Masters Program in Geospatial Technologies

USING SENTINEL-1 TIME SERIES FOR MONITORING DEFORESTATION IN REGIONS WITH HIGH PRECIPITATION RATE
Study Case: Chocó-Colombia

Jhon Eric Aunta Duarte

Dissertation submitted in partial fulfilment of the requirements
for the Degree of *Master of Science in Geospatial Technologies*

NOVA IMS Information Management School
Universitat Jaume I
Westfälische Wilhelms-Universität Münster

USING SENTINEL-1 TIME SERIES FOR MONITORING DEFORESTATION IN REGIONS WITH HIGH PRECIPITATION RATE

Study Case: Chocó-Colombia

Dissertation supervised by

Joel Dinis Baptista Ferreira da Silva, PhD

Universidade Nova de Lisboa

Co-supervised by:

Filiberto Pla, PhD.

Universitat Jaume I

Castellón de la Plana, Spain

Co-supervised by

Torsten Prinz, PhD.

WWU Münster University

February 2023

DECLARATION OF ORIGINALITY

I declare that the work described in this document is my own and not from someone else. All the assistance I have received from other people is duly acknowledged and all the sources (published or not published) are referenced. This work has not been previously evaluated or submitted to NOVA Information Management School or elsewhere.

Lisbon, 25.02.2023
Jhon Eric Aunta Duarte

[digital signature]

OR

[the signed original has been archived by the NOVA IMS services]

ACKNOWLEDGMENTS

First of all, I want to thank God for giving me the health and the opportunity to achieve this goal, I would like to extend my gratitude to my academic supervisors Dr. Joel Baptista Ferreira da Silva, Dr. Torsten Prinz and Dr. Filiberto Pla for their feedback and time. Also, I would like to thank Dr. Marco Painho, I thought that you always have the right words in the right moment, I really appreciated your support during the master program.

To dear fiends I made during the Master course, I really enjoyed the hours that expend with you, doing homework, projects and laughing.

To professors in the universities, thank you for sharing your experience and knowledge and pushing us in our life objectives.

Lastly, I want to express my love and gratitude to my wife Caterine, my parents Blanca and Benigno, my sister, Natalie, who were always there to listen to me.

USING SENTINEL-1 TIME SERIES FOR MONITORING DEFORESTATION IN REGIONS WITH HIGH PRECIPITATION RATE

Study Case: Chocó-Colombia

ABSTRACT

Despite nowadays there are many optical sensors out there, meteorological conditions in some places on the Earth makes very difficult to have access to images without clouds. Some of those places have unique ecosystems and landscape with natural forest that should be taken care of. SAR images has proven its capabilities for monitoring deforestation since the first sensors were deployed. Sentinel-1 allows to have free access to SAR data with high temporal resolution. Therefore, this study explores the use of SAR data for monitoring deforestation in places where the precipitation rate is too high. A time-series approach is used as framework to detect forest disturbances; the work tests if performing a combination of the Sentinel-1 bands through a modified version from the RFDI gets better results than the original bands; two methods for detecting changes along the time focus on deforestation are compared. The results show that VH band is the best input with similar overall accuracy with the two methods, around 80%, the mRFDI showed acceptable results but it does not prove any improvement on the deforestation events detected. It was concluded that with a workflow optimization, it can be used to overcome the optical images problem to monitor deforestation events.

KEYWORDS

SAR Time series

Deforestation

Change detection.

Monitoring System

KEYWORDS (In other language)

Série Temporal SAR

Desmatamento

Detecção de mudança

Sistema de monitoramento

ACRONYMS

ESA – European Spatial Agency

EO – Earth Observation

GLAD-L – Global Land Analysis and Discovery - Landsat

GLAD-S2 – Global Land Analysis and Discovery – Sentinel-2

GEE – Google Earth Engine

GRD – Ground Range Detected

GFW – Global Forest Watch

IDEAM – Hydrology, Meteorology and Environmental Studies of Colombia

mRFDI – Modified Radar Forest Degradation Index

NDVI – Normalized Difference Vegetation Index

NICFI – Norway's International Climate and Forests Initiative Satellite Data Program

RADD – Radar for Detecting Deforestation

RFDI – Radar Forest Degradation Index

REDD+ – Reducing of Emissions from Forest Degradation and Deforestation

RVI – Radar Vegetation Index

S1 – Sentinel-1

SAR – Synthetic Aperture Radar

SMByC – The Forest and Carbon Monitoring System

UNFCCC – United Nations Framework Convention on Climate Change

VH – Vertical Horizontal

VV – Vertical Vertical

INDEX OF THE TEXT

INDEX OF TABLES

INDEX OF FIGURES

1 INTRODUCTION

1.1. Context

Deforestation is a problem that Colombian Government has tracked since 2012 (Cabrera et al., 2014). The last report about deforestation in 2020 showed that it increased 8% more than in 2019, a total of 171685 ha has been deforested (IDEAM, 2021).

To monitor that problem deserves the attention of many government offices, since Natural Forest cover 52.3% of the Colombia inland territory (IDEAM, 2021).

In the framework of the well-known strategy: Reducing of Emissions from Forest Degradation and Deforestation REDD+, Colombia established its own path to fit the policies from the framework created by United Nations Framework Convention on Climate Change UNFCCC (United Nations Framework Convention on Climate Change, n.d.).

The Forest and Carbon Monitoring System SMByC is the project in charge of the IDEAM (stands for Institute of Hydrology, Meteorology and Environmental Studies in Spanish) which fits the REDD+ framework. The project aims for monitoring deforestation and the improvement of carbon stocks (Cabrera et al., 2014).

The specific objectives from the SMByC are, the official report for natural forest cover, develop methodologies for measuring the annual deforestation rate, create deforestation alerts every three months, model and simulate the driving factors for deforestation and degradation, monitoring the forests carbon stocks (IDEAM, n.d.-b).

The availability of Satellite Earth Observation EO data allows to measure deforestation and the changes in the forest land cover which has been one of the main fields of study through the time, the natural forest is part of the more common land cover surfaces.

Synthetic Aperture Radar SAR is one of the sources of those EO data and it exists many years ago, but usually it is not the first option when remote sensing images are needed. Deforestation monitoring systems using only SAR are very few around the world. For instance, Brazil has developed an option of Near Real Time deforestation

system based on SAR sentinel-1 images, every time new S1 image is available the system triggers the algorithms behind the system (Doblas et al., 2022).

Monitoring means to observe or check the progress of something through the time, in the framework of this thesis, monitoring deforestation means to check the progress of the forest disturbances in the natural forest land cover in the region selected as study area. The time lapse selected for this study goes from 01.01.2021 up to 30.09.2022.

1.2. Problem Statement and Motivation

Deforestation is a serious problem in Colombia, it is an illegal activity. Several government offices have projects which aim to monitor and track deforestation but mainly in Amazon and Orinoquia Regions.

The Colombian Pacific region has the highest precipitation rate in the country. (See figure 1., left side of the country). Chocó is the largest administrative division in the Pacific Region and its precipitation rate is between 5000 mm and 9000 mm per year in some specific places. Then the sky almost always has clouds, the region landscape is dominated by tropical forest, small towns small towns on the banks of the rivers, few roads, usually the rivers and streams are the way that people use to commute between places.

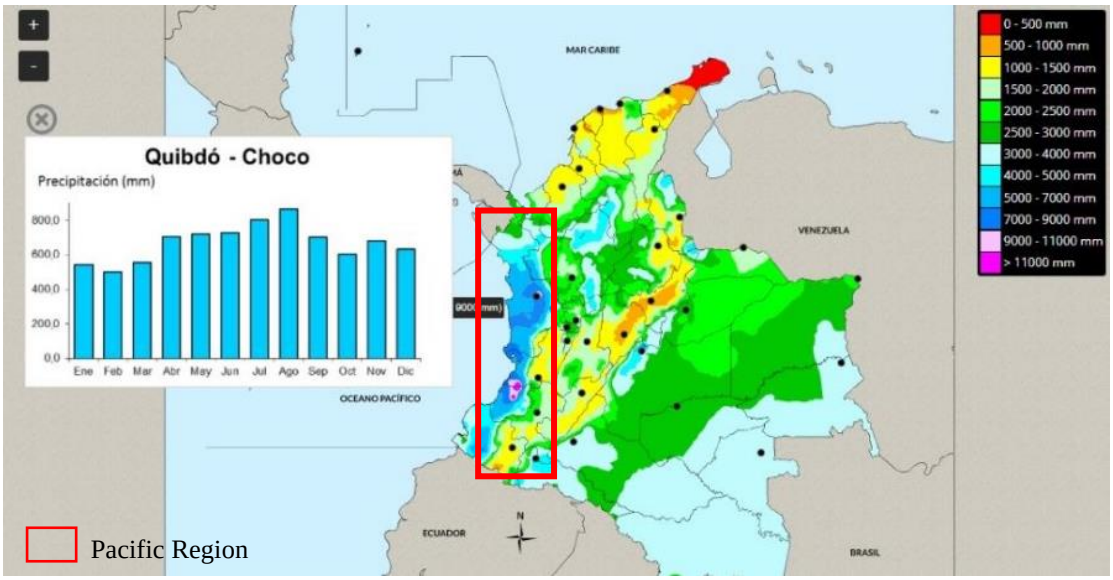

Figure 1. Colombian Annual Precipitation rate, adapted from (IDEAM, n.d.-a).

SAR images have proven their potential in tropical areas for monitoring forest and land cover changes as it has been mentioned by Ballère et al. (2021), SAR is capable to work in any weather conditions, even at nights (Bouvet et al., 2018) . Also, Bouvet et al. (2018) highlighted that benefit and declared it as one of the most promising remote sensing tools for near-real time mapping forest disturbances.

Colombia has a complete methodology and framework for monitoring the forest disturbances, but the public documents available about the technical guidelines are focus on optical sensors, they do not mention how monitor the deforestation in regions where rains almost every day like Chocó.

SMByC is also in charge of early warnings of deforestation, but the weather conditions in regions like Chocó, makes very difficult and challenging the forest cover using only optical images.

Sentinel 1 sensor from the European Space Agency ESA grants free access to SAR data with high temporal resolution (every 12 days a new image in the region where Chocó is located).

Thus, the motivation of this thesis is: to overcome the lack of information from optical sensors using Sentinel-1 information in order to contribute to the early warning alerts in a deforestation monitoring system. The solution will use tools and open software that could be potentially deployed/used in easy way.

1.3. Research Questions

The main objective of this study is to get a reliable result for monitoring deforestation using only SAR images. There are places on earth where the rain appears almost every day and the optical sensors do not overcome that situation.

To do the above implies follow the changes of the forest land cover over time, then it will be necessary to work with large datasets, timeseries of SAR data.

The SAR backscatter has a well-known behaviour for forest land cover, most of the studies for detecting changes on forest land cover uses VV and VH bands values directly like Reiche et al. (2018); just some studies like the work from Hirschmugl et

al. (2020) uses ratios between VV and VH bands showing positive contribution in their objectives.

In optical sensors is very common to use ratios in methodologies for change detection, Hirschmugl et al. (2020) use time series of NDVI for mapping forest disturbances. NDVI as other vegetation indices decreases the effect relief at mountainous places and the changes due to seasonality.

Hence, the current study has the following research questions:

- Is it possible to get a reliable result for monitoring deforestation using only radar images, in regions with very high precipitation rate?
- Does the modified version of Radar Forest Degradation Index RFDI for Sentinel-1 have better results for mapping deforestation?
- Is it feasible to use the workflow proposed in this study in a production environment?

1.4. Expected Contributions

The main expected contribution is to demonstrate if the SAR index: mRFDI can perform better results than the use of the individual bands. Taking into consideration than the index is the result from mathematical operation between two polarizations VH and VV bands and it is the combination of both bands under the index result.

Complementary contributions are to show that SAR images can be used as the unique input for monitoring deforestation, and that monitoring process can be achieved with time-series datasets as inputs.

1.5. Document Organization

This document is organised into five chapters. The first chapter, I introduce the context, motivation, objectives, research questions and expected contributions.

The second chapter consist of the literature review, which briefly explains the basic concepts of SAR, afterwards you will find a summary about how scientists and researchers have used SAR data for monitoring and mapping deforestation.

In the chapter three I describe the data used, a small description of the study area and the software used to solve the research questions. Also, I further present the methodology workflow about the steps for processing SAR data until to get the updated Forest/Non-Forest layer.

The chapter four is the result analysis, it describes based on comparisons and quality assessment the results, chapter five summarizes what I found, the contributions, limitations and future work that could be done using this study.

2 LITERATURE REVIEW

2.1. SAR Concepts

SAR sensors are active sensors, it means they have their own source of energy, they emit that energy towards the earth surface, the energy is reflected but the different objects in the surface and the sensor collect the reflected energy. In the specific case of the SAR, the sensor antenna transmits a microwave pulses, those pulses illuminate specific area on the ground while the sensor is moving following its orbit in the space. The pulses have different reflection type on the ground depending on the object type (see figure 2).

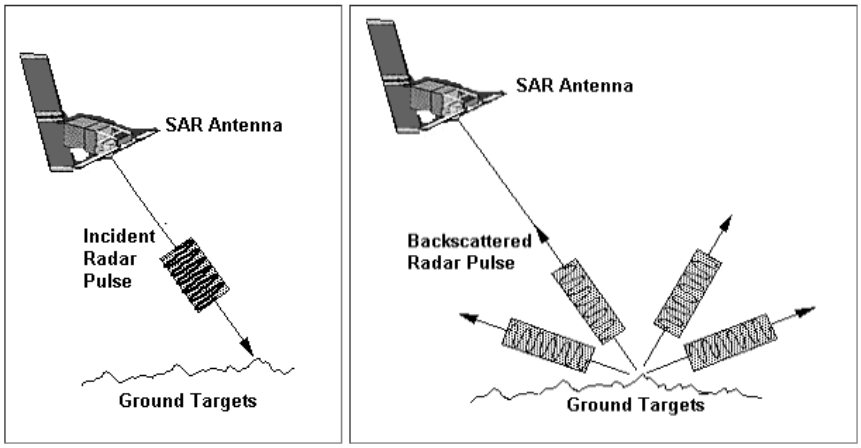

Figure 2. SAR principles; Left: radar pulse transmitted to the ground. Right: Radar pulse scattered by the ground objects (Centre for Remote imaging, 2001).

SAR images are result of different parameters which affect them, the SAR system parameters are:

Wavelength or frequency

In the SAR technology there are ranges in the electromagnetic spectrum where it operates. That spectral range is subdivided in bands, the studies developed in those bands have created trends in the use of a particular band, the table 1 shows the bands and their usual applications.

Band	Wavelength	Typical Application
Ka	0.8 – 1.1 cm	Rarely used for SAR (airport surveillance)
K	1.1 – 1.7 cm	Rarely used (H_2o absorption)
Ku	1.7 – 2.4 cm	Rarely used for SAR (Satellite altimetry)
X	2.4 – 3.8 cm	High-resolution SAR (urban monitoring, ice and snow, little penetration into vegetation cover, fast coherence decay in vegetated areas)
C	3.8 – 7.5 cm	SAR workhorse (global mapping, change detection, monitoring of areas with low to moderate vegetation, improved penetration, higher coherence, ice, ocean, maritime navigation)
S	7.5 – 15 cm	Little but increasing use for SAR-based Earth observation, agriculture monitoring (NISAR will carry an S-band channel, expands C-band applications to higher vegetation density)
L	15 – 30 cm	Medium resolution SAR (Geophysical monitoring: biomass and vegetation mapping; high penetration, InSAR)
P	30 – 100 cm	Biomass. First P-Band spaceborne SAR will be launched – 2020, vegetation mapping and assessment. Experimental SAR

Table 1. Designation of microwave bands and typical applications (Flores et al., 2019).

Polarization

Due to SAR is an active sensor, one of its features is to control the signal polarization, it can be done when the signal is transmitted to the ground and when the signal is received by the antenna. The polarization is the oscillation plan of the signal, the usual oscillation planes are vertical "V" and horizontal "H" (Flores et al., 2019).

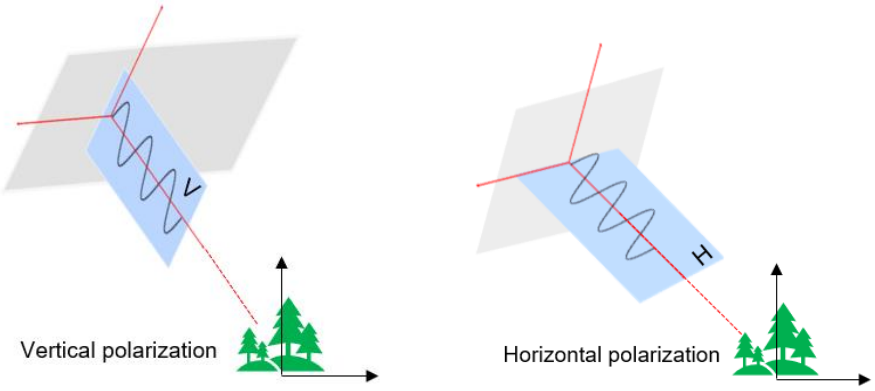

Figure 3. SAR Polarizations; Left: radar pulse in Horizontal polarization. Right: radar pulse in vertical polarization.

The common convention for SAR uses two letters for indicating the planes of polarization, the first letter stands is for the transmitted signal and the second letter for the received signal.

VV, HH indicates co-polarization.

HV, VH indicates cross-polarization.

Noise

SAR data has a well-known source of noise commonly called speckle, the speckle is "… a result of interference from the many scattering echoes within a resolution cell"(Flores et al., 2019). In pragmatic terms, the speckle gives to the SAR images and mottled effect when you visualized them. But also, it is a feature that can affect any processing workflow.

There are several filters which perform operations at pixel level, trying to keep the meaning pixels (details) while reduce the speckle noise, the most used are the Lee filter and modifications of it.

There are other filters such as the Quegan Yu filter, which has been used in time series approaches, it preserves the changes over the time while reduces the speckle (Flores et al., 2019).

Looks

SAR data acquisition has a very different principle than the optical sensors, the radar antenna is always side-looking, in optical sensors usually the sensor points to nadir. That feature implies that SAR data has a special treatment for geometry corrections and also the way in which each SAR scene is created.

The looks are the number of times a target was seen from the antenna, in every of those looks the target has a different "perspective". The multilooking process merges those looks and provides a image with better geometric characteristics.

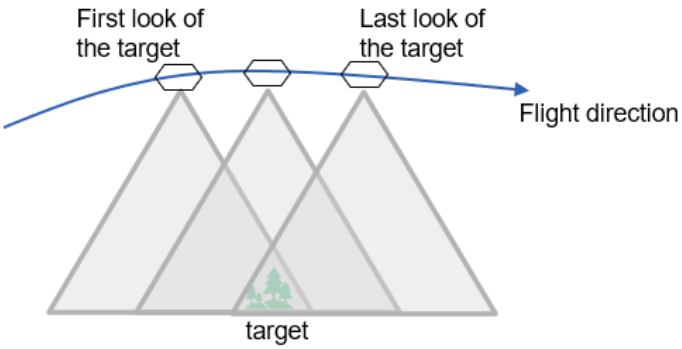

Figure 4. Multilooking process.

Bandwidth

The bandwidth is related to the spatial resolution of SAR data, it depends on two aspects, the Azimuth resolution and the range resolution. The azimuth is the direction parallel to the sensor flight path and the range is the direction perpendicular to the sensor flight.

The azimuth resolution depends on the size antenna, in the SAR system, it means the time that the target is observed. The range resolution varies across the range direction, and it depends on the indent angle.

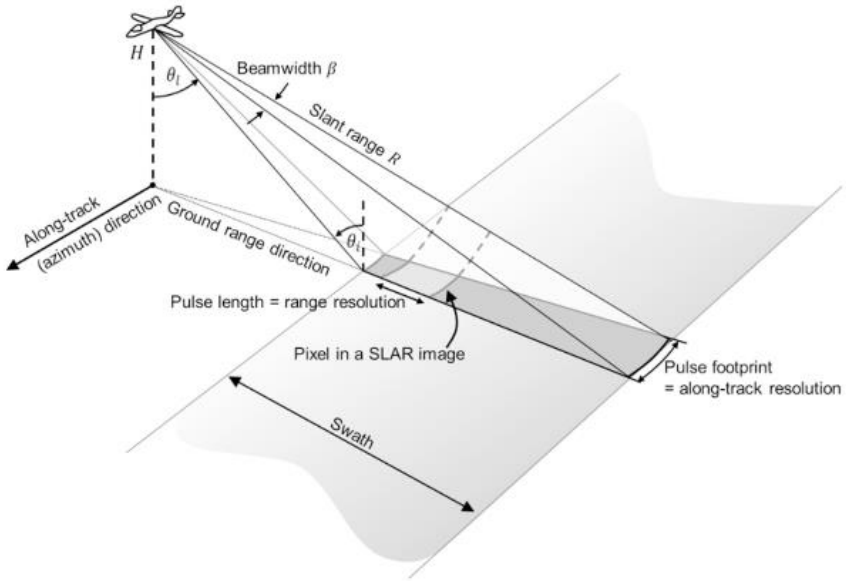

Figure 5. Bandwidth, spatial resolution of SAR data (Flores et al., 2019).

Incidence Angle

The incidence angle can affect the radar signal and the backscatter from the targets, steeper angles can increase the backscatter from the same targets with larger incidence angles.

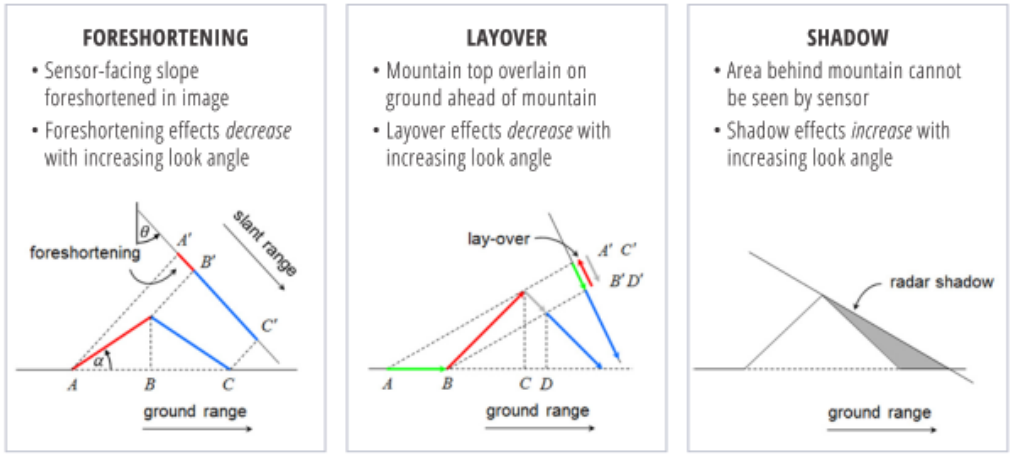

Figure 6. Most common geometric distortions on SAR images (Flores et al., 2019).

However, the aspect to highlight are the distortions on the geometry due to the side looking observation. The figure 6 summarizes the most relevant.

Basically, in the foreshortening distortions the distances between the bottom and the top of structures facing the sensor are compressed. The layover makes that the top mountains appear ahead of the bottom mountains, it inverts the mountain relief.

Usually, the distortions correction is part of the pre-processing steps for SAR data.

Orbit (ascending or descending)

The Orbit indicates the side looking direction of the sensor, most of the radar sensors, and in specific Sentinel-1 has a polar orbit, it means the sensor crosses the equator in each orbit around the earth. If the sensor is crossing from north pole towards south pole, that is the descending orbit. At the vice versa mode from south to north pole that is the ascending orbit.

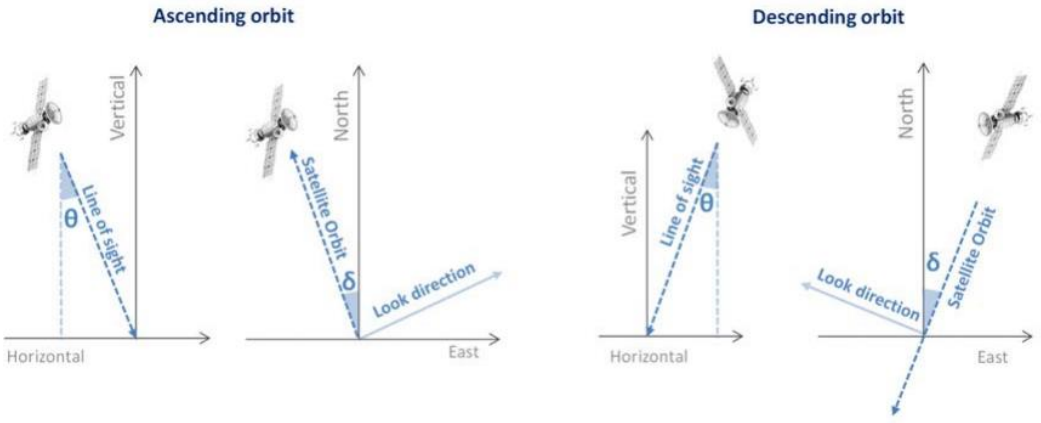

Figure 7. Line of sight acquisition in descending and ascending orbit, adapted from (Mora et al., 2016).

There are other parameters of the SAR system, Antenna size, pulse repetition frequency and imaging mode which were not explained here as they are not relevant for this thesis.

2.2. SAR for monitoring deforestation

Two group of characteristics affect the backscatter values, the first one is the sensor features, the second the target characteristics. In the first group, the frequency (radar band), incidence angle and polarization are the most important. From the second group the structural features and moisture content of the target objects influence on the backscatter values (Flores et al., 2019).

The figure 8 shows typical kind of scattering signal from forest is called *volume*, the signal is reflected several times on the branches and leaves of the vegetation. Although, depending on the radar band the radar signal can penetrate even until the soil.

Thanks to the groups of features explained about the SAR images, different approaches has been developed for deforestation mapping and tracking disturbances on that land cover.

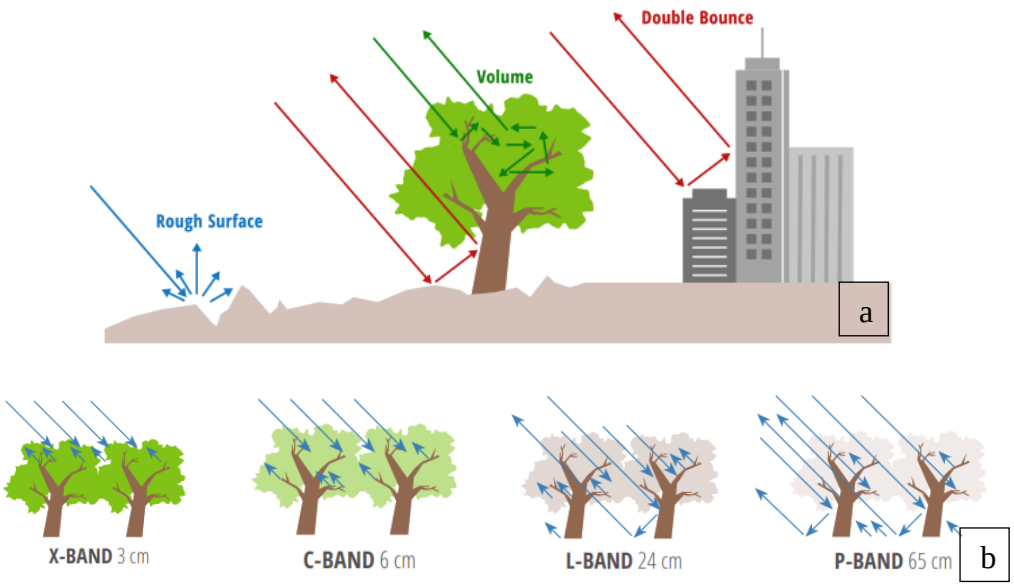

Figure 8. SAR scattering and penetration; a) Typical types of scattering, b) Capabilities of penetration of SAR at different wavelengths (Flores et al., 2019).

Sentinel-1 looks not to have the characteristics for monitoring deforestation as the radar signal for band C cannot penetrate the canopy, but some studies of Bouvet et al (2018) and Ballère et al. (2021) performed their workflows using Sentinel-1 data.

Bouvet et al. (2018) developed a new approach to detect deforestation, they looked for the shadows at the borders between the deforested area and the forest, with that approach they were capable to detect patches in the forests and establish a near-real time alert system for deforestation.

In the case of Ballère et al. (2021) they created a system for monitoring deforestation was built for the French Guiana in a country scale using sentinel-1 images, they assessed the results against ground data and the Global Land Analysis and Discovery (GLAD) Forest Alerts produced by the University of Maryland; the system was able to detect deforestation types with a overall accuracy of 97.4%, the result was 4% higher than the deforestation detected based on the optical dataset. In the same path Doblas et al. (2022) developed a full operational Near-Real time forest disturbance system in Brazil.

Other authors like Hirschmugl et al. (2020) combined optical and SAR data for detecting forest changes over the time in Gabon and Perú. Basically, they used as methodology two separate workflows, one using optical data and another using SAR data, at the end they combine the results for the final product from each workflow. So, at the sites where they have optical and SAR data they had a double check in the forest changes, but for the places where they only have SAR data, they validated the results based on the double check from the previous ones.

Detecting deforestation can be obtained from different approaches, for instance Mermoz & le Toan (2016) used fuzzy classification for detecting deforestation and regrowth of the forest using Alos Palsar images. Gašparović & Dobrinić (2020) developed a comparison using six machine learning techniques for classification, they wanted to detect the vegetation in the urban environments using SAR images. Something to highlight, they got better results with an overall accuracy between 70% and 93% when they used multitemporal stack layers instead of single scenes or just one pair of scenes.

So far from literature detecting deforestation in SAR images in a context of monitoring system, the best way to do it is to use time series data, which implicates large datasets.

Nowadays scientist and researchers have access to large datasets and cloud computing resources when the want to use time series, Canty et al. (2020) they present the relevance of the sentinel-1 dataset in the Google Earth Engine (GEE) platform, they tested and algorithm for multitemporal change detection in different places on the planet Earth. Platforms such as GEE provides options for automatizing processes and workflows using two coding languages JavaScript or Python.

Into the timeseries there are still several options to process them, but Flores et al.(2019), Ruiz-Ramos et al. (2020), Ygorra et al. (2021) and Aquino et al. (2022) highlight the use of Cumulative Sums Cusum as method for change detection with SAR data, the method has demonstrated good results for detecting changes in the backscatter values related to the forest disturbances and deforestation.

Another changed detection method in SAR data time series with promising results is the Log-ratio, Ajadi et al. (2016) used this method as first part of their workflow, then they complete their proposal with probability calculations for mapping deforestation. The method is also mentioned by Flores et al. (2019) as one way to have fast mapping changes in the forest using SAR data.

The methods above share a subjective issue, as the SAR data has an inherent ambiguity, false positives are present frequently in the change detection methods, then it is very common to stablish threshold values in order to identify the range of values where the actual deforestation is happening. The work from Motohka et al. (2014) and Reiche et al. (2018) are studies where subjective threshold values in the timeseries is adopted for detecting changes.

Finally, an interesting approach for change detection in SAR data are the band indices, the Dual Polarization SAR Vegetation Index DPSVI is an index for monitoring degradation in the vegetation cover, dos Santos et al. (2021) modified the DPSVI and got better discrimination of the vegetation than the original index.

Other radar indices such as Radar vegetation index RVI and radar forest degradation index RFDI were used in a methodology for land cover mapping, the methodology

also includes the use of optical data and indices derived from the optical images as inputs for running a random forest classification de Luca, N Silva, et al. (2022) obtained an overall accuracy of 90% and they claimed an improvement in 2.53% with the SAR data.

The SAR vegetation indices can be used in a multitemporal context, the RFDI with Alos Palsar images was the input of an algorithm that allows to detect repeated disturbances on forest developed by Joshi et al. (2015).

After present a summary of some of the available research literature about detecting deforestation using SAR data, the next chapter will elaborate the details of the methodology adopted for this thesis.

3 DATA AND METHODOLOGY

3.1. Study Area

The study area is located in pacific region in Colombia, the coordinates that illustrate the rectangle used as Area of Interest AOI are the following: upper left corner [05°37’38.35”N, 76°53’6.54”W] and bottom right corner [04°57’40.55”N, 76°33’49.40”W].

The study area is inside the Choco department (administrative division), the region has a complex climatology, the dense vegetation, the allocation of mountain ranges and valleys and the air movements are part of the factor that influence that in some places of the region the annual precipitation reaches more than 10.000 mm (Díaz Merlano & Gast, 2009).

Historically the rainy periods go from April to May and October to December, but most of the region has precipitations during the whole year (Díaz Merlano & Gast, 2009).

The relief in the study area is mainly alluvial plains formed by the riverbeds, the attitude over it is between 60 to 150 m above the sea level.

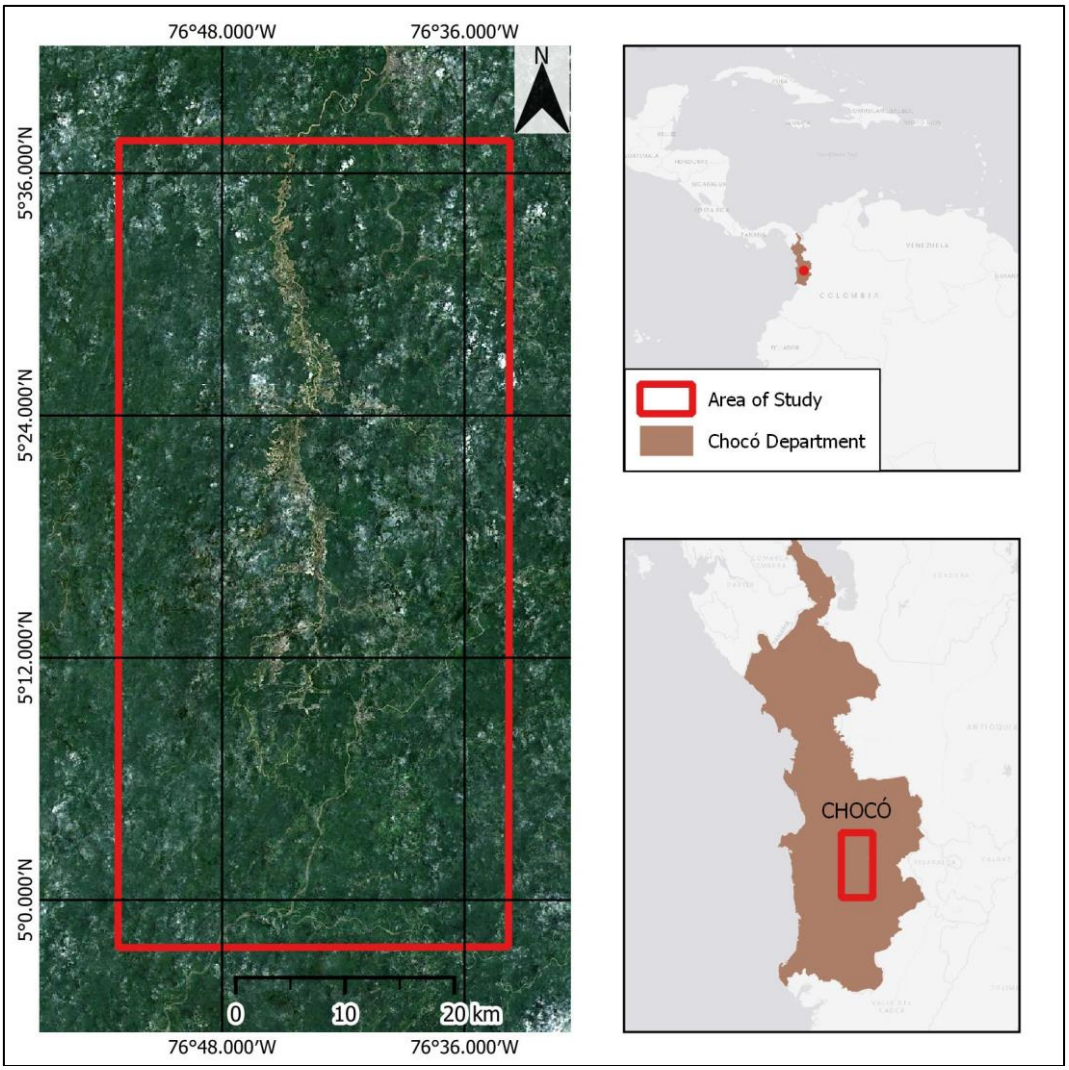

Figure 9. Study Area Map

The land cover of the region is mainly natural forest, under the ecosystem classification is called humid basal forest, in the Chocó department more than 80% of its extension belongs to that ecosystem (Instituto de Hidrología et al., 2017).

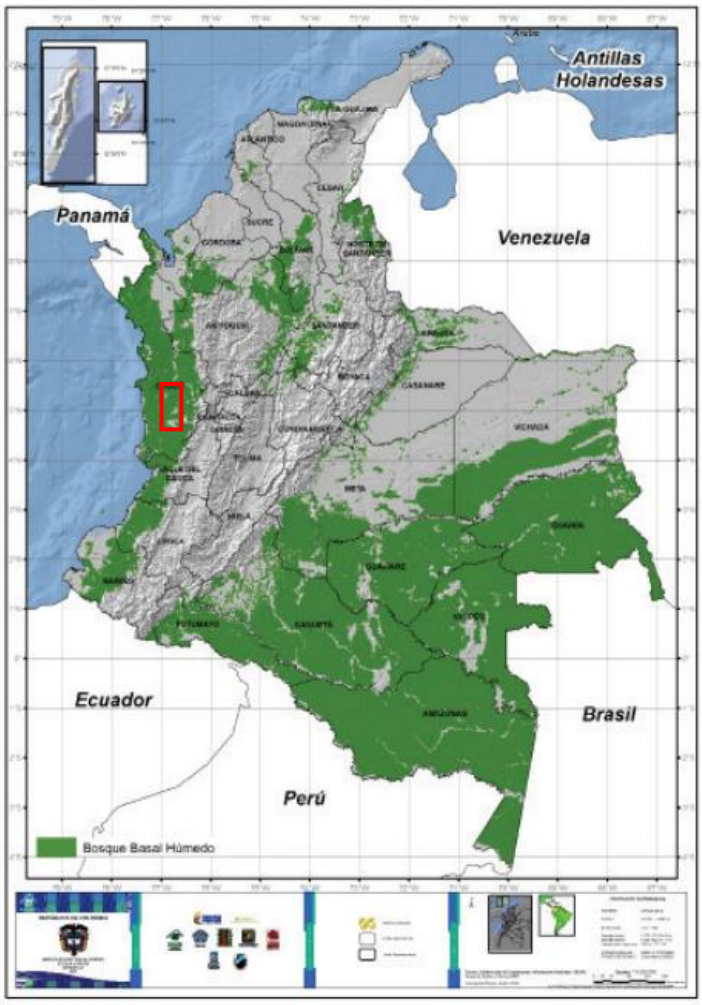

Figure 10. Humid basal forest ecosystem(Instituto de Hidrología et al., 2017).

Finally, it is important to highlight the study area (the red rectangle in the figure 10) frames the Quito river basin which has been affected illegal gold mining through the time.

3.2. Sentinel-1 data

Data used in this project were obtained from the Copernicus Program Sentinel through the Google Earth Engine python API.

The Sentinel-1 mission provides SAR images from dual polarization in the C band of the radar microwaves, GEE python API provides access to the collection of images,

the collection of images cover dates from 2014 up to today and it is possible to do any workflow without download SAR input images to your own computer.

Sentinel-1 (S1) mission has different level of processing, for this study we selected the S1 Ground Range Detected GRD, SAR images usually has 3 main components the Intensity, amplitude and the phase, the S1 GRD collection has only the data related to the intensity and amplitude of the microwave signal.

According to the GEE webpage, the S1 GRD collection in its platform contains all the GRD scenes processed using the Sentinel-1 toolbox and the final product is a ortho-corrected scene (Google Earth Engine, n.d.).

The S1 collection includes 4 possible band combinations which correspond to the band polarization [VV, HH, VV +VH, HH+HV], the available band combinations depend on the region over the planet Earth, some places have access to all band combinations, others only have two of them.

For the study area in this thesis the combination bands available are [VV, VV + VH], it means that the signal is transmitted in vertical polarization, and it is received in vertical and horizontal polarization.

Another advantage of using the GEE S1 GRD collection is that the scenes have already some pre-processing, each scene already has thermal noise removal, radiometric calibration and terrain correction using the SRTM 30 DEM, the final corrected values in the bands are in decibels after conversion from linear to logarithmic scale.

The band values obtained from the GEE python API are called σ^0, sigma naught or sigma zero, it means the backscatter values in decibels with radiometric terrain correction.

However, according to Flores et al. (2019) for forest disturbance and biomass applications, the backscatter in γ^0 **(**gamma zero**)** is recommended, thus other pre-processing steps are necessary over the data.

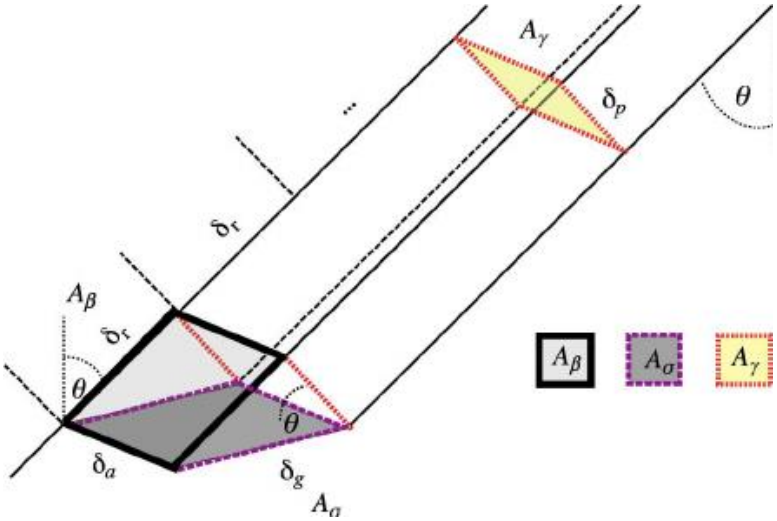

Figure 11. The three different backscatter SAR values, betta, sigma, gamma (Atwood et al., 2012).

The figure 11 illustrates the differences on the plane regarding to the backscatter definition, in the case of gamma zero (γ^0), it takes the plane perpendicular to the local look direction and the literature refers this plane as the most suitable for volume scatterers, which is the case of the forest.

The original pixel size from S1 GRD collection in GEE is 10 meters, but for this study, during the pre-processing steps, before to retrieve the images the pixel size is resampled to 30 meters to speed-up the retrieving operation and fit the scale selected for this study, only deforestation events larger than 1 ha will be considered.

3.3. Validation data

As validation data there are two datasets, the first one is provided by the Norway's International Climate and Forest Initiative Satellite Data Program NICFI, which allow us to access the monthly basemaps mosaics from planet images for free after a request for academic purposes.

The planet basemaps are derived product from Planet doves satellites, they are build using the best scene during the month (less clouds, less haze, etc). the spatial resolution is around 3 meters (Planet Labs PBC, n.d.).

According to the documentation planet constellation is capable of revisiting the same place over the Earth each day, then they have many opportunities to obtain images without clouds.

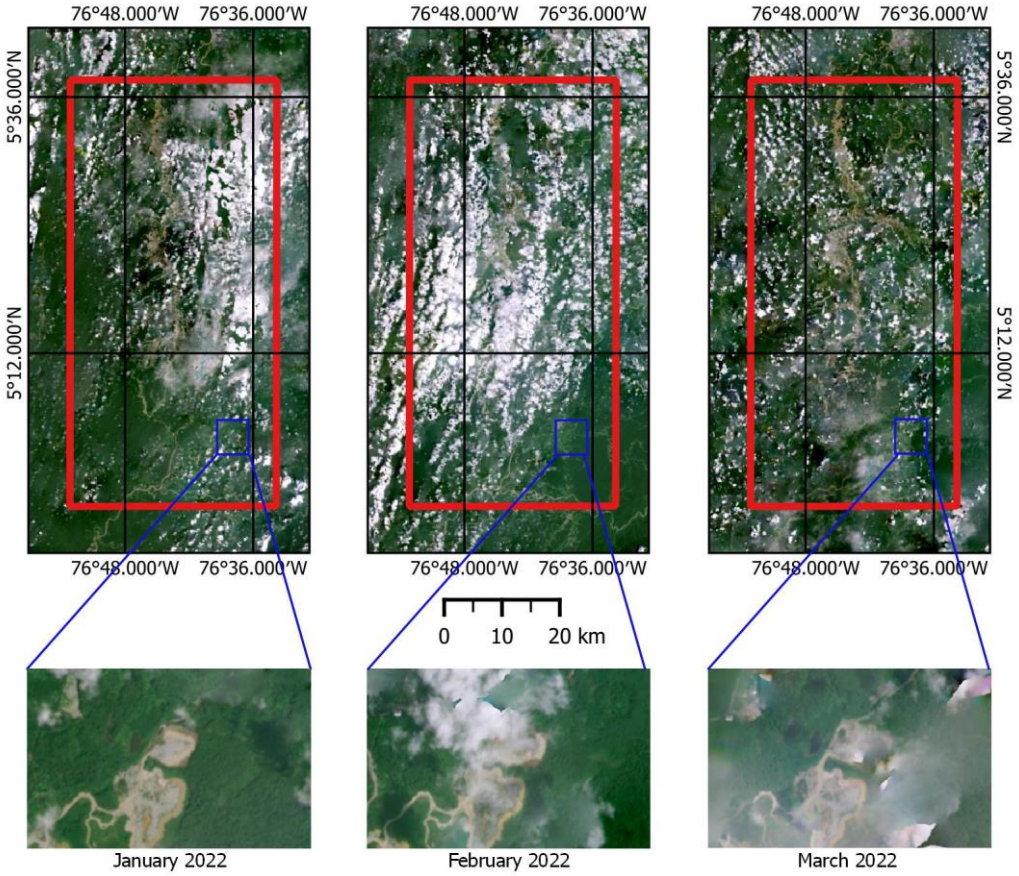

Figure 12. Example of Planet basemaps over the study area mosaic for January, February and March 2022.

The figure 12 illustrates some examples of the monthly mosaics, despite the daily revisit capability of the constellation satellites, the study area still have clouds and haze.

The second one is provided by Global Forest Watch GFW, an open online platform with data and tools for monitoring deforestation. From the platform was selected the dataset called "Integrated deforestation alerts", the dataset gathers information from three alert systems GLAD-L, GLAD-S2 and RADD. GLAD-L uses Landsat data, GLAD-S2 Sentinel-2 data and RADD Sentinel-1 data. Those alerts are collected for worldwide in the tropical regions.

3.4. Software Description

The development of this thesis was using open-source software and open online platforms. The use of those platforms diminishes the processing times and reduce the resources needed in the personal computer.

As it is going to be detailed in the next point, the methodology relays on time series approach. Those kind of analysis demands computational resources which are difficult to have in a laptop or personal computer.

Thus, it was decided to use SEPAL, which is a cloud-based platform that can be used for free prior to a registration process and acceptance of your purpose. SEPAL has several already geospatial tools focus on land monitoring sponsored by Food and Agriculture organization of the United Nations FAO, Openforis and NIFCI, also has the support of several institutions such as ESA, NASA, World Bank Group and others.

Inside the platform you find a combination of tools such as GEE, orfeo toolbox, python, jupyter notebooks, GDAL, R, Rstudio, SNAP toolkit and OpenForis Geospatial toolkit. Also, there are access to virtual machines with up to 64GB of RAM memory (SEPAL development team, n.d.).

For post-processing purposes and mapping was used QGIS version 3.16.

3.5. Detection of deforestation

The detection of deforestation is the main objective of this thesis, the methodology which support it includes three main stages, the figure 13 depicts them. The pre-processing which aims to retrieve and perform some corrections on the SAR images.

Right after the second stage begins with the analysis of the time-series datasets using two approaches for change detection, the first one uses the cumulative sums algorithm (Cusum), the second approach performs the Log-ratio.

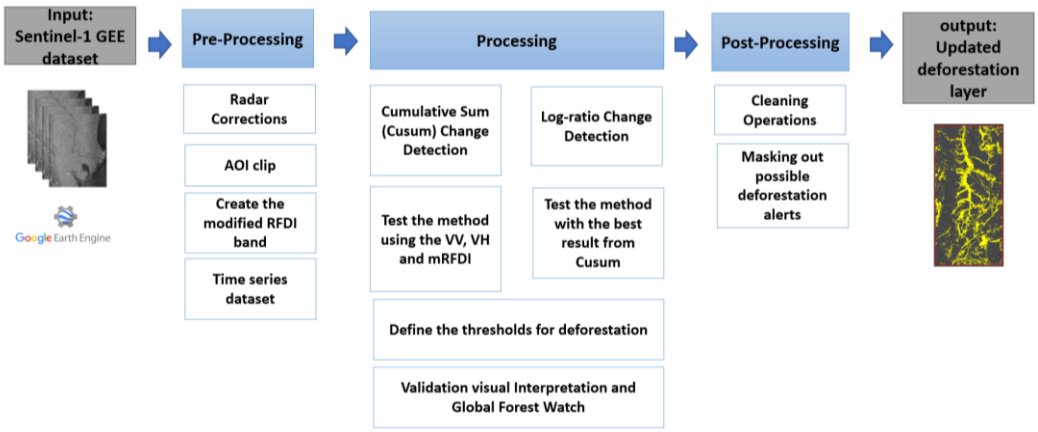

Figure 13. General methodology.

The post-processing stage includes cleaning operations in order to remove isolated pixels and fulfil blank spaces between pixel regions. At the end, the result is a dataset which represents the deforestation alerts in the monitoring period.

The processes in each stage are illustrated in the figure 14 and the following paragraphs explain in detail each of them.

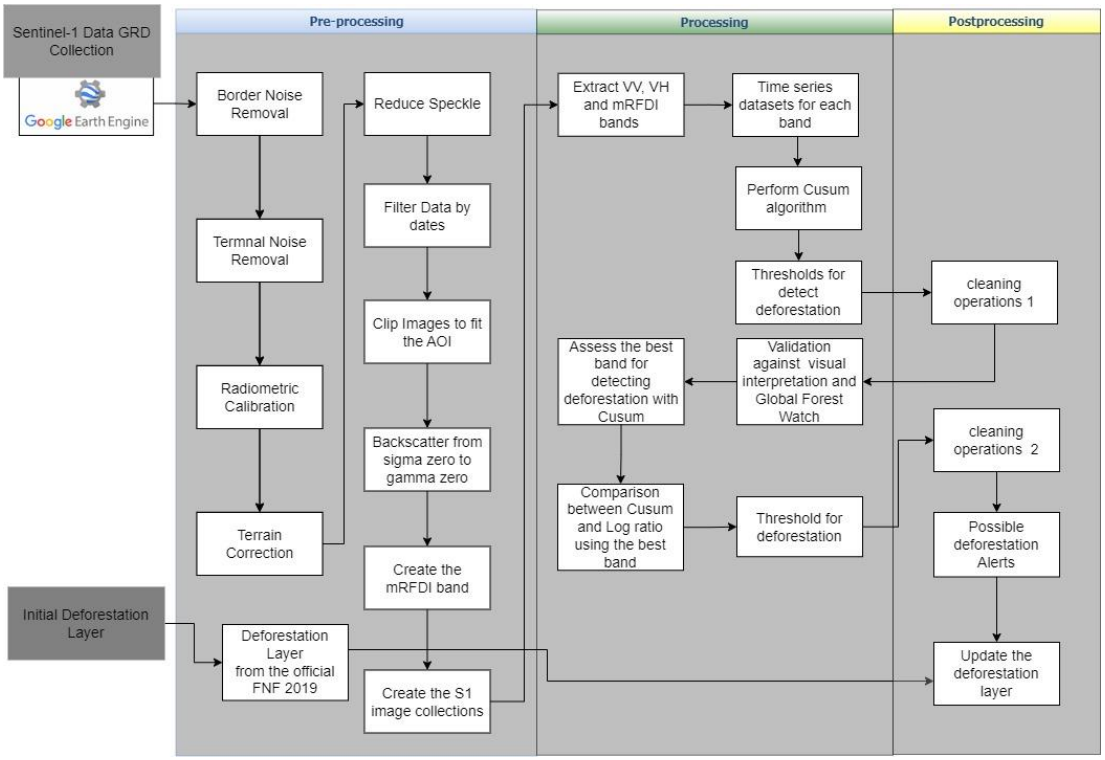

Figure 14. Methodology Workflow in detail

Pre-processing:

S1 GRD collection from GEE has already some corrections over the images: border noise removal, thermal noise removal, radiometric calibration, and terrain correction. Then for the other steps in the pre-processing stage it was performed using some available scripts in python. Due to the orbit direction has influence in the backscatter values only scenes in descending orbit were used.

The speckle reduction step aims to decrease the variation of the backscatter values, the speckle effect is an inherent characteristic from the SAR images. The Refined Lee filter which is a filter based on directional kernels that reduces the noise and preserve the spatial resolution and texture, the algorithm used in this study comes from Mullissa et al. (2021) and its github account.

The following step delimitates the time period for the images, for this study the dates go from 2021/01/01 up to 2022/09/01. That period includes 116 scenes in 71 dates. Although the official documentation from S1 informs a revisit period of 12 days for Colombia, some of the retrieved images have 6 days between them. That fact does not influence the results of this work, the figure 15 illustrates the dates from the S1 scenes retrieved for this work in a timeline chart, each point represents a scene.

The size scene for S1 images is bigger than the AOI then a clipped operation is performed using the boundary box showed in the

figure **9**, the boundary box follows the shape of the watershed from the “Rio Quito” river.

The next step is to convert the backscatter values from sigma zero to gamma zero, for that purpose it was applied the same approach that Doblas et al. in (2022), a small part of their code modified for this step.

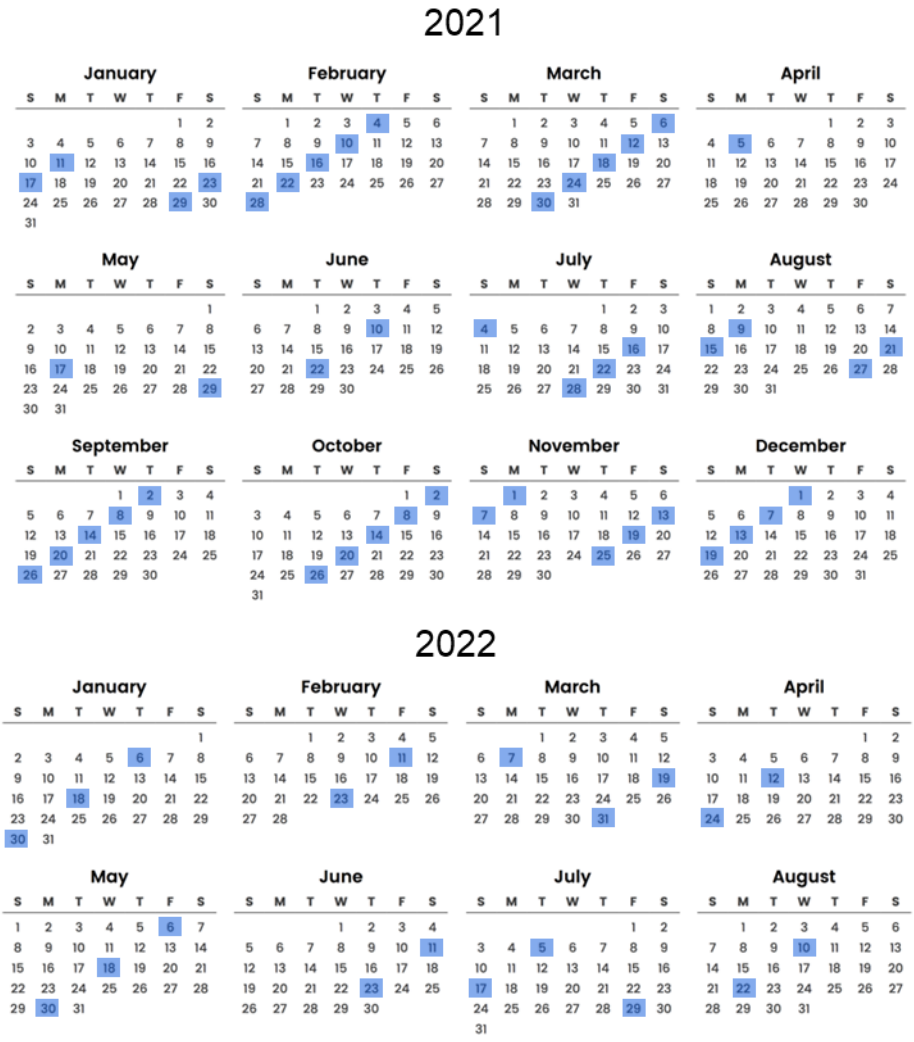

Figure 15. Dates of the S1 scenes for this study.

The following step is the most important in the pre-processing, the creation of the modified radar forest degradation index mRFDI, the latter is an index focus on vegetation degradation which uses the VV and VH bands from S1, it was proposed by Nicolau et al. in (2021) and other authors such as de Luca, M. N. Silva, et al. (2022) have used it in their works. See equation Eq. (1).

$$mRDFI = \frac{\gamma^0 vv - \gamma^0 vh}{\gamma^0 vv + \gamma^0 vh} \qquad \text{Eq. (1)}$$

The final step is to create image collection for each band, VV, VH and the mRFDI.

Processing:

The processing stage is focus on the change detection methods for time series analysis, two concepts are behind the two methods selected for this thesis. The first concept is:

Cumulative Sums, Cusum is an algorithm widely used in financial sector, due to its beneficial results as a tool for log time series analysis (Flores et al., 2019; Ruiz-Ramos et al., 2020).

In the geospatial field, the Cusum algorithm has become an option for time series analysis as the latter has become more popular thanks the availability of large datasets (bigdata) such as data cubes and others big datasets.

The Cusum performs the cumulative sums of values from a variable that can be measured it over the time. The literature has proved that SAR backscatter values decrease if a logging or deforestation event happens on a previous forest land cover. So, it is expected to have a stable trend in the backscatter values if the forest remains along the time, but if a forest disturbance happens in some point of the time, the backscatter values will decrease, and a graphic trend of the backscatter will show that event.

Applying the above to the time-series of SAR images, the equation Eq. (2) illustrates the Cusum algorithm,

$$s = \sum_{i=1}^{n} \gamma_{i,j}^{0} - \overline{\gamma_j} \qquad \text{Eq. (2)}$$

Where *S* is the cumulative Sum, γ_i^0 is the backscatter value take it on the image *i* on pixel *j*, $\bar{\gamma}$ is the backscatter mean over pixel *j* along the whole time series, *n* is the total number of images of the time series.

Basically, the Cusum is applied to the differences between each value at each pixel in the time series and the mean ("expected value") at each pixel along the whole time series.

The second concept is the suppression of background using ratio from images from different dates:

Log-ratio algorithm for change detection is based on the common knowledge about that the ratios between images from different dates can remove the common background between the two images and keep the features where the changes happen.

In order to identify possible changes between SAR images, the ratio is performed between a referenced image and a new image (Flores et al., 2019; Ruiz-Ramos et al., 2020). Equation Eq. (3) illustrates the mathematical approach:

$$r = \frac{x_i}{x_r} = xR \qquad \text{Eq. (3)}$$

Where r is the observed intensity x_i the new images, x_r is the referenced image, x is the multiplicative noise contribution and R the true intensity ratio. (Flores et al., 2019; Ruiz-Ramos et al., 2020).

Due to the grey values distribution of the SAR images are non-normal, performing a logarithmic scaling operation, contributes to ease the analyse in a subsequent thresholding process to find the changes to highlight. Then the Equation 3 can be written as in the Equation Eq. (4).

$$r = \log_{10}\left(\frac{x_i}{x_r}\right) \qquad \text{Eq. (4)}$$

Once the change detection methods are performed, the next step is to select the threshold where the changes appear, based on the statistical analysis of the Cusum curve and the histogram from the change image in the case of Log-ratio.

The last step is the validation process, it follows three perspectives:

a) Comparison between the results and the integrated alerts for deforestation dataset from GFW using a random sampling points. 300 hundred points were created using equalized stratified sampling strategy, it means half of the points target the deforestation pixels and the other half pixels labelled as not deforestation, the reference map used as ground truth was the dataset from GFW.

b) Comparison between the results and the integrated alerts for deforestation dataset from GFW using an overlay operation. An intersection between

polygons was performed. If the intersection area was greater than 0.5 ha then the polygons is validated as “true positive” for deforestation alert. Polygons from the reference map or ground truth comes from the GFW dataset, polygons from the results which do not intersect the refence layer are “false positives”, polygons from GFW dataset which do not intersect are “false negatives”. True negatives are not considered, it is not the objective to identify areas which remains in forest land cover.

c) Comparison between the results and visual interpretations of S1 images and optical images Planet monthly mosaics. 30% of the polygons obtained as deforestation events are selected randomly, a visual interpretation using as ground truth the planet monthly mosaics and the SAR visual compositions were interpreted in order to calculate the confusion matrix.

Post-Processing:

This stage consists of a process of cleaning operations in order to remove isolated pixels and update the layer of deforestation alerts.

Once the possible deforestation alerts from the Cusum and Log Ratio change detection methods is retrieved, it is necessary to apply the following spatial tools.

a) Raster overlay: Removes the deforestation events that intersect the initial deforestation layer. The “initial deforestation layer” masks out water bodies, urban areas, and the official deforestation quantification for the year 2019.

b) Reclassify: Pixels values are converted to integer, masking out all the possible deforestation alert pixels with the same pixel value.

c) Sieve Filter: Removes raster polygons smaller than a provided threshold size (in pixels) and replaces them with the pixel value of the largest neighbour polygon (QGIS project, n.d.). The size selected was 1 ha, we only select as deforestation events pixel regions or groups with are greater than 1 hectare.

d) Binary Morphological Operation Filter: Performs a morphological operation on a monoband image, in this case, it was a closing operation with 1 pixel as

radius, which contributes to fill blank, or no data pixels surrounded by pixels labelled as deforestation.

The final step is to update the initial deforestation alerts layer by adding the new pixels detected as deforestation events.

4 ANALYSIS AND RESULTS

4.1. Deforestation alerts with Cusum change detection.

The period where the changes will be monitor goes from 01/01/2022 up to 01/09/2022. However, in order to check the changes in the year 2022, it is necessary to have an observation period, the latter goes from 01/01/2021 up to 31/12/2021, the time series datasets included both.

Time series can reduce the moisture effect on the forest cover and also reduce the impact on the backscatter (Flores et al., 2019). As Cusum performs over time series to look for the point that global trend changes, the bigger the time series is, the results will be better. Then, the observation period contributes as container of historical data or reference.

The change detection with Cusum was done in python and similar approach that Aquino et al. in (2022). Once the collections are separated by bands VV, VH and mRFDI, each of them is converted to timeseries data using a multidimensional array format.

The mean value of each pixel is measured along the whole period of time. Afterwards, the residuals in each pixel are calculated for each time, in this case 71 dates. The figure 16 illustrates the backscatter for the VV and its mean for a specific place.

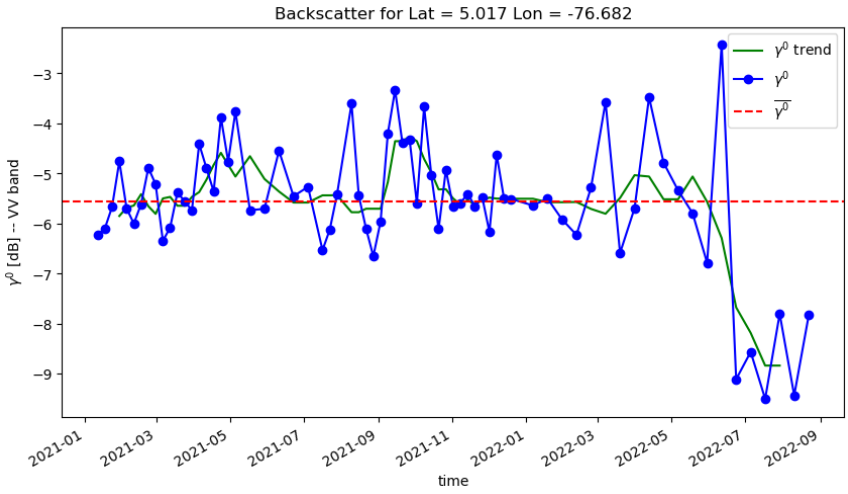

Figure 16. Backscatter VV band, backscatter mean and trend for specific location with deforestation activity during the time series.

The residuals are the values from the subtraction between the mean and each backscatter value for each image or time.

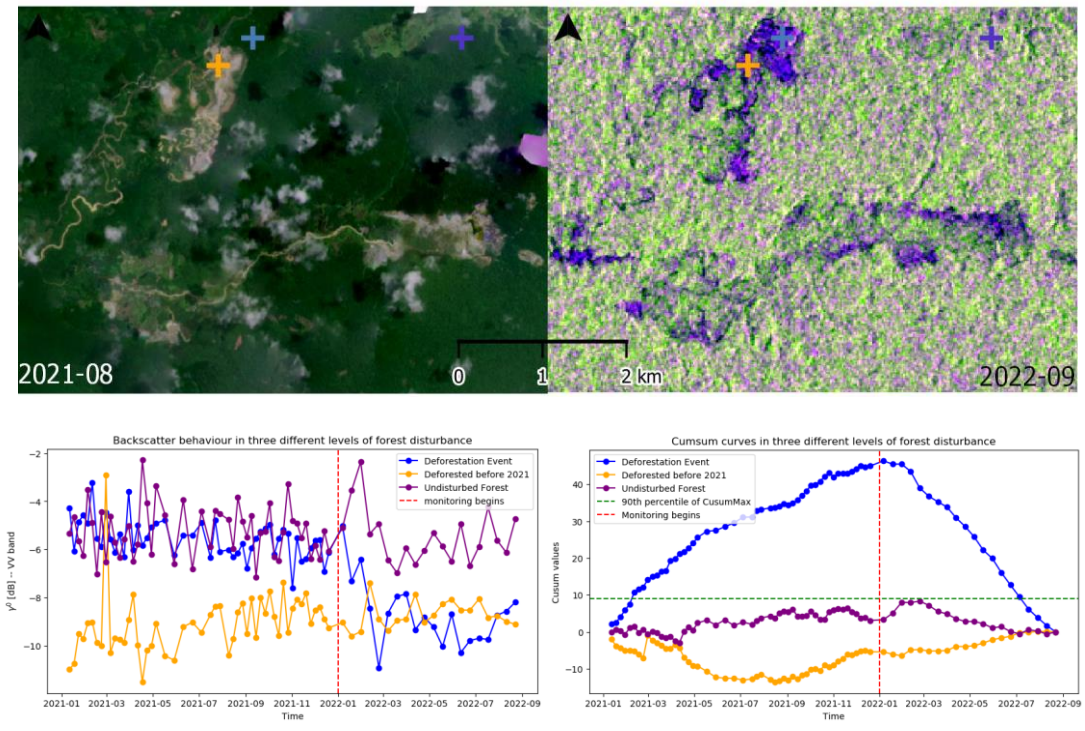

Figure 17. VV Backscatter and Cusum curve for three levels of forest disturbance. Upper left: Planet basemap monthly mosaic august 2021. Upper right: RGB composite (VV, VH, ratio VV/VH) S1 September 2022. Lower left: backscatter curves on the VV band for each place along the study period. Lower right: Cusum curve for each place.

After applying the Equation Eq. (2), it is obtained a value for each pixel which represents the cumulative sums along the time-series. When a plot of the Cusum curve is performed, the behaviour of the curve from pixels which have decrease the backscatter due to deforestation is a mountain-shape where the maximum value of the Cusum is an inflection point. (See figure 17 lower right), that maximum value occurs close to the date of the deforestation event, then the algorithm also provides information about "when" the deforestation happened.

The figure 17 also shows the optical and SAR images for a subset from the AOI, the purple colour is for a place with undisturbed forest during whole time study period, the orange colour for a deforested place before this study which also remains as deforested place and the blue colour represents a place with deforestation event during the monitoring period (year 2022).

Similar plots are obtained with the VH band (see the figure 18), the difference with the VV band is that the backscatter values ranges between – 10 and -22 in the VH band and between -6 to -10in the VV band, the places are the same depicted in the figure 17, the colours represent the same level of forest disturbance than the previous figure.

To check every Cusum curve for each pixel allows to identify the value that can be used as limit for masking out the deforestation events, the value should be near to the peak of the inflection point.

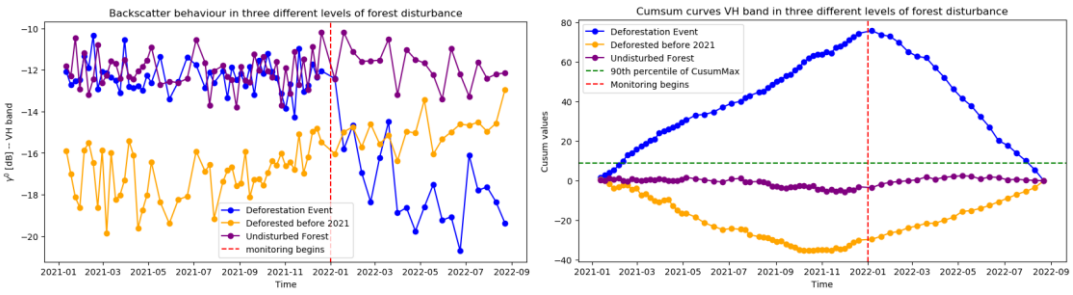

Figure 18. VH Backscatter and Cusum curve for three levels of forest disturbance. Left: backscatter curves on the VV band for each place along the study period. Right: Cusum curve for each place.

Based on the statistical analysis from the Cusum curves and try-and-error process, it was selected the percentile 90th of the maximum value from the Cusum curves as

limit or threshold for masking out the pixels which shows possible deforestation events.

Therefore, any pixel with maximum value greater than the percentile 90 of the maximum values becomes a possible deforestation event. If there are pixel values less than the value of the percentile 90, those pixels are labelled as “no deforestation events” during the monitoring period. In the case of VH band, the figure 18 right side, the green dotted line represents the threshold selected, around 10, The latter is used for the whole AOI P_{90}.

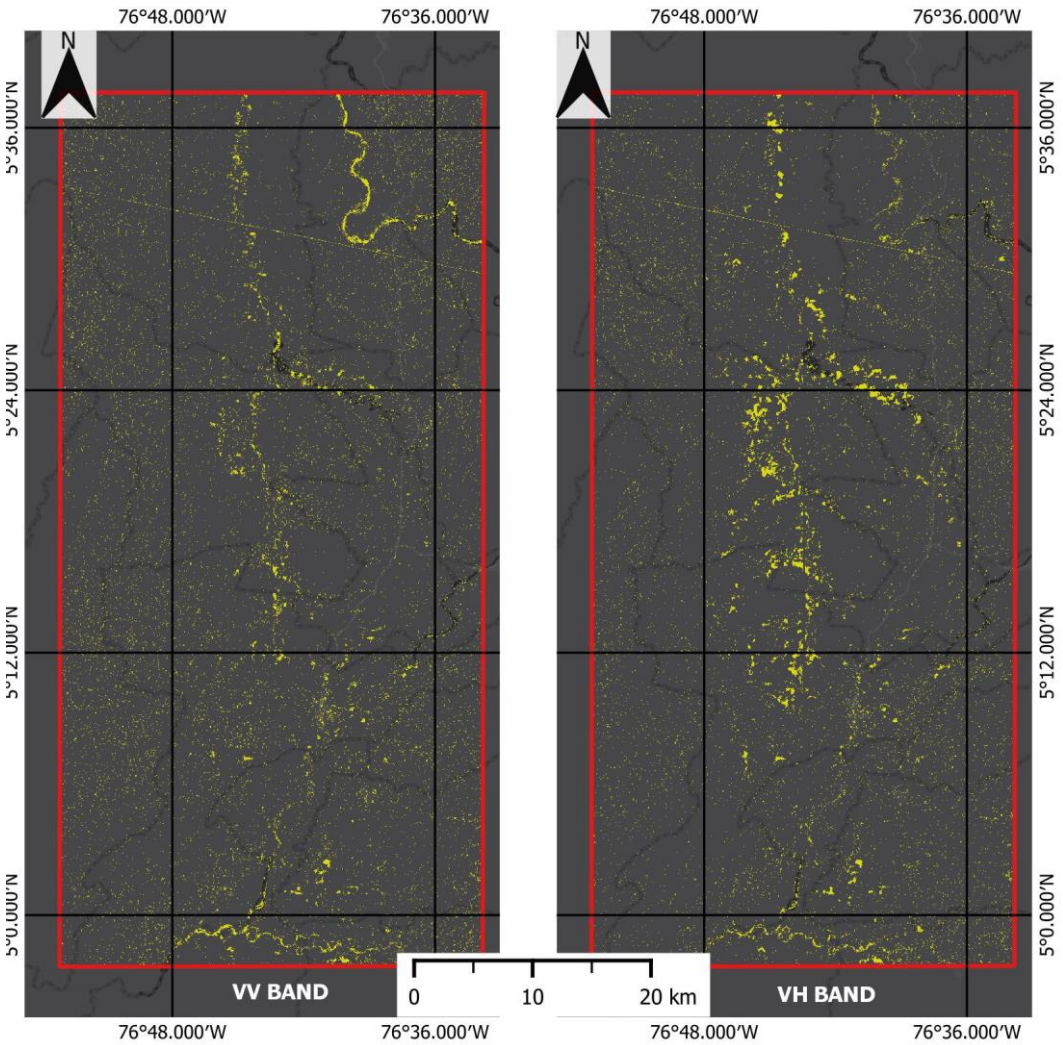

Figure 19. Cusum change detection preliminary results, the yellow polygons indicate possible deforestation events on the monitoring period; left: Using VV band; right: Using VH band.

Applying the threshold value, a preliminary result is obtained, the figure 19 depicts the results from Cusum change detection using VV and VH bands, before postprocessing stage.

Similar argumentation was used with the mRFDI, as index, it goes from 0 to 1, if value is less than 0.3 indicates dense forest, values between 04 and 0.6 indicate degraded forest, if the value is greater than 0.6 indicates deforestation. The figure 20 shows the mRFDI from a subset of the AOI at the beginning and end of the monitoring period.

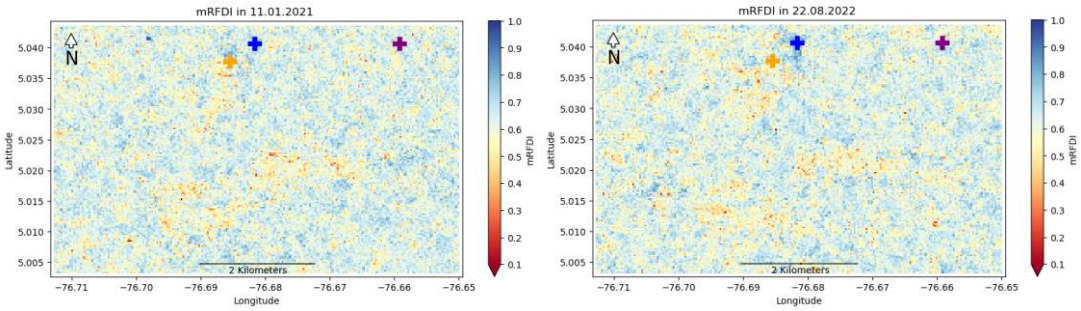

Figure 20. mRFDI for a subset on the left the first image of the time series (11-01-2021), on the right the last image of the time series (08-22-2022).

The index interpretation can be confused, due to the range values of the index, time series plot can illustrate better how the mRFDI behaves for each level of forest disturbances, as values near to 1 indicates deforestation, it means whether the mRFDI value increases, then the forest cover is presenting disturbances, as consequence the shape of the Cusum curve is a U-shape.

Therefore, with the mRFDI, it is necessary to check the minimum values of the Cusum as indicator of deforestation. The figure 21 lower left and lower right, illustrates the mRFDI plot over the time period and the Cusum curve with the mRFDI band. The plot of the mRFDI depicts a limitation of the index as indicator for deforestation events, the variability of the mRFDI along the time. Also, no matter the disturbance level, the plot shows that deforested places, undisturbed places and places suffering a deforestation event share the mRFDI values overlap between them.

After statistical analysis and try-and-error process, the pixels selected as possible deforestation events are those which minimum value from Cusum are less than the

percentile 10th (around -0.60). Pixels values greater than the threshold selected are labelled as no deforestation event.

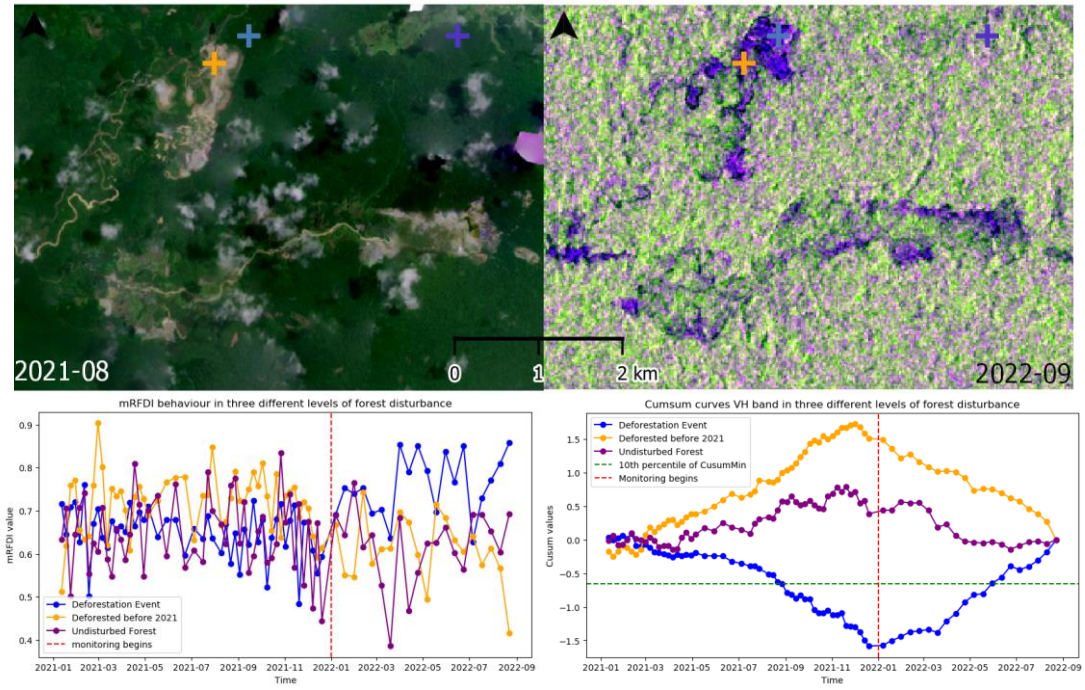

Figure 21. mRFDI and Cusum curve for three levels of forest disturbance. Upper left: Planet basemap monthly mosaic august 2021. Upper right: RGB composite (VV, VH, ratio VV/VH) S1 September 2022. Lower left: mRFDI curves for each place along the study period. Lower right: Cusum curve for each place.

In similar way with the previous figures, the figure 21 represents in blue colour a place with deforestation event on the monitoring period, in purple colour a place with undisturbed forest and orange colour a place deforested before this study which also remains deforested.

Once the post-processing was applied for the VV, VH and mRFDI, it was performed the accuracy assessment based on the three points of view mentioned in 3.5, the table 2 summarizes the results.

Band	Confusion matrix 300 random points		Intersection polygons		Visual Interpretation random 30% of results	
	Overall Accuracy	Kappa	Precision	Recall	Overall Accuracy	Kappa
VV band	0.51	0.02	0.05	0.079	0.68	0.35
VH band	0.63	0.26	0.24	0.55	0.84	0.69
mRFDI	0.65	0.31	0.24	0.12	0.84	0.69

Table 2. Accuracy of the results using Cusum.

In the same way, the table 3 presents the number of regions and area labelled as deforestation alert in each band5.

Band	number of Polygons	Area (ha)
VV band	218	737.2
VH band	319	1695.5
mRFDI	75	297.2

Table 3. Area detected as deforestation events in the monitoring period.

Although, the accuracy between the VH band and mRFDI looks similar, the table 3 shows that the Cusum change detection with the VH band was capable to capture more places or polygons which represent deforestation event in the monitoring period than the VV band and the mRFDI. In fact, from the extension perspective the area detected as deforestation with the VH band is twice than the area using VV and five times the area detected with the mRFDI.

The figure 22 depicts the deforestation alerts detected from each band, in the zoom-in window can be noticed that some places where detected as deforestation event in the three bands, but the size or extension of those events is not the same.

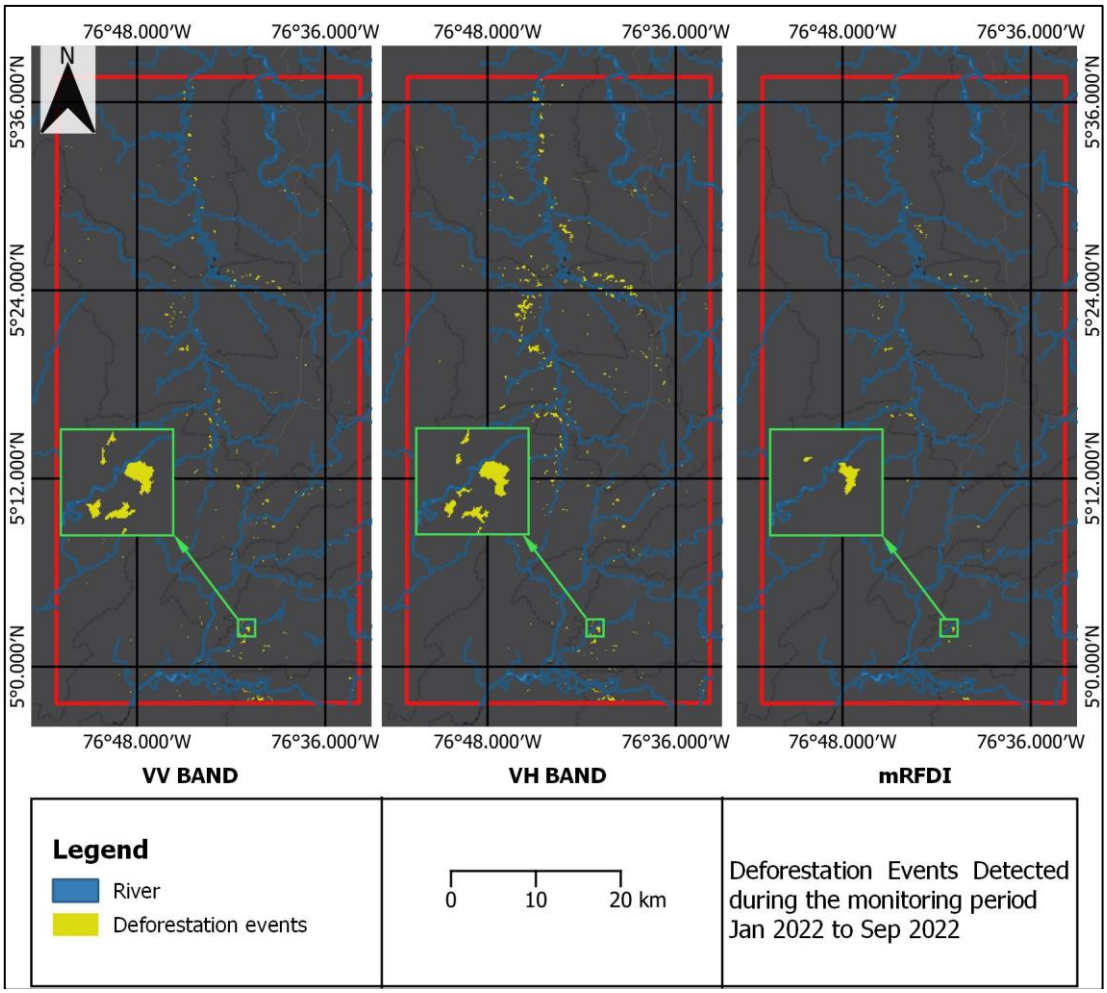

Figure 22. Deforestation events from 2022/01/01/ up to 2022/08/31, left: using VV band, centre: using VH band, right: using mRFDI.

The size of the deforestation alerts was also analysed, the figure 23 illustrates the violin plot about the size of the deforestation events with every band and the mRFDI.

While the VV band and mRFDI allowed to detect deforestation events with size up to 20 and 25 hectares, the VH band allowed to detect deforestation events with size greater than 90 hectares. However, all of them show that most of the deforestation events size has less than 5 hectares.

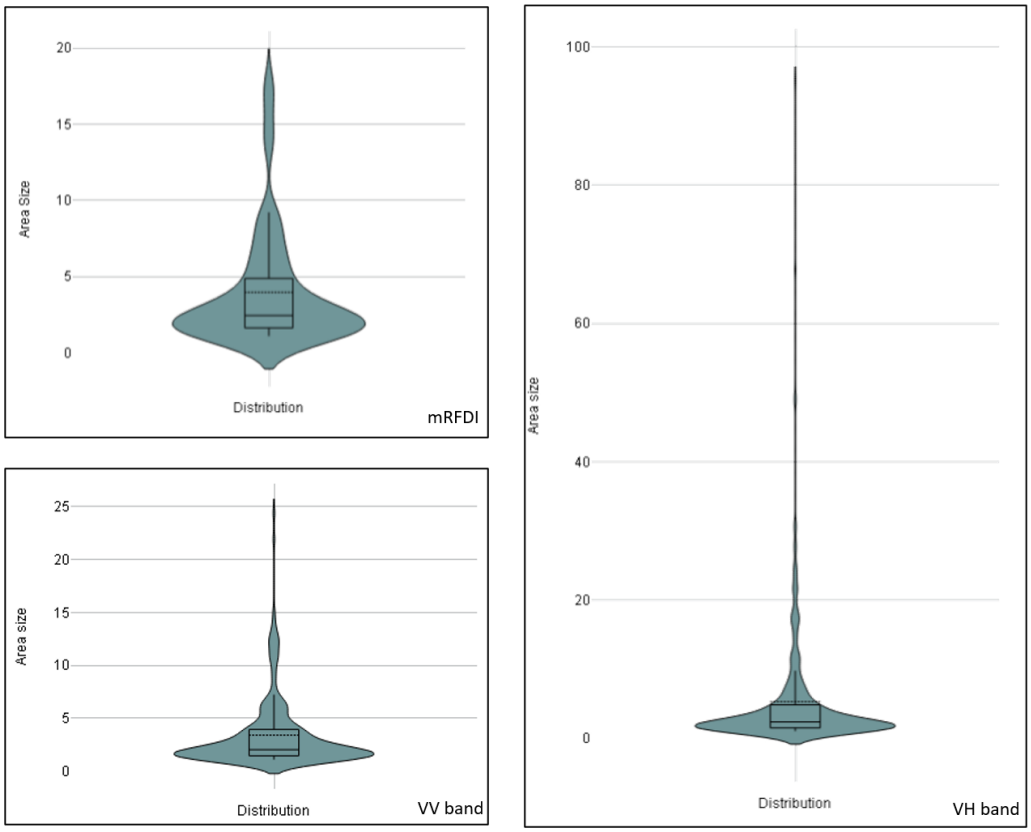

Figure 23. Violin plot Area size from the deforestation events detected during the monitoring period.

Moreover, a comparison based on selection by location was performed to assess if the deforestation events detected using each band and the mRFDI were the same places, the table 4 shows that 94% of the events detected with the mRFDI intersect deforestation events from the use of the VH band, and the 66% of the events detected with the VV band crosses the deforestation events from the VH band.

Selection based on spatial relationship	mRFDI compared to VV band	mRFDI compared to VH band	VV band compared to VH band
Are within	1 (1,33%)	8 (10,6%)	3 (1.37%)
Intersect	38 (50.6%)	71 (94.6%)	146 (66.9%)
Disjoint	37 (49.3%)	4 (5.3%)	72 (33.02%)

Table 4. Spatial relationships between polygons detected as deforestation using VV, VH bands and mRFDI.

Based on the analysis from the results, the VH band was chosen as the best input which allows to detect deforestation events using Cusum change detection on the

monitoring study. Then the next subchapter will show the results using the VH band as input but the Log ration methodology for detection deforestation alerts.

4.2. Deforestation alerts with Log-Ratio change detection.

In the same way the change detection with Cusum, the period where the changes will be monitor goes from 01/01/2022 up to 01/09/2022, and the reference period goes from 01/01/2021 up to 31/12/2021.

The change detection with Log-ratio only was applied to the band which demonstrated best result from the Cusum approach (VH band). The Log-ratio relies on the ratio operation between two images. According to Ajadi et al. and Flores et al. (2016; 2019) the log ratio should be applied using images on the same season.

The figure 24 shows the average backscatter in the VH band for the whole AOI for each date in the time series, and it does not show evidence of seasonality as the backscatter is very similar in each date; then it was decided to run the Log-ratio using as reference the mean of the images from 2021 and as new image the mean of the images from the monitoring period, the images from 2022.

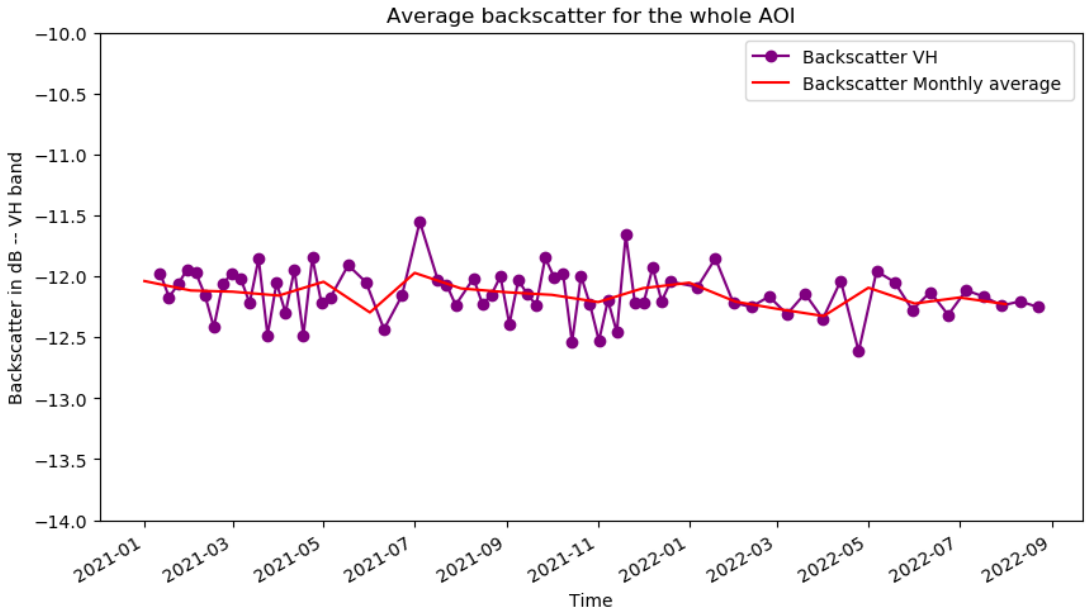

Figure 24. Average backscatter VH band in the AOI.

Afterwards, it was performed a rescaling operation converting the decibel values to the power domain, following the equation Eq. (5).

$$\gamma^0_{pwr} = 10^{\frac{\gamma^0_{dB}}{10}} \qquad \text{Eq. (5)}$$

Where γ^0_{pwr} is the backscatter in power domain and γ^0_{dB} is the backscatter in logarithmic scale in decibels. After the conversion, the operation in the Equation Eq. (4) is performed and the ratio image is obtained.

Plotting the histogram from the result (See figure 25), the unchanged pixels will be in the middle of the histogram, on the right tale the pixels where the backscatter increases between the periods, on the left tale the pixels where the backscatter decreases.

Focus on the left tale, another try-and-error process was performed in order to choose a right threshold for masking out the possible deforestation events. Based on the statistical measures such as the mean and standard deviation (See figure 25), the threshold was selected.

$$da = \mu - 3\sigma \qquad \text{Eq. (6)}$$

Where da stands for deforestation alerts, μ is the mean and σ is the standard deviation from the ratio image obtained.

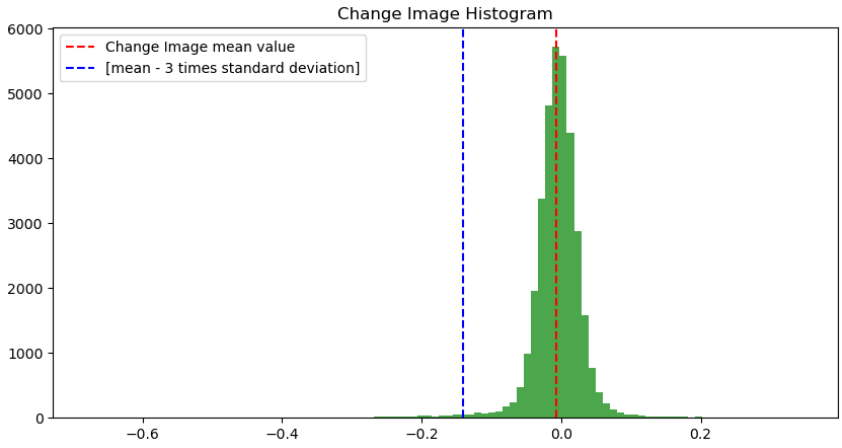

Figure 25. Histogram from the result of the Log-ratio operation.

Equation Eq. (6) was the appropriate equation found for detecting deforestation trying to reduce the false positives (pixels detected as deforestation but them do not represent any deforestation event during the monitoring period).

Thus, the pixel values which are less than the selected threshold become in places with deforestation events and the pixel values greater than the threshold are labelled as no deforestation event.

4.3. Comparison between Cusum and Log Ratio

After the post-processing stage, it was performed the accuracy assessment in the same way that was applied before with Cusum approach, but this time it was compared the performance of detecting deforestation events using the VH band between the Cusum method and the Log-ratio method, the table 5 illustrates the accuracy against the three points of view used.

Band	Confusion matrix 300 random points		Intersection polygons		Visual Interpretation random 30% of results	
	Overall Accuracy	Kappa	Precision	Recall	Overall Accuracy	Kappa
VH band Log Ratio	0.62	0.24	0.26	0.29	0.80	0.61
VH band Cusum	0.63	0.26	0.24	0.55	0.84	0.69

Table 5. Comparison Accuracy using Log-ratio and Cusum with the VH band.

Additionally, the table 6 presents the number of regions and area labelled as deforestation alert.

Band	number of Polygons	Area (ha)
VH band Log Ratio	155	786.4
VH band Cusum	319	1695.5

Table 6. Area detected as deforestation events in the monitoring period from the two methods.

Although, the accuracy between the two methods is similar, the table 6 shows that the accuracy from the Cusum are greater than Log ratio by 4%, also the kappa index in both cases indicate good agreement for the results as the value is between 0.60 and 0.70.

However, it is very interesting that the area found as deforestation events with Cusum is more than twice the area from Log ratio method. Then, it potentially indicates that Cusum method can detect more deforestation events than the Log ratio and still perform a similar overall accuracy.

The figure 26 depicts the deforestation alerts detected from each method using the VH band, there is a zoom-in window that shows the differences for a specific place, while the Cusum method found 6 places with deforestation events, the Log-ratio method found three places.

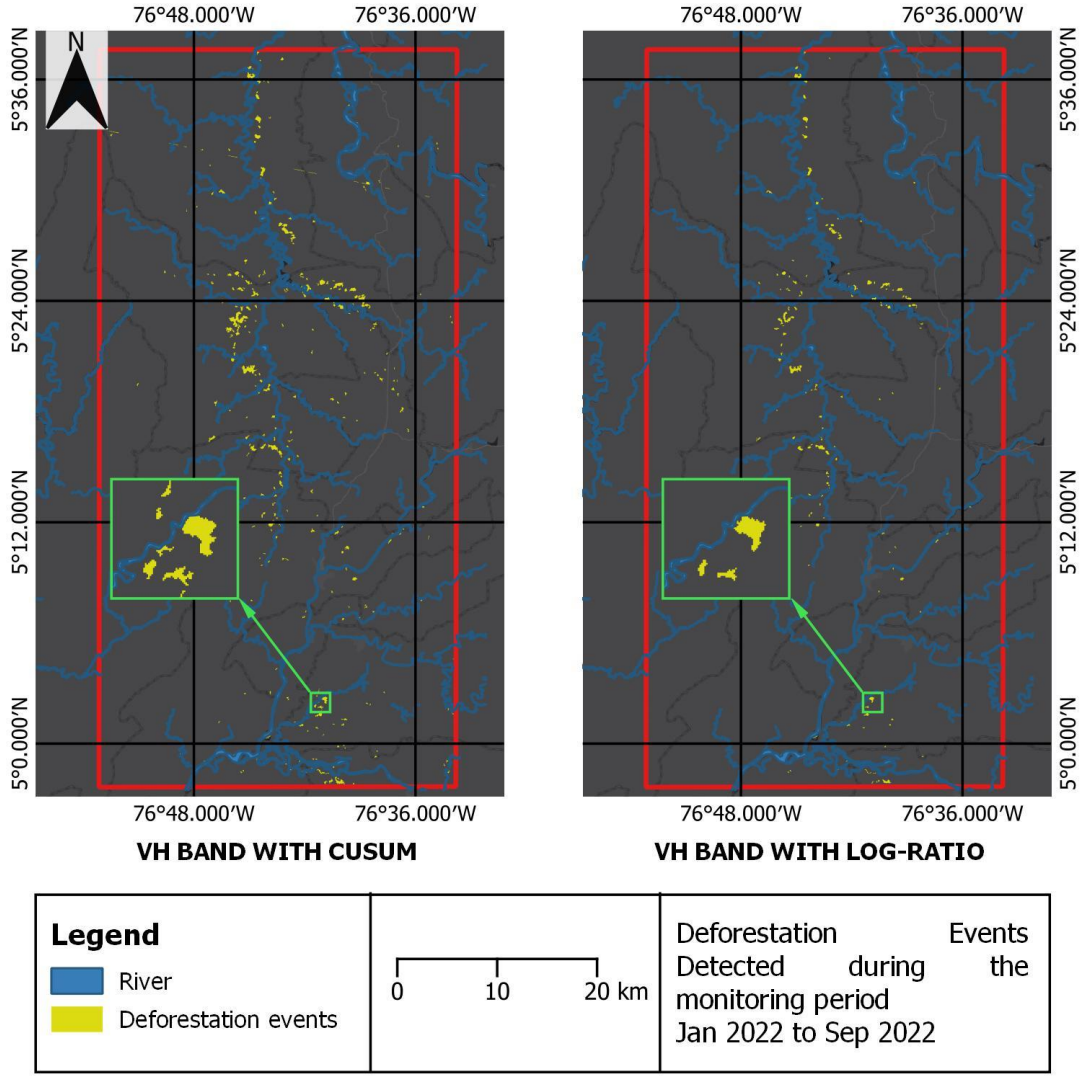

Figure 26. Deforestation events from 2022/01/01/ up to 2022/08/31, left: Cusum method using VH band, right: using Log/ratio method suing VH band.

On the other hand, the figure 27 illustrates the violin plots about the size of the deforestation events with each method, the plots deploy that most of the deforestation events have area extension less than 5 hectares.

Also, the biggest area found with Log ratio method have around 50 hectares, while with Cusum the biggest deforestation event has around 95 hectares.

As in the 4.1 subchapter, a comparison based on selection by location was performed in order to assess if the deforestation events detected using each method cover the same places or not.

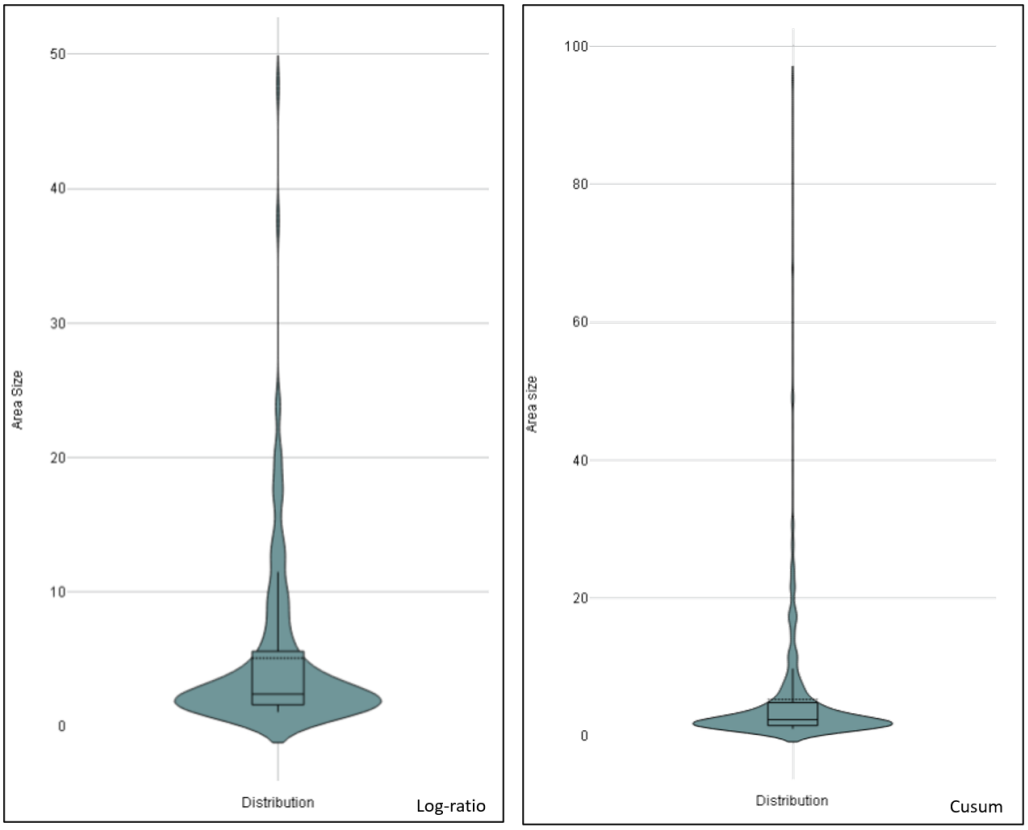

Figure 27. Violin plot Area size from the deforestation events detected during the monitoring period.

The table 7 shows an interesting fact from the results, all the results using Log-ratio method intersects the results from Cusum method, and 5 polygons from the alerts found in Log ratio method are completely inside polygons from alerts detected performing Cusum.

Selection based on spatial relationship	Log ratio compared Cusum
Are within	5 (3.22%)
Intersect	155 (100%)
Disjoint	0 (0%)

Table 7. Spatial relationships between polygons detected as deforestation using VH band.

Besides the possible deforestation events detected, another product obtained from Cusum method is the spatial representation regarding the dates when the deforestation happened or was detected.

The figure 28 depicts the dates about the deforestation was detected during the monitoring period. Those dates illustrates when the Cusum curve reaches the maximum value, but also serves as date marker about when the deforestation event began to be noticeable enough that the trend backscatter changed and the cusum curve got an inflection point and change of the slope direction.

The maximum value from cusum curve is highly dependent on the average backscatter value and the global statistics from the time series used as parameters for the threshold, then, it could happen that the deforestation event occurred a little bit earlier or later than the date calculated with the methodology; the algorithm cannot find the first date of logging activity, it will be necessary more images that confirm the backscatter decrease.

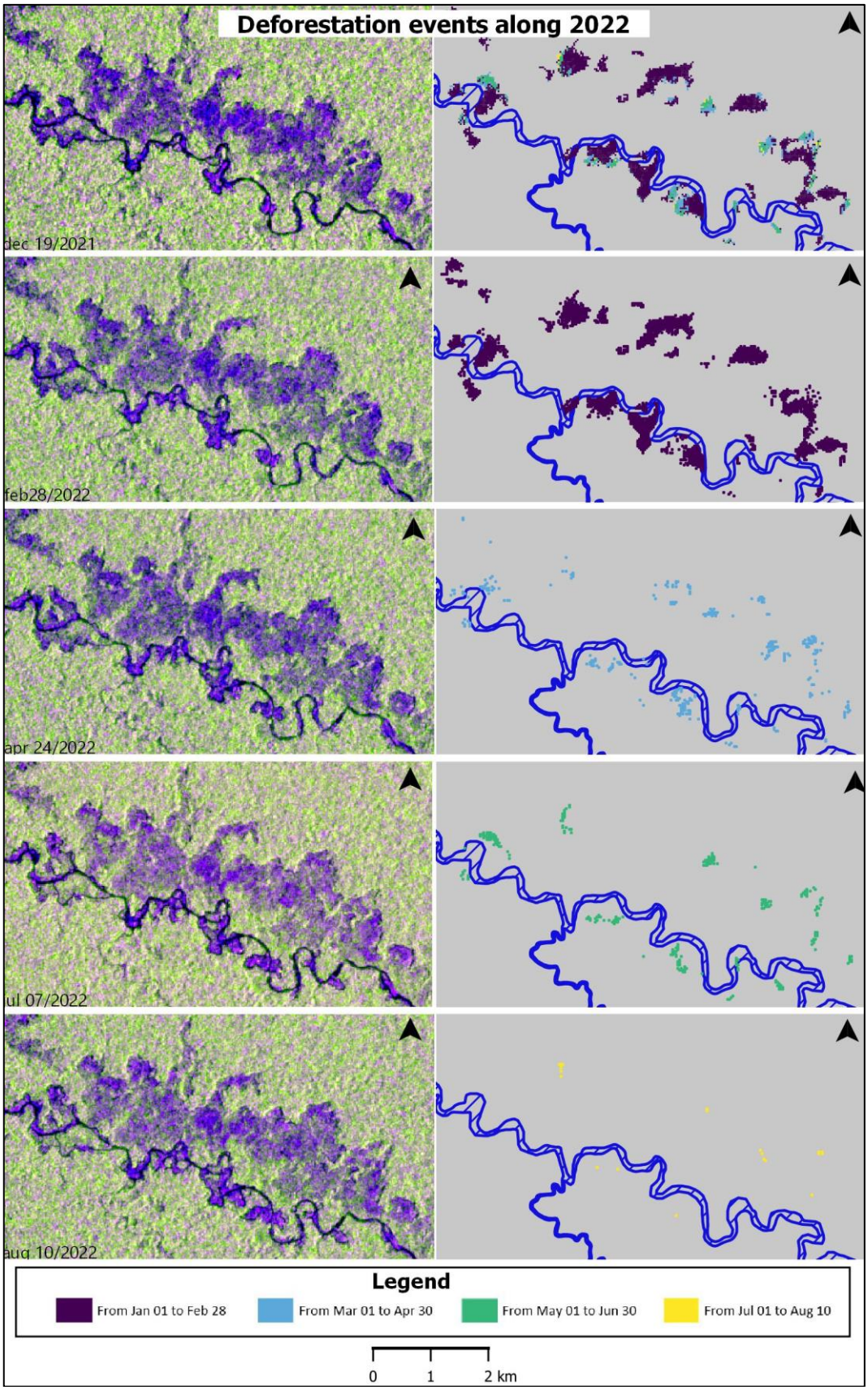

Figure 28. Dates map of the Deforestation events in small subset from the AOI.

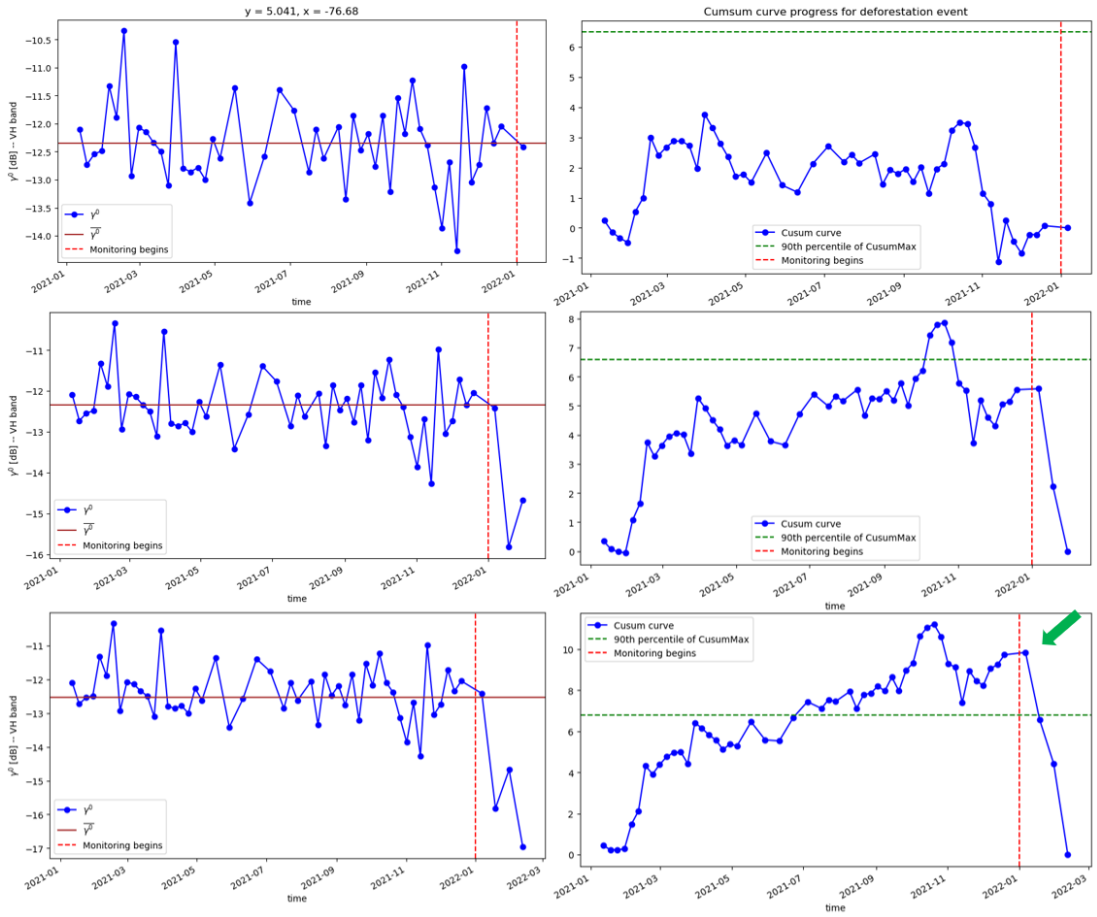

Figure 29. Necessary Time to detect deforestation events.

The figure 29 depicts the progress on the Cusum curve in order to find the deforestation event, it demonstrates for that specific place, it was necessary 4 consecutive S1 images where the backscatter was under the mean, and the backscatter decreasing was enough to pull the statistics and find the first inflection point on the Cusum curve in the monitoring period, pointed it out by the green arrow.

Finally, assuming that the result obtained from the deforestation events with the VH band and Cusum can be added directly to the initial deforestation layer, the table 8 indicates how much the deforestation increased in the study area. The figure 30 depicts the updated forest/non-Forest map.

Non-Forest land cover in the AOI	Area (ha)
Initial deforestation layer	51809.42
Updated layer	53504.92

Table 8. Updated size non-forest land cover in the AOI.

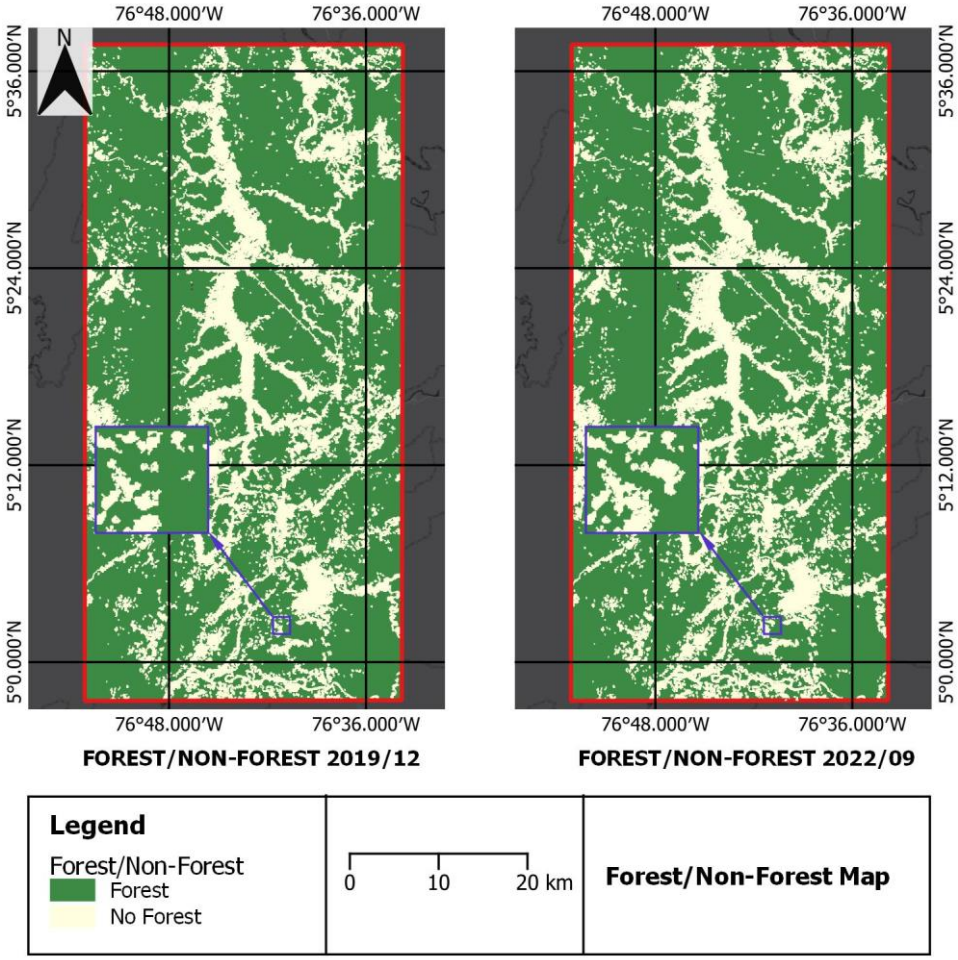

Figure 30. Updated forest/non-forest map for the AOI.

4.4. Contribution of the mRFDI for detecting deforestation

The results from the mRFDI as input for the Cusum algorithm depicted a good performance according to the overall accuracy from the table 3, but low capacity to discriminate the deforestation on the monitoring period, compared with the VH band as input of the Cusum, it only got 5 times less than results from VH band.

As the mRFDI indicates values between 0 and 1, the time series depicts less variation on the pixel values, the pixel values of disturbances have decimal order and can influence the cusum curve, making it a curve with several fluctuations along the time series. The figure 31 illustrates that behaviour for different places on the AOI. Even the upper right and lower right images were not detected as deforestation events and

event when the images show the opposite, the cusum curves have shapes with several peaks.

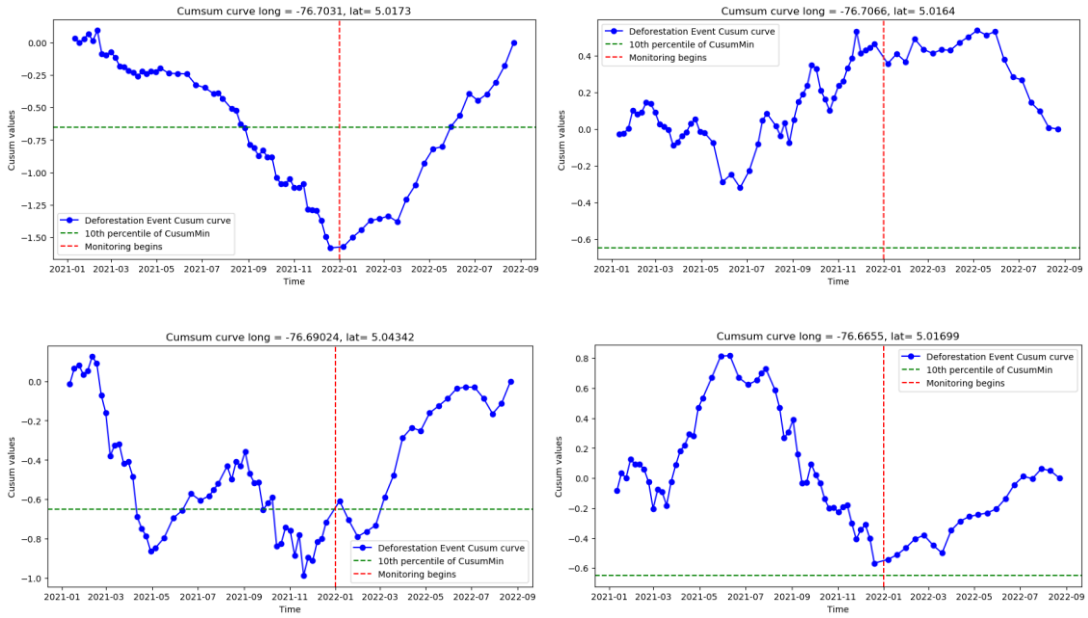

Figure 31. Comparison of cusum curves for places with deforestation on the monitoring period using the mRFDI

Thus, the mRFDI did not add additional features that can contribute to a best performance that the use of VH and VV band for detecting deforestation, the above into the framework proposed for this work. However, in other frameworks such as those presented by Luca, N Silva, et al. (2022) and by Joshi et al. (2015). The mRFDI and the original RFDI demonstrated success as inputs for classification land cover and forest disturbances respectively.

5 DISCUSSION AND CONCLUSIONS

5.1. Conclusion

This study performed two methodologies for time series analysis focus on detecting deforestation events. The backscatter decreases when a logging event happens and based on follow that behaviour Cusum algorithm can be used for detecting those changes. The log-ratio is a method for highlighting the changes between two times and this study used it to look for the deforestation events.

This thesis offers its results of applicability of SAR Images on places where the precipitation rate is very high as reliable when the possibility of optical images is narrowed to small temporalities over the year. Strategies such as REDD+ relies on the use of EO data and in the specific case of Colombia the use of remote sensing images are focus on optical images for methodologies that track the deforestation, degradation, and improvement of the carbon stocks.

The results show that using Cusum with the VH band polarization has better accuracy than using the VV band polarization, 84% over 65% respectively, it means and also confirms the literature that cross polarizations bands (VH) are more sensible to the volume scattering than the co-polarized bands (VV) even in SAR sensors such as S1(works on band C) which is not the more appropriate to vegetation mapping according to applications according Flores et al. (2019).

The main drawback from the Cusum algorithm is the selection of the monitoring period, if the latter is too long then the possible deforestation events can be covered, for instance, if there was a logging event, and after the logging operation the forest was replaced by crops with fast growing and the monitoring period is too long the deforestation of the forest will be difficult to catch. Then a possible recommendation is to select fixed monitoring periods.

Same disadvantage has the Log-ratio method applied here, if an event happened just once for short time, like a fire on the forest, it can be covered by the average of the other images used. So, if a short deforestation event happened during the monitoring period, that event was covered by the average of same pixel during that time. But if the log ratio is applied to short period, it can catch more changes, but the image used

as reference also will change in short period that can cause double detections or more false positives.

Overall, the answers to the research questions are the following:

Is it possible to get a reliable result for monitoring deforestation using only radar images, in regions with very high precipitation rate?

The subchapter 4 expressed the results, depicting good performances against two of the three points of view for the quality assessment. Highlighting that the VH band produces the best result in both methodologies, Cusum and Log-ratio, 63% and 62% of accuracy using the reference deforestation alerts from GFW, 84% and 80% of accuracy using the visual interpretation as reference.

Documents such as the official methodology for quantify deforestation in Colombia (Cabrera et al., 2014) mentions a whole process of quality assessment, but it does not mention the accuracy value or kappa coefficient that can be considered as reliable.

Thus, for the purpose of this study, it was obtained values of overall accuracy of 84% and kappa coefficient of 0.69, it can be considered as reliable result since the objective is to detect or monitor deforestation events. Those events should be checked before to create a final report or map about deforestation in a specific region.

Additionally, this study does not have into account concepts such as forest regeneration or regrowing. Then the results cannot be summarized directly to a global deforestation value or rate.

Does the modified version of Radar Forest Degradation Index RFDI for Sentinel-1 have better results for mapping deforestation?

Although, there was high expectation on the use of mRFDI as input of the methodologies for detecting deforestation events, the results evidence there is not benefit or advantage using the mRFDI instead of the VV or VH bands. Despite, the index performs acceptable accuracy, the result from VH band includes most than 90 % of the deforestation events detected with mRFDI and the VH band allowed to find 5 times more deforestation events than the mRFDI.

Thus, the mRFDI does not show better results for mapping deforestation events, at least with the methodologies applied to this study.

Is it feasible to use the workflow proposed in this study in a production environment?

The methodology workflow for this study was developed using open-source software, the pre-processing and processing stages were performed using python libraries and jupyter notebooks, the post-processing stage was performed using QGIS tools.

The main inputs are the S1 collection images from google earth engine which have free access through google earth engine account which is also free.

Thus, the methodology workflow can serve as framework and be optimized especially on the post-processing stage, in the step related to the validation of the deforestation events. In fact, there are examples on the literature about it, Doblas et al (2022) developed a full operational near real detection forest disturbances based on S1 images retrieved from google earth engine.

Regarding the hardware, it would be necessary to assess the required capabilities in order to check the power computing for running the algorithms.

5.2. Limitations and future work

This study was focus on the detection of deforestation events and mapping them. But usually quantify deforestation includes also to check regeneration or vegetation regrowing. Detecting vegetation regrowing was not part of the methodology due to the complexity of classifying levels of regeneration, in the monitoring period for this study a dense forest will not appear, but deforested land can show grass growing and bushes in short time due to the ecosystem is humid basal forest. If that is the case, the backscatter reduction will not be enough to be selected in the threshold from the Cusum curves.

The Cusum algorithm needs that the deforestation event remains for some time to be detected by it, it means, Cusum is not able to show deforestation from the last image from the time series. The best results (figure 28) show deforestation events up to August 10 in a time series which cover up to September 1st.

The initial deforestation layer used as mask for removing comes from the Forest/non-Forest map for the year 2019, which is the last public available for the study area, it can influence in some way the final results, as the study period is the whole 2021 and part of the 2022, but 2022 is the monitoring period, it was not measured the deforestation during 2020 and 2021.

Regarding the future work several paths can be carry out, to monitor the recovery or vegetation regeneration is the most direct step; instead of methods based on statistical measurements for selecting the thresholds where the deforestation appears, it can be used other methods such as probabilistic methods, where function can be tested if there is a probability that a pixel represents a deforestation event. The latter also can be tested with machine learning approaches.

Another possibility is further work on the very simple but powerful approach Log-ratio method for almost near real time monitoring as this method allows to highlight changes between a reference image and a new image, having into account the main drawback of the method which is that it will be necessary to select thresholds from the change image.

BIBLIOGRAPHIC REFERENCES

Ajadi, O. A., Meyer, F. J., & Webley, P. W. (2016). Change Detection in Synthetic Aperture Radar Images Using a Multiscale-Driven Approach. *Remote Sensing*, *8*(6). https://doi.org/10.3390/rs8060482

Aquino, C., Mitchard, E. T. A., McNicol, I. M., Carstairs, H., Burt, A., Puma Vilca, B. L., Obiang Ebanéga, M., Modinga Dikongo, A., Dassi, C., Mayta, S., Tamayo, M., Grijalba, P., Miranda, F., & Disney, M. (2022). Reliably mapping low-intensity forest disturbance using satellite radar data. *Frontiers in Forests and Global Change*, *5*. https://doi.org/10.3389/ffgc.2022.1018762

Atwood, D., Small, D., & Gens, R. (2012). Improving PolSAR Land Cover Classification With Radiometric Correction of the Coherency Matrix. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *5*, 848–856. https://doi.org/10.1109/JSTARS.2012.2186791

Ballère, M., Bouvet, A., Mermoz, S., le Toan, T., 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 Sensing of Environment*, *252*, 112159. https://doi.org/10.1016/J.RSE.2020.112159

Bouvet, A., Mermoz, S., Ballère, M., Koleck, T., & le Toan, T. (2018). Use of the SAR shadowing effect for deforestation detection with Sentinel-1 time series. *Remote Sensing*, *10*(8). https://doi.org/10.3390/RS10081250

Cabrera, E., Galindo, G., & Vargas, D. M. (2014). Protocolo de Procesamiento Digital de Imágenes para la Cuantificación de la Deforestación en Colombia, Nivel Nacional Escala Gruesa y Fina. Instituto de Hidrología, Meteorología, y Estudios Ambientales-IDEAM. In *Investigación agraria. Sistemas y recursos forestales* (Vol. 84, Issue 33).

Canty, M. J., Nielsen, A. A., Conradsen, K., & Skriver, H. (2020). Statistical analysis of changes in sentinel-1 time series on the Google earth engine. *Remote Sensing*, *12*(1). https://doi.org/10.3390/RS12010046

Centre for Remote imaging, sensing & processing. (2001). *Tutorial Microwave Remote Sensing*. https://crisp.nus.edu.sg/~research/tutorial/mw.htm

de Luca, G., M. N. Silva, J., di Fazio, S., & Modica, G. (2022). Integrated use of Sentinel-1 and Sentinel-2 data and open-source machine learning algorithms for land cover mapping in a Mediterranean region. *European Journal of Remote Sensing*, *55*(1), 52–70. https://doi.org/10.1080/22797254.2021.2018667

de Luca, G., N Silva, J. M., di Fazio, S., & Modica, G. (2022). *European Journal of Remote Sensing ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tejr20 Integrated use of Sentinel-1 and Sentinel-2 data and open-source machine learning algorithms for land cover mapping in a Mediterranean region*. https://doi.org/10.1080/22797254.2021.2018667

Díaz Merlano, J. M., & Gast, F. (2009). El Chocó biogeográfico de Colombia /. In B. de O. C. (Colombia) (Ed.), *Banco de Occidente Credencial.* Banco de Occidente Credencial.

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 Sensing*, *14*(15). https://doi.org/10.3390/rs14153658

dos Santos, E. P., da Silva, D. D., & do Amaral, C. H. (2021). Vegetation cover monitoring in tropical regions using SAR-C dual-polarization index: seasonal and spatial influences. *International Journal of Remote Sensing*, *42*(19), 7581–7609. https://doi.org/10.1080/01431161.2021.1959955

Flores, A., Herndon, K., Thapa, R., & Cherrington, E. (2019). *The SAR Handbook: Comprehensive Methodologies for Forest Monitoring and Biomass Estimation*. https://doi.org/10.25966/nr2c-s697

Gašparović, M., & Dobrinić, D. (2020). Comparative Assessment of Machine Learning Methods for Urban Vegetation Mapping Using Multitemporal Sentinel-1 Imagery. *Remote Sensing*, *12*(12). https://doi.org/10.3390/rs12121952

Google Earth Engine. (n.d.). *Sentinel-1 Algorithms*. Retrieved December 16, 2022, from https://developers.google.com/earth-engine/guides/sentinel1

Hirschmugl, M., Deutscher, J., Sobe, C., Bouvet, A., Mermoz, S., & Schardt, M. (2020). Use of SAR and optical time series for tropical forest disturbance mapping. *Remote Sensing*, *12*(4). https://doi.org/10.3390/RS12040727

IDEAM. (n.d.-a). *Atlas Climatológico de Colombia*. Retrieved July 2, 2022, from http://atlas.ideam.gov.co/visorAtlasClimatologico.html

IDEAM. (n.d.-b). *MONITOREO Y SEGUIMIENTO DE BOSQUES*. Retrieved July 2, 2022, from http://www.ideam.gov.co/web/ecosistemas/deforestacion-colombia

IDEAM. (2021). *Resultados del monitoreo de deforestación año 1. 2020 2. primer trimestre 2021*. http://www.ideam.gov.co/documents/10182/113437783/Presentacion_Deforestacion2020_ SMByC-IDEAM.pdf/8ea7473e-3393-4942-8b75-88967ac12a19

Instituto de Hidrología, M. y E. A. (Ideam), Instituto de Investigación de Recursos Biológicos Alexander von Humboldt (Instituto Humboldt), Instituto de Investigaciones Marinas y Costeras José Benito Vives de Andréis (Invemar), & Instituto Geográfico Agustín Codazzi (IGAC). (2017). *Memoria técnica. Mapa de ecosistemas continentales, costeros y marinos de Colombia (MEC), escala 1:100.000*.

Joshi, N., Mitchard, E., Woo, N., Torres, J., Moll-Rocek, J., Ehammer, A., Collins, M., Jepsen, M., & Fensholt, R. (2015). Mapping dynamics of deforestation and forest degradation in tropical forests using radar satellite data. *Environmental Research Letters*, *10*, 034014.

Mermoz, S., & le Toan, T. (2016). Forest disturbances and regrowth assessment using ALOS PALSAR data from 2007 to 2010 in Vietnam, Cambodia and Lao PDR. *Remote Sensing*, *8*(3). https://doi.org/10.3390/RS8030217

Mora, O., Ordoqui, P., Iglesias, R., & Blanco, P. (2016). Earthquake Rapid Mapping Using Ascending and Descending Sentinel-1 TOPSAR Interferograms. *Procedia Computer Science*, *100*, 1135–1140. https://doi.org/10.1016/j.procs.2016.09.266

Motohka, T., Shimada, M., Uryu, Y., & Setiabudi, B. (2014). Using time series PALSAR gamma nought mosaics for automatic detection of tropical deforestation: A test study in Riau, Indonesia. *Remote Sensing of Environment*, *155*, 79–88. https://doi.org/https://doi.org/10.1016/j.rse.2014.04.012

Mullissa, A., Vollrath, A., Odongo-Braun, C., Slagter, B., Balling, J., Gou, Y., Gorelick, N., & Reiche, J. (2021). Sentinel-1 SAR Backscatter Analysis Ready Data Preparation in Google Earth Engine. *Remote Sensing*, *13*(10). https://doi.org/10.3390/rs13101954

Nicolau, A. P., Flores-Anderson, A., Griffin, R., Herndon, K., & Meyer, F. J. (2021). Assessing SAR C-band data to effectively distinguish modified land uses in a heavily disturbed Amazon forest. *International Journal of Applied Earth Observation and Geoinformation*, *94*. https://doi.org/10.1016/j.jag.2020.102214

Planet Labs PBC. (n.d.). *Visual Basemaps*. Retrieved December 18, 2022, from https://developers.planet.com/docs/data/visual-basemaps/

QGIS project. (n.d.). *QGIS Documentation*. 24.2.1. Raster Analysis.

Reiche, J., Verhoeven, R., Verbesselt, J., Hamunyela, E., Wielaard, N., & Herold, M. (2018). Characterizing Tropical Forest Cover Loss Using Dense Sentinel-1 Data and Active Fire Alerts. *Remote Sensing*, *10*(5). https://doi.org/10.3390/rs10050777

Ruiz-Ramos, J., Marino, A., Boardman, C., & Suarez, J. (2020). Continuous forest monitoring using cumulative sums of sentinel-1 timeseries. *Remote Sensing*, *12*(18). https://doi.org/10.3390/RS12183061

SEPAL development team. (n.d.). *System for earth observation, data access, processing, analysis for land monitoring*. Retrieved October 16, 2022, from https://docs.sepal.io/en/latest/

United Nations Framework Convention on Climate Change. (n.d.). *What is REDD+?* Retrieved July 2, 2022, from https://unfccc.int/topics/land-use/workstreams/redd/what-is-redd

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. *International Journal of Applied Earth Observation and Geoinformation*, *103*. https://doi.org/10.1016/j.jag.2021.102532

2023	***MONITORING DEFORESTATION IN REGIONS WITH HIGH PRECIPITATION RATE*** ***Case of study Chocó- Colombia***	Jhon Eric Aunta Duarte

Masters Program in **Geospatial Technologies**

Supported by:
Education and Culture
ERASMUS MUNDUS