0
	10%	92.7	96.9	97.3	97.3	49.7	90.5	93.1

Table B.9
Spatial accuracy metrics for deforested polygons in the 0.1–1 ha range over Cerrado. T_{poly} defined in (1). Labels 'C1'–'C4' refer to different BOCD configurations. p: precision, s: sensitivity, F_1: F1-score, TP: true positive, FP: false positive, FN: false negative. Metrics expressed as percentages.

	T_{poly}	BOCD C1	BOCD C2	BOCD C3	BOCD C4	GLAD-L
$p = \frac{TP}{TP+FP}$	75%	100	100	100	100	100
	50%	100	100	100	100	100
	30%	99.8	99.8	99.8	99.8	100
	10%	98.8	98.7	98.8	99.2	99.2
$s = \frac{TP}{TP+FN}$	75%	38.2	39.1	39.9	41.3	1.0
	50%	67.7	70.3	71.1	72.3	3.7
	30%	83.9	85.1	86.6	87.3	9.7
	10%	92.7	93.2	95.2	95.7	20.0
$F_1 = \frac{2ps}{p+s}$	75%	55.2	56.2	57.0	58.5	19.0
	50%	80.8	82.5	83.1	83.9	7.1
	30%	91.2	91.8	92.8	93.1	17.7
	10%	95.7	95.8	97.0	97.4	33.3

Table B.10
Confusion matrix for deforested polygons ≥1 ha over Amazonia. T_{poly} defined in (1). $A_1 - A_3$ are the size categories defined in Table 1. 'TOT' indicates the performance for categories $A_1 - A_3$ jointly.

True positive [%]																
T_{poly}	BOCD C3				RADD				GLAD-L				GFW			
	TOT	A_1	A_2	A_3	TOT	A_1	A_2	A_3	TOT	A_1	A_2	A_3	TOT	A_1	A_2	A_3
75%	32.2	39.2	25.7	20.6	33.3	36.6	29.2	31.0	11.6	10.8	12.4	12.7	68.3	69.9	66.2	67.5
50%	65.3	73.3	59.0	49.2	62.2	68.2	58.1	48.4	24.9	24.8	23.4	29.4	86.0	86.7	85.8	83.3
30%	79.9	86.9	73.7	67.5	76.9	77.8	66.4	58.9	32.7	33.6	31.9	32.9	91.5	91.9	88.5	84.9
10%	95.9	97.2	94.5	94.4	90.3	93.7	92.1	91.9	40.0	43.8	41.7	43.7	96.1	97.8	96.3	95.4
False positive [%]																
75%	0	0	0	0	0	0	0	0	0	0	0	0	0.4	0.8	0.4	0
50%	0	0	0	0	0	0	0	0	0.8	0.8	1.2	0	4.0	4.2	8.6	3.2
30%	0	0	0	0	0.4	0.5	1.1	0	1.2	1.2	1.2	1.2	16.0	15.9	16.0	16.3
10%	0.8	1.4	0.5	0	2.8	2.7	2.7	3.3	3.4	3.6	3.2	3.3	52.6	52.7	52.6	52.2

Table B.11
Confusion matrix for deforested polygons ≥1 ha over Cerrado. T_{poly} defined in (1). $A_1 - A_3$ are the size categories defined in Table 1. 'TOT' indicates the performance for categories $A_1 - A_3$ jointly.

True positive [%]								
T_{poly}	BOCD C3				GLAD-L			
	TOT	A_1	A_2	A_3	TOT	A_1	A_2	A_3
75%	64.2	63.5	68.1	59.1	1.6	1.5	2.2	1.7
50%	73.8	74.5	76.2	68.7	6.4	6.5	6.5	6.1
30%	86.6	88.5	89.7	78.3	12.2	12.0	12.4	12.2
10%	94.4	97.5	93.5	90.4	30.0	30.0	29.7	30.4
False positive [%]								
75%	0	0	0	0	0.4	0.8	0	0
50%	0	0	0	0	0.4	0.8	0	0
30%	0	0	0	0	0.4	0.8	0	0
10%	2.4	0.8	2.7	6.25	2.4	2.4	2.7	2.1

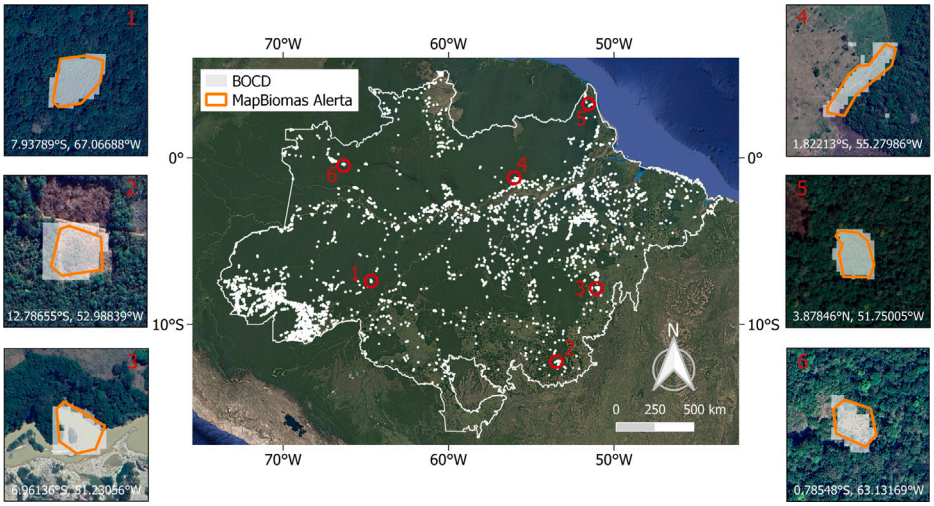


Fig. B.16. Map of BOCD C3 detections over the MapBiomas Alerta small-clearings reference dataset (0.1–1 ha) in Amazonia. Reference polygons in orange; BOCD detections in light gray. Optical background: Google Earth (© 2024 Google).

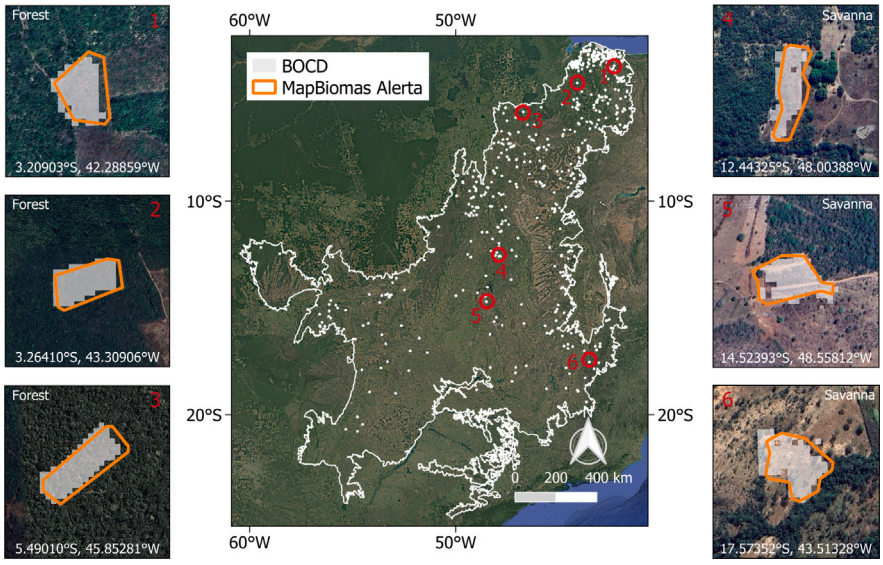


Fig. B.17. Map of BOCD C3 detections over the MapBiomas Alerta small-clearings reference dataset (0.1–1 ha) in the Cerrado. Reference polygons in orange; BOCD detections in light gray. Optical background: Google Earth (© 2024 Google).

Data availability

Both the Sentinel-1 input data and the MapBiomas Alerta reference dataset are freely accessible online.

References

Adams, R.P., MacKay, D.J., 2007. Bayesian online changepoint detection. arXiv:0710.3742. URL https://doi.org/10.48550/arXiv.0710.3742.

Asner, G.P., Llactayo, W., Tupayachi, R., Luna, E.R., 2013. Elevated rates of gold mining in the amazon revealed through high-resolution monitoring. Proc. Natl. Acad. Sci. 110 (46), 18454–18459, URL https://doi.org/10.1073/pnas.1318271110.

Assunção, J., Gandour, C., Rocha, R., Rocha, R., 2019. The effect of rural credit on deforestation: Evidence from the Brazilian amazon. Econ. J. 130, URL https://doi.org/10.1093/ej/uez060.

Ballère, M., Bouvet, A., Mermoz, S., Toan, T.L., Koleck, T., Bedeau, C., André, M., Forestier, E., Frison, P.-L., Lardeux, C., 2021. SAR data for tropical forest disturbance alerts in French Guiana: Benefit over optical imagery. Remote Sens. Environ. 252, 112159, URL https://doi.org/10.1016/j.rse.2020.112159.

Balling, J., Herold, M., Reiche, J., 2023. How textural features can improve SAR-based tropical forest disturbance mapping. Int. J. Appl. Earth Obs. Geoinf. 124, 103492, URL https://doi.org/10.1016/j.jag.2023.103492.

Barry, D., Hartigan, J.A., 1993. A Bayesian analysis for change point problems. J. Amer. Statist. Assoc. 88 (421), 309–319, URL https://doi.org/10.1080/01621459.1993.10594323.

Bourgoin, C., Ceccherini, G., Girardello, M., Vancutsem, C., Avitabile, V., Beck, P.S.A., Beuchle, R., Blanc, L., Duveiller, G., Migliavacca, M., Vieilledent, G., Cescatti, A., Achard, F., 2024. Human degradation of tropical moist forests is greater than previously estimated. Nature 631, 570–576, URL https://10.1038/s41586-024-07629-0.

Bouvet, A., Mermoz, S., Ballère, M., Koleck, T., Toan, T.L., 2018. Use of the SAR shadowing effect for deforestation detection with sentinel-1 time series. Remote. Sens. 10 (8), URL https://doi.org/10.3390/rs10081250.

Carstairs, H., Mitchard, E.T.A., McNicol, I., Aquino, C., Chezeaux, E., Ebanega, M.O., Dikongo, A.M., Disney, M., 2022. Sentinel-1 shadows used to quantify canopy loss from selective logging in gabon. Remote. Sens. 14 (17), URL https://doi.org/10.3390/rs14174233.

Chakravarti, I., Laha, R., Roy, J., 1967. Handbook of Methods of Applied Statistics, Volume I. Wiley, pp. 392–394.

Cox, D.R., 2001. Renewal theory. Handbook of Statist. 19, URL https://doi.org/10.1016/S0169-7161(01)19015-2.

de Freitas Bussinguer, J., de Mello Baptista, G.M., Sano, E.E., Leal, F., 2024. Understanding the spatio-temporal behavior of sentinel-1 SAR vegetation indices over the Brazilian Savanna. IEEE Trans. Geosci. Remote Sens. 62, 1–18, URL https://doi.org/10.1109/TGRS.2024.3381468.

Doblas, J., Lima, L., Mermoz, S., Bouvet, A., Reiche, J., Watanabe, M., Anna, S.S., Shimabukuro, Y., 2023. Inter-comparison of optical and SAR-based forest disturbance warning systems in the amazon shows the potential of combined SAR-optical monitoring. Int. J. Remote Sens. 44 (1), 59–77, URL https://doi.org/10.1080/01431161.2022.2157684.

Doblas, J., Reis, M.S., Belluzzo, A.P., Quadros, C.B., Moraes, D.R.V., Almeida, C.A., Maurano, L.E.P., Carvalho, A.F.A., Sant'Anna, S.J.S., Shimabukuro, Y.E., 2022. DETER-r: An operational near-real time tropical forest disturbance warning system based on sentinel-1 time series analysis. Remote. Sens. 14 (15), URL https://doi.org/10.3390/rs14153658.

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. 12 (23), URL https://doi.org/10.3390/rs12233922.

Fearnhead, P., Liu, Z., 2007. On-line inference for multiple changepoint problems. J. R. Stat. Soc. Ser. B Stat. Methodol. 69 (4), 589–605, URL https://doi.org/10.1111/j.1467-9868.2007.00601.x.

Feller, W., 1968. An Introduction to Probability Theory and its Applications: Volume I. Wiley.

Finer, 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 360, 1303–1305, URL https://doi.org/10.1126/science.aat1203.

Forzieri, G., Dakos, V., McDowell, N.G., Ramdane, A., Cescatti, A., 2022. Emerging signals of declining forest resilience under climate change. Nature 608, 534–539, URL https://doi.org/10.1038/s41586-022-04959-9.

GFW, 2024. World resources institute - global forest watch. https://www.globalforestwatch.org/. (Accessed 26 November 2024).

Goodman, J.W., 1976. Some fundamental properties of speckle*. J. Opt. Soc. Am. 66 (11), 1145–1150, URL https://doi.org/10.1364/JOSA.66.001145.

Haddad, N., Brudvig, L., Clobert, J., Davies, K., Gonzalez, A., Holt, R., Lovejoy, T., Sexton, J., Austin, M., Collins, C., Cook, W., Damschen, E., Ewers, R., Foster, B., Jenkins, C., King, A., Laurance, W., Levey, D., Margules, C., Melbourne, B., Nicholls, A., Orrock, J., Song, D., Townshend, J., 2015. Habitat fragmentation and its lasting impact on earth's ecosystems. Sci. Adv. 1, URL https://doi.org/10.1126/sciadv.1500052.

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 (3), 034008, URL 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, URL https://doi.org/10.1038/s41558-020-00976-6.

Hoang, N.T., Kanemoto, K., 2021. Mapping the deforestation footprint of nations reveals growing threat to tropical forests. Nat. Ecol. Evol. 5, 845–853, URL https://doi.org/10.1038/s41559-021-01417-z.

Hoekman, D., Quiñones, M., 2000. Land cover type and biomass classification using AirSAR data for evaluation of monitoring scenarios in the Colombian amazon. IEEE Trans. Geosci. Remote Sens. 38, 685–696, URL https://doi.org/10.1109/36.841998.

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

Magagi, R., et al., 2022. Potential of L-and C-bands polarimetric SAR data for monitoring soil moisture over forested sites. Remote. Sens. 14, 5317, URL https://doi.org/10.3390/rs14215317.

MapBiomas, 2024. MapBiomas alerta - deforestation alert validation and refinement system with high-resolution images. https://alerta.mapbiomas.org/en. (Accessed 26 November 2024).

Mermoz, S., Bouvet, A., Koleck, T., Ballère, M., Toan, T.L., 2021. Continuous detection of forest loss in Vietnam, Laos, and Cambodia using sentinel-1 data. Remote. Sens. 13 (23), URL https://doi.org/10.3390/rs13234877.

Miranda, J., Börner, J.J., Kalkuhl, M., Soares-Filho, B.S., 2019. Land speculation and conservation policy leakage in Brazil. Environ. Res. Lett. 14, URL https://doi.org/10.1088/1748-9326/ab003a.

Miranda, N., Meadows, P.J., 2015. Radiometric Calibration of S-1 Level-1 Products Generated by the S-1 IPF. Technical Note ESA-EOPG-CSCOP-TN-0002, ESA, URL https://sentinel.esa.int/documents/247904/685163/S1-Radiometric-Calibration-V1.0.pdf.

Mullissa, A., Reiche, J., Herold, M., 2023. Deep learning and automatic reference label harvesting for sentinel-1 SAR-based rapid tropical dry forest disturbance mapping. Remote Sens. Environ. 298, 113799, URL https://doi.org/10.1016/j.rse.2023.113799.

Mullissa, A., Saatchi, S., Silva, R., Erickson, T., Provost, N., Osborn, F., Ashary, A., Moon, V., Melling, D., 2024. LUCA: A sentinel-1 SAR-based global forest land use change alert. Remote. Sens. 16, 2151, URL https://doi.org/10.3390/rs16122151.

Murphy, K.P., 2007. Conjugate Bayesian Analysis of the Gaussian Distribution. Technical Report, The University of British Columbia, URL https://www.cs.ubc.ca/~murphyk/Papers/bayesGauss.pdf.

Picard, G., Toan, T.L., Quegan, S., Caraglio, Y., Castel, T., 2004. Radiative transfer modeling of cross-polarized backscatter from a pine forest using the discrete ordinate and eigenvalue method. IEEE Trans. Geosci. Remote Sens. 42, 1720–1730, URL https://doi.org/10.1109/TGRS.2004.831229.

RAD2023, 2024. RADD2023 :Annual Deforestation Report of Brazil 2023. Technical Report, São Paulo, Brazil, p. 154, URL https://doi.org/10.1088/1748-9326/ac5193.

Reiche, J., Balling, J., Pickens, A., Masolele, R., Berger, A., Weisse, M., Mannarino, D., Gou, Y., Slagter, B., Donchyts, G., Carter, S., 2024. Integrating satellite-based forest disturbance alerts improves detection timeliness and confidence. Environ. Res. Lett. 19, URL https://doi.org/10.1088/1748-9326/ad2d82.

Reiche, J., Hamunyela, E., Verbesselt, J., Hoekman, D., Herold, M., 2018a. Improving near-real time deforestation monitoring in tropical dry forests by combining dense sentinel-1 time series with landsat and ALOS-2 PALSAR-2. Remote Sens. Environ. 204 (5), 147–161, URL https://doi.org/10.1016/j.rse.2017.10.034.

Reiche, J., Mullissa, A., Slagter, B., Gou, Y., Tsendbazar, N.-E., Odongo-Braun, C., Vollrath, A., Weisse, M.J., 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 (2), 024005, URL https://doi.org/10.1088/1748-9326/abd0a8.

Reiche, J., Verhoeven, R., Verbesselt, J., Hamunyela, E., Wielaard, N., Herold, M., 2018b. Characterizing tropical forest cover loss using dense sentinel-1 data and active fire alerts. Remote. Sens. 10, 777, URL https://doi.org/10.3390/rs10050777.

Robb, R., Cochran, G., 1963. Sampling techniques (john wiley & sons, 2nd edition). Proc. Edinb. Math. Soc. 102532, URL https://doi.org/10.1017/S0013091500025724.

Rodrigues, A.D.A., Macedo, M., Silvério, D., Maracahipes, L., Coe, M., Brando, P., Shimbo, J.Z., Rajão, R., Filho, B., Bustamante, M., 2022. Cerrado deforestation threatens regional climate and water availability for agriculture and ecosystems. Global Change Biol. 28, URL https://doi.org/10.1111/gcb.16386.

Shimabukuro, Y., Santos, J., Formaggio, A., Duarte, V., Rudorff, B., 2016. The Brazilian Amazon Monitoring Program: PRODES and DETER Projects. pp. 153–169, URL https://doi.org/10.1201/b13040-9.

Slagter, B., Reiche, J., Marcos, D., Mullissa, A., Lossou, E., Peña-Claros, M., Herold, M., 2023. Monitoring direct drivers of small-scale tropical forest disturbance in near real-time with sentinel-1 and -2 data. Remote Sens. Environ. 295, 113655, URL https://doi.org/10.1016/j.rse.2023.113655.

Small, D., 2011. Flattening Gamma: Radiometric terrain correction for SAR imagery. IEEE Trans. Geosci. Remote Sens. 49, 3081–3093, URL https://doi.org/10.1109/TGRS.2011.2120616.

Stevenson, M., 2007. An Introduction to Survival Analysis. Technical Report, EpiCentre, IVABS, Massey University, URL http://www.biecek.pl/statystykaMedyczna/Stevenson_survival_analysis_195.721.pdf.

Tacconi, L., Rodrigues, R., Maryudi, A., 2019. Law enforcement and deforestation: Lessons for Indonesia from Brazil. For. Policy Econ. 108, 101943, URL https://doi.org/10.1016/j.forpol.2019.05.029.

Trigueiro, W.R., Nabout, J.C., Tessarolo, G., 2020. Uncovering the spatial variability of recent deforestation drivers in the Brazilian Cerrado. J. Environ. Manag. 275, 111243, URL https://doi.org/10.1016/j.jenvman.2020.111243.

Vancutsem, C., Achard, F., Pekel, J.-F., Vieilledent, G., Carboni, S., Simonetti, D., Gallego, J., Aragão, L.E., Nasi, R., 2021. Long-term (1990–2019) monitoring of forest cover changes in the humid tropics. Sci. Adv. 7, URL https://doi.org/10.1126/sciadv.abe1603.

Verbesselt, J., Zeileis, A., Herold, M., 2012. Near real-time disturbance detection using satellite image time series. Remote Sens. Environ. 123, 98–108, URL https://doi.org/10.1016/j.rse.2012.02.022.

Watanabe, M., Koyama, C.N., Hayashi, M., Nagatani, I., Shimada, M., 2018. Early-stage deforestation detection in the tropics with L-band SAR. IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens. 11, 2127–2133, URL https://doi.org/10.1109/JSTARS.2018.2810857.

Watanabe, M., Koyama, C.N., Hayashi, M., Nagatani, I., Tadono, T., Shimada, M., 2021. Refined algorithm for forest early warning system with ALOS-2/PALSAR-2 scansar data in tropical forest regions. Remote Sens. Environ. 265, 112643, URL https://doi.org/10.1016/j.rse.2021.112643.

Ygorra, B., Frappart, F., Wigneron, J.-P., Moisy, C., Catry, T., Baup, F., Hamunyela, E., Riazanoff, S., 2021. Monitoring loss of tropical forest cover from sentinel-1 time-series: A cusum-based approach. Int. J. Appl. Earth Obs. Geoinf. 103, 102532, URL https://doi.org/10.1016/j.jag.2021.102532.

Zhao, K., Wulder, M.A., Hu, T., Bright, R., Wu, Q., Qin, H., Li, Y., Toman, E., Mallick, B., Zhang, X., Brown, M., 2019. Detecting change-point, trend, and seasonality in satellite time series data to track abrupt changes and nonlinear dynamics: A Bayesian ensemble algorithm. Remote Sens. Environ. 232, 111181, URL https://doi.org/10.1016/j.rse.2019.04.034.