UCLouvain

AGRO LOUVAIN

Faculté des bioingénieurs

Mapping the Tropical Moist Forests Extent Based on Combination of Sentinel-1 and Sentinel-2 Satellites

Study area: Tshopo province, Democratic Republic of Congo

Auteur: Sourav Karmakar

Promoteur(s): Pierre Defourny (UCLouvain)

Baptiste Delhez (UCLouvain)

Lecteur(s) : Sophie Bontemps (UCLouvain)

Evelyn Uuemaa (University of Tartu)

Année académique 2022-2023

Mémoire de fin d'études présenté en vue de l'obtention du diplôme de Master SAIV – MSc Erasmus+ GEM : Geo-information Science and Earth Observation for Environmental Modelling and Management

Acknowledgement

I would first and foremost like to express my profound sense of gratitude to my respected supervisor Professor Pierre Defourny, for his indispensable and expert guidance, direct concern, constructive criticism, and inspiration throughout this thesis. Then, I would like to express my immense gratitude to my co-supervisor Baptiste Delhez for his continuous support, prompt replies, technical expertise, and pragmatic look which made it possible to develop the framework of this research.

Next, I warmly thank Thibauld Collet for his suggestions and support regarding data accessibility and processing. I express my gratitude to Nicolas Deffense for developing the 'eo-toolbox' which made it easier to conduct my analysis in python and GEE platform. I am also thankful to the other researchers from the Geomatics laboratory for their arguments, criticisms, and suggestions.

I would like to thank Erasmus Mundus Joint Masters scholarship for hosting me conducting my master's under the GEM program. I am grateful to my current institution 'Université catholique de Louvain' and my former educational institutions 'University of Tartu' and 'University of Chittagong' for raising my foundation in the field of GIS and remote sensing.

Later, I would like to thank my beloved wife for her continuous mental support and motivation throughout this journey. I also emphasize the enthusiastic support of my parents, younger sister, friends, and relatives from Bangladesh whose communication always encouraged me to conduct this work.

Finally, I am particularly grateful to myself for the passion, dedication and perseverance I put into this research. Without my self-motivation and commitment, completing this thesis would be difficult.

Table of Contents

List of Figures

List of Tables

List of Acronyms

AWF	African Wildlife Foundation
BAIS2	Burned Area Index for Sentinel-2
CBFP	Congo Basin Forest Partnership
CIFOR	Center for International Forestry Research
COMIFAC	Central Africa Forest Commission
COP	Conference Of Parties
DEM	Digital Elevation Model
DRC	Democratic Republic of Congo
ECMWF	European Center for Medium-Range Weather Forecasts
ERS	European Remote Sensing
ESA	European Space Agency
ETM	Enhanced Thematic Mapper
EUMETSAT	European Organization for the Exploitation of Meteorological Satellites
FCDM	Forest Canopy Disturbance Monitoring Tool
GEE	Google Earth Engine
GFC	Global Forest Change
GFW	Global Forest Watch
GLCM	Grey-Level Co-Occurrence Matrix
GRD	Ground Range Detected
GRFM	Global Rain Forest Mapping
IRI	Interfaith Rainforest Initiative
IW	Interferometric Wide Swath
KSAT	Kongsberg Satellite Services
LULC	Landuse and Landcover
MODIS	MODIS (or Moderate Resolution Imaging Spectroradiometer
MSI	Multi-Spectral Instrument
NBR	Normalized Burn Ratio
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NGO	Non-governmental Organization
NICFI	Norway's International Climate & Forests Initiative
NIR	Near-infrared
OLI	Operational Land Imager
RADD	Radar for Detecting Deforestation

REDD	Reduce Emissions from Deforestation and Degradation
RVI	Radar Vegetation Index
S1	Sentinel-1
S1S2	Sentinel-1 and Sentinel-2
S2	Sentinel-2
SAR	Synthetic Aperture Radar
SRTM	Shuttle Radar Topography Mission
SWI	Surface Waterproofing Index
SWIR	Short Wave Infrared
TM	Thematic Mapper
TMF	Tropical Moist Forest
TREES	Tropical Ecosystem Environment Observation by Satellites
VH	Vertical transmit/Horizontal receive
VV	Vertical transmit/Vertical receive
WRI	World Resource Institute
WWF	World Wildlife Fund

1. Introduction

Tropical forests are one of the most important biomes for their floral and faunal diversity and wide ecological importance. They provide incessant ecosystem services including carbon sequestration, water regulation, biodiversity conservation and livelihood support for millions of people (Luyssaert et al., 2008; Vancutsem et al., 2021; Watson et al., 2018). However, tropical forests have been threatened by deforestation and forest degradation driven by increasing population growth, agricultural expansion, shifting cultivation, logging and mining activities in an unsustainable manner (Tyukavina et al., 2018; Umunay et al., 2019).

Remote sensing has been considered as a cost-effective and efficient way of monitoring and managing tropical forests. It paves the way for estimating forest coverage using a combination of different satellite data and methodologies. Despite the availability of global forest extent maps based on remote sensing data (Hansen et al., 2013; Shimada et al., 2014), these maps often lack the adequate spatial and temporal resolution to capture the dynamics of tropical forests at the local scale (Sannier et al., 2016). Moreover, carbon assessment relying on those maps can lead to possible errors and uncertainties (Cerretelli et al., 2021) which can hinder climate mitigation strategies. Therefore, there is a growing importance for up-to-date forest extent maps at a landscape level with a higher level of detail and accuracy. With the availability of ESA's Sentinel-1 and Sentinel-2 satellites, there has been a heightened interest in combining optical and Synthetic Aperture Radar (SAR) data to improve classification accuracy and address the limitations of individual sensors (Heckel et al., 2020; Hirschmugl et al., 2018). In the context of tropical moist forests, where frequent cloud cover can limit continuous monitoring by optical data, SAR data can penetrate clouds and provide complementary information on forest structure and distribution.

In this context, this study will explore the potential of combining Sentinel-1 and Sentinel-2 data for tropical moist forest mapping at a landscape level. The study area is chosen T-shopo province, Democratic Republic of Congo for its significant woodland coverage. This thesis will explore the regional context, review relevant scientific literature, describe the materials and methods used to achieve forest extent classification. As a final result, forest extent classification maps at the landscape level will be presented, discussed and evaluated following accuracy assessment protocol to ensure their validity and reliability. The findings of this research will contribute to the development of more reliable methods for better monitoring and management of tropical forests.

2. Literature Review

This chapter provides a comprehensive overview of current state-of-the-art knowledge in the mapping of forest extent within tropical moist forests. Section 2.1 describes the regional context of Congo basin tropical forest in DRC. Subsequently, section 2.2 focuses on discussing the remote sensing programs and products that have been implemented in tropical forest monitoring. Section 2.3 presents relevant information regarding the utilization of Copernicus satellite data for this purpose. Furthermore, section 2.4 explores the insights gained from combination of data from both Sentinel-1 and Sentinel-2 satellites. Finally, in section 2.6, description of classification algorithms employed in this study are provided.

2.1 Regional context

2.1.1 Congo basin tropical moist forest

Congo basin forest is the second largest dense tropical rain forest after Amazon, covering a forest area of around 2 million square kilometers and spreading across most of the parts of six African countries, namely: Democratic Republic of Congo, Cameroon, Equatorial Guinea, Central African Republic, Gabon and Republic of Congo (Somorin et al., 2012). Congo river basin is itself the second largest basin after Amazon, covering a watershed of 3.7 million km^2 (Modu & Herbert, 2014). The basin excludes Gabon and Equatorial Guinea but extends to western Zambia, northern part of Angola, Burundi, Rwanda, South Sudan and Tanzania. Congo basin has Sahara Desert at the north, Atlantic Ocean at the south and west, and East African lakes at the western side (Gana & Herbert, 2014). Although it is mainly rainforest, the central zone includes mangroves, the rest areas are dryland rainforest including woodland and wooded savanna (De Wasseige et al., 2015). Congo basin forest constitutes almost 18% of world's tropical forest that provides resources and livelihood for around 80 million people (Bergen, 2019). Moreover, it is home to the largest number of world's tropical forest vertebrates with other 400 mammal species, 1000 bird species, 700 types of fish, and over 10,000 plant species. Among them, around 3000 of floral and 360 of faunal species are endemic to the region (Bele et al., 2015). On top of that, the regulating service of Congo basin ecosystem is indispensable. An estimated 29 billion tons of carbon are stored in the peat swamp forest including 1.5 billion tons absorbed by the basin every year (UNEP, 2023).

2.1.2 Forest dynamics and deforestation drivers of Congo basin tropical moist forest

Notwithstanding its importance, Congo basin forest has undergone rapid degradation in recent decades. In between 2000 to 2014, this forest has lost 16 million ha of forest area in total, around 85% of which is caused by small-scale agricultural clearing (Tyukavina et al., 2018). Before permanent deforestation, around 11% of regional deforestation (650,000 ha) between 2015 and 2020 are found to be fragmented, suggesting an increase in small holder subsistence activities (El-Sheikh, 2022). Historically, forest loss in this area is influenced by rural population growth and subsistence agriculture (Megevand & Mosnier, 2013). However, industrial and commercial logging, road expansion, natural resource extraction, settlement expansion, and poaching for bush meat are reported as other drivers of forest loss in recent years (Reiche et al., 2021; Tyukavina et al., 2018; Umunay et al., 2019). Another important driver of forest loss is the need for fuel wood in the regional area. Wood constitutes 75% of industrial and domestic heating in tropical African megacities like Kinshasa or Brazzaville. As a result, a supply chain has been established to meet the requirement, majority of which are taken from clearing of sifting cultivation and the rest are directly or indirectly taken from the forest (Collet, 2022). Forest products are used to combat poverty by supplying foods and livelihoods to the poor (Bele et al., 2015; Nkem et al., 2013; Scherr et al., 2004). Most of the forest loss in the Congo basin is reported unsustainable and illegal, as a clear reflectance of their economic condition, limited capacity, political instability and poorly equipped material used (Kleinschroth et al., 2019; Tegegne et al., 2016; Umunay et al., 2019).

Compared to other tropical forest in Amazon and Southeast Asia, the deforestation rate in Congo basin forest is relatively low. However, recent ongoing trends of habitat fragmentation and forest loss is posing great threat to the irreplaceable biodiversity of this forest since 2020 (El-Sheikh, 2022). As the Congo basin forest is shared among six countries, forest dynamics and drivers reflect the geographical, economical, policy measures and demographical variation among the countries (Tyukavina et al., 2018).

2.1.3 Regional context of DRC

Democratic Republic of Congo (DRC) holds 60% of Congo basin forest which is equivalent to 1.3 million square kilometers. More than 67% of total land use of the country is forest (Schure et al., 2014). DRC offers various landscapes including mountain, savanna, grassland and plateau which provides unique habitat for some endemic like Okapi, Mountain gorilla, Chimpanzees. In addition,

around 23 Giga tonnes of Carbon has been estimated in aboveground and belowground biomass of DRC forested areas (Xu et al., 2017). Despite the importance, DRC experienced 14.6 million hectares loss of forest land in between 2001 to 2019 which is the highest among the Congo basin countries. Global Forest Watch (GFW) has reported 17.1 million hectares tree cover loss from 2001 to 2021 in DRC which is equivalent to 10.5 giga tonnes of CO_2 emission.

There are several challenges DRC has faced over decades which thwart natural forest conservation. An estimated 40 million people directly depend on forests for their food, medicine, timber, and other ecosystem services in DRC. Total 40% of local inhabitants residing near the forested area of DRC use forest as a primary source of their income. (Damania et al., 2016). Historically domestic conflicts and war have created refugees and soldiers, therefore increased forest dependency for food and other resources (Draulans & Krunkelsven, 2002). Another research studied conservation and conflict in this region found significant forest loss during conflict. It also stated the contribution of mining concession to forest loss, but during the war forest loss from mining has reduced (Butsic et al., 2015).

By studying land-use context from 2000 to 2015, found that small scale subsistence agriculture is the main reason of deforestation (Molinario et al., 2020). Other reasons reported are large-scale agroindustry and resources extractions such as mining and logging. As part of Interfaith Rainforest Initiative (IRI), UNEP reported three main drivers of forest loss in DRC: Agriculture and Charcoal production, Illegal logging, and conflict, of which 90% of loss was attributed by agriculture and charcoal production (UNEP, 2019).

Protecting tropical forest from deforestation and degradation is one of the main climate change mitigation strategies recognized in Paris Agreement (Umunay et al., 2019). Therefore, it has vital importance to preserve and protect the remaining tropical forest for sustainable economic growth without compromising its ecosystem services. In DRC, a total of 29 international convention and 40 national laws were signed to reduce the forest loss. Since 2009, DRC came under the umbrella of Reduce Emissions from Deforestation and Degradation (REDD+) to limit its ongoing deforestation and degradation. DRC have been working with National and international partnership to reduce forest loss; to name a few Congo Basin Forest Partnership (CBFP), Central AfriUA Forest Commission (COMIFAC), World Wildlife Fund (WWF), Center for International Forestry Research (CIFOR), African Wildlife Foundation (AWF) (Ngounou, 2020). In 2012, DRC

undertook national REDD+ framework with a vision of increasing forest cover 63.5% by 2030. Again, in 2021 COP 26, Glasgow, a total 145 countries including Congo basin countries has been agreed to reduce deforestation rate 10% every year from 2021. DRC not only failed to keep the promise, instead, it has experienced 5% rise in deforestation rate (El-Sheikh, 2022) **.** Considering the deforestation scenarios and need for conservation mentioned above, continuous monitoring of forest extents and changes is very important for DRC.

2.1.4 Need for Forest extent map

A forest extent map also known as forest, non-forest (F/NF) map shows the size, shape and spatial distribution of forested and non-forested area in a particular area or landscape. Up-to-date information of forest dynamics is crucial for national forest inventory, carbon stock assessment, biodiversity conservation and restoration, global environmental assessment and local forest management planning (Sharma et al., 2018; Waser et al., 2015). In the dynamic tropical forest, where forest change is very rapid, an accurate, spatially detailed, updated forest extent map is very important. Traditional forest extent maps are produced from visual photo interpretations with field survey, which is cumbersome and limited to small scale. But in this era, various remote sensing approaches has been successfully delineated forest non forest boundaries ranging from regional to global scale (Dostálová et al., 2016; Hansen et al., 2013; Kempeneers et al., 2012).

2.2 Remote sensing of tropical forest monitoring

Considering the mitigation strategies of deforestation and degradation taken by different partnership organizations, government, and NGOs, updated data on forest cover is vital. For the vast forested area like DRC, Remote sensing-based forest quantification could be cost-effective solution to provide large scale data in shortest time. It is considered as one of the major tools to identify and quantify forest loss (Kim et al., 2014). Consecutive advancement in remote sensing and satellite imageries made it possible to map and monitor tropical forest dynamics for a long period. Following this, a long-term forest monitoring initiative from 1990 to 2019 was carried out focusing on degradation, deforestation, and recovery of tropical forests over three decades (Vancutsem et al., 2021). Satellite observation has been successfully used to detect near-real time newly damaged tropical forest areas and served a wide range of government and private users, contributing to sustainable land management and taking action against illegal forest practices

(Reiche et al., 2021). Some global remote sensing initiatives and products concentrated on tropical forests monitoring are described below.

2.2.1 Global initiatives

2.2.1.1 Tropical Ecosystem Environment Observation by Satellites (TREES)

In 1990, TREES program has been initiated jointly by the Commission of the European Communities and European Space Agency to monitor pan-tropical deforestation and to develop new satellite remote sensing techniques. TREES program first experimentally used European Remote Sensing (ERS)- 1 and ERS-2 satellites with a Synthetic Aperture Radar (SAR) and found SAR tool useful for cloudy tropical forests (ESA & JRC, 1991). TREES program provided tropical forest map with 1.1km spatial resolution for the first time and made geo-referenced information available at a regional level. However, without proper accuracy assessment, this coarse-resolution data is not trustworthy (Achard et al., 2001)

2.2.1.2 Global Rain Forest Mapping (GRFM)

In 1995, the GRFM project was launched with the objective of mapping tropical rainforests at a spatial resolution of 100m. This project was led by National Space Development Agency of Japan (NASDA) in collaboration with National Aeronautics and Space Administration (NASA), Joint Research Centre of the European Commission (JRC) and Japanese Ministry of International Trade and Industry (MITI) (Rosenqvist et al., 2000). The project used L-band SAR image and made the products freely available for scientific learning purposes (Rosenqvist et al., 2004). While this project opens a new window for vegetation mapping and inundated area mapping over the pan-tropical region, it faced several technical challenges (De Grandi et al., 2000).

2.2.1.3 Global Forest Watch (GFW)

GFW program was initiated by World Resource Institute (WRI) in 1997 as a part of Forest Frontiers Initiative to update forest states across the world. GFW collects data from MODIS, LandSat, GlobCover and crowdsourcing initiatives in collaboration with Google, ESRI and other local government, and non-governmental sectors. Since 2014, it provides open-access online data (available at www.globalforestwatch.org) for global forest monitoring to help its user to measure forest loss, to preserve natural resources, to stop illegal deforestation and fires. GFW uses cutting-edge technology to generate near real-time alerts visualizing vulnerable regions to recent tree cover

loss. GFW is enriched with several datasets regarding forest change, land cover, land-use, climate and biodiversity (GFW, 2022).

In addition to these global initiatives, long-term forest monitoring products based on remote sensing are developed for specific purposes and specific areas of interest. Some of the existing optical and SAR-based forest monitoring products are described below.

2.2.2 Operational forest monitoring systems

2.2.2.1 Global Forest Change (GFC)

GFC map was introduced by University of Maryland in partnership with GFW program to map spatially and temporally detailed observation of worldwide forest change. It provides a percent tree cover map using Landsat time series analysis at a 30 m* 30m grid (Hansen et al., 2013). Annual loss and gain are calculated from tree cover map as a binary layer. Forest loss refers to the complete or comprehensive removal of tree cover, where trees are defined as vegetation above 5m in height. Forest gain is defined as complete or comprehensive recovery, conversion of non-forest to forest state. The first version of GFC dataset was released in 2013 which provided a high-resolution map of tropical forest degradation from 2000 to 2012. Upon release, GFC estimated 2.3 million km^2 global forest loss over the period, 0.8 million km^2 gain on the contrary. Unlike other climate regions, statistically significant forest loss was observed only in tropical forest domain, 2101 km^2 forest loss annually (Hansen et al., 2013). Over the period GFC used different optical sensors ranging from MSI Landsat 5 TM, Landsat 7 ETM+, Landsat 8 OLI sensors in different versions. The first version of GFC reported 83.1% producer accuracy with an overall accuracy of 99.5%.

From version 1 to version 2 GFC updates and incorporated reprocessing of data, improved training methods, validation methods, quality assessment and algorithm adjustment. Version 1.9 provides the dataset from 2011 to 2021 which used Landsat 8 from 2013 to 2021 for improved forest detection. However, there are some known inconsistencies described in the usage notes. It used different sensor technology in different periods; therefore some recent years have better signals compared to other years. A reduced collection has been mentioned during the decommission of Landsat 5 and launch of Landsat 8 in 2012. Algorithm adjustment is also responsible for year-to-year inconsistencies in the dataset.

Guild et al. (2022) used the product 'treecover2000′ and 'lossYear' from version 1.9 and used them with Planet/NICFI and Forest Canopy Disturbance Monitoring Tool (FCDM), found that forest loss conforms well with drought and fire monitoring. However, after a pixel is classified as lost, this dataset does not take into account the possibility of regrowth (Guild et al., 2022). Another study compared GFC dataset with nationally derived in-situ data from Guyana and found 99.34% of overall accuracy and 94.35% of producer's accuracy from 2001 to 2017. However yearly comparison resulted in overestimation and underestimation in several years. The study concluded GFC dataset can be used after calibration, but for detecting small-scale changes, solely this dataset should not be used without reference data (Galiatsatos et al., 2020). Another suitability study of GFC data also reported similarly. Although the overall accuracy was found to be more than 96%, the classification error was reported as non-negligible (Sannier et al., 2016).

2.2.2.2 Tropical Moist Forest (TMF)

Tropical Moist Forest (TMF) product is developed by European Commission's Joint Research Centre specifically to monitor tropical forests (Vancutsem et al., 2021). Landsat 40-year time series observation is used in TMF products with 30m spatial resolution. TMF provides detailed forest disturbance information by specifying forest loss as deforestation and degradation. Deforestation is defined as a pixel transformation from undisturbed TMF to another land cover where degradation refers to temporarily disturbed forest observed within 2.5 years (Vancutsem et al., 2021). TMF data also provides forest regrowth information by using two complementary layers: transition map and annual change from 1990 to 2021. The transition map provides 70 transition classes which can be further reclassified according to their description. One big advantage of TMF is deforestation and degradation can be separated by observing and comparing long-term and short-term degradation trends. The first version of TMF was available from 1990 to 2019 and estimated a 17% disappearance of tropical moist forests since 1990 (Vancutsem et al., 2021). In the latest version of 2021, updated commodity mask and smallholder plantation dataset was incorporated by visual delineation of missing plantation areas. In addition, differentiation between deforestation and degradation is well established by classifying direct deforestation, deforestation after degradation and multiple degradation.

The overall accuracy of TMF products is reported 91.4% with 9.4% omissions and 8.1% false detection by Vancutsem et al. (2021). Later, another study used TMF products for mapping and

monitoring cocoa deforestation from 2000 to 2019 found 62.2% precision and 82.9% recall (Renier et al., 2023)

2.2.2.3 Radar for Detecting Deforestation (RADD)

Wageningen University, with the collaboration of GFW, ESA, UMD and Deltares developed RADD alert system (WUR, 2020). It is based on Sentinel-1 C band SAR data which provides gap-free observation during persistent cloud cover. RADD was the first publicly accessible 10m resolution monitoring system that uses radar sensors to provide near real-time forest disturbance monitoring (Reiche et al., 2021). In tropical forests it has the real advantage of rapid detection of small-scale forest disturbance with a revisiting period of 6 to 12 days. Currently, this alert covers 13 south American countries, 6 central American countries, 25 African and 6 countries in Asia pacific, within the latitude range 30°N to 20°S. A probability-based disturbance detection algorithm is used, where disturbance is triggered from the latest SAR image and confirmed as high (> 97.5%) or low confidence (> 85%) within 90 days following subsequent observation. With a minimum mapping unit of 0.1 ha, forest disturbance is defined as the complete or partial removal of tree cover within a 10m pixel level by RADD alert. However, it did not differentiate between anthropogenic forest disturbance and natural disturbance. For small-scale disturbance, RADD performs well but long patches of deforestation may take a longer time to reach higher probability to be confirmed as a disturbance when influenced by wet soil or woody debris. Short wavelength-based C-band radar is highly sensitive to moisture, therefore resulting in chances of false detection in swamp or tidal forests. In addition, usage of global forest baseline may raise inconsistencies in the dataset when used at the local scale. RADD alerts are freely available at GEE and GFW platforms, alerts contain the date of deforestation upon availability of sentinel 1 image. The first version is released in 2019, achieved user and producer's accuracy greater than 95% for the area bigger than 0.2 ha (Reiche et al., 2021). The latest version in 2021 reduced the minimum mapping unit to 0.1 ha with an improvement of area-based forest masks and improved swamp forest detection. A feasibility study was carried out to compare RADD detection with Planet mosaics in 50 logging sites at Congo basin and found that RADD accurately estimated the logging concession. However, this study concluded that the area with few disturbances has a relatively lower reliability and aggregation level should be carefully chosen (Welsink et al., 2023).

GFC, TMF and RADD provide global scale and pan-tropical scale datasets are designed for large area coverage. In every region these datasets do not provide similar accuracy. Therefore, solely depending on those datasets for a specific region of interest could lead to some inaccuracies. In addition, these datasets are useful for detecting long-term trend change, using them for specific year or season could create underestimation or overestimation. Therefore, there is a need to use further satellite observation to validate those datasets for smaller regions of interest for short-term periods.

2.3 Satellite data from Copernicus program

Copernicus is the earth observation program managed and coordinated by European commission in partnership with European Space Agency (ESA), European Organization for the Exploitation of Meteorological Satellites (EUMETSAT), European Center for Medium-Range Weather Forecasts (ECMWF), EU agencies and Mercator Ocean. Copernicus programs use both satellite and ground stations' data to be used in diversified applications: environmental conservation, urban sprawl, agriculture, transport planning and tourism as the dataset is freely available for all users. Several dedicated Sentinel satellites and contributing missions are the main source of Copernicus dataset (ESA, 2022). For our specific use of interest, Sentinel-1 and sentinel-2 datasets are described in the following sections.

2.3.1 Sentinel-1 (S1)

Sentinel-1 (S1) mission consists of two polar-orbiting satellites, S1A and S1B launched in 2014 and 2016 respectively, mission ended for S1B in 2022. This mission facilitates freely available high-resolution radar data for the first time. These two satellites are based on C-band SAR which offer continuous measurement under the presence of persistent cloud coverage at a spatial resolution of 10 m.

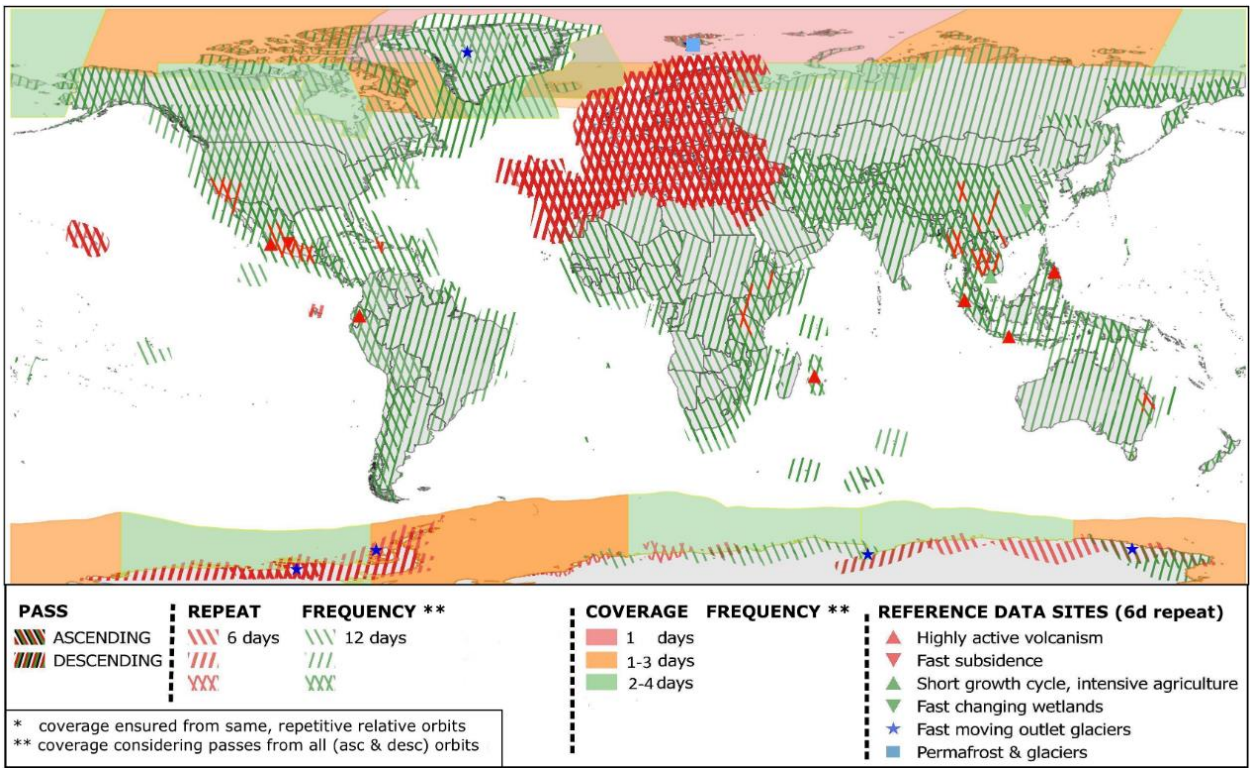


Figure 1: Sentinel-1 observation scenario (ESA,2023)

Using the two satellites, most of the places are covered in 6 days. However, permafrost and glacier regions along with Europe and Canada have a coverage frequency of 1-2 days. The majority of tropical regions are continuously covered every 12 days repeating frequency. Democratic Republic of Congo region is covered every 12 days, but only by descending orbit (Figure 1).

S1 collects data in 4 operational modes namely Strip-Map (SM), Interferometric Wide Swath (IW), Extra-Wide Swath and Wave-Mode. Different operational modes have specific swath, spatial resolution and polarization modes which made them useful for specific regions of the globe. IW mode is the main mode for terrestrial data acquisition and forest remote sensing with a swath of 250km, a spatial resolution of 5m by 20m, with polarization mode HH+HV, VV+VH, HH and VV. Besides, S1 products offer three levels of data collection from raw data to preprocessed and directly usable data. Level 0 is the compressed dataset containing all the information collected, where Level 1 data are georeferenced and Ground Range Detected (GRD), on the other hand, level 2 data is highly pre-processed and delivered as environmental variables. Level 1 data is widely used for advanced remote sensing applications. ESA provides guidelines, toolbox, algorithms and tutorials for better management of interferometry, polarimetry and interference of S1 data (ESA, 2022).

Recent studies have examined the effectiveness of Sentinel-1 in distinguishing between forest and non-forest areas. Bouvet et al. (2018) conducted a study in Peruvian Amazon that reported 95% successful detection of deforestation based on ascending and descending SAR observation. Another study conducted by Hoekman et al. (2020) in six study sites to examine the ability of C-band SAR data to separate the forest from the non-forest area found that time series of SAR data increases the accuracy compared to a single scene. The classification accuracy was found between 81.6% to 88.6% for tropical rainforests and further increased by using only co-polarized (VV) and cross-polarized (VH) backscattering annual mean and standard deviation. Apart from VV and VH, Radar Vegetation Index (RVI) is found less sensitive to changing environmental conditions, therefore can be useful for vegetation monitoring (Sahadevan et al., 2013).

2.3.2 Sentinel-2 (S2)

Sentinel-2 mission, developed by ESA as a part of Copernicus program, carries two sun-synchronous satellites S2A launched in 2015 and later S2B was launched in 2017 to complement the first satellite. It provides optical imagery collected from a 290 km wide swath, high-resolution, multispectral sensor. It provides 13 multispectral bands with a resolution ranging from 10m to 60 m and has a revisiting frequency of 5 days while using two satellites together.

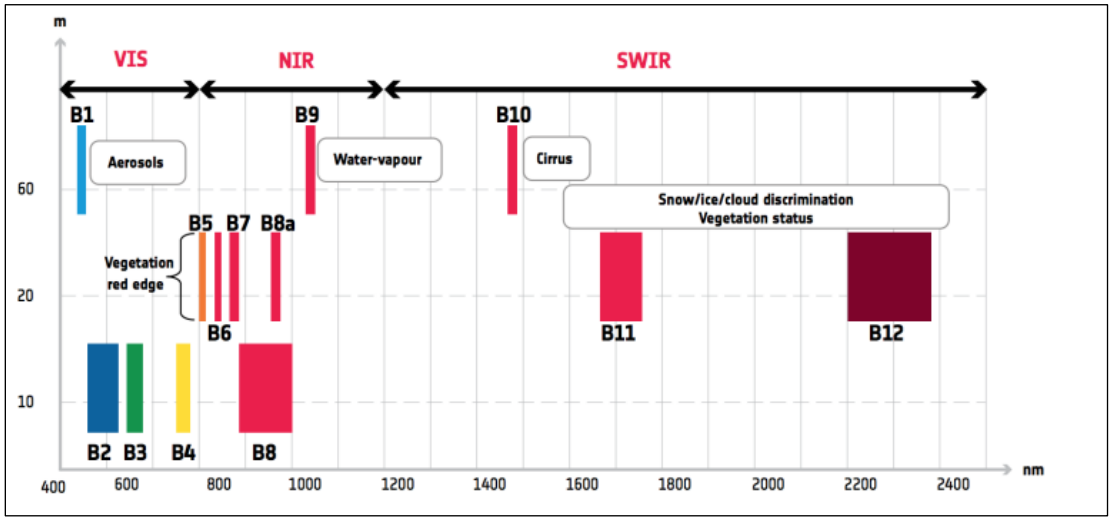


Figure 2: Spectral resolution and band wavelength of Sentinel-2 bands (ESA, 2023)

Three visible bands B2, B3, B4 and one near-infrared (NIR) B8 band are available at 10 m resolution supporting the need for high-resolution earth observation. Three vegetation red edge bands B5, B6, B7 and one narrow NIR band B8a are available at 20m resolution. These bands are very useful for detecting vegetation characteristics and plant biophysical parameters. In addition,

S2 provides two Short Wave Infrared (SWIR) bands at 20-meter-resolution and one SWIR at 60-meter-resolution which are useful for atmospheric corrections and cloud masking.

Immitzer et al. (2016) demonstrated the capabilities of pre-operational S2 data for crop type and tree species mapping. With an overall classification accuracy ranging from 65% to 76% for six crops and seven tree species categories, this study confirmed the higher value of red-edge and SWIR bands for vegetation monitoring.

Puletti et al. (2018) also tested the potentiality of S2 data for forest-type mapping in Mediterranean environments. This study calculated four multispectral indices from available bands with 10m to 20m resolution and mapped different forest categories and four forest types with an accuracy of 83%. The study concluded that using multitemporal images in different phenological periods provided better detection compared to a single observation. The improvement of statistical probability of classification has been observed in other studies by using multispectral indices (Mahdianpari et al., 2019; Tavares et al., 2019).

Phiri et al. (2020) reviewed S2 data for land cover mapping and explored the performance of S2 in vegetation mapping, urban area mapping, agriculture and natural hazard monitoring and mentioned the limited applicability of S2 for continuous monitoring in the persistent cloud-prone zone. Similarly, several studies mentioned the influence of cloud cover affecting the performance of disturbance monitoring in tropical forests either by limiting the observation or preventing spectral analysis by signal perturbation (N. Chen et al., 2021; Coluzzi et al., 2018; Nazarova et al., 2020). On the other hand, S1-based SAR data can be used irrespective of cloud coverage, and promising to detect small-scale change in tropical forests (Reiche et al., 2021). Therefore Sentinel-1 and Sentine-2 can complement each other for continuous forest monitoring in tropical regions.

2.4 Combination of S1 and S2

Previous studies have demonstrated better results when combining two different satellite products for land monitoring (Anchang et al., 2020; N. Chen et al., 2021; Heckel et al., 2020; Hu et al., 2020; J. Li & Roy, 2017; Shimizu et al., 2019). Tavares et al. (2019) focused on integration of S1 and S2 by image stacking of two S1 images and one S2 scene of close date. Beside combining specific bands, they also combined the textural indices of S1 and spectral indices of S2. Overall

91.07% accuracy was found by combining S1 and S2 which was higher than using the individual sensors.

Hirschmugl et al. (2018) also combined optical (S2) and SAR (S1) time series data for forest and LULC mapping in Malawi by applying both data-based integration and result-based integration from 15 S2 scenes and 14 S1 GRD images. The data-based combination merged all the preprocessed bands of S2 and statistics stack of S1 before undergoing the classification steps, whereas the result-based combination used the Bayesian probabilistic principle and reliability measures. Overall a 5% improvement is reported when combining both S1 and S2 for LULC mapping and a 1.5% improvement is observed for forest mapping. They found data-based integration slightly outperforms the result-based integration approaches. De Luca et al. (2022) used three optical and two SAR-based indices for land cover mapping in Mediterranean region following the time series analysis. This study also combined the coherence of SAR interferometry and optical biophysical parameters for establishing a correlation with canopy structure. The result improved by 2.53% when optical and SAR indices are combined rather than using them individually. A combination of S1 SAR and S2 MSI data were used to map mangroves in China, found S2 more effectively measure the mangrove, however fusion of both produced the best overall accuracy. This study combined 7 spectral indices from S2, two S1 bands and 3 texture indices from Grey-Level Co-Occurrence Matrix (GLCM) (Hu et al., 2020). Anchang et al. (2020) mapped woody canopies in tropical savannas from 16 satellite-derived mastics, where the monthly median composite of VV and VH bands from S1 is combined with the monthly greenest pixel composite derived from S2 for 2-year time span. This study concluded that the fusion of optical reflectance and radar backscatter data can be used for effectively mapping tree and shrub canopies. Several studies confirmed that, when two different sensors are used, the maximum revisiting duration increased (J. Li & Roy, 2017; Shimizu et al., 2019). Dense observation frequency helps to understand ground scenarios more precisely because in tropical forests, disturbed forests sometimes experience vegetation regrowth within a short duration, which would otherwise not be captured (Shimizu et al., 2019).

Dostálová et al., (2016) used S1A SAR data for forest mask delineation with a spatial resolution of 10 m and 500m2 minimum mapping unit. Overall accuracy was obtained at 92% with a kappa of 0.81. Another study aimed to provide a detailed forest cover map using S2 data and aerial

images. Their result showed 95.21% overall accuracy when compared with the national forest inventory (Ganz et al., 2020). Combination of S1 and S2 is also used for binary mangrove forest mapping. Multispectral vegetation indices from S2 and Radar Vegetation index (RVI) are combinedly used in a random forest classification and observed acceptable performance using the combination of S1 and S2 images (Sharifi et al., 2022). These studies provide motivation to combine S1 and S2 data for forest cover classification in tropical forest regions.

2.5 Classification algorithms

Machine learning classification algorithm has been used in the field of remote sensing based on supervised and unsupervised algorithms. Unsupervised algorithms are self-taught that use clustering for grouping spectral classes, whereas supervised algorithms train the classifier from a labeled training dataset (Rodriguez-Galiano et al., 2012). Random Forest (RF) is one of the most widely used supervised learning algorithms that build on the concept of decision trees. Decision trees create strategies to make a decision based on randomly selected features and observe the features. Then all decision trees are used to predict the class values (Breiman, 2001). Several studies have determined the effectiveness of RF classifier for forest non-forest classification and found it efficient (Bolyn et al., 2018; C. Li et al., 2014; Purwanto et al., 2023). In terms of binary forest mapping, the RF classifier has to be trained from field-collected samples or highly accurate secondary data sources of forest and non-forest. After the model is trained, it applies to the wider area to make the decision based on the trained dataset and assess its accuracy. Depending on the provided dataset, RF classifier can learn from optical images, radar images and vegetation indices or textural features. However, the algorithm is heavily dependent on the input data, therefore cloud cover and other atmospheric effects may result in wrong classification classes (Ienco et al., 2018).

These atmospheric problems do not impact the SAR dataset as the signal can penetrate through the cloud. However, SAR backscattering is also influenced by shadowing effects, saturation effects and higher noise during the ongoing rainfall during data collection. Therefore, using S1 and S2 observation can be complementary in terms of forest extent classification of tropical forests.

3. Objectives

This master thesis aims to evaluate the effectiveness of optical and SAR-based satellite sensors for mapping tropical moist forest extent in the Congo basin. The primary objective is to develop forest extent map considering an adaptive method of combining optical and SAR sensors regarding the landscape types for the study area (Tshopo province, Democratic Republic of Congo).

To achieve the mentioned objective, two specific objectives are taken into consideration. The first specific objective is to assess the performance of the existing operational products in order to determine the most effective deforestation product in terms of accuracy and area coverage in the study area. The study will compare and analyze the outputs of existing deforestation detection systems: Global Forest Watch and Tropical Moist Forest (both based on optical time series), and SAR-based RADD alert for different forest landscape patterns.

The second specific objective is to develop a forest extent classification map from high-resolution satellite data building on the comparison at landscape level. The ability of SAR images to monitor persistent cloud-covered areas could be complimentary with Sentinel-2-based optical data. For this, Sentinel-1, Sentinel-2 and their combination will be tested for their applicability to different landscapes.

Finally, the developed classification maps will undergo a validation protocol to evaluate their reliability and limitations.

4. Materials and Methods

This chapter outlines the methodology employed in this study. Section 4.1 describes the materials used in two subsections: Study area and Data source. Subsequently, Section 4.2 describes the methodology for comparing existing optical and SAR products and conducting forest and non-forest classification within the context of this study.

4.1 Materials

4.1.1 Study area

The study site is located in Tshopo province in the Democratic Republic of Congo (DRC) covering an area of 12372 km^2. The study area covers the full administrative boundary of Kisangani and partially covers Ubundu, Banalia and Bafwasende cities. The region has been an important commercial hub for DRC since the 19th century when it was the commercial capital. In terms of tropical woodland coverage, Kisangani is the largest city in DRC.

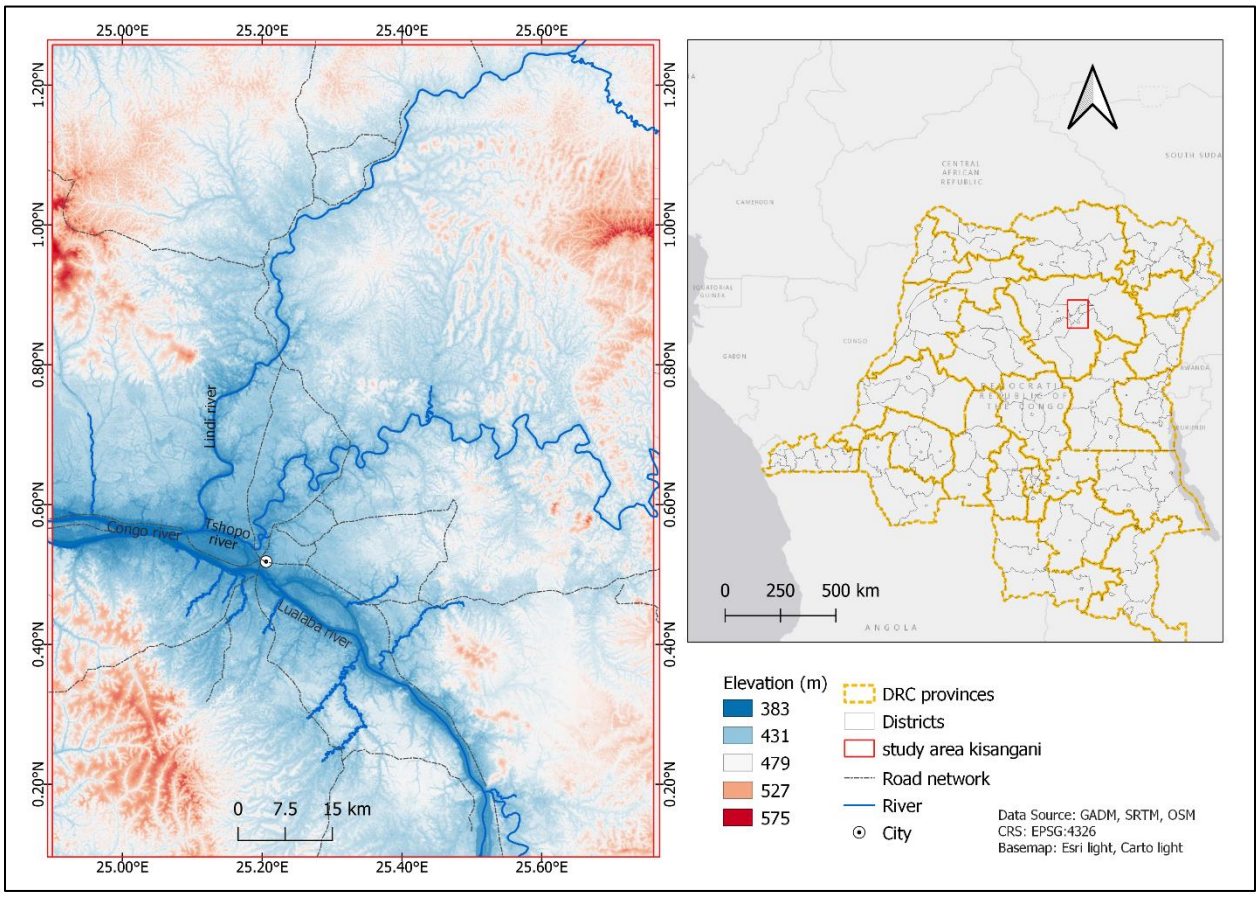


Figure 3: Study area map, Tshopo province, Democratic Republic of Congo (Background: SRTM 30m DEM)

At the southern end, the Congo river crosses the study area from south to west. Lindi river is one of the affluents of the Congo river, flows from the north-eastern part of the study area and empties into the Congo river near Kisangani. Another small river Tshopo flows from the central-eastern part of the area and falls into the Lindi river just before the confluence with the Congo river. Apart from those, there are several small canals and creeks, distributed throughout the study area, ensuring navigation and water demands of the ecosystem. Tshopo province is located mostly in flat relief, with altitude ranging from 200 to 600 m. Our study area falls within the elevation range of 383m to 575m (Figure 3). Average annual rainfall ranges from 1600 mm to 2000 mm, with a dry season from April to October and a rainy season from November to March (World Bank, 2021). The area falls under a humid equatorial climate zone and is dominated by forest area (approximately 90% of its total land area). The main forest types in this area include moist evergreen and moist semi-deciduous forests along with swamp and riparian forests (Dalimier et al., 2022). Agriculture is the main source of income for Tshopo, mainly focusing on cassava, rice, cowpea, coffee, cocoa, oil palm, rubber etc. (South Pole, 2020).

Five homogenous landscapes are drawn inside the study area to compare the performance of optical and SAR products at a landscape level. However, in this study area, it is difficult to find homogenous landscapes, therefore landscapes are demarcated based on their dominant land cover. Sentinel-1 and Sentinel-2 tiles are also taken into consideration during the landscape delineation, each subset falls under single S1 or S2 tiles. Figure 4 represents our delineated study subsets used for comparing the existing products.

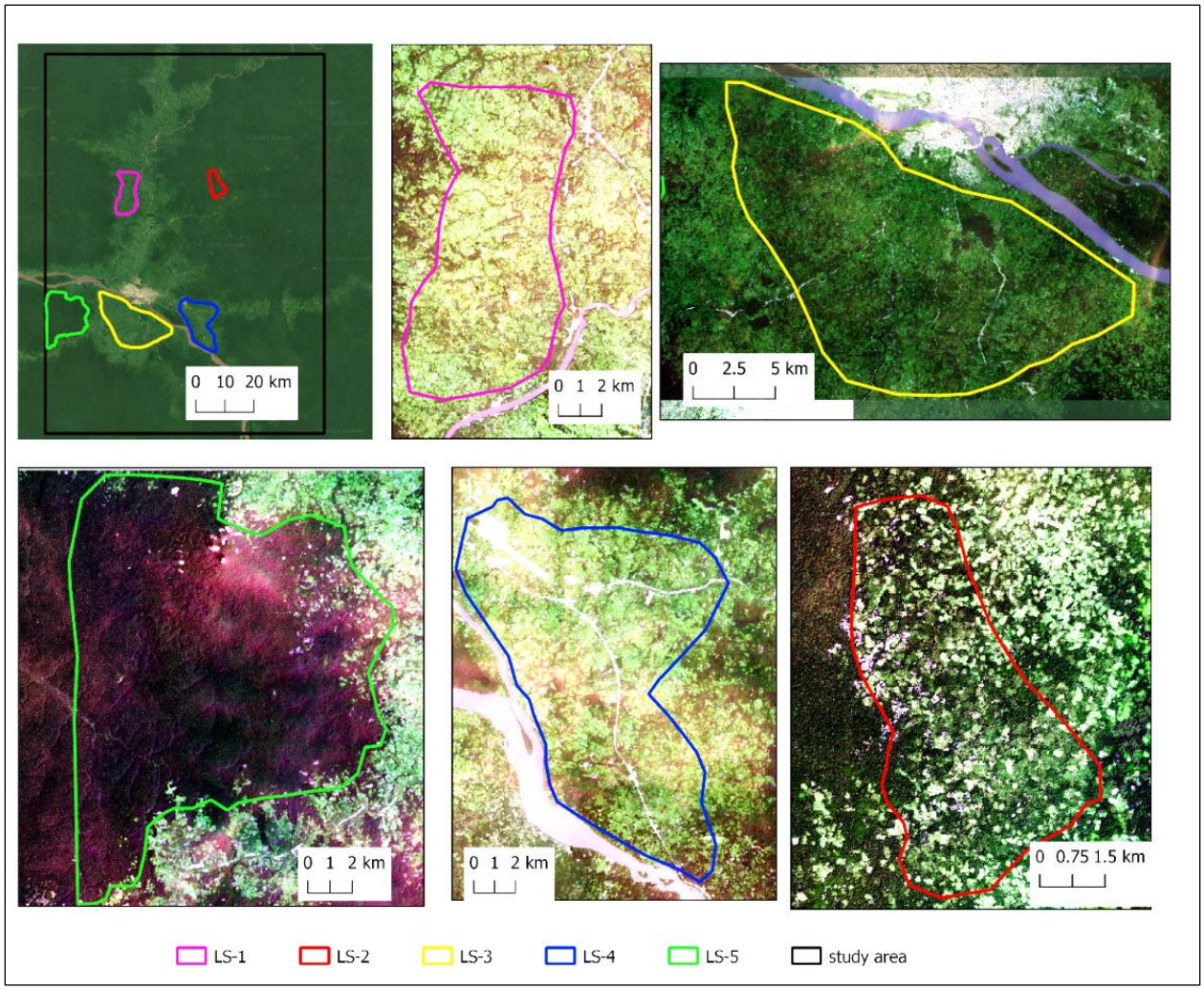


Figure 4: Study area and landscape subsets (Background: Planet monthly mosaic December 2020)

LS-1: Landscape dominated by agriculture and sparse vegetation: LS-1 is located in the north-western part of Kisangani. Located on the western bank of the Lindi river, this area covers mainly the region Badile, Bengamisa, and Boliambe. According to GFW, from 2001 to 2021, it has lost around 71.3 Km^2 of tree cover, equivalent to a 76% decrease since 2000 tree cover. Observing it from Google Earth and Planet images, LS-1 has a mixture of permanent agriculture, shifting cultivation, and sparse vegetation and includes some settlements of locals. However, it is dominated by agriculture and sparse vegetation covering an estimated area of 84 km^2.

LS-2: Landscape dominated by degraded forest: LS-2 is located near Badumbi, central Kisangani, Tshopo province. This subset represents the degraded forest adjacent to the intact forest. According to GFW, from 2001 to 2021 this place lost 13 km^2 of forest cover, most of which

took place between 2015 to 2020. In 2020, LS-2 experienced the highest rate of tree cover loss. Among the study subsets, LS-2 is the smallest, covering an area of 30 km^2.

LS-3: Landscape dominated by agriculture and palm-tree plantation: LS-3 is located in the southern part of the Kisangani municipal area. It has the biggest size of study subsets, around 256 km^2. Located on the southern bank of the Lualaba river, this landscape covers some parts of the Babama, Batiagai, Banangau and Batiakaba regions. According to GFW, this area is a part of Central Congolian lowland forests which has lost 154 km^2 of forest cover from 2001 to 2021. Land cover of this area includes palm tree plantation, agriculture, heterogeneous vegetation cover, and built area. However, among others, this landscape is dominated by Palm tree plantations and other agriculture.

LS-4: Landscape dominated by mixed landcover: LS-4 is located on the south-eastern part of Kisangani having a total area of 137 km^2. The area covers Kisangani airport and includes the surroundings of Madula, Batiapanga and Bakapoke regions. It is dominated by mixed land covers such as settlement, built areas, agriculture, palm tree plantation and sparse vegetation. According to the GFW, this landscape falls under Northeast Congolian lowland forests and Eastern Congolian swamp forests. Between 2001 to 2021, it has lost 96.4 km^2 of forest cover, (72% decrease in tree cover since 2000).

LS-5: Landscape dominated by intact forest: LS-5 is in the south-western part of Kisangani which covers Yalokombe and Yaune regions. This landscape is dominated by intact forest with an area 189 km^2. However, forest disturbance has been noticed near the eastern buffer area of this area. According to GFW, a total of 17.8 km^2 of tree cover loss has been observed from 2001 to 2021, most of which has taken place after 2013.

4.1.2 Data sources

Two types of datasets were used in this study. Existing optical- and SAR- based deforestation products were used as a reference dataset which was used for training samples after proper accuracy assessments. Three types of satellite imagery were used for the classification of forest and non-forest areas and validation purposes.

4.1.2.1 Reference datasets

In this study, three different reference datasets were utilized. To capture deforestation patterns through optical imagery, two time-series products based on Landsat data, namely: GFC and TMF, were employed. Additionally, a deforestation alert system derived from the SAR dataset (RADD alert) was employed. The comparison and analysis of these three reference datasets were performed at the landscape level aiming to identify their most suitable combination for creating effective training samples.

4.1.2.1.1 GFC

Global Forest Change products (Hansen et al., 2013) are freely available in the GFW platform, GLAD API and GEE collection. The latest version V1.9 was used in this study that covers from 2000 to 2021. Global forest extent and change are divided into 10*10-degree tiles and calculated from the Landsat time series. GFC dataset provides several layers namely, Tree cover 2000, Global Forest cover gain, Year of gross forest cover loss, Data mask, and Landsat cloud-free composite for year 2000 and 2021. For our analysis, we used the Gross forest loss event (lossyear) band which is available at 30m*30m resolution. This band measures the yearly forest loss from 2000 to 2021 as a stand-replacement disturbance or a shift from a forest to a non-forest state.

4.1.2.1.2 TMF

Tropical Moist Forest (TMF) product (Vancutsem et al., 2021) is freely available on EU science hub for download and also available as a web explorer, GEE and WMS layer. The dataset is only available for tropical areas and distributed in 10*10 degree tiles. Several bands are available for downloads: Undisturbed and degraded TMF, Transition map with sub classes, Transition map with main classes, Annual change collection (1990-2021), Degradation year and Deforestation year. First all bands are downloaded and clipped to our study area. For our analysis, we used the product ‘annual change collection’ to separate degradation and deforestation years. The degradation year band represents the year when forest cover experienced degradation for the first time following the transition classes, whereas the deforestation year denotes the yearly forest cover transition from forest to non-forest classes.

4.1.2.1.3 RADD

RADD alerts (Reiche et al., 2021) calculate forest disturbance for each acquisition of C-band SAR image of Sentinel-1. Wageningen University and research provide a GEE script to download

RADD alert data and export it to GeoTiff file. The latest RADD alerts since 2019 were downloaded for this analysis. Both high confidence alerts (forest disturbance probability > 97.5%) and low confidence alerts (probability > 85%) are selected to detect forest disturbance in 2020. After exporting the product, all the alerts denoting Day of the Year (DoY) for the year 2020 were selected.

4.1.2.2 Satellite imagery

Three types of satellite products are used in this study. Sentinel-2 and Sentinel-1 C band SAR datasets were used for classification while high-resolution Planet monthly mosaics were used for validation purposes.

4.1.2.2.1 Sentinel-1 (S1)

For the S1 metrics calculation, the full available time series of 2020 was used. S1 mission provides C-band Synthetic Aperture Radar (SAR) data at 5.405 GHz. The collection contains Ground Range Detected (GRD) scenes of 10m resolution. Only descending orbits were available in our study area and Interferometric Wide Swath (IW) mode was chosen with a swath of 250 km. As band combination, single co-polarization bands of Vertical transmit/ Vertical receive (VV) and Vertical transmit/Horizontal receive (VH) were available. S1 GRD collection in GEE is already preprocessed using the Sentinel-1 toolbox (B. Chen et al., 2017; Hu et al., 2020). Main preprocessing steps are the application of orbit file, border and thermal noise removal, radiometric calibration, terrain correction using SRTM 30m DEM, and conversion of backscatter coefficient into dB. Overall, 32 S1 scenes are found within our study area for the year 2020.

4.1.2.2.2 Sentinel-2 (S2)

Sentinel-2 L2A images were downloaded from Copernicus sci-hub (scihub.copernicus.eu/). The whole study area is covered by two S2 tiles: T35NKA and T35NLA. To calculate the annual change of 2020, two cloud-free images were chosen at the start of 2020 and 2021. The first cloud-free image was found on 14th January 2020 for tiles T35NKA and 29th January 2020 for T35NLA. However, at the end of 2020, no optical cloud-free images were available. Therefore, 2nd February 2021 was used as a closer date for both tiles. Each image contains 13 spectral bands with three different resolutions from 10m, 20m to 60m. Later, these bands were resampled and clipped for different indices calculations.

4.1.2.2.3 Planet

High-resolution Planet images are distributed through NICFI data program with the help of KSAT and Airbus. This made available the monthly mosaic of Planet base maps at the tropical areas at a spatial resolution of 4.77m. From December 2015 to August 2020, only bi-annual mosaics were available, after September 2020 onward monthly mosaics were available in the collection. For this study two bi-annual mosaics were used: one from 12/2019 to 05/2020, and another from 06/2020 to 08/2020 in addition to a single monthly mosaic from September 2020 to January 2021. Planet data were used to validate our comparison of optical and SAR-based deforestation products. Also, the Planet dataset was combinedly used with Google earth images to delineate the landscape subsets inside our study area.

4.2 Methods

To determine the best forest extent map at the landscape level, first existing optical and SAR forest disturbance products are compared for different landscapes. Then a classification was carried out which combines optical and SAR images based on the analysis at the landscape level. Figure 5 presents the overall approach of this thesis implemented for 5 different landscapes.

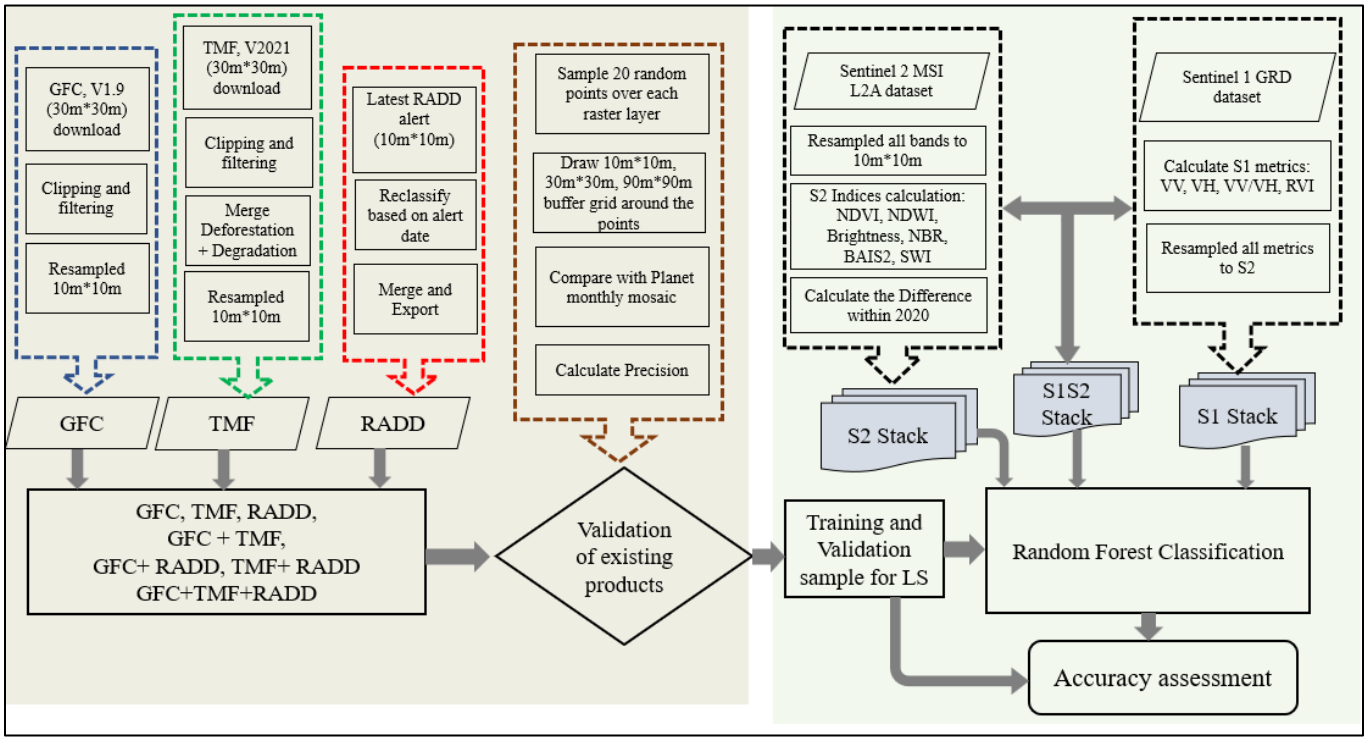


Figure 5: Schematic workflow of forest extent classification (on left: comparison of existing products, on right: forest-extent classification, the dotted line represents processing steps)

4.2.1 Comparison among existing optical and SAR-based deforestation products

To determine the effectiveness of optical and SAR-based deforestation products, a comparison was carried out from the optical-based existing products: GFC, TMF and SAR-based product: RADD alert. The main steps are described below.

4.2.1.1 Pre-processing of existing products: As we intend to compare three different products, pre-processing is done to maintain the same spatial and temporal configuration of the three products. The GFC dataset was available from 2000 to 2021, whereas TMF products are available from 1990 to 2021. On the other hand, RADD alert products are more recent, available from 2019. Therefore, the year 2020 has been chosen for the year of study as it is the middle year between 2019 and 2021 when all three products were available. Another reason to choose 2020 as a year of

study was, this year has been reported as the third most challenging year of tropical deforestation (Ngeunga, 2021). The year 2021 is also used for the temporal comparisons of optical products with SAR products of 2020. For the TMF dataset, both degradation year and deforestation year were merged to determine all the forest disturbances recorded by TMF. Both GFC and TMF are resampled to 10m*10m resolution using nearest neighbor methods to compare with RADD products. For the RADD products, Day of the Year (DoY) was reclassified and aggregated year 2020 before the comparison.

4.2.1.2 Creation of intersecting layers for each landscape

After all the products are pre-processed into the same spatial and temporal configurations, three Geotiff layers are created and labeled as 'GFC', 'TMF', and 'RADD' accordingly. From these three layers, 4 intersection layers are created by using the raster calculator. Intersection is labeled using '+' sign for simplicity. Therefore, we have four intersecting layers namely 'GFC+TMF', 'GFC+RADD', 'TMF+RADD' and finally 'GFC+TMF+RADD'. Then the area corresponding to each of the seven layers was calculated. In the following, the layers are displayed in an overlay where the layer with larger areas is positioned in the background and the layer with smaller areas is positioned in the foreground. From the intersection layers, the agreement and disagreement between the optical and SAR-based sensors are visually compared. Then, the intersecting layers were clipped for each landscape and underwent a validation process.

As GFC and TMF-based optical products are based on cloud-free observation, persistent cloud cover in our study area resulted in the delayed observation of the optical products. Therefore, to take into account the temporal discrepancies, the intersection of 'GFC +TMF' layer of 2021 with RADD 2020 product was also considered.

4.2.1.3 Validation of intersecting layers

To complement the visual observation, a quantitative accuracy assessment was further carried out by comparing those products with Planet monthly mosaic images. The validation was done for every seven layers and in each landscape. For doing so, 20 points are randomly sampled over each layer. For each landscape, there are 20*7=140 points and for all five landscapes, in total there are 140*5=700 points that are chosen for validation. As our products are based on 10m and 30m spatial resolution, validations of pixels corresponding to only points could lead to inaccuracies. Therefore,

to avoid such geometric mismatch, three square grids, sizes of 10m*10m, 30m*30m, 90m*90m were drawn around the random points serving as validation samples (Figure 6)

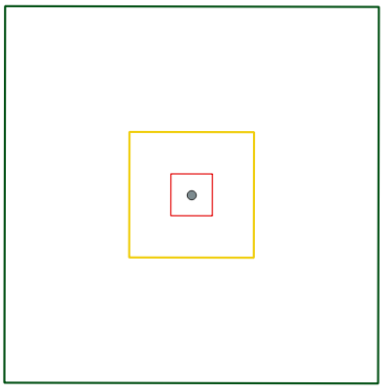


Figure 6: Response design for validation of optical and SAR-based deforestation products, all random points are surrounded by 10m*10m grid (red), 30m*30m grid (yellow) and 90m*90m grid (green)

As the product was filtered to 2020, high-resolution monthly Planet mosaic (less than 5m*5m) for the month January and December 2020 were used for validation. In the absence of images from those starting and ending months, the nearest available Planet images within a range of 2 months were used. By comparing the first and last Planet monthly mosaic, all the points and associated grids are labeled as the followings, 1: confirmed deforestation (Forest in starting month, non-forest in ending month), 2: not deforestation, (Forest in starting month, forest in ending month), 3: old deforestation (non-forest in both starting and ending month), 4: Ambiguous (uncertain visual observation).

To validate the points, we observed the pixel of the Planet mosaic of that specific point and assigned a label accordingly. For the square grid level, the majority of pixels within the grid were considered to determine the label of the square. For example, if most of the pixel within 90m*90m square falls under deforestation, only then it was labeled as confirmed deforestation. Then the Precision/User Accuracy (UA) of confirmed deforestation in each landscape was calculated as an accuracy indicator which represents the fraction of correctly classified reference deforestation points concerning all points of forest disturbance. The validation results of existing products were used to identify the best complementary layer in terms of accuracy and area coverage of deforestation for each landscape. Later, this layer was used to take the training points for the classification.

4.2.2 Forest extent classification

4.2.2.1 Training dataset

In this study, training samples are collected based on the analysis carried out from the best combination layer of GFC, TMF, RADD deforestation products that explained the forest disturbance for each landscape. Non-forest training points were also randomly sampled over the best-combined layers. For example, if GFC+TMF explains the best confident deforested areas in a landscape from the accuracy assessment, this raster layer was used to generate non-forest training points. For the intact forest raster, GFC, TMF, and RADD-based deforestation products were merged, then 'other land covers' from TMF 'annual change' dataset were added. After that, the intact forest layer was generated by subtracting the added layer from the landscape boundary layer. Then forest training points are randomly sampled from this layer. We respectively sampled 400 and 200 training points from the forest and non-forest classes in each landscape (Figure 7.A).

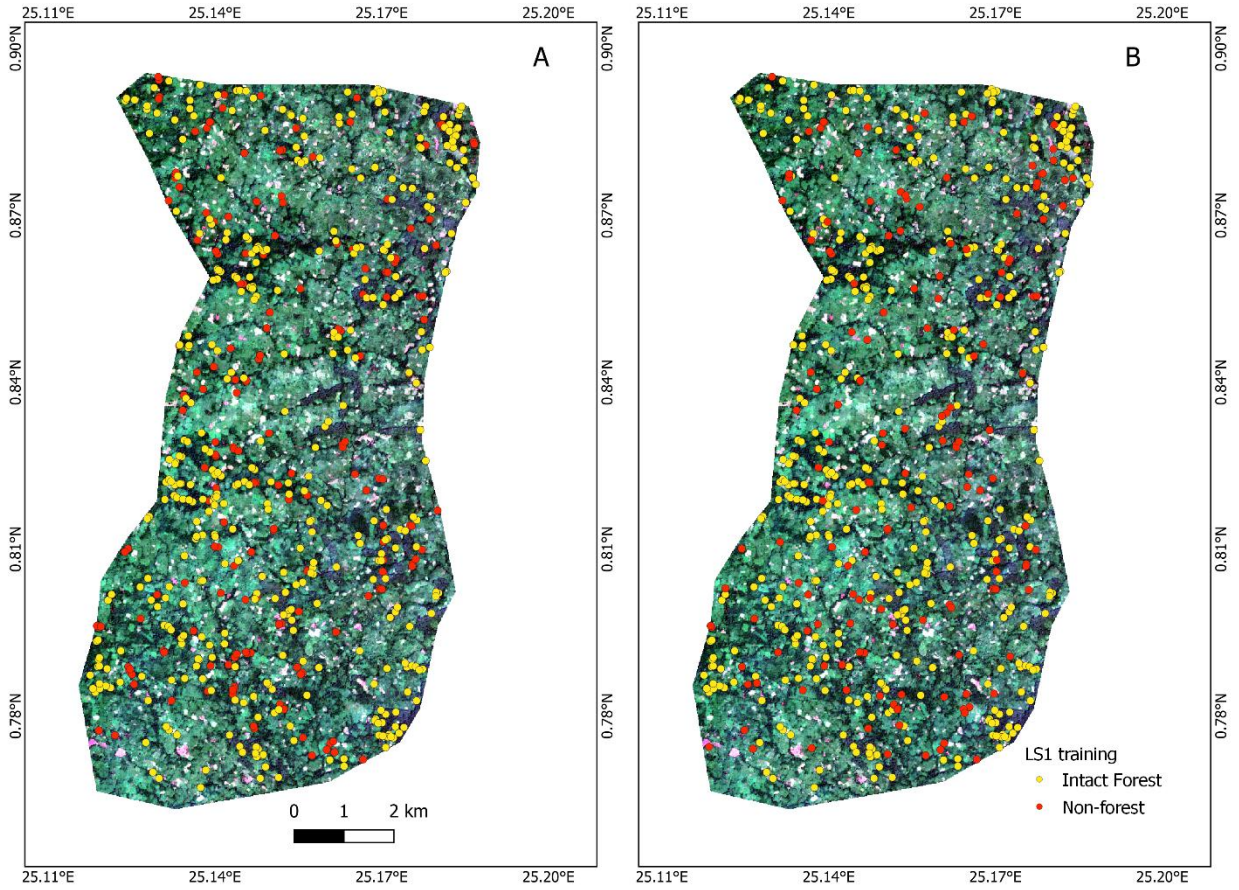


Figure 7: Example of training samples for LS-1 (A. training samples based on two land cover classes, B. training samples derived from three land cover classes)

The training samples were evenly distributed throughout the study landscapes (Figure 7, See more in Appendix 9.1). In the landscape dominated by mixed land cover, an additional 200 points were sampled from other land cover classes. Then deforestation (non-forest training samples) and other land cover training samples were merged into a single class and subsequently 200 samples were randomly taken from this merged layer as a non-forest training sample (Figure 7.B). Finally, classification was performed based on both two-class-based training samples and three-class-based training samples in those landscapes.

4.2.2.2 S1 metrics

The backscattering coefficient of S1 SAR data may have speckle and seasonal variation throughout the year, potentially producing inaccuracies in a yearly composite. To address this, trimestral median composites were first computed from both VV and VH polarization, then four composites were produced. Later standard deviation was calculated for each polarization as a representative of the yearly composite. To comply with vegetation indices from S2, Radar Vegetation Index (RVI) was used. As this study used duel polarized SAR data, therefore modified RVI index (Equation 1) from quad-polarized SAR to dual-polarized proposed by (Nasirzadehdizaji et al., 2019) was used.

$$RVI \approx 4*VH/(VH+VV) \quad \ldots\ldots(1)$$

RVI was calculated for both VV and VH backscatter values of each image, then yearly standard deviation of RVI was calculated. Veloso et al., (2017) found that ratio of both bands can be complementary when combined with VV and VH backscatter. Therefore, we also computed yearly standard deviation of VV and VH ratio from trimestral median composites. In total, four SAR metrics-VV, VH, VV/VH and RVI were used from the S1 images (Appendix 9.2).

4.2.2.3 S2 indices calculation

Puletti et al., (2018) observed that multispectral indices can significantly discriminate forest areas from non-forest, compared to S2 bands. Similarly, Kobayashi et al., (2020) also found higher overall accuracy while using spectral indices for classification. Therefore, in this study, multispectral indices from S2 were used instead of using S2 multispectral bands. Six multispectral indices were computed from S2 cloud-free observation, as shown in Table 1. The calculation process involved computing all the indices at the first and last date in each landscape. Then the differences between the first and last dates were calculated. As a result, for each landscape, there

is a raster of the six indices' difference within one year which was derived from 12 multispectral indices (Appendix 9.3).

Table 1: Indices calculated from Sentinel-2 images

Indices	Formula	Reference
Normalized Difference Vegetation Index (NDVI)	$\frac{(\mathrm{NIR} - \mathrm{RED})}{(\mathrm{NIR} + \mathrm{RED})} = \frac{(\mathrm{B08} - \mathrm{B04})}{(\mathrm{B08} + \mathrm{B04})}$	(Tucker, 1979)
Normalized Difference Water Index (NDWI)	$\frac{(\mathrm{NIR} - \mathrm{SWIR})}{(\mathrm{NIR} + \mathrm{SWIR})} = \frac{(\mathrm{B08} - \mathrm{B11})}{(\mathrm{B08} + \mathrm{B11})}$	(Gao, 1996)
BRIGHTNESS	$\sqrt{(Green^2 + Red^2 + NIR^2 + SWIR^2)}$ $= \sqrt{(\mathrm{B03}^2 + \mathrm{B04}^2 + \mathrm{B08}^2 + \mathrm{B11}^2)}$	(Marzialetti et al., 2020)
Normalized Burn Ratio (NBR)	$\frac{(\mathrm{NIR} - \mathrm{SWIR})}{(\mathrm{NIR} + \mathrm{SWIR})} = \frac{(\mathrm{B08} - \mathrm{B12})}{(\mathrm{B08} + \mathrm{B12})}$	(Key & Benson, 2006)
Burned Area Index for Sentinel-2 (BAIS2)	$\left(1 - \sqrt{\frac{\mathrm{B06} * \mathrm{B07} * \mathrm{B8A}}{\mathrm{B4}}}\right) *$ $\left(\frac{\mathrm{B12} - \mathrm{B8A}}{\sqrt{\mathrm{B12} + \mathrm{B8A}}} + 1\right)$	(Filipponi, 2018)
Sentinel-2 Water Index (SWI)	$\frac{(\mathrm{NIR} - \mathrm{SWIR})}{(\mathrm{NIR} + \mathrm{SWIR})} = \frac{(\mathrm{B05} - \mathrm{B11})}{(\mathrm{B05} + \mathrm{B11})}$	(Jiang et al., 2020)

NDVI measures the differences in reflectance between NIR and red bands. A higher NDVI value indicates a higher presence of healthy vegetation, making it a valuable index for forest monitoring. Although NDWI provides information about the presence of water body, lower value represents vegetation and the differences in vegetation cover can be identified. Brightness index highlights the area with higher reflectance across all spectral bands, therefore, can be useful for determining forest, and non-forest areas. NBR is useful to identify the burn areas, a higher value indicates healthy vegetation whereas a lower value indicates bare ground and burnt areas. Similar to NBR, BAIS2 also provides information on burn areas, also useful for describing vegetation properties. SWI focuses on the areas with higher moisture contents therefore distinct between forest and non-forest areas.

4.2.2.4 Band composite

In this study, we aimed to integrate both S1 and S2 data at the landscape level to improve the accuracy of forest extent classification. To achieve this, data-based integration approaches were used, where both S1 and S2 features are merged in a band stack (S1+S2). First, we calculated six multispectral indices from S2, and then S1 metrics are resampled to the S2 indices before the integration. Table 2 represents the band combination that was used for classification steps. According to the result obtained from the comparison of optical and SAR-based products, we used different band combinations for different landscapes.

Table 2: Combination of S1 and S2 features used for the classification

Combination	Features
S1	VV, VH, VV/VH, RVI
S2	NDVI, NDWI, Brightness, NBR, BAIS2, SWI
S1 + S2	VV, VH, VV/VH, RVI, NDWI, NDWI, Brightness, NBR, BAIS2, SWI

Only S2 indices are used for LS-1, LS-3 and LS-4 where the RADD product was assessed with the lowest performance. On the other hand, for landscapes LS-2 and LS-5, we used all three combinations: S1, S2 and S1+S2.

4.2.2.5 Random Forest classification

The random forest classification was carried out in GEE environment. We trained five random forest models for each landscape. The training samples were randomly split into two collections, the training collection and validation collection, where approximately 80% of the samples were used for training and the rest 20% were used for validation. Initially, 100 decision trees were set for training the RF classifier, which was later increased to 200 trees for different landscape levels. Based on our requirements, the combined band stack was then classified using the trained models. The resulting classified images were visually examined and then underwent accuracy assessment using the validation collection.

4.2.2.6 Accuracy assessment

In order to assess the accuracy of the classification, a confusion matrix is generated in GEE using the validation samples. The accuracy matrix is used as a quality indicator of classification which compares the predicted values to the actual values. The matrix is produced with True Positive (TP),

True Negative (TN) and False Positive (FP), False Negative (FN) based on actual and predicted values. TP and TN indicate the correctly classified pixel whereas FP and FN refer to those pixels that are misclassified as another class. Later Producers' Accuracy (PA) and Users' Accuracy (UA) were calculated from the accuracy matrix.

Table 3: Accuracy assessment matrix

		Actual values	
		Forest	Non-Forest
Predicted values	Forest	**TP** (True positive)	**FP** (False positive)
	Non-Forest	**FN** (False negative)	**TN** (True negative)

PA is the ratio of the number of correctly classified reference pixels and the total number of reference pixels of that class while UA is the ratio of the correctly identified pixel of a class and the total number of predicted pixels in that class. Then the Overall Accuracy (OA) of classification is calculated from the sum of TP and TN divided by the total number of validation points. F-score is another performance indicator of classification calculated from UA and PA. Sometimes accuracy can be misleading when one class has a higher number of samples than others. In that case, although accuracy is higher, models may poorly represent the minority class. F-score combines UA and PA into a single metric thereby providing a balanced evaluation. Table 4 represents the performance indicator used in this study and their formula.

Table 4: Performance indicators used in this study (Fawcett, 2006)

Performance indicator	**Formula**
PA of the Forest class	TP / (TP + FN)
PA of the Non-Forest class	TN / (TN + FP)
UA of the Forest class	TP / (TP + FP)
UA of the Non-Forest class	TN / (TN + FN)
Overall accuracy (OA)	(TP + TN) / Total sample size
F-score of each class	2/{(1/UA)+(1/PA)}

In classification, two types of errors are considered. Omission error indicates the reference pixel that was omitted from the correct class (100% - PA) and commission error indicates the pixels that were wrongly classified (100% -UA).

5. Results

This chapter presents the results achieved in this study. Section 5.1 provides the comparison results of existing optical and SAR-based deforestation products at the landscape level. In section 5.2 the classification scenarios in different landscapes are presented. Following the result, the next chapter will provide a discussion, analyzing the implication and significance of the findings.

5.1 Comparison of Optical and SAR-based deforestation products

To compare the spatial agreement and disagreement, three products and their intersecting layers are compared. Overall, within the whole study area, TMF covers the biggest chunk of forest disturbances compared to all other layers (Table 5, see in Appendix 9.4). This is because both degradation year and deforestation year are merged to include all the disturbances TMF detected in 2020. GFC products cover 28% of the total deforested area recorded by all layers. RADD only covers 9.65% of the total area. Only 4.47% of the total deforested area is found to be covered by the intersection of three layers. From visual interpretation, it has been observed that GFC and TMF cover the central deforested areas whereas RADD alert covers mostly the buffer region near the intact forest. As the existing products work differently in various landscapes (See in Appendix 9.5, 9.6), therefore, we have drawn five different landscapes to observe how three existing products detected deforestation under various landscape conditions. At the landscape level, it was found that SAR-based products has a better signal in forest-dominated landscapes (LS-2 and LS-5) while in landscape (LS-1, LS-3, LS-4) deforestation products based on optical sensors have better area coverage (Figure 8).

Table 5: Deforested area calculation of all three products and their intersecting layers for the whole study area

Layers	pixel count	Area (km^2)	% of total deforestation
GFC	1889736	188	28.24
TMF	2178765	216	32.56
RADD	645881	64.1	9.65
GFC+TMF	938146	93.2	14.02
GFC+RADD	363151	36.1	5.43
TMF+RADD	376591	37.4	5.63
GFC+TMF+RADD	298787	29.7	4.47
total	6691057	665	100

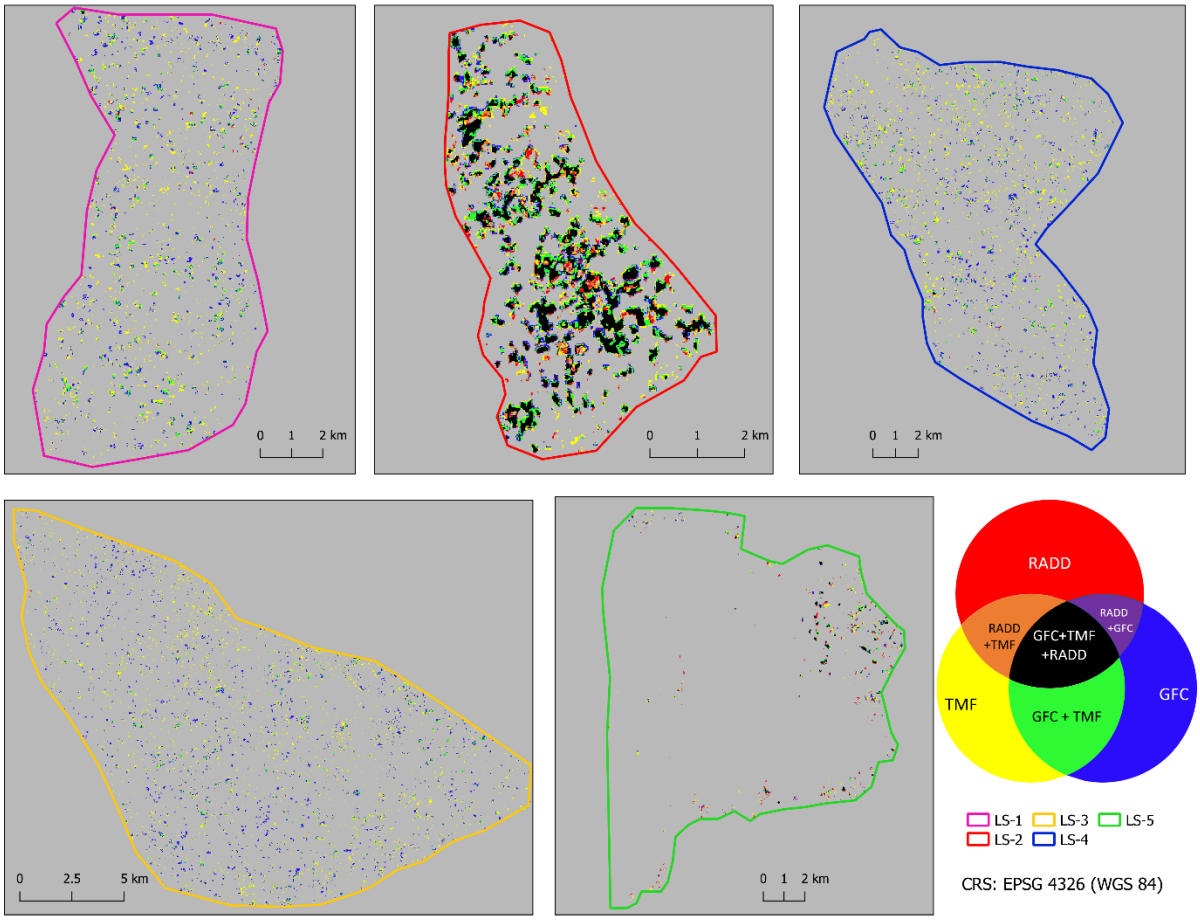


Figure 8: Agreement and disagreement between optical and SAR-based deforestation products at the landscape level

5.1.1 LS-1

In landscape-1, TMF (48%) and GFC (35%) layers are the most dominant layer observed in terms of deforested areas. Intersection of GFC and TMF cover 13.38% of total deforested area. RADD alert alone covers 1.47% of total forest disturbance (Figure 9 and Table 6) while all combined layers with RADD represent less than 1% of total disturbed area.

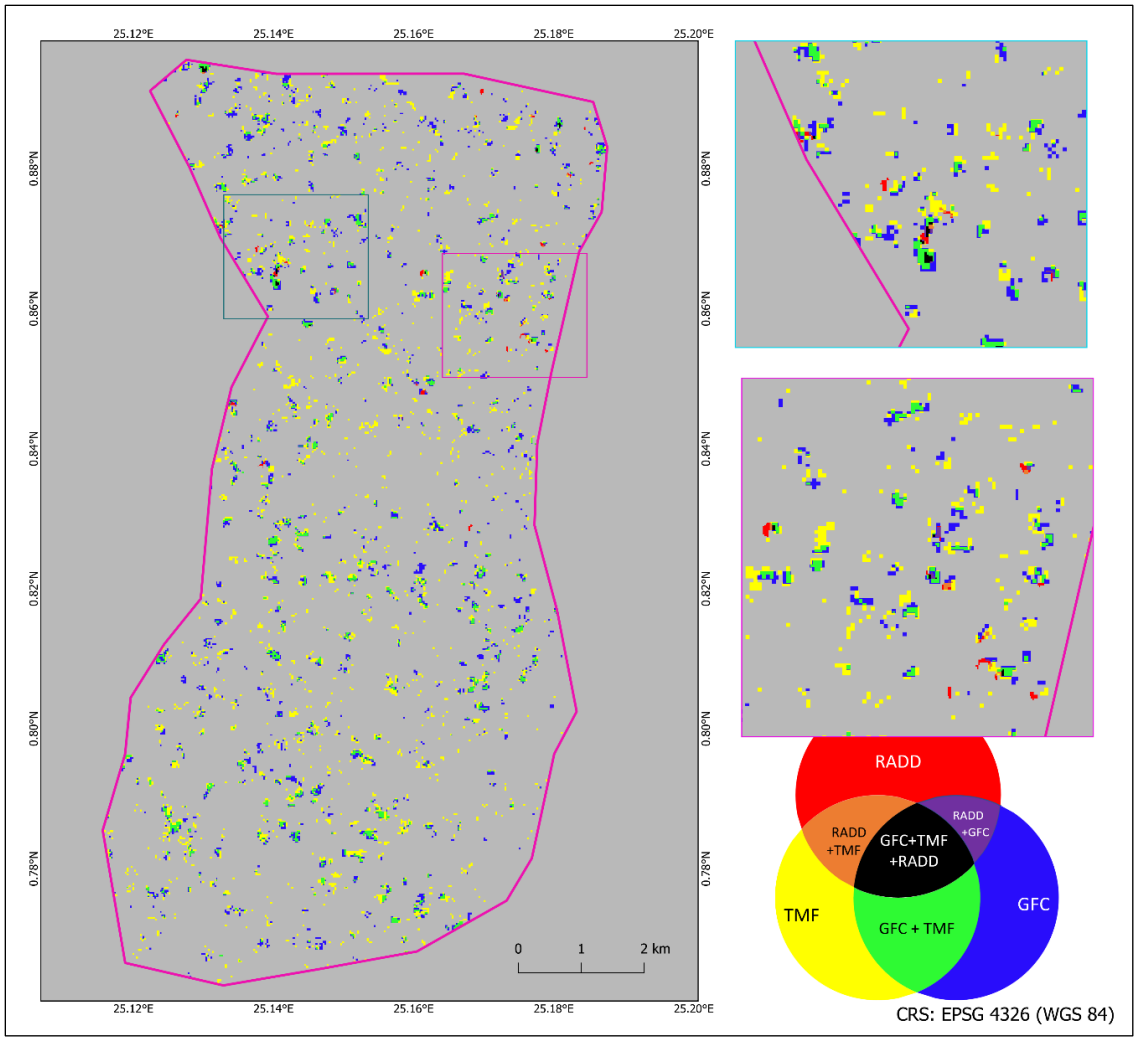


Figure 9:Agreement and disagreement between optical and SAR-based deforestation products at Landscape-1

Table 6: LS-1, deforested area as mapped by optical and SAR products

LS-1	Pixel count	% of total detected deforestation
TMF	36668	48.37
GFC	26855	35.43
GFC+TMF	10140	13.38
RADD	1111	1.47
GFC+RADD	395	0.52
TMF+RADD	395	0.52
GFC+TMF+RADD	238	0.31

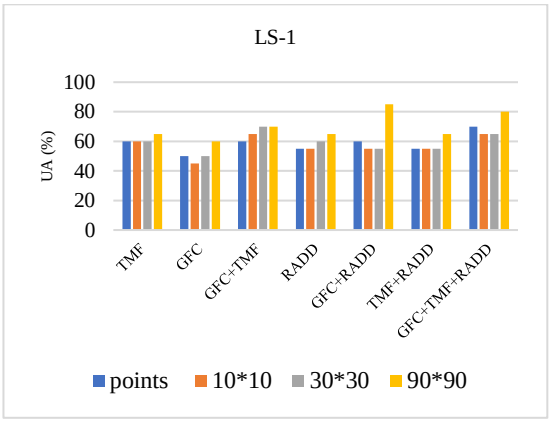


Figure 10: Users' accuracy of the intersection layers in LS 1

In terms of accuracy assessment (Figure 10), UA is increased from points to 90m*90m buffer reference area in every layer, except for GFC layer. For GFC layers, one of our validation points falls under the deforestation zone based on point pixel, however the maximum pixels under 10m*10m, 30m*30m buffer do not correspond to deforested pixels. In this landscape, lower UA is observed for GFC (50-60%) whereas for TMF, it is slightly higher (60-65%). When both products are combined their intersection has better UA ranges from 60%-75%. And when optical products are combined with SAR-based products, we achieved the highest UA range (70-80%). However, as RADD alert covers only a few places (Figure 9), the validation points are closely located and contributing to the highest accuracy. Therefore, RADD achieved higher UA compared to GFC and TMF, but only covers a few deforested events in this landscape. Therefore, the intersection of both optical sensors (GFC+TMF) was considered as having better reliability in terms of deforestation area coverage and users' accuracy in this landscape. Later this layer was used to collect non-forest training samples in this landscape.

5.1.2 LS-2

Landscape-2 is dominated by degraded forests where both optical and SAR-based deforestation products performed better (Figure 11) compared to other landscapes. According to Table 7, the highest convergence (9.62%) among the GFC, TMF and RADD products has been observed within this area in comparison to all other landscapes. TMF alone has the highest coverage of deforested area (20.36% of the total) followed by GFC alone (19.48%). As RADD alerts alone predicted better forest disturbance here (13.46%), the three products combinedly produced more than 9% of the total disturbed area, whereas for other landscapes it was smaller than 9%. In terms of validation (Figure 12), the highest users' accuracy (UA) (85%-95%) has been observed in the intersection of the three layers where the lowest was found for the GFC alone. UA increased from points to 90m*90m buffer zone, except for three intersection layers: GFC+TMF, TMF+RADD and GFC + RADD layer. In those layers, we observed 2 points, 1 point and 2 points respectively fell near the boundary of deforested area, therefore for the 90m*90m buffer area, the majority of pixels do not correspond to deforestation. GFC and TMF obtained between 60% and 80% UA individually but when combined with RADD alert, the accuracy increased from 80% to 90%. RADD alert itself has a UA of 75% to 80% individually, but its combination with optical-based product increases the accuracy. Therefore, in this landscape, the intersection of optical and RADAR products has

better area coverage and the highest users' accuracy, which was later used for training non-forest samples in this landscape.

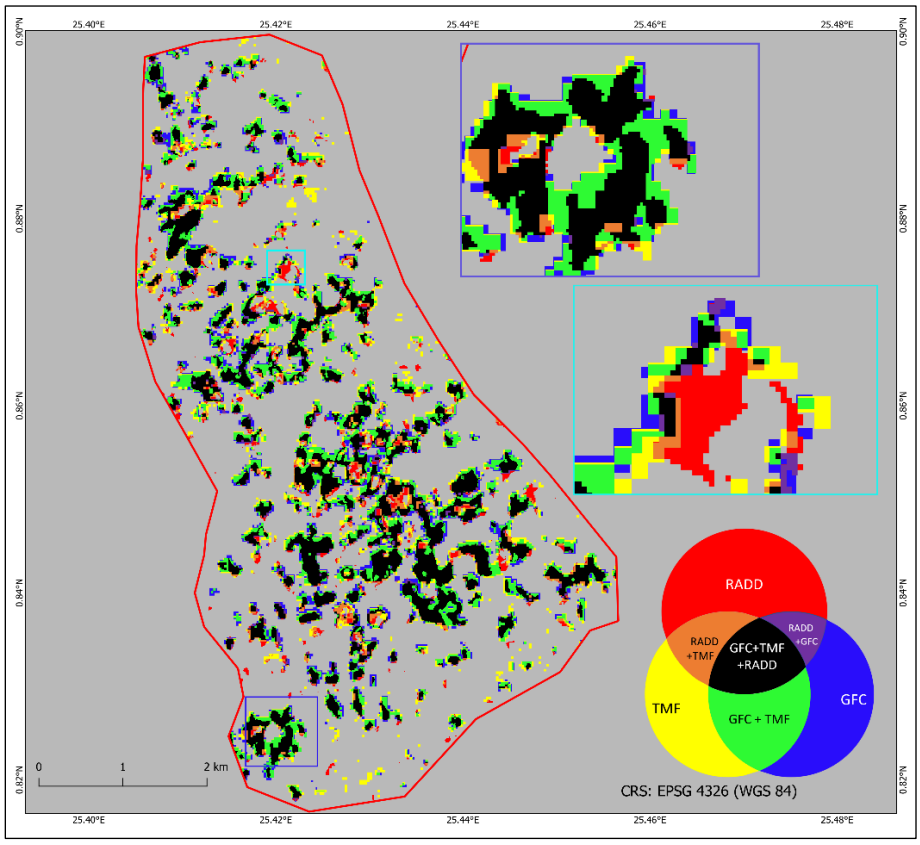


Figure 11: Agreement and disagreement between optical and SAR-based deforestation products at Landscape-2

Table 7: LS-2, deforested area as mapped by optical and SAR products

LS-2	Pixel count	% of total detected deforestation
TMF	76597	20.36
GFC	73276	19.48
GFC+TMF	58301	15.50
RADD	50621	13.46
TMF+RADD	42201	11.22
GFC+RADD	38954	10.36
GFC+TMF+RADD	36191	9.62

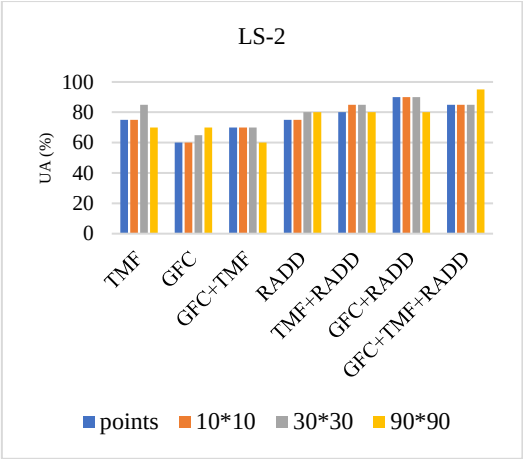


Figure 12: Users' accuracy of the intersection layers in Landscape-2

5.1.3 LS-3

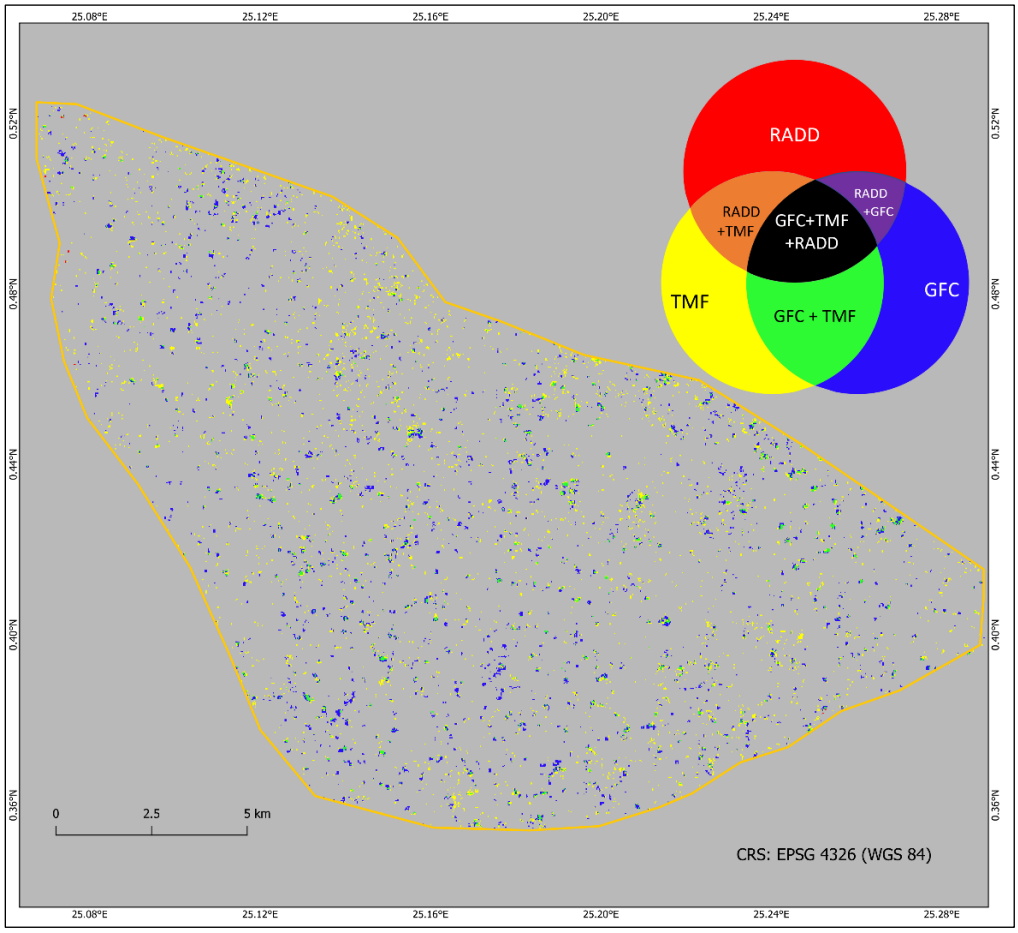


Figure 13: Agreement and disagreement between optical-based products and SAR-based products at Landscape 3

Table 8: LS-3, deforested area as mapped by optical and SAR products

LS-3	pixel count	% of total detected deforestation
TMF	71494	48.33
GFC	60721	41.05
GFC+TMF	15523	10.49
RADD	196	0.13
GFC+TMF+RADD	0	0
GFC+RADD	0	0
TMF+RADD	0	0

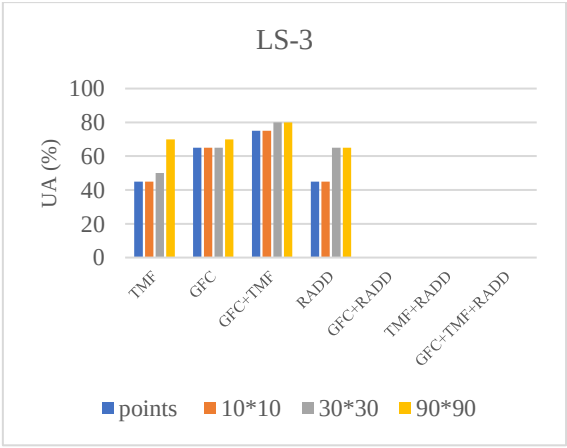


Figure 14: Users' accuracy of the intersection layers at Landscape-3

In Landscape-3 (Figure 13), TMF covers 48% of total disturbed areas calculated by all seven layers which is the highest among all layers (Table 8). GFC products also dominated this landscape (41% of the total). Around 10.5% of the total deforested area is covered by the intersection of GFC and TMF products. On the other hand, RADD only covers 196 pixels in this area which denotes the inability of RADD alert to detect deforestation in this region. As a result, no converging pixel was found when optical products are combined with RADD alert. According to Figure 14, the highest UA has been observed for GFC intersecting with TMF layer (75% to 80%), whereas the lowest was found for RADD (45% to 65%). Although TMF covers a bigger area than GFC, the accuracy is observed lower for TMF (GFC 65%-70% vs TMF 45%-70%). Therefore, GFC+TMF layer is complementary in terms of area coverage and accuracy. This layer is used to derive the training samples in this landscape.

5.1.4 LS-4

In Landscape-4, almost half of the total deforested area is covered by TMF products, whereas GFC covers 36% of the total disturbed pixel (Table 9). Their intersection covers 12.48%. In this mixed landscape, RADD alert only counted 542 pixels which are equivalent to 0.45% of the total deforested area counted by all the 7 layers. Regarding the validation (Figure 16), UA increases in every layer while moving from points to 90m*90m buffer region. RADD only covers 0.45% of total degraded area, which is located only in a few specific plots (Figure 15). Therefore, random validation points coincide with each other while using RADD. In this layer, 8 out of 20 points are located in the same area (the biggest chunk of disturbance recorded by RADD) (shown in Appendix 9.7) which increases the RADD's accuracy compared to GFC and TMF products. As those points are not spatially independent, a higher level of accuracy was found, from 95% to 100% while optical sensor-based products combined with S1-based RADD alert. Albeit their higher accuracy, the area coverage was below 1%. Therefore, GFC+TMF layer is considered here as the complementary layer regarding the deforested area coverage and accuracy and is later used for training and validation samples.

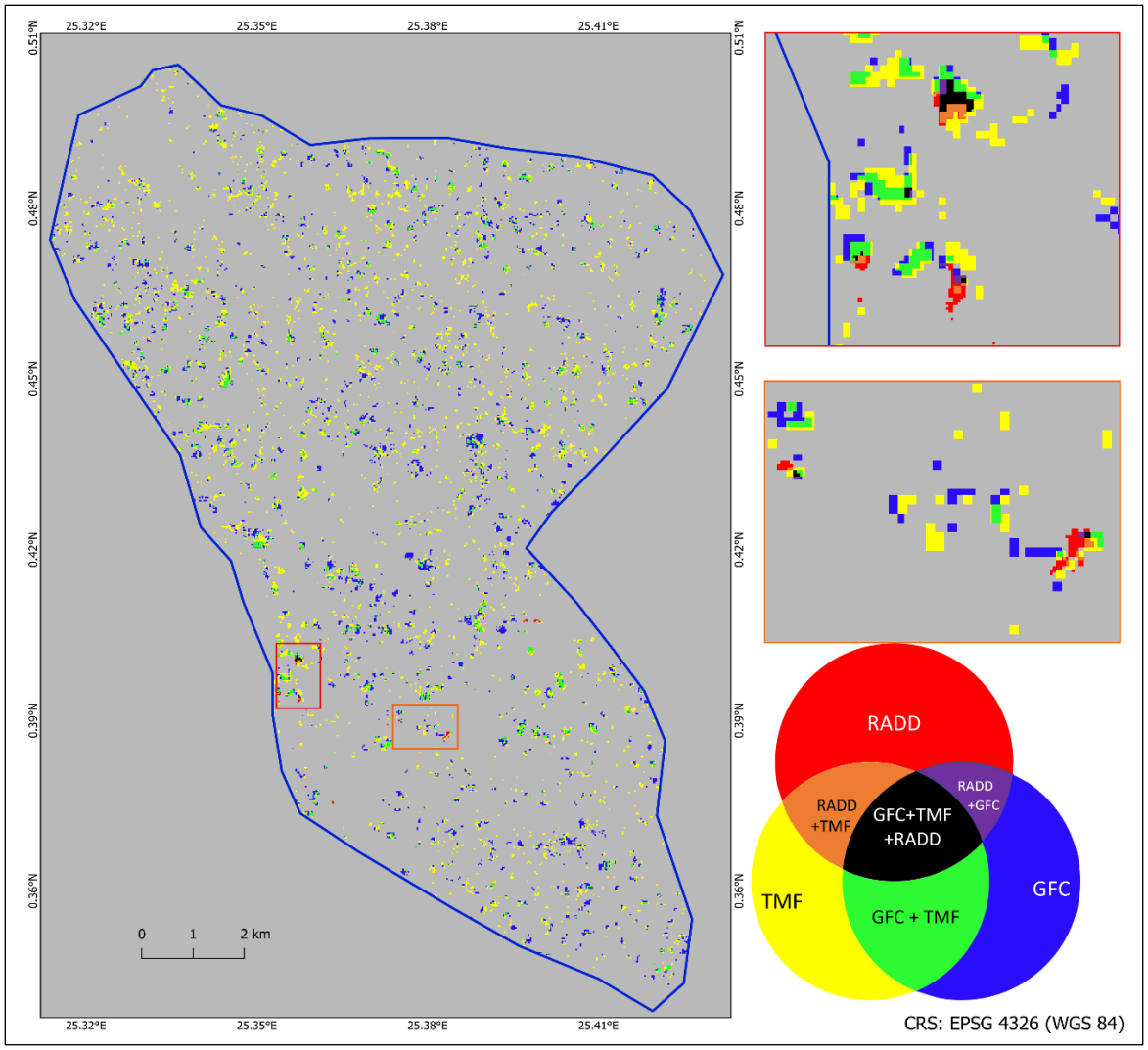


Figure 15:Agreement and disagreement between optical and SAR-based deforestation products at Landscape 4

Table 9: LS-4, deforested area as mapped by optical and SAR products

LS-4	Pixel count	% of total detected deforestation
TMF	60253	49.93
GFC	44307	36.72
GFC+TMF	15056	12.48
RADD	542	0.45
TMF+RADD	219	0.18
GFC+RADD	179	0.15
GFC+TMF+RADD	121	0.10

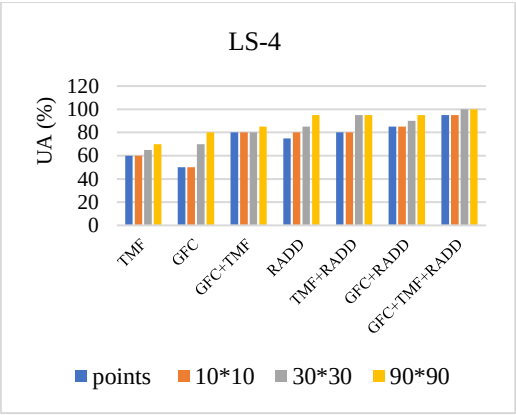


Figure 16: Users' accuracy of the intersection layers in Landscape-4

5.1.5 LS-5

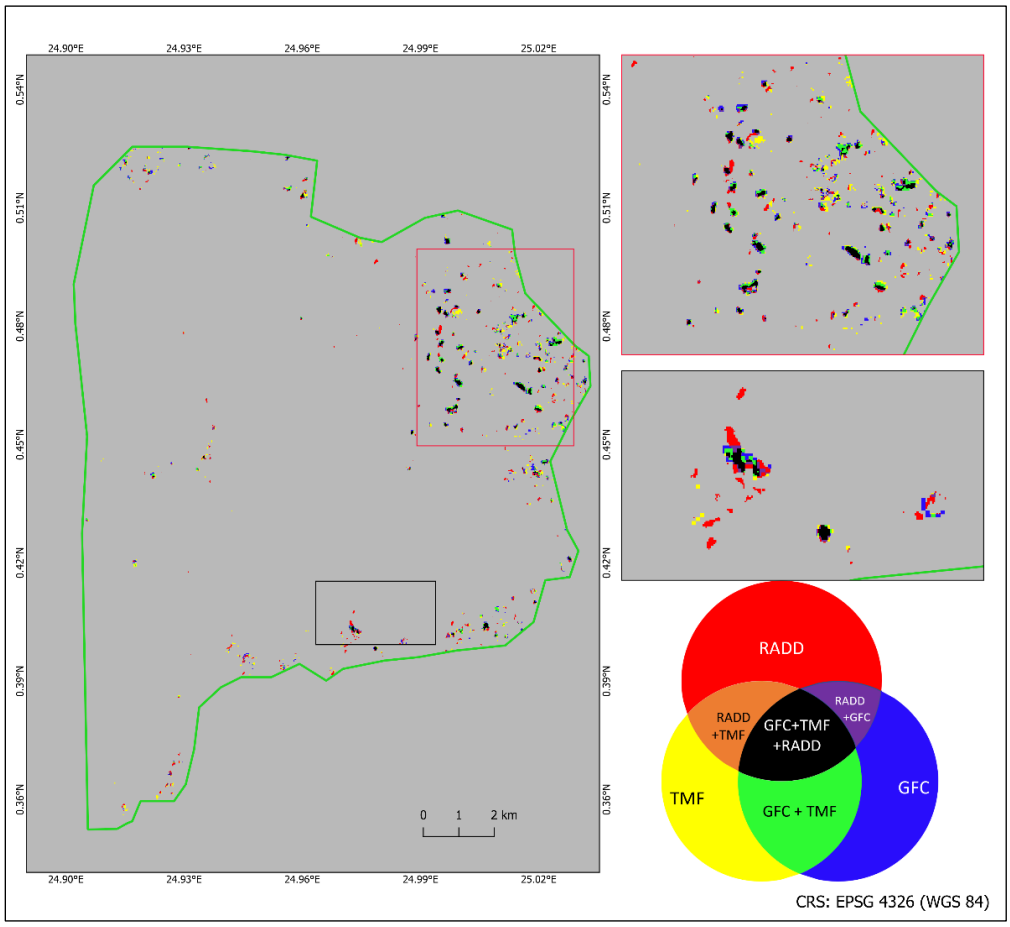


Figure 17: Agreement and disagreement between optical and SAR-based deforestation products at Landscape-5

Table 10: LS-5, deforested area as mapped by optical and SAR products

LS-5	Pixel count	% of total detected deforestation
TMF	16325	25.14
GFC	13319	20.51
RADD	12077	18.60
GFC+TMF	8197	12.62
TMF+RADD	5571	8.58
GFC+RADD	5253	8.09
GFC+TMF+RADD	4195	6.46

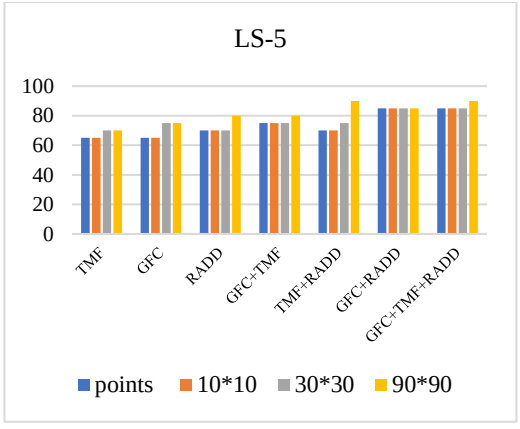


Figure 18: Users' accuracy of the intersection layers in Landscape-5

Landscape-5 is dominated by intact forests where recent forest disturbances have been noticed. Here, RADD alert covers deforestation almost a similar area as GFC (18.6% and 20.51%) (Table 10). TMF alone covers 25.14% of the total area whereas the intersection with GFC covers 12.62% of the total. Combined product with RADD covers more than 6% of total area. In terms of accuracy assessment (Figure 18), the UA of RADD alert was higher than that of GFC and TMF (70% to 80% vs 60% to 70%). Although TMF covers the larger percentage of deforested areas than GFC, the UA of TMF was found to be lower than that of GFC. Combined products with RADD obtained higher levels of UA rather than using a single product. In all layers, UA has increased from points to 90m*90m grid buffer zone. In this landscape, the intersection among the three layers obtained the highest level of users' accuracy (90%) and covered more than 6% of total deforested areas. Although the area coverage of this layer was lower, those pixels represent the confidence forest disturbance pixels where three of the existing products agreed. Therefore, in this landscape, the intersection of optical and SAR-based products can be used for taking training and validation samples.

5.1.6 Temporal comparison

As GFC and TMF products are based on the latest cloud-free observation, persistent cloud cover in our study area sometimes resulted in delayed observations. We found 20004 pixels where 'GFC +TMF' layer of 2021 intersects with RADD 2020 product. However, this represents only 0.4% of the total disturbed area where it was 4.47% for the spatial layer (Table 11). Therefore, these temporal discrepancies are hard to identify from visual interpretation (Appendix 9.9). To better understand the temporal pattern, we split the RADD layer monthly and use different combinations of RADD with GFC and TMF intersections. Instead of using RADD yearly alert from 2020, adding only the last two months November to December 2020 increased the convergence with GFC21+TMF21 layer up to 65% (Table 12). By increasing the duration, 92% of total (GFC 21 +TMF 21 + RADD 20) converging areas were found when considering RADD alerts from June to December 2020. However, the more we go back from November and December, the harder it gets to identify those pixels visually. Therefore, the best visual discrepancies were observed for the months November and December.

Table 11: Spatial and temporal agreement of optical and SAR-based deforestation products

	GFC 20, TMF 20, Radd 20			GFC 21, TMF 21, RADD 20		
Layers	pixel count	Area (Km^2)	% of total area	pixel count	Area (Km^2)	% of total area
GFC	1889736	188	28.24	1825813	181	36.25
TMF	2178765	216	32.56	1655244	164	32.87
RADD	645881	64.1	9.65	645881	64.1	12.83
GFC+TMF	938146	93.2	14.02	809903	80.4	16.08
GFC+RADD	363151	36.1	5.43	42085	4.18	0.83
TMF+RADD	376591	37.4	5.63	37068	3.68	0.74
GFC+TMF+RADD	298787	29.7	4.47	**20004**	1.99	0.4

Table 12: Using different combination of RADD 2020 DoY combined with GFC and TMF 2021

Temporal	pixel count	% of GFC21 TMF21 RADD 20
GFC 21+ TMF21 + RADD 20	20004	100
GFC21 +TMF 21 + RADD 20319-20355	13105	**65.51**
GFC21 +TMF 21 + RADD 20283-20355	14030	70.14
GFC21 +TMF 21 + RADD 20247-20355	15459	77.28
GFC21 +TMF 21 + RADD 20163-20355	18380	91.88

Figure 19 illustrates that RADD alerts counted those new deforested areas within 14th November 2020 to 20th December, 2020. However, due to the absence of optical cloud-free imagery, the optical products (GFC and TMF) incorporated those new deforestation areas that occurred in 2020 into their 2021 datasets. Since GFC and TMF provide only yearly products, it was not possible to identify the specific months in 2021 when the new deforestation of 2020 was included in the optical products. Therefore, by combining optical and SAR-data, we can detect the ground conditions at least two months earlier than relying solely on optical data.

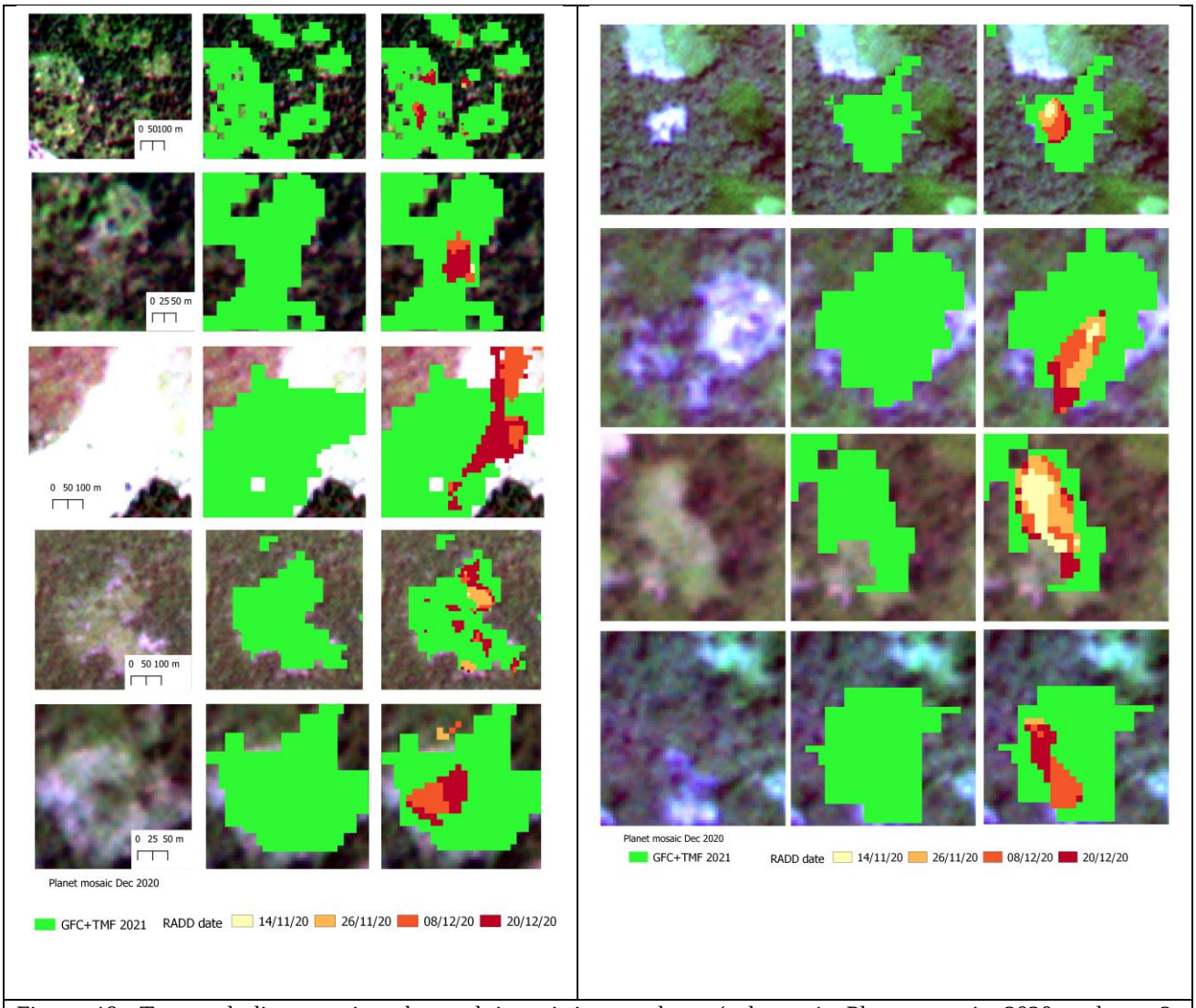


Figure 19: Temporal discrepancies observed in existing products (column 1: Planet mosaic 2020, column 2: GFC21+TMF 21, column 3: GFC21+TMF21+RADD20)

5.2 Classification in different landscapes

All five landscapes have been classified using different combinations of band stacks. In landscape LS-1, LS-3 and LS-4, classification was carried out based on only S-2 indices while in landscape LS-2 and LS-5, classification was carried out based on S1, S2 and their combination. All the confusion matrix of the classification is given in Appendix 9.8.

5.2.1 LS-1 classification (S2)

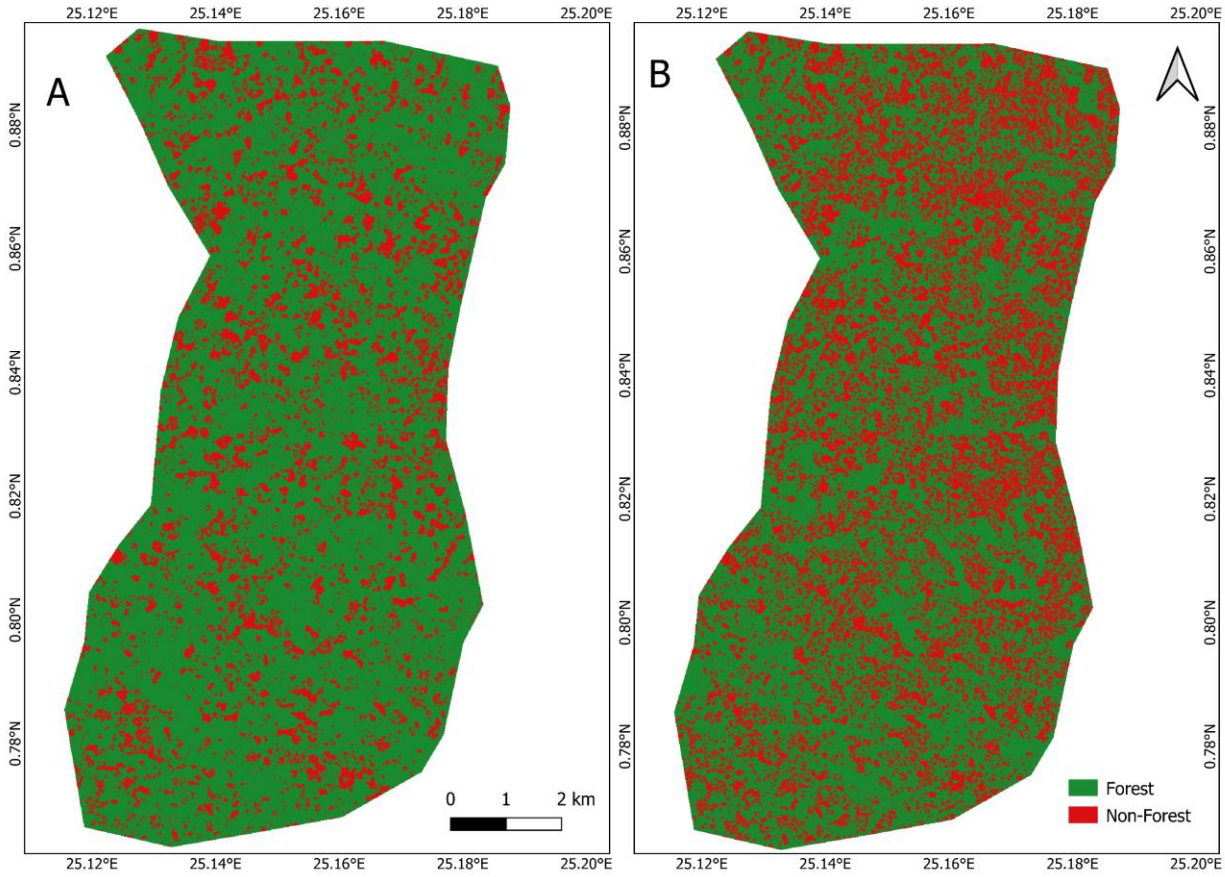


Figure 20: Forest extent classification in Landscape-1 (A. Training samples derived from two landcover classes: forest, non-forest; B. Training sample based on 3 landcover classes: forest, non-forest, others)

Table 13: Accuracy assessment of classification in LS-1 (OA=Overall Accuracy, PA=Producer's accuracy, UA=Users' accuracy, F=Forest, NF=Non-Forest)

LS-1 (%)	OA	PA (F)	PA (NF)	UA (F)	UA (NF)	F-score (F)	F-score (NF)
S2 2 classes	84	83	85	93	67	88	74
S2 3 classes	83	84	81	93	62	88	70

Landscape 1 was dominated by agriculture and sparse vegetation. According to the comparison of optical and SAR products, we found that in this landscape optical products have good area coverage and accuracy for deforested areas. Therefore, S2-based classification was carried out in this landscape, as shown in Figure 20. First, classification was carried out based on two training classes: forest and non-forest. OA of 84% was obtained, F-score of 88% was obtained for the forest class (Table 13). PA was found to be higher for non-forest areas (83% and 85% for forest and non-forest respectively), which represents a higher percentage of non-forest reference pixels that were correctly classified compared to forest pixels. On the other hand, UA of non-forest was found 67% compared to 93% for forest which represents a higher level of commission error in the non-forest class. However, the visual interpretation of classification results shows an overestimation of forest class. Some other agricultural lands and built-up areas are also classified as forests. Thus, an additional training and validation dataset was prepared including a 3rd land cover class 'other area' and then training points from deforestation and other areas are merged to a single class 'non-forest'. Then the classification was carried out considering the forest and modified non-forest training samples. Classification with three land cover training classes reduced the overestimation of forest classes. It also slightly reduced the performance indicators of the classification, OA reduced to 83%, PA of non-forest reduced to 81%, UA of non-forest reduced to 62%. While F-score for forest class remained the same, F-score of non-forest class reduced by 4%, suggesting a decline in models' ability to accurately classify the non-forest areas. While comparing the existing deforestation products, lower UA was obtained in this landscape even with the best combination, which resembles the lower UA obtained by non-forest classes.

5.2.2 LS-2 classification (S1, S2, S1+S2)

In landscape-2 (area dominated by damaged forest), we found better agreement of optical and SAR-based products. Therefore, classification was carried out for both S1 and S2, and their combined layer as presented in Figure 21. The Lowest accuracy was obtained for S1-based classification. With S1, overall accuracy was obtained 75% with a F-score of 83% for the forest class (Table 14). UA of 92% was found for the forest class while only 41% UA was obtained for the non-forest layer, which represents a higher commission error for the non-forest layer. While using S1 only, the model highly misclassified non-forest areas as forests. On the other hand, S2-based classification achieved a good OA of 87%, (13% higher than S1). With an F-score of 90%, and 79% for forest and non-forest classes, it represents better agreement with classification

products. Also, the UA of the non-forest area was increased to 74% which represent the model based on S2 classified reduced the commission error. The best OA was found when optical and SAR data are combined (Figure 21, C). With an OA of 88%, and F-score of 91% for forest class, the combination of S1 and S2 represent the best classification in this degraded forest-dominated landscape. The OA was 13% higher in comparison to S1 based classification. Compared to individual products, S1+S2 represents 1% improvement in OA. In terms of PA, 1% improvement was found in the forest class while 3% improvement was observed for the non-forest class. The combination of S1 and S2 also improved the UA by 2% for the forest class while UA remained the same as S2 for the non-forest classes.

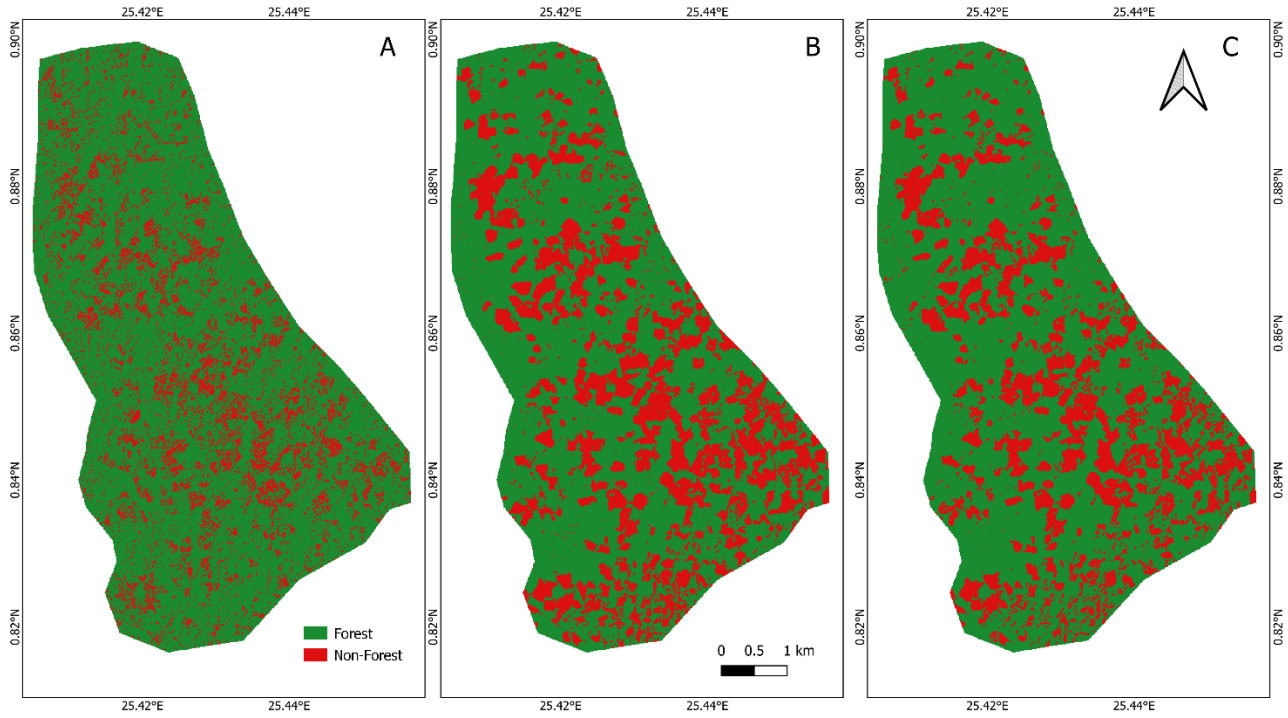


Figure 21: Forest extent classification in Landscape-2 (A. Classification based on S1, B. Classification based on S2, C. Classification based on combination of S1 and S2)

Table 14: Accuracy assessment of classification in LS-2 (OA=Overall Accuracy, PA=Producer's accuracy, UA=Users' accuracy, F=Forest, NF=Non-Forest)

LS-2 (%)	OA	PA (F)	PA (NF)	UA (F)	UA (NF)	F-score (F)	F-score (NF)
S1	75	75	73	92	41	83	52
S2	87	87	85	93	74	90	79
S1+S2	88	88	88	95	74	91	80

5.2.3 LS-3 classification (S2)

Landscape-3 is dominated by agriculture and palm tree plantation. According to our analysis, we found that optical products were reliable in this landscape in terms of deforested area coverage and users' accuracy. Therefore, forest extent classification was carried out based on S2, where training samples with two and three training classes (Figure 22) were used. Classification based on two training classes obtained an OA of 85%, with an F-score of 89% for the forest class and 74% for the non-forest class (Table 15). Where with three classes OA was reduced to 83% and F-score of forest was reduced to 88%. Although the OA was higher for classification with two classes, it showed an overestimation of forest classes. It includes built-up areas, agricultural landscapes, and grasslands as forested areas. While using three training classes, the model can distinguish between forest and other land cover well although the OA and F-score obtained is comparatively lower. In this area, we also observed lower confidence in detecting deforestation while comparing the existing deforestation products. Therefore, the classification based on those training samples results in lower accuracy as reflected in the results.

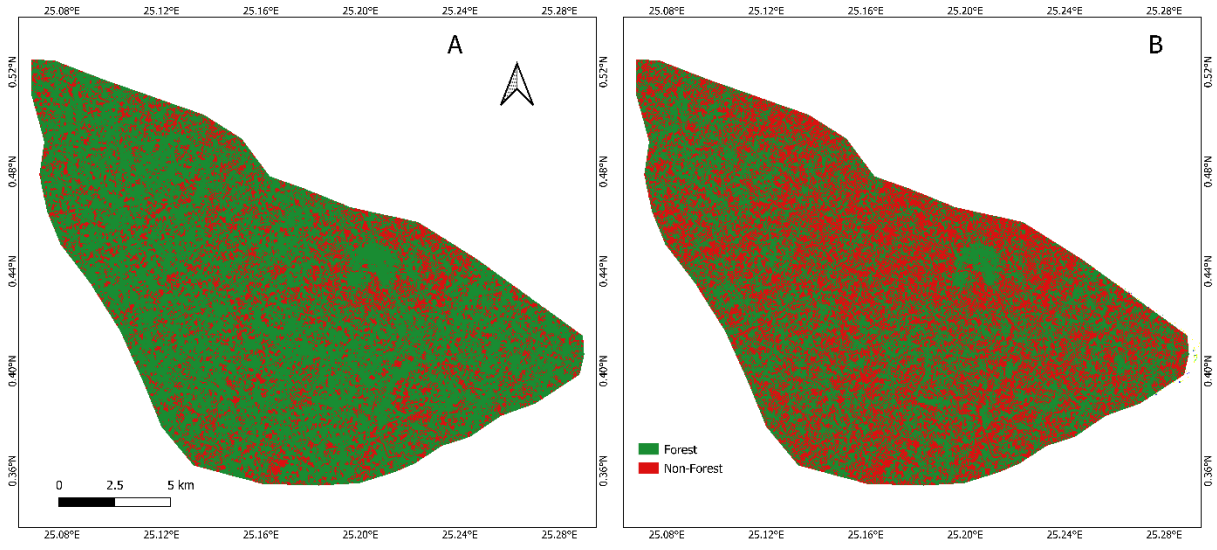


Figure 22: Forest extent classification in Landscape-3 (A. Training samples derived from two landcover classes: forest, non-forest; B. Training sample based on 3 landcover classes: forest, non-forest, others)

Table 15:Accuracy assessment of classification in LS-3 (OA=Overall Accuracy, PA=Producer's accuracy, UA=Users' accuracy, F=Forest, NF=Non-Forest)

LS-3 (%)	OA	PA (F)	PA (NF)	UA (F)	UA (NF)	F-score (F)	F-score (NF)
S2 2 classes	85	82	93	97	62	89	74
S2 3 classes	83	83	82	93	67	88	73

5.2.4 LS-4 classification (S2)

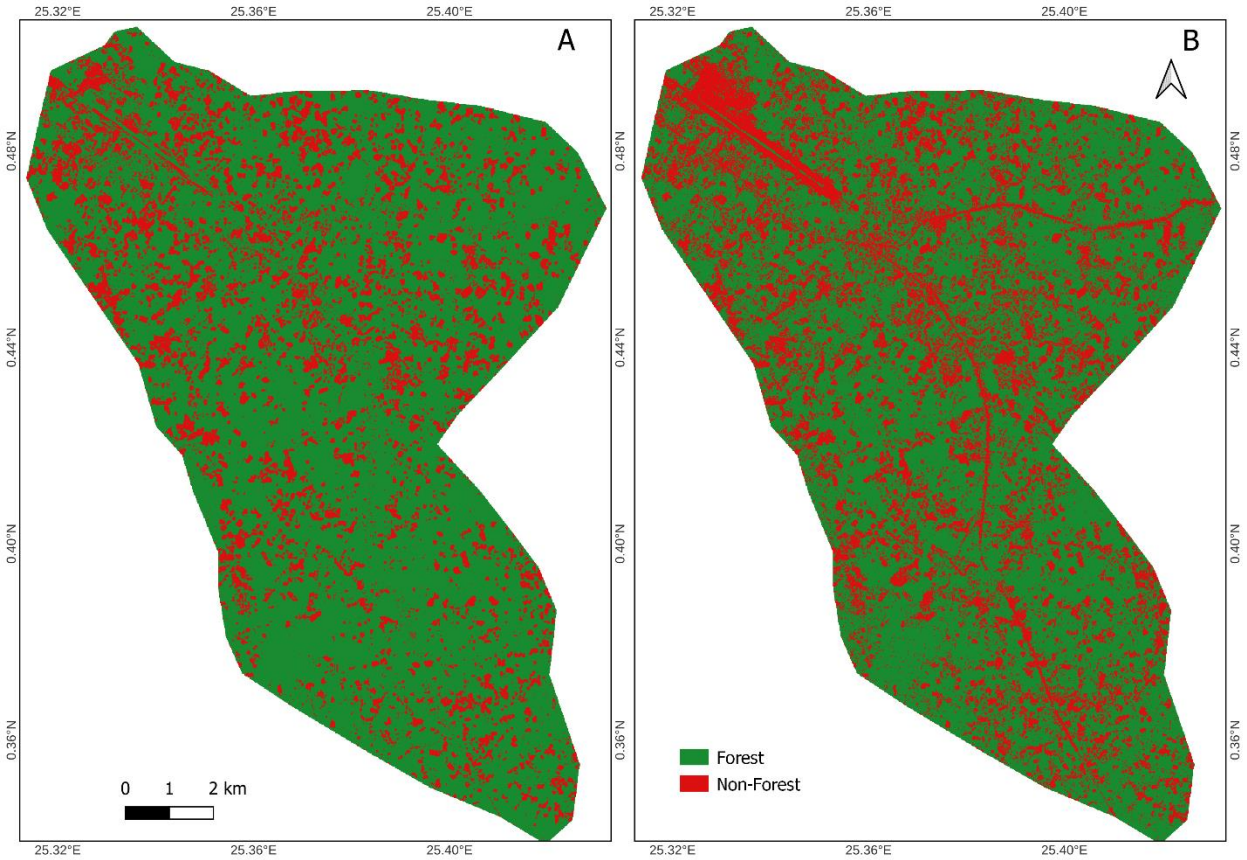


Figure 23: Forest extent classification in Landscape-4 (A. Classification based on 2 training classes: forest, non-forest; B. Classification based on 3 training classes: forest, non-forest, others)

Table 16: Accuracy assessment of classification in LS-4 (OA=Overall Accuracy, PA=Producer's accuracy, UA=Users' accuracy, F=Forest, NF=Non-Forest)

LS-4 (%)	OA	PA (F)	PA (NF)	UA (F)	UA (NF)	F-score (F)	F-score (NF)
S2 2 classes	78	85	64	84	66	84	65
S2 3 classes	84	89	75	88	76	88	75

Landscape 4 is dominated by mixed land cover where we found better reliability for optical products. As a result, only S2 was used for the classification in this landscape. Initially, a classification approach based on forest and non-forest training samples (Figure 23.A) was employed, by which OA of 78% was obtained with 84% F-score for the forest area (

Table 16). Although the PA of the forest class is good (85%), the classification tends to overestimate the forested areas. This was due to misclassification of Roads, settlements, and other

built-up areas as forested areas. To address this issue, a second classification was carried out where training points also included other land cover areas as non-forest areas. With three land cover classes, OA increased to 84% with a rise of 4% and 10% rise in the F-score of forest and non-forest respectively. Classification where the training sample also includes other land cover areas along with deforested areas as a non-forest training sample provided a better estimation of the forest cover in this landscape. By including the other land cover types in the classification process, the misclassification of non-forest areas as forests was reduced.

5.2.5 LS-5 classification (S1, S2, S1+S2)

Landscape-5 is dominated by intact forests. This area represents how intact forests are getting degraded from the buffer zones. Here, we found better convergence among optical and SAR-based products. Consequently, the classification was conducted using all three combinations: S1-based, S2-based, combination of S1- and S2-based (Figure 24). In the case of S1-based classification, we achieved an OA of 80% with an F-score of 86% for the forest class (Table 17). However, the PA and UA of non-forest were found to be poor (77% and 57% respectively). Although, this classification correctly identified the large deforested area near the buffer side, it generated false non-forest areas in the central part of the landscape. However, using optical-based S2 reduces the problem (Figure 24.B) and OA increased by 11%, with better F-score for both forest and non-forest classes (93% and 86% respectively).

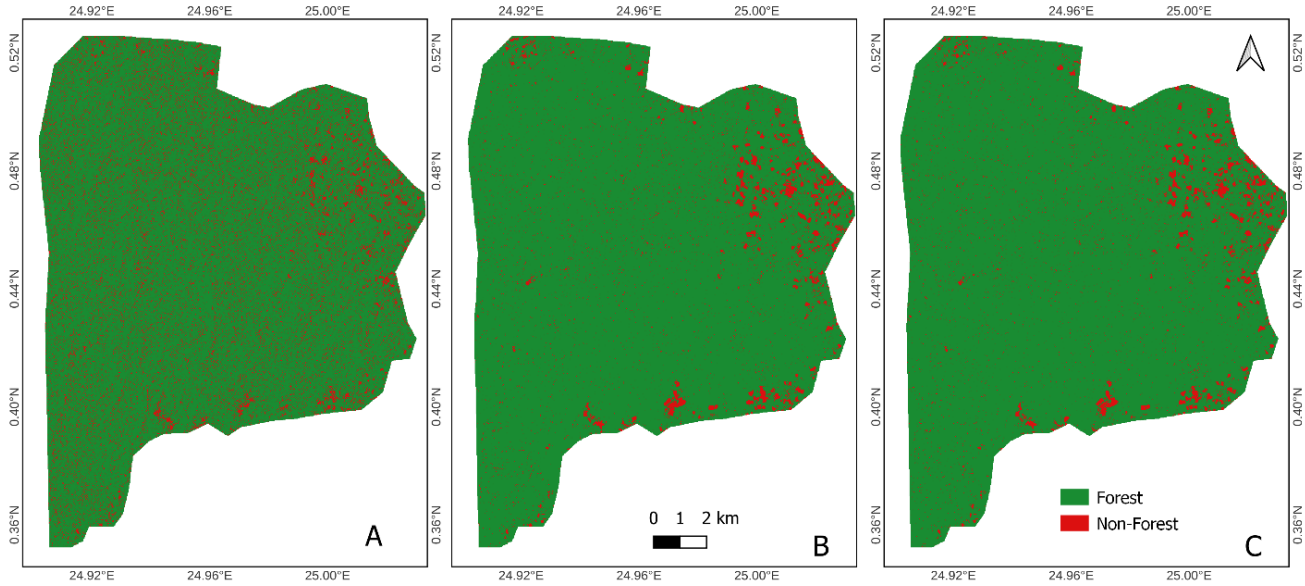


Figure 24: Forest extent classification in Landscape-2 (A. Classification based on S1, B. Classification based on S2, C. Classification based on combination of S1 and S2)

Table 17: Accuracy assessment of classification in LS-4 (OA=Overall Accuracy, PA=Producer's accuracy, UA=Users' accuracy, F=Forest, NF=Non-Forest)

LS-5 (%)	OA	PA (F)	PA (NF)	UA (F)	UA (NF)	F-score (F)	F-score (NF)
S1	80	81	77	92	57	86	66
S2	91	92	90	95	83	93	86
S1+S2	93	93	92	96	86	94	89

By combining the optical and SAR-based classification, the highest OA was achieved compared to individual products. The combined approach showed a 13% improvement in OA compared to the S1-based classification and a 2% improvement compared to the S2-based classification. This indicates the best agreement among the classification products. Additionally, combined use of S1 and S2 resulted in a 1% to 2% improvement in PA of forest and non-forest respectively, while in terms of UA the improvement was observed by 1% and 3%. Overall, the combined approach yielded a 2% overall improvement rather than using the individual S1 or S2 alone for classification, emphasizing the effectiveness of utilizing multiple data sources.

6. Discussion

This chapter provides a discussion based on the obtained result. Section 6.1 focuses on the discussion of the comparison of existing optical and SAR-based deforestation products. In section 6.2, classification results are thoroughly discussed. Following this, section 6.3 provides a description of the best classification maps for each landscape. Finally, section 6.4 highlights the limitation and perspectives of this study, providing insights into areas for further research and potential improvements.

6.1 Comparison of existing deforestation products

According to the comparison carried out, the performance of existing deforestation products was found landscape-specific. Various optical and SAR-based products exhibited varying levels of performance in determining the deforestation of different landscapes. SAR-based RADD alert performed well in the vicinity of forest buffer regions. The highest contribution of RADD alerts has been observed in landscapes dominated by intact or degraded forests, where the intersection layers also worked well. However, RADD didn't perform effectively in areas dominated by heterogeneous land cover, such as agriculture, plantation and mixed land cover areas. In such regions, non-forest areas are small-scale and closely located to other land cover types, making it challenging to determine those areas by SAR backscattering. On the other hand, optical sensor-based products GFC and TMF complement those areas where RADD didn't perform well. Nevertheless, the accuracy of optical sensor-based products is relatively lower in those landscapes. In all cases (except LS-3) the highest accuracy has been observed in the convergence of GFC, TMF and RADD although the area covered by their intersection has the lowest area coverage. Although GFC and TMF delineate the bigger size of forest disturbance, the accuracy was not always good especially in the absence of RADD alert. In almost every case, when GFC and TMF are combined with RADD alert, the UA is better than that of their individual products. Therefore, we can conclude that a combination of optical images and radar images can observe forest disturbance with better accuracy.

A total of 20004 pixels are found where RADD detected the changes in 2020 but GFC and TMF detected in 2021. 65% of those pixels are recorded by RADD within November and December alone. Therefore, SAR-based RADD alerts can observe forest disturbance earlier than GFC and TMF while optical products are impacted by cloud cover and show a delay in detection. Therefore,

in areas where both optical and SAR-based products can be combined, the temporal scale of observation has to be reduced. However, it was not possible to determine the exact date of 2021 when GFC and TMF calculated those areas of November and December 2020 as both optical products are annual products.

From our observation, it has been found that optical and SAR-based products combinedly have a higher level of UA, and the temporal scale of observation is reduced. Training points over those areas would likely cover high-confidence forest disturbance areas. But not all landscapes are well covered by SAR-based alerts in terms of deforested areas. According to our result, in the case of landscape near the forest-dominated areas (LS-2 and LS-5), the intersection of optical and SAR-based products has been used for training samples while in heterogenous landscapes dominated by agriculture, built-up area and sparse vegetation (LS-1, LS-3, LS-4), in lack of SAR coverage, the intersection between optical products was only used. Landscapes LS-1, LS-3, LS-4 are located on the bank of Lindi river and Lualaba river, unlike LS-2 and LS-5. In those areas, SAR alerts can be influenced by moisture variation. Reiche et al., (2021) also mentioned the higher sensitivity of C-band SAR data to moisture variation can induce false detections in the Congo basin.

6.2 Classification

In our study, we combined data from S1 and S2 satellites in two of our landscapes, LS-2 and LS-5. For the remaining three landscapes, we relied solely on S2 imagery for classification, as it proved to be more effective in terms of accurately determining deforested areas in those landscapes. Consequently, S2-based classification was performed in all five landscapes. When considering S2 imagery, LS-1, LS-3, LS-4 had an OA ranging from 78% to 85%, while LS-2 and LS-5 demonstrated higher OA, achieving 87% and 91% respectively. This indicates that overall accuracy was relatively higher in landscapes predominantly covered by forests compared to those dominated by agriculture or mixed land cover. This disparity can be attributed to the spectral signatures of the forest class being confused by other land cover types and experiencing overestimation in areas dominated by agriculture or plantations when using training data based on two classes. Consequently, the number of false negatives increased for the forest classes. To mitigate this issue, we opted to classify those landscapes by using training data derived from three land cover classes. The inclusion of the third class slightly reduced the OA in LS-1 and LS-3 but increased the OA in LS-4. In every case, the overestimation of forest areas was reduced when

using training data from three land cover classes. Comparatively lower OA in LS-1, LS-3 and LS-4 can be justified considering the lower level of UA observed in those landscapes during the comparison of optical and SAR-based deforestation products. It was found that the best reliability was achieved by combining only optical products, with a maximum UA of 80%. In contrast, the forest-dominated areas showed the highest UA of 95% when optical and SAR data were combined. Moreover, the use of non-forest training samples derived only from optical deforestation products led to a decrease in classification accuracy. This suggests the importance of incorporating diverse training samples from different data sources to improve the classification.

The landscapes dominated by forest areas (LS-2 and LS-5), exhibited the highest accuracy when combining data from S1 and S2. In both landscapes, classification with only S1 achieved the lowest overall accuracy 75% and 80% respectively for LS-2 and LS-5. We used only four SAR features as a band combination of SAR images, the inclusion of more SAR metrics would potentially increase the effectiveness of S1-based classification. However, these findings align well with the previous studies (De Luca et al., 2022; Hu et al., 2020; Tavares et al., 2019). Travares et al. (2019) combined S1 textural indices with S2 multispectral indices and found the lowest OA of 56% for S1. Similar results were observed by Heckel et al. (2020) where integration of multi-temporal S1 data with single-timestep S2 data yielded the lowest overall accuracy of 87.5% when using S1 alone. Although the SAR product has the advantage of penetrating the clouds, it has the limitation of distinct various feature classes (Tavares et al., 2019). On the other hand, usage of S2 multispectral indices produced better results in those two landscapes compared to S1, which is also consistent with De Luca et al. (2022) and Hu et al. (2020), where S2 achieved almost 8% and 3% higher OA than S1. In our study, we observed 12% and 11% higher overall accuracy calculated by S2 at landscape LS-2 and LS-5 respectively compared to S1. Also, it was observed that PA and UA of S2 were also higher compared to S1, suggesting that classification based on S2 data reduced the tendency to overestimate forest area when compared to S1. However, the availability of S2 cloud-free images in tropical regions poses a significant constraint.

There is a synergistic effect of combining optical and SAR data, OA was improved when SAR-based products are combined with optical-based products. We achieved an overall improvement of OA by 1% and 2% in our landscape LS-2 and LS-5 respectively. Although the improvement is not really significant, this finding is converging with some recent publications. By integrating two

S1 images and one S2 scene, Tavares et al. (2019) found an overall improvement of 1.54%. Another study combined optical and SAR products for F/NF and LULC classification found an overall 5% of improvement in OA for LULC mapping, however, only a 1.5% improvement was reported in terms of F/NF classification (Hirschmugl et al., 2018). De Luca et al. (2022) also observed a 2.53% improvement in land cover mapping in the Mediterranean regions by using S1 and S2 time series. By integrating S1 with S2, around 1.58% of average improvement was also reported in an attempt at crop type classification (Sun et al., 2020).

On the other hand, some other studies achieved higher levels of improvement when combining optical and radar sensors by using advanced fusion approaches. Blickensdörfer et al. (2022) used S2 and Landsat dense time series for crop sequence mapping and reported a 6% to 10% improvement while adding SAR feature in combination with an optical sensor and incorporating environmental factors. Another study obtained a 2% to 5% improvement by using fusing S1 and S2 multitemporal images for urban surface mapping (Shrestha et al., 2021). Using object-based classification, an overall 5% improvement was found in an attempt to paddy mapping by using optical and SAR time series (Cai et al., 2019). By incorporating Geographic-Object-Based Image Analysis (GEOBIA) and data mining together with optical and SAR data for wetland tree classification, a 4% improvement was found (Ruiz et al., 2021). Therefore, in future research, advanced fusion approaches will be employed to combine S1 and S2 data to assess the level of improvement. The knowledge acquired from this research will serve as a foundation for implementing advanced techniques in the future.

6.3 Best classification maps for each landscape

In our study, training and validation datasets are randomly generated from the existing deforestation products. Therefore, classification is computed from those secondary datasets meaning that inaccuracies and misclassification of the existing products can influence the classification. With regard to the low number and low quality of the validation dataset, the quantitative assessment from the confusion matrix is limited. In addition to the quantitative accuracy assessment, a qualitative assessment was carried out by visual interpretation of the Planet image to determine the best classification results for each landscape. The best classification in each landscape was used for forest and non-forest area calculation. The classification products are

zoomed in and compared with Planet imageries to look for the misclassification areas and strengths and weaknesses of the classification.

LS-1:

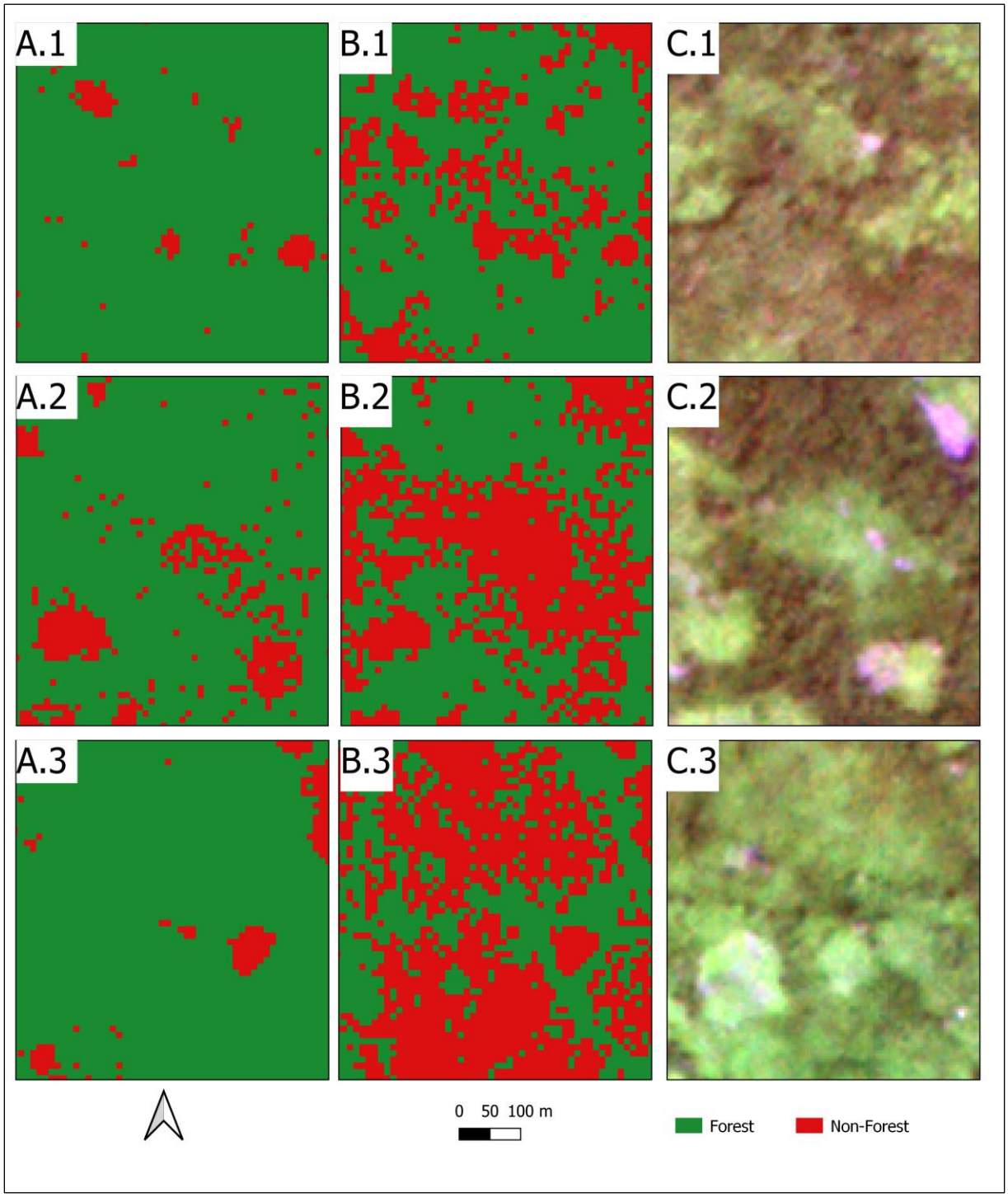


Figure 25: Comparison of different classifications with Planet image in LS-1 (A. Training samples derived from two landcover classes, B. Training sample based on three landcover classes, C. Planet monthly mosaic of November 2020)

In landscape-1, two classifications were carried out including two and three training classes and obtained an OA of 84% and 83% respectively. Although the quantitative assessment suggested training samples based on two classes as the best products, photointerpretation with Planet imagery shows an overestimation of forest area with a two-class-based classification. According to the observation of Figure 25, classification based on two classes (Column A) miscalculated agricultural areas as forests, while using 3 classes for the classification (Column B), it estimated agricultural area as other land covers which reduces the overestimation of forested area. Therefore, the classification based on three land covers training classes is considered as the best classification result in this landscape.

LS-2:

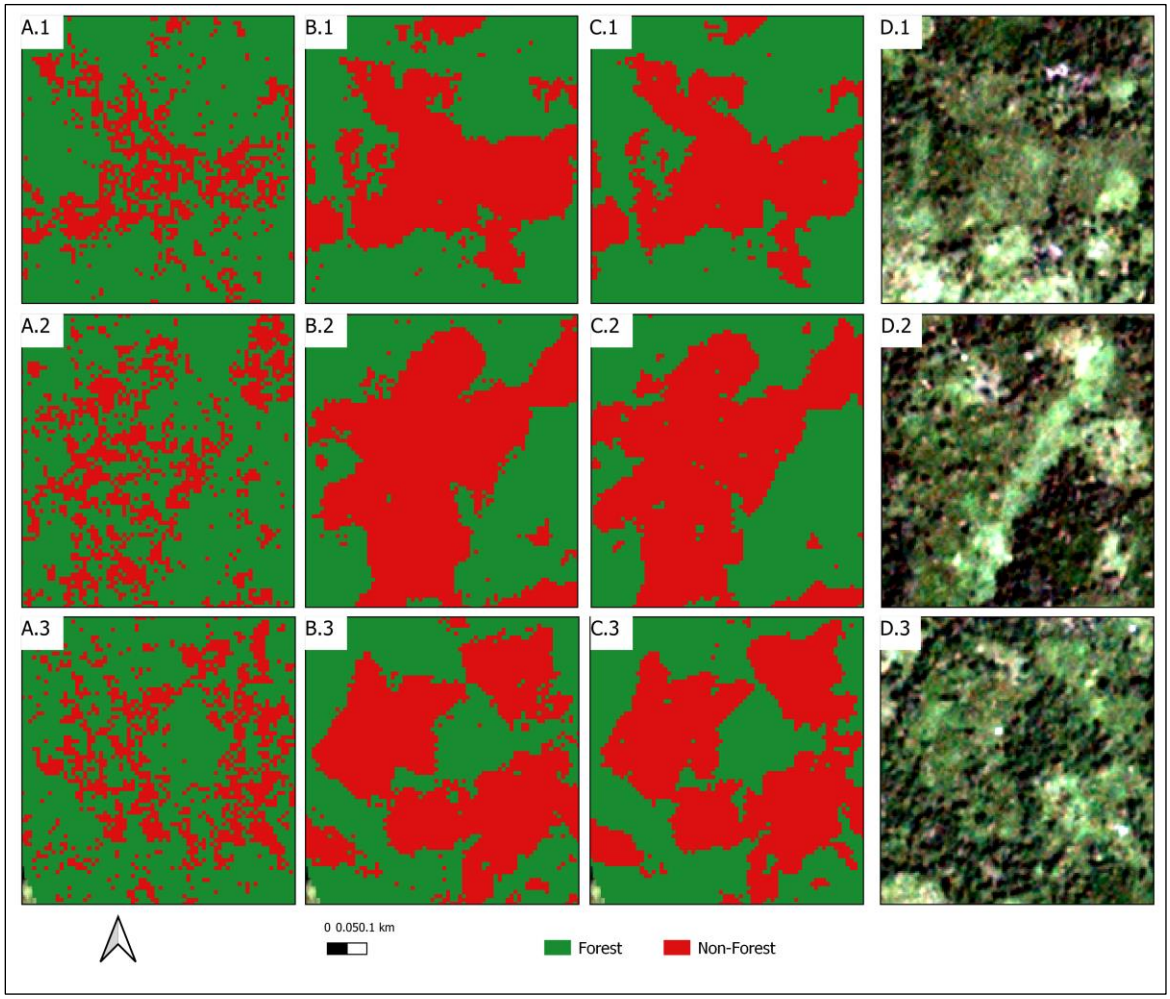


Figure 26: Comparison of different classifications with Planet image in LS-2 (A. S1-based classification, B. S2-based classification, C. classification based on combination of S1 and S2, D. Planet monthly mosaic of December 2020)

For landscape 2 (dominated by forest area) we classified based on S1, S2, and the combination of S1 and S2 and the OA of 75%, 87% and 88% were found respectively. While using S1 alone, the classification result was the worst, it fails to determine the non-forest areas efficiently as we can observe scattered pixels (Figure 26, Column A) around the disturbed forest zone. While using S2, the speckle is reduced, and forest loss was visible in an efficient manner.

However, when S1 and S2 are combined, smaller scattering pixels in B2 and B3 were somewhat reduced in C2, and C3. As the OA was obtained higher for S1+S2, the classification based on the combination of S1 and S2 is considered as the best products in this landscape.

LS-3:

In Landscape 3, we obtained an OA of 85% when using the classification approach based on two land cover training classes and an OA of 83% for three land cover training classes. However, according to the visual interpretation (Figure 27), the classification based on three land cover training classes provided better results for this landscape. The classification based on 2 classes overestimated the forest area (Column A) and misclassified other land cover types such as built-up areas and agricultural fields. On the other hand, when the classification was run with 3 land cover classes, the overestimation was reduced and a better distinction between forest and other land cover types (Column B) was achieved. Specially, in B.1 and B.3 the classification correctly identified the built-up area, while in B.2, it includes the agricultural field and removed it from the forest area. Therefore, the classification based on three land cover classes is considered as the best product despite having a lower accuracy.

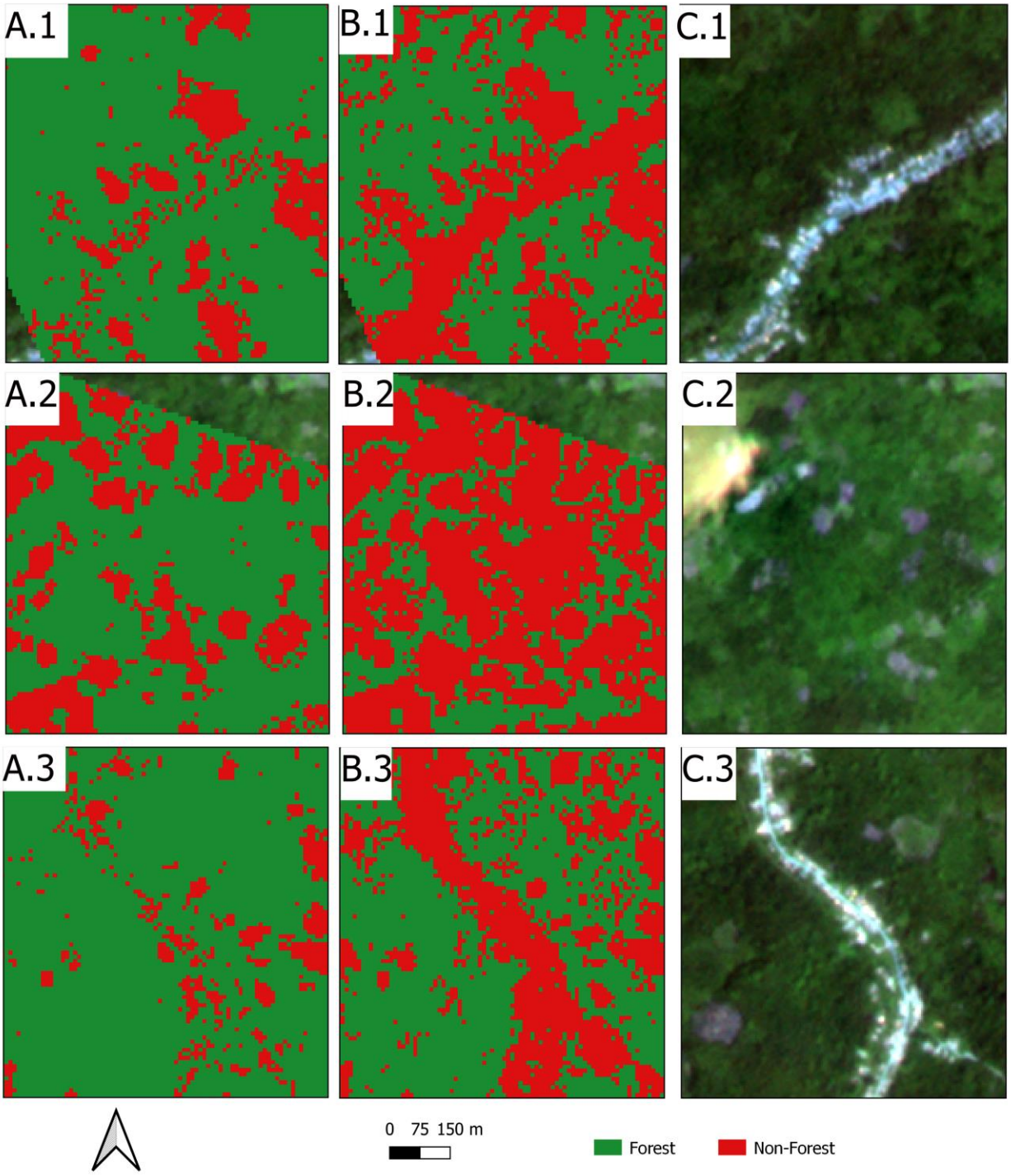


Figure 27: Comparison of different classifications with Planet image in LS-3 (A. Training samples derived from two landcover classes, B. Training sample based on three landcover classes, C. Planet monthly mosaic of November 2020)

LS-4:

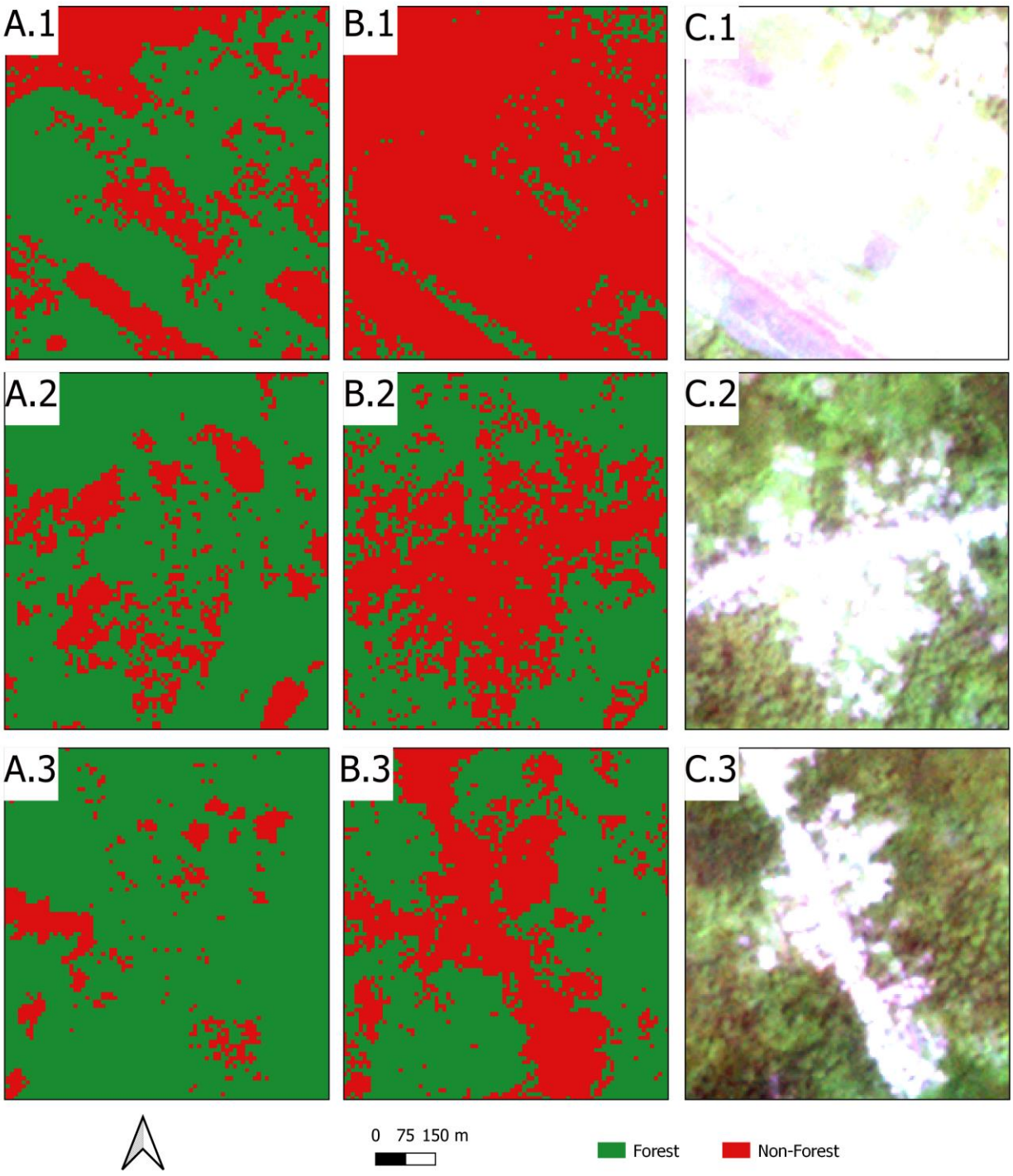


Figure 28: Comparison of different classifications with Planet image in LS-4 (A. Training samples derived from two landcover classes, B. Training sample based on three landcover classes, C. Planet monthly mosaic of November 2020)

For LS-4, the two-class-based classification approach failed to correctly identify non-forest land cover areas, as illustrated in Figure 28. Specifically, some portions of the other land cover such as built-up areas are misclassified as forest. In Figure 28.A1, the two-class based classification incorrectly classified some parts of Kisangani airport as forest, however, this problem was resolved while using three landcover classes for training. Similar trends were observed in A2 and A3, where classification based on two classes misclassified built-up areas as forests. Conversely, classification based on three classes (Figure 28, column B) successfully detected these areas, thereby reducing the overestimation. Therefore, in this landscape, classification based on three classes yielded the best results from both quantitative and qualitative accuracy assessment.

LS-5:

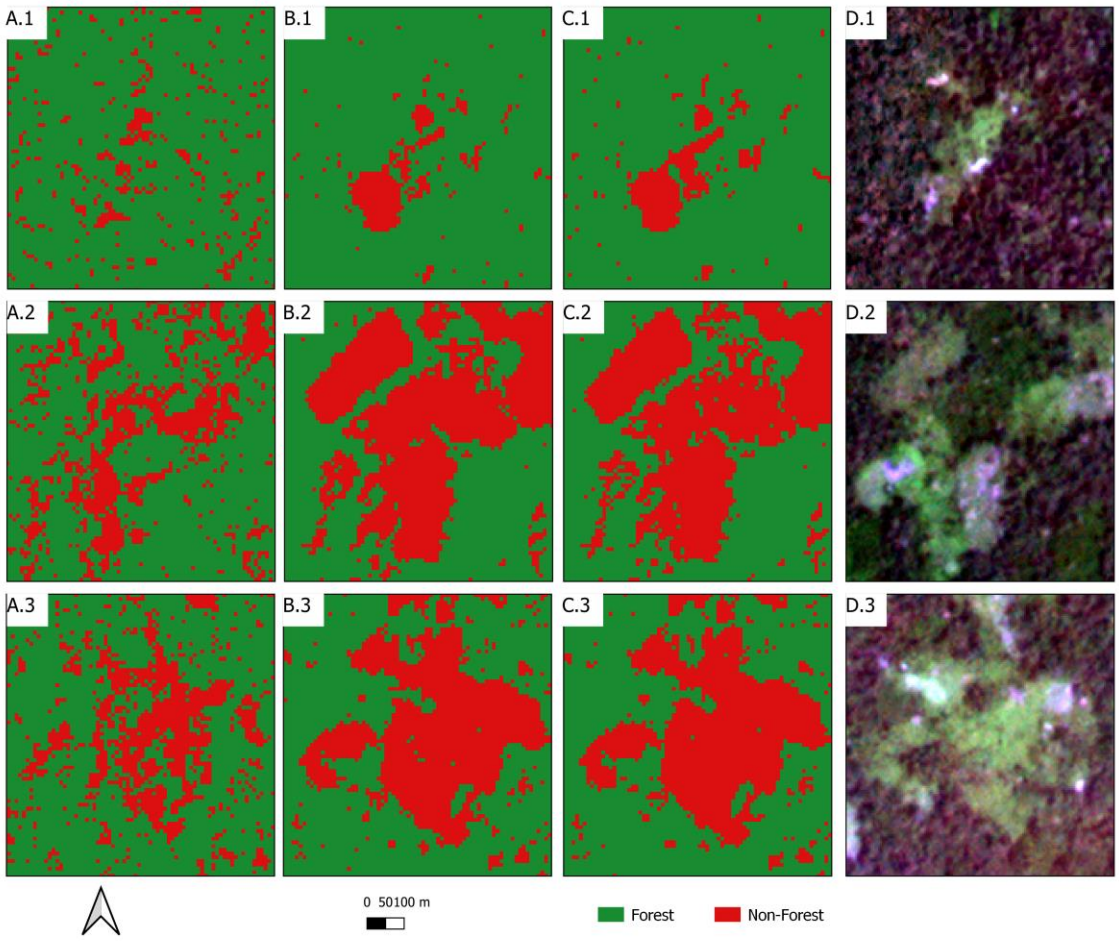


Figure 29: Comparison of different classifications with Planet image in LS-5 (A. S1-based classification, B. S2-based classification, C. classification based on combination of S1 and S2, D. Planet monthly mosaic of December 2020)

For LS-5, we found that S1 alone could not determine the non-forest areas (Figure 29) well. The performance of S1 was relatively better for large-scale forest loss (A2 and A3) compared to small-scale forest loss (A1). S2 produces better results in this landscape, however, it is also influenced by small misclassification. When S1 and S2 are combined, the model performs slightly better than S2 alone, however, it was difficult to identify the improvement in visual interpretation. Therefore, considering the quantitative accuracy, we considered S1+S2 to be the best combination of F/NF classification in this landscape.

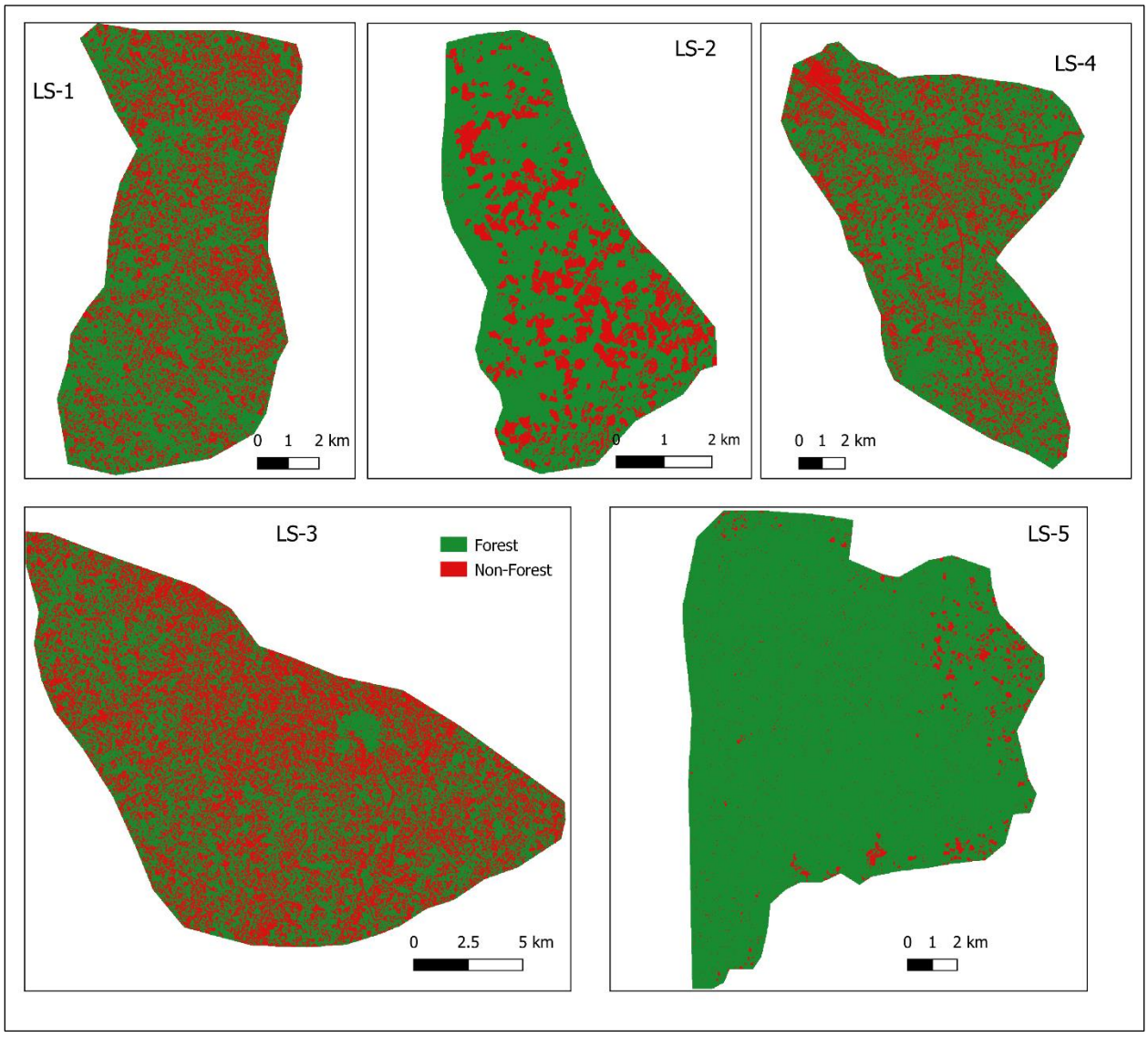


Figure 30: The final forest extent classification maps for each landscape

After selecting the best product for each landscape, the forest and non-forest areas of each landscape are calculated based on it (Figure 30).

Table 18: Area calculation for the best classification result and comparison with TMF (Vancutsem et al., 2021) global maps

Landscape	Forest (km^2)	non-Forest (km^2)	Total area (km^2)	Forest TMF (km^2)	% of forest cover by our study	% of forest cover by TMF
LC1	46.32	37.78	84.10	49.74	55.08	59.14
LC2	22.59	8.06	30.65	26.08	73.69	85.09
LC3	140.02	116.71	256.73	144.70	54.54	56.36
LC4	92.15	45.12	137.27	95.77	67.13	69.77
LC5	181.84	7.77	189.61	186.44	95.90	98.33

In landscape 1, we found more than half of forest coverage (55%) in 84 km^2 of area (Table 18). Among our landscapes, the highest forest coverage (95%) has been observed in LS-5 which was dominated by intact forest. The lowest forest coverage has been observed in landscape 3 which was an agriculture-dominated landscape. We compared our area coverage by the best classification result with TMF annual forest cover of 2020. The result shows the convergence between our result and TMF. However, we found that TMF products overestimated the forest coverage by 2% to 11% compared to our study. During the comparison of existing deforestation products, we also noticed an overestimation by TMF products by comparing it with the Planet images. Our study combined S1 and S2 data within 1 year at the landscape level where TMF is based on Landsat multiyear time series for pan-tropic scale. This overestimation can be attributed to the difference in spatial resolution and data processing techniques between our study and the study conducted by JRC TMF.

6.4 Limitations and Perspectives

This study used S1 time series and S2 cloud-free images for forest extent classification based on the comparison of existing optical and SAR-based products. For training and validation samples this study depends on the accuracy and area coverage of existing deforestation products. Any inconsistencies regarding the spatial and temporal resolution, area coverage, and workflow may influence our training and validation sample. In addition, although we achieved higher accuracy

when combining optical and SAR-based products, their area coverage was smaller compared to individual products. Therefore, the training and validation dataset are sampled over those pixels which sometimes produces training samples close to each other. In that case, fully independent training and validation samples taken from very high-resolution imageries or ground truthing would improve the results. According to our comparison of the existing deforestation products, this study calculated S1 and S2 combinations for classification in two landscapes. In the future, with more reliable training data, the combination of S1 and S2 would be carried out in all different landscapes to assess their improvement for forest extent classification. This improvisation in training sample would make our method more reliable for future use.

Regarding the S1 and S2 band combinations, the strengths and weaknesses of individual bands were assessed through only visual interpretation of deforested areas with Planet mosaic. In the case of S1, we found that individual polarization of VV and VH provided a better explanation of yearly forest and non-forest areas compared to RVI and VV/VH (Appendix 9.2). As for S2, we observed that spectral indices of NDVI, NDWI, NBR and BAIS2 are better for determining the forest and non-forest areas compared to Brightness and SWI (Appendix 9.3). In the future, a sensitivity analysis would be conducted to better understand the relative importance of individual bands in determining the classification outcome and to consolidate our understanding of which bands should be included and excluded.

Regarding the S1-based classification, we used a trimestral median composite. However, other data aggregation approaches like min, max, mean, percentile, etc., would produce different results. Also, some other S1 metrics such as VV+VH, moisture index, modified burned area index, etc., could be tested in different combinations to assess the improvement of the classification. Another way of improving SAR data processing could be to use GLCM textural features in different moving windows to extract several features from S1 backscattering. S1 data used in this study was already preprocessed from GEE environment. Further incorporating environmental factors such as elevation, temperature, precipitation etc. for regional scale can be undertaken to improve the accuracy of S1-based products (Blickensdörfer et al., 2022). Future research is needed to explore different SAR bands, polarimetric composition, preprocessing approaches, normalization techniques and filtering approaches to understand to what extent SAR can effectively support forest non-forest mapping.

In future studies, there is room for methodological improvement to enhance the significance level of combining S1 and S2 data. The focus of this study was on the year 2020 when all three existing products were available. However, incorporating dense time series could potentially improve the accuracy of the classifications. Previous studies have successfully used multiyear time series when integrating S1 and S2 and achieved a higher level of performance (Blickensdörfer et al., 2022; Wang et al., 2020). For instance, the use of three-year time series of S1 and S2 data led to a 6-10% improvement in performance during crop type mapping (Blickensdörfer et al., 2022). Therefore, considering longer time series could be beneficial for future analysis.

Another potential improvement could involve incorporating advanced deep learning models, for example, TWINNS (TWIn Neural Networks for Sentinel data) (Ienco et al., 2019), Residual combined Squeeze-and-Excitation and Non-local Network (RSNNet) (Zhou et al., 2022), to effectively combine S1 and S2 data. While this study used data-based combination approaches, some studies have used result-based Bayesian probability approaches for combining optical and SAR sensors (Hirschmugl et al., 2017; Reiche et al., 2018, 2016). Therefore, future research could explore the use of result-based approaches to determine the most suitable methods for tropical forest extent mapping. By leveraging these advanced techniques, the effectiveness of combining S1 and S2 data can be further improved.

7. Conclusion

In this study, the objective was to explore the effectiveness of optical and SAR imagery for mapping forest extent at the landscape level. For this purpose, first existing optical and SAR-based products are compared, and their accuracy was estimated for different landscapes to collect the training samples. Then classification was carried out using S1 and S2 data according to the analysis at the landscape level. Finally, the classifications were compared to determine the best classification for specific landscapes. Based on the analysis, several conclusions can be made.

Utilizing available optical- and SAR- based GFC, TMF and RADD deforestation products is a viable approach for gathering training samples for forest and non-forest areas. However, it is more appropriate to focus on their intersection at the landscape level to identify the confident forest disturbances. In cases where SAR products are not available and only optical products are, a lower UA was obtained. Moreover, it is crucial to ensure a robust validation protocol before using those existing products. Here we proposed the creation of buffer squares of varying sizes around the randomly sampled points and counting their majority pixel to avoid confusion between forest and non-forest. However, obtaining fully independent training and validation samples that derive from ground-truth data or very high-resolution satellite imagery can yield better results.

The best classification result was achieved when the combination of S1 and S2 was used, while the worst result was observed with S1 data as the only input suggesting its limitation when used alone. The inclusion of three land cover training classes reduces the overestimation and makes a better distinction compared to the classification based on two training classes for S2 alone. We used a data-based combination of S1 and S2 and observed an overall 1% to 2% improvement, which converged with the literature studied regarding forest cover classification. However, more investigations are needed to explore and validate various data- and result-based fusion techniques combining optical and SAR data to fully leverage their potential benefits.

In densely cloud-covered areas like tropical forests, the combination of optical and SAR data has the real advantage for gap-free observation. With some limitations, this research can contribute to the current knowledge of combining S1 and S2 data in tropical forest regions. Higher accuracy and more detailed up-to-date forest information are crucial for REDD+ mitigation strategies. As it is difficult to obtain long-term cloud-free optical data in the tropics, therefore we recommend the synergistic use of optical and radar-based sensors for better forest extent classification.

8. References

Achard, F., Eva, H., & Mayaux, P. (2001). Tropical forest mapping from coarse spatial resolution satellite data: Production and accuracy assessment issues. *International Journal of Remote Sensing*, *22*(14), 2741–2762. https://doi.org/10.1080/01431160120548

Anchang, J. Y., Prihodko, L., Ji, W., Kumar, S. S., Ross, C. W., Yu, Q., Lind, B., Sarr, M. A., Diouf, A. A., & Hanan, N. P. (2020). Toward Operational Mapping of Woody Canopy Cover in Tropical Savannas Using Google Earth Engine. *Frontiers in Environmental Science*, *8*. https://www.frontiersin.org/articles/10.3389/fenvs.2020.00004

Bele, M. Y., Sonwa, D. J., & Tiani, A.-M. (2015). Adapting the Congo Basin forests management to climate change: Linkages among biodiversity, forest loss, and human well-being. *Forest Policy and Economics*, *50*, 1–10. https://doi.org/10.1016/j.forpol.2014.05.010

Bergen, M. (2019). *Congo Basin Deforestation Threatens Food and Water Supplies Throughout Africa*. https://www.wri.org/insights/congo-basin-deforestation-threatens-food-and-water-supplies-throughout-africa

Blickensdörfer, L., Schwieder, M., Pflugmacher, D., Nendel, C., Erasmi, S., & Hostert, P. (2022). Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany. *Remote Sensing of Environment*, *269*, 112831. https://doi.org/10.1016/j.rse.2021.112831

Bolyn, C., Michez, A., Gaucher, P., Lejeune, P., & Bonnet, S. (2018). Forest mapping and species composition using supervised per pixel classification of Sentinel-2 imagery. *BASE*. https://doi.org/10.25518/1780-4507.16524

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), 8. https://doi.org/10.3390/rs10081250

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Butsic, V., Baumann, M., Shortland, A., Walker, S., & Kuemmerle, T. (2015). Conservation and conflict in the Democratic Republic of Congo: The impacts of warfare, mining, and protected areas on deforestation. *Biological Conservation*, *191*, 266–273. https://doi.org/10.1016/j.biocon.2015.06.037

Cai, Y., Lin, H., & Zhang, M. (2019). Mapping paddy rice by the object-based random forest method using time series Sentinel-1/Sentinel-2 data. *Advances in Space Research*, *64*(11), 2233–2244. https://doi.org/10.1016/j.asr.2019.08.042

Cerretelli, S., Poggio, L., Yakob, G., Boke, S., Habte, M., Coull, M., Peressotti, A., Black, H., & Gimona, A. (2021). The advantages and limitations of global datasets to assess carbon stocks as proxy for land degradation in an Ethiopian case study. *Geoderma*, *399*, 115117. https://doi.org/10.1016/j.geoderma.2021.115117

Chen, B., Xiao, X., Li, X., Pan, L., Doughty, R., Ma, J., Dong, J., Qin, Y., Zhao, B., Wu, Z., Sun, R., Lan, G., Xie, G., Clinton, N., & Giri, C. (2017). A mangrove forest map of China in 2015: Analysis of time series Landsat 7/8 and Sentinel-1A imagery in Google Earth Engine cloud computing platform. *ISPRS Journal of Photogrammetry and Remote Sensing*, *131*, 104–120. https://doi.org/10.1016/j.isprsjprs.2017.07.011

Chen, N., Tsendbazar, N.-E., Hamunyela, E., Verbesselt, J., & Herold, M. (2021). Sub-annual tropical forest disturbance monitoring using harmonized Landsat and Sentinel-2 data.

International Journal of Applied Earth Observation and Geoinformation, *102*, 102386. https://doi.org/10.1016/j.jag.2021.102386

Collet, Thibauld. (2022). *Tropical forest type and land cover mapping from SAR-C multi-year time series (Sentinel-1)* [Faculté des bioingénieurs, Université catholique de Louvain]. http://hdl.handle.net/2078.1/thesis:35837

Coluzzi, R., Imbrenda, V., Lanfredi, M., & Simoniello, T. (2018). A first assessment of the Sentinel-2 Level 1-C cloud mask product to support informed surface analyses. *Remote Sensing of Environment*, *217*, 426–443. https://doi.org/10.1016/j.rse.2018.08.009

Dalimier, J., Achard, F., Delhez, B., Desclée, B., Bourgoin, C., Eva, H., Gourlet-Fleury, S., Hansen, M., Kibambe, J.-P., Mortier, F., Ploton, P., Réjou-Méchain, M., Vancutsem, C., Langner, A., Sannier, C., Ghomsi, H., Jungers, Q., & Defourny, P. (2022). *Distribution of forest types and changes in their classification*.

Damania, R., Barra, A. F., Burnouf, M., & Russ, J. D. (2016). *Transport, economic growth, and deforestation in the Democratic Republic of Congo*.

De Grandi, G., Mayaux, P., Rauste, Y., Rosenqvist, A., Simard, M., & Saatchi, S. S. (2000). The Global Rain Forest Mapping Project JERS-1 radar mosaic of tropical Africa: Development and product characterization aspects. *IEEE Transactions on Geoscience and Remote Sensing*, *38*(5), 2218–2233. https://doi.org/10.1109/36.868880

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 Wasseige, C., Tadoum, M., Eba'a Atyi, R., & Doumenge, C. (2015). *The forests of the Congo Basin—Forests and climate change* (Afrique centrale). Weyrich. https://agritrop.cirad.fr/578904/

Dostálová, A., Hollaus, M., Milenković, M., & Wagner, W. (2016). Forest area derivation from Sentinel-1 data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, III–7*, 227–233. https://doi.org/10.5194/isprs-annals-III-7-227-2016

Draulans, D., & Krunkelsven, E. V. (2002). The impact of war on forest areas in the Democratic Republic of Congo. *Oryx*, *36*(1), 35–40. https://doi.org/10.1017/S0030605302000066

El-Sheikh, S. (2022). *Deforestation in the Congo Basin Increased by 5% in 2021, Destruction of World's Second Largest Tropical Forest and Largest Carbon Sink Puts Climate Goals at Risk—Forest Declaration*. https://forestdeclaration.org/press-release-new-report-deforestation-in-the-congo-basin-increased-by-5-in-2021-destruction-of-worlds-second-largest-tropical-forest-and-largest-carbon-sink-puts-climate-goals-at-risk/

ESA. (2022). *Europe's Copernicus programme*. https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Europe_s_Copernicus_programme

ESA, & JRC. (1991). *TREES. Tropical Ecosystem Environment Observations by Satellites.* (Strategy proposal 1991-93. Part 1 : AVHRR data collection and analysis.; TREES SERIES A: Technical Document N°1). Institute of Remote Sensing Application.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Filipponi, F. (2018). BAIS2: Burned Area Index for Sentinel-2. *Proceedings*, *2*(7), 7. https://doi.org/10.3390/ecrs-2-05177

Galiatsatos, N., Donoghue, D. N. M., Watt, P., Bholanath, P., Pickering, J., Hansen, M. C., & Mahmood, A. R. J. (2020). An Assessment of Global Forest Change Datasets for National Forest Monitoring and Reporting. *Remote Sensing*, *12*(11), 11. https://doi.org/10.3390/rs12111790

Gana, M. B., & Herbert, B. (2014). Spatial analysis from remotely sensed observations of Congo basin of east African high land to drain water using gravity for sustainable management of low laying Chad basin of Central Africa. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*.

Ganz, S., Adler, P., & Kändler, G. (2020). Forest Cover Mapping Based on a Combination of Aerial Images and Sentinel-2 Satellite Data Compared to National Forest Inventory Data. *Forests*, *11*(12), 12. https://doi.org/10.3390/f11121322

Gao, B. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, *58*(3), 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3

GFW. (2022). *Global Forest Watch*. https://www.globalforestwatch.org/about/

Guild, R., Wang, X., & Russon, A. E. (2022). Tracking Deforestation, Drought, and Fire Occurrence in Kutai National Park, Indonesia. *Remote Sensing*, *14*(22), 22. https://doi.org/10.3390/rs14225630

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., & Loveland, T. R. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, *342*(6160), 850–853. https://doi.org/10.1126/science.12446

Heckel, K., Urban, M., Schratz, P., Mahecha, M. D., & Schmullius, C. (2020). Predicting Forest Cover in Distinct Ecosystems: The Potential of Multi-Source Sentinel-1 and -2 Data Fusion. *Remote Sensing*, *12*(2), 2. https://doi.org/10.3390/rs12020302

Hirschmugl, M., Deutscher, J., Gutjahr, K.-H., Sobe, C., & Schardt, M. (2017). Combined use of SAR and optical time series data for near real-time forest disturbance mapping. *2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp)*, 1–4. https://doi.org/10.1109/Multi-Temp.2017.8035208

Hirschmugl, M., Sobe, C., Deutscher, J., & Schardt, M. (2018). Combined Use of Optical and Synthetic Aperture Radar Data for REDD+ Applications in Malawi. *Land*, *7*(4), 4. https://doi.org/10.3390/land7040116

Hoekman, D., Kooij, B., Quiñones, M., Vellekoop, S., Carolita, I., Budhiman, S., Arief, R., & Roswintiarti, O. (2020). Wide-Area Near-Real-Time Monitoring of Tropical Forest Degradation and Deforestation Using Sentinel-1. *Remote Sensing*, *12*(19), 19. https://doi.org/10.3390/rs12193263

Hu, L., Xu, N., Liang, J., Li, Z., Chen, L., & Zhao, F. (2020). Advancing the Mapping of Mangrove Forests at National-Scale Using Sentinel-1 and Sentinel-2 Time-Series Data with Google Earth Engine: A Case Study in China. *Remote Sensing*, *12*(19), 19. https://doi.org/10.3390/rs12193120

Ienco, D., Gaetano, R., Interdonato, R., Ose, K., & Ho Tong Minh, D. (2018). *Combining Sentinel-1 and Sentinel-2 Satellite Image Time Series Via RNN for Object-based Land Cover Classification*.

Ienco, D., Interdonato, R., Gaetano, R., & Ho Tong Minh, D. (2019). Combining Sentinel-1 and Sentinel-2 Satellite Image Time Series for land cover mapping via a multi-source deep

learning architecture. *ISPRS Journal of Photogrammetry and Remote Sensing*, *158*, 11–22. https://doi.org/10.1016/j.isprsjprs.2019.09.016

Immitzer, M., Vuolo, F., & Atzberger, C. (2016). First Experience with Sentinel-2 Data for Crop and Tree Species Classifications in Central Europe. *Remote Sensing*, *8*(3), 3. https://doi.org/10.3390/rs8030166

Jiang, W., Ni, Y., Pang, Z., He, G., Fu, J., Lu, J., Yang, K., Long, T., & Lei, T. (2020). A New Index for Identifying Water Body from Sentinel-2 Satellite Remote Sensing Imagery. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *V-3–2020*, 33–38. https://doi.org/10.5194/isprs-annals-V-3-2020-33-2020

Kempeneers, P., Sedano, F., Seebach (Reithmaier), L., Strobl, P., & San-Miguel-Ayanz, J. (2012). Data Fusion of Different Spatial Resolution Remote Sensing Images Applied to Forest-Type Mapping. *Geoscience and Remote Sensing, IEEE Transactions On*, *49*, 4977–4986. https://doi.org/10.1109/TGRS.2011.2158548

Key, C. H., & Benson, N. C. (2006). Landscape assessment (LA). *FIREMON: Fire Effects Monitoring and Inventory System*, *164*, LA-1.

Kim, D.-H., Sexton, J. O., Noojipady, P., Huang, C., Anand, A., Channan, S., Feng, M., & Townshend, J. R. (2014). Global, Landsat-based forest-cover change from 1990 to 2000. *Remote Sensing of Environment*, *155*, 178–193. https://doi.org/10.1016/j.rse.2014.08.017

Kleinschroth, F., Laporte, N., Laurance, W. F., Goetz, S. J., & Ghazoul, J. (2019). Road expansion and persistence in forests of the Congo Basin. *Nature Sustainability*, *2*(7), 628–634.

Kobayashi, N., Tani, H., Wang, X., & Sonobe, R. (2020). Crop classification using spectral indices derived from Sentinel-2A imagery. *Journal of Information and Telecommunication*, *4*(1), 67–90. https://doi.org/10.1080/24751839.2019.1694765

Li, C., Wang, J., Hu, L., Yu, L., Clinton, N., Huang, H., Yang, J., & Gong, P. (2014). A Circa 2010 Thirty Meter Resolution Forest Map for China. *Remote Sensing*, *6*(6), 6. https://doi.org/10.3390/rs6065325

Li, J., & Roy, D. P. (2017). A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring. *Remote Sensing*, *9*(9), 9. https://doi.org/10.3390/rs9090902

Luyssaert, S., Schulze, E., Börner, A., Knohl, A., Hessenmöller, D., Law, B. E., Ciais, P., & Grace, J. (2008). Old-growth forests as global carbon sinks. *Nature*, *455*(7210), 213–215.

Mahdianpari, M., Salehi, B., Mohammadimanesh, F., Homayouni, S., & Gill, E. (2019). The First Wetland Inventory Map of Newfoundland at a Spatial Resolution of 10 m Using Sentinel-1 and Sentinel-2 Data on the Google Earth Engine Cloud Computing Platform. *Remote Sensing*, *11*(1), 1. https://doi.org/10.3390/rs11010043

Marzialetti, F., Di Febbraro, M., Malavasi, M., Giulio, S., Acosta, A. T. R., & Carranza, M. L. (2020). Mapping Coastal Dune Landscape through Spectral Rao's Q Temporal Diversity. *Remote Sensing*, *12*(14), 14. https://doi.org/10.3390/rs12142315

Megevand, C., & Mosnier, A. (2013). *Deforestation trends in the Congo Basin: Reconciling economic growth and forest protection*. World Bank Publications.

Modu, B., & Herbert, B. (2014). Spatial analysis from remotely sensed observations of Congo basin of East African high Land to drain water using gravity for sustainable management of low laying Chad basin of Central Africa. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *40*, 279–286.

Molinario, G., Hansen, M., Potapov, P., Tyukavina, A., & Stehman, S. (2020). Contextualizing Landscape-Scale Forest Cover Loss in the Democratic Republic of Congo (DRC) between 2000 and 2015. *Land*, *9*(1), 1. https://doi.org/10.3390/land9010023

Nasirzadehdizaji, R., Balik Sanli, F., Abdikan, S., Cakir, Z., Sekertekin, A., & Ustuner, M. (2019). Sensitivity Analysis of Multi-Temporal Sentinel-1 SAR Parameters to Crop Height and Canopy Coverage. *Applied Sciences*, *9*(4), 4. https://doi.org/10.3390/app9040655

Nazarova, T., Martin, P., & Giuliani, G. (2020). Monitoring Vegetation Change in the Presence of High Cloud Cover with Sentinel-2 in a Lowland Tropical Forest Region in Brazil. *Remote Sensing*, *12*(11), 11. https://doi.org/10.3390/rs12111829

Ngeunga, M. (2021). *Primary Forest Loss in Congo Basin Escalates in 2020*. https://pulitzercenter.org/stories/primary-forest-loss-congo-basin-escalates-2020

Ngounou, B. (2020). *DRC: the challenges of implementing the Redd+ process on forest management*. Afrik21. https://www.afrik21.africa/en/congo-the-challenges-of-implementing-the-redd-process-on-forest-management/

Nkem, J. N., Somorin, O. A., Jum, C., Idinoba, M. E., Bele, Y. M., & Sonwa, D. J. (2013). Profiling climate change vulnerability of forest indigenous communities in the Congo Basin. *Mitigation and Adaptation Strategies for Global Change*, *18*(5), 513–533.

Phiri, D., Simwanda, M., Salekin, S., Nyirenda, V. R., Murayama, Y., & Ranagalage, M. (2020). Sentinel-2 Data for Land Cover/Use Mapping: A Review. *Remote Sensing*, *12*(14), 14. https://doi.org/10.3390/rs12142291

Puletti, N., Chianucci, F., & Castaldi, C. (2018). Use of Sentinel-2 for forest classification in Mediterranean environments. *Ann. Silvic. Res*, *42*(1), 32–38.

Purwanto, A. D., Wikantika, K., Deliar, A., & Darmawan, S. (2023). Decision Tree and Random Forest Classification Algorithms for Mangrove Forest Mapping in Sembilang National Park, Indonesia. *Remote Sensing*, *15*(1), 1. https://doi.org/10.3390/rs15010016

Reiche, J., Bruin, S. de, Verbesselt, J., Hoekman, D. H., & Herold, M. (2016). *Near Real-time Deforestation Detection using a Bayesian Approach to Combine Landsat, ALOS PALSAR and Sentinel-1 Time Series*. Living Planet Symposium 2016. https://research.wur.nl/en/publications/near-real-time-deforestation-detection-using-a-bayesian-approach-

Reiche, J., Hamunyela, E., Verbesselt, J., Hoekman, D., & Herold, M. (2018). 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 Sensing of Environment*, *204*, 147–161. 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. (2021). Forest disturbance alerts for the Congo Basin using Sentinel-1. *Environmental Research Letters*, *16*(2), 024005.

Renier, C., Vandromme, M., Meyfroidt, P., Ribeiro, V., Kalischek, N., & Ermgassen, E. K. H. J. Z. (2023). Transparency, traceability and deforestation in the Ivorian cocoa supply chain. *Environmental Research Letters*, *18*(2), 024030. https://doi.org/10.1088/1748-9326/acad8e

Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *67*, 93–104. https://doi.org/10.1016/j.isprsjprs.2011.11.002

Rosenqvist, A., Shimada, M., Chapman, B., Freeman, A., De Grandi, G., Saatchi, S., & Rauste, Y. (2000). The Global Rain Forest Mapping project—A review. *International Journal of Remote Sensing*, *21*(6–7), 1375–1387. https://doi.org/10.1080/014311600210227

Rosenqvist, A., Shimada, M., Chapman, B., McDonald, K., De Grandi, G., Jonsson, H., Williams, C., Rauste, Y., Nilsson, M., Sango, D., & Matsumoto, M. (2004). An overview of the JERS-1 SAR Global Boreal Forest Mapping (GBFM) project. *IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium*, *2*, 1033–1036 vol.2. https://doi.org/10.1109/IGARSS.2004.1368587

Ruiz, L. F. C., Guasselli, L. A., Simioni, J. P. D., Belloli, T. F., & Barros Fernandes, P. C. (2021). Object-based classification of vegetation species in a subtropical wetland using Sentinel-1 and Sentinel-2A images. *Science of Remote Sensing*, *3*, 100017. https://doi.org/10.1016/j.srs.2021.100017

Sahadevan, D. K., Sitiraju, S., & Sharma, J. (2013). *Radar Vegetation Index as an Alternative to NDVI for Monitoring of Soyabean and Cotton*.

Sannier, C., McRoberts, R. E., & Fichet, L.-V. (2016). Suitability of Global Forest Change data to report forest cover estimates at national level in Gabon. *Remote Sensing of Environment*, *173*, 326–338. https://doi.org/10.1016/j.rse.2015.10.032

Scherr, S. J., White, A., & Kaimowitz, D. (2004). *A new agenda for forest conservation and poverty reduction: Making forest markets work for low-income producers*. Cifor.

Schure, J., Levang, P., & Wiersum, K. F. (2014). Producing Woodfuel for Urban Centers in the Democratic Republic of Congo: A Path Out of Poverty for Rural Households? *World Development*, *64*, S80–S90. https://doi.org/10.1016/j.worlddev.2014.03.013

Sharifi, A., Felegari, S., & Tariq, A. (2022). Mangrove forests mapping using Sentinel-1 and Sentinel-2 satellite images. *Arabian Journal of Geosciences*, *15*(20), 1593. https://doi.org/10.1007/s12517-022-10867-z

Sharma, R. C., Hara, K., & Tateishi, R. (2018). Developing Forest Cover Composites through a Combination of Landsat-8 Optical and Sentinel-1 SAR Data for the Visualization and Extraction of Forested Areas. *Journal of Imaging*, *4*(9), 9. https://doi.org/10.3390/jimaging4090105

Shimada, M., Itoh, T., Motooka, T., Watanabe, M., Shiraishi, T., Thapa, R., & Lucas, R. (2014). New global forest/non-forest maps from ALOS PALSAR data (2007–2010). *Remote Sensing of Environment*, *155*, 13–31. https://doi.org/10.1016/j.rse.2014.04.014

Shimizu, K., Ota, T., & Mizoue, N. (2019). Detecting Forest Changes Using Dense Landsat 8 and Sentinel-1 Time Series Data in Tropical Seasonal Forests. *Remote Sensing*, *11*(16), 16. https://doi.org/10.3390/rs11161899

Shrestha, B., Ahmad, S., & Stephen, H. (2021). Fusion of Sentinel-1 and Sentinel-2 data in mapping the impervious surfaces at city scale. *Environmental Monitoring and Assessment*, *193*(9), 556. https://doi.org/10.1007/s10661-021-09321-6

Somorin, O. A., Brown, H. C. P., Visseren-Hamakers, I. J., Sonwa, D. J., Arts, B., & Nkem, J. (2012). The Congo Basin forests in a changing climate: Policy discourses on adaptation and mitigation (REDD+). *Global Environmental Change*, *22*(1), 288–298. https://doi.org/10.1016/j.gloenvcha.2011.08.001

South Pole. (2020). *JNR Scenario 1 Tshopo Province, DRC* (VCS Version 3; JNR Baseline Description). https://verra.org/wp-content/uploads/2020/04/VCS-JNR-Baseline-Description_Tshopo_9MAR2020.pdf

Sun, L., Chen, J., Guo, S., Deng, X., & Han, Y. (2020). Integration of Time Series Sentinel-1 and Sentinel-2 Imagery for Crop Type Mapping over Oasis Agricultural Areas. *Remote Sensing*, *12*(1), 1. https://doi.org/10.3390/rs12010158

Tavares, P. A., Beltrão, N. E. S., Guimarães, U. S., & Teodoro, A. C. (2019). Integration of Sentinel-1 and Sentinel-2 for Classification and LULC Mapping in the Urban Area of Belém, Eastern Brazilian Amazon. *Sensors*, *19*(5), 5. https://doi.org/10.3390/s19051140

Tegegne, Y. T., Lindner, M., Fobissie, K., & Kanninen, M. (2016). Evolution of drivers of deforestation and forest degradation in the Congo Basin forests: Exploring possible policy options to address forest loss. *Land Use Policy*, *51*, 312–324. https://doi.org/10.1016/j.landusepol.2015.11.024

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, *8*(2), 127–150. https://doi.org/10.1016/0034-4257(79)90013-0

Tyukavina, A., Hansen, M. C., Potapov, P., Parker, D., Okpa, C., Stehman, S. V., Kommareddy, I., & Turubanova, S. (2018). Congo Basin forest loss dominated by increasing smallholder clearing. *Science Advances*, *4*(11), eaat2993. https://doi.org/10.1126/sciadv.aat2993

Umunay, P. M., Gregoire, T. G., Gopalakrishna, T., Ellis, P. W., & Putz, F. E. (2019). Selective logging emissions and potential emission reductions from reduced-impact logging in the Congo Basin. *Forest Ecology and Management*, *437*, 360–371.

UNEP. (2019). *Democratic Republic of the Congo: A primer on deforestation for Religious leaders and faith communities*. https://www.interfaithrainforest.org/s/Interfaith_CountryPrimer_DRC.pdf

UNEP. (2023). *Critical ecosystems: Congo Basin peatlands*. https://www.unep.org/news-and-stories/story/critical-ecosystems-congo-basin-peatlands

Vancutsem, C., Achard, F., Pekel, J.-F., Vieilledent, G., Carboni, S., Simonetti, D., Gallego, J., Aragão, L. E. O. C., & Nasi, R. (2021). Long-term (1990–2019) monitoring of forest cover changes in the humid tropics. *Science Advances*, *7*(10), eabe1603. https://doi.org/10.1126/sciadv.abe1603

Wang, J., Xiao, X., Liu, L., Wu, X., Qin, Y., Steiner, J. L., & Dong, J. (2020). Mapping sugarcane plantation dynamics in Guangxi, China, by time series Sentinel-1, Sentinel-2 and Landsat images. *Remote Sensing of Environment*, *247*, 111951. https://doi.org/10.1016/j.rse.2020.111951

Waser, L. T., Fischer, C., Wang, Z., & Ginzler, C. (2015). Wall-to-Wall Forest Mapping Based on Digital Surface Models from Image-Based Point Clouds and a NFI Forest Definition. *Forests*, *6*(12), 12. https://doi.org/10.3390/f6124386

Watson, J. E. M., Evans, T., Venter, O., Williams, B., Tulloch, A., Stewart, C., Thompson, I., Ray, J. C., Murray, K., Salazar, A., McAlpine, C., Potapov, P., Walston, J., Robinson, J. G., Painter, M., Wilkie, D., Filardi, C., Laurance, W. F., Houghton, R. A., … Lindenmayer, D. (2018). The exceptional value of intact forest ecosystems. *Nature Ecology & Evolution*, *2*(4), 4. https://doi.org/10.1038/s41559-018-0490-x

Welsink, A.-J., Reiche, J., De Sy, V., Carter, S., Slagter, B., Requena Suarez, D., Batros, B., Pena Claros, M., & Herold, M. (2023). Towards the use of satellite-based tropical forest disturbance alerts to assess selective logging intensities. *Environmental Research Letters*. https://doi.org/10.1088/1748-9326/acd018

World Bank. (2021). *Climate Change Knowledge Portal*. https://climateknowledgeportal.worldbank.org/country/congo-dem-rep/climate-data-historical

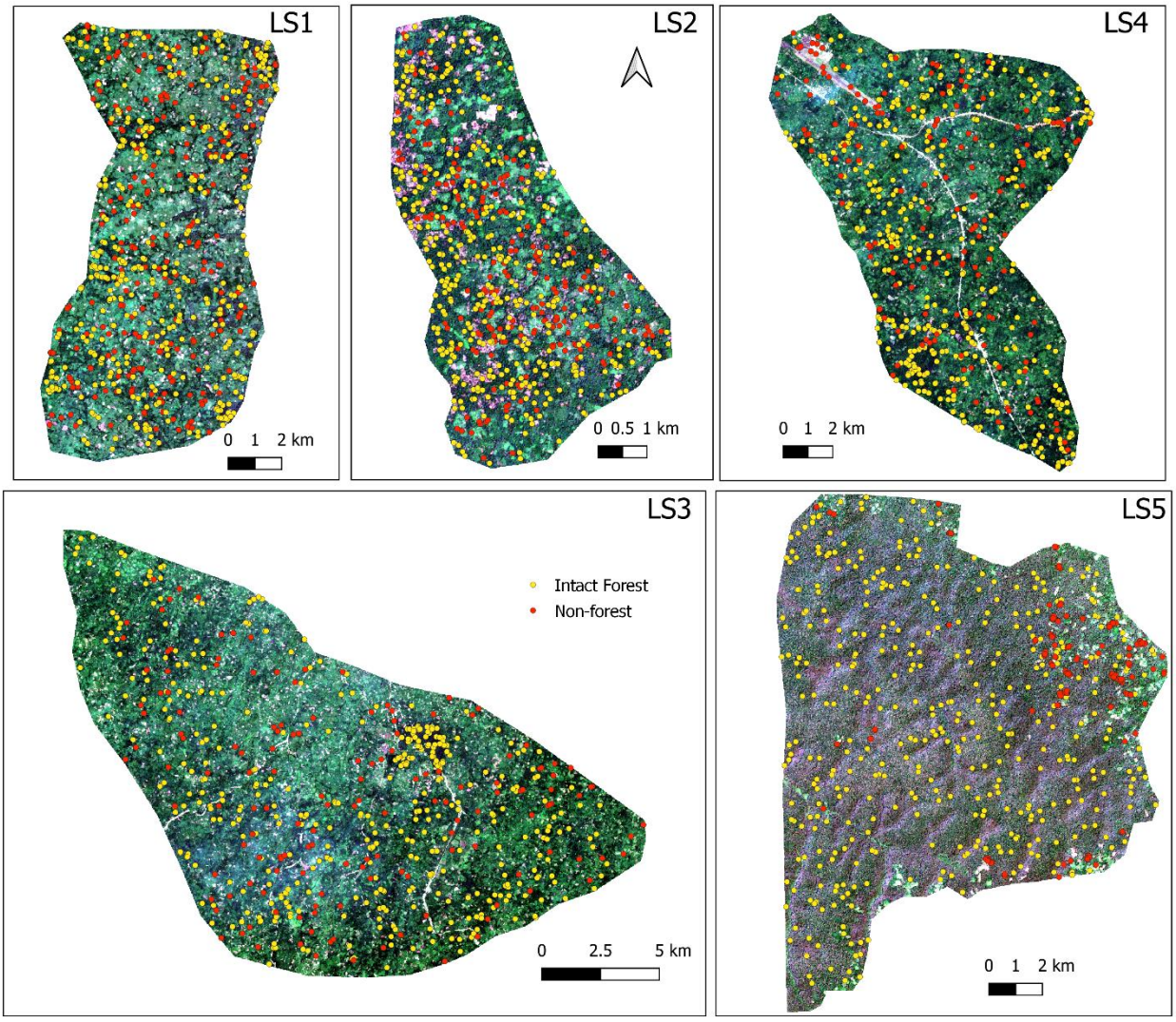
WUR. (2020). *RADD Forest Disturbance Alert*. Wageningen University & Research. https://www.wur.nl/en/research-results/chair-groups/environmental-sciences/laboratory-of-geo-information-science-and-remote-sensing/research/sensing-measuring/radd-forest-disturbance-alert.htm

Xu, L., Saatchi, S. S., Shapiro, A., Meyer, V., Ferraz, A., Yang, Y., Bastin, J.-F., Banks, N., Boeckx, P., Verbeeck, H., Lewis, S. L., Muanza, E. T., Bongwele, E., Kayembe, F., Mbenza, D., Kalau, L., Mukendi, F., Ilunga, F., & Ebuta, D. (2017). Spatial Distribution of Carbon Stored in Forests of the Democratic Republic of Congo. *Scientific Reports*, *7*(1), 1. https://doi.org/10.1038/s41598-017-15050-z

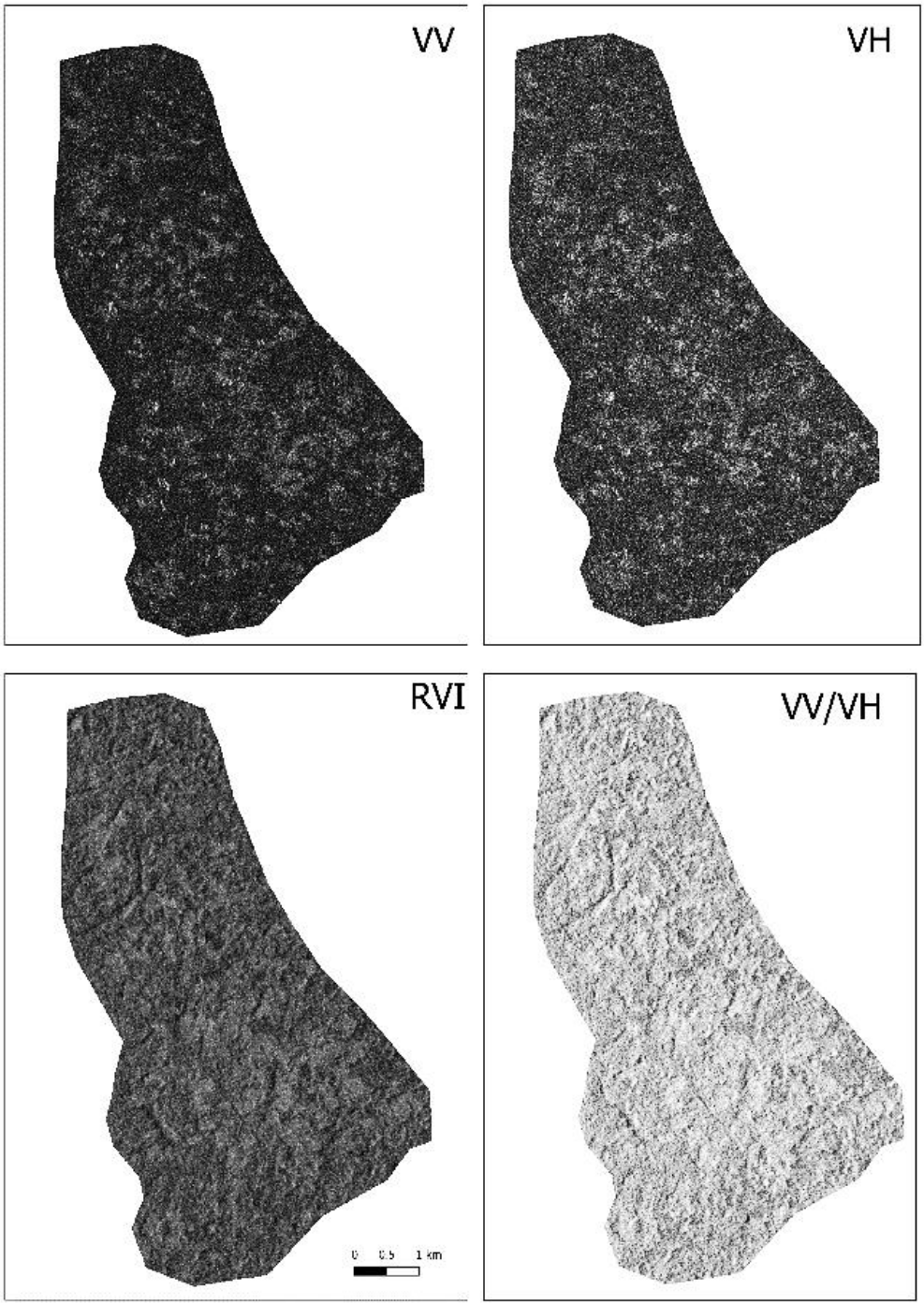
Zhou, L., Shao, Z., Wang, S., & Huang, X. (2022). Deep learning-based local climate zone classification using Sentinel-1 SAR and Sentinel-2 multispectral imagery. *Geo-Spatial Information Science*, *25*(3), 383–398. https://doi.org/10.1080/10095020.2022.2030654

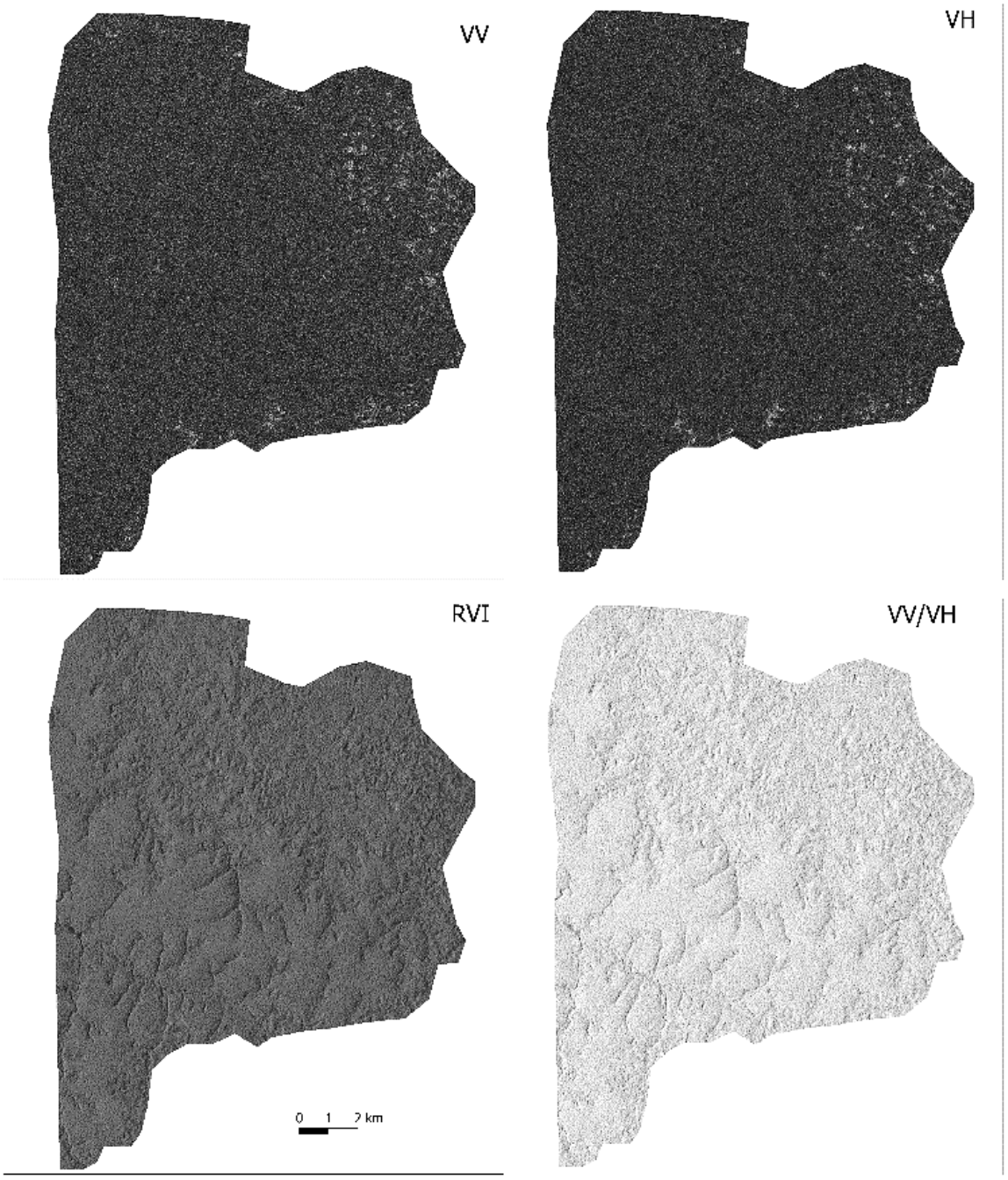
9. Appendices

9.1 Training samples in different landscapes

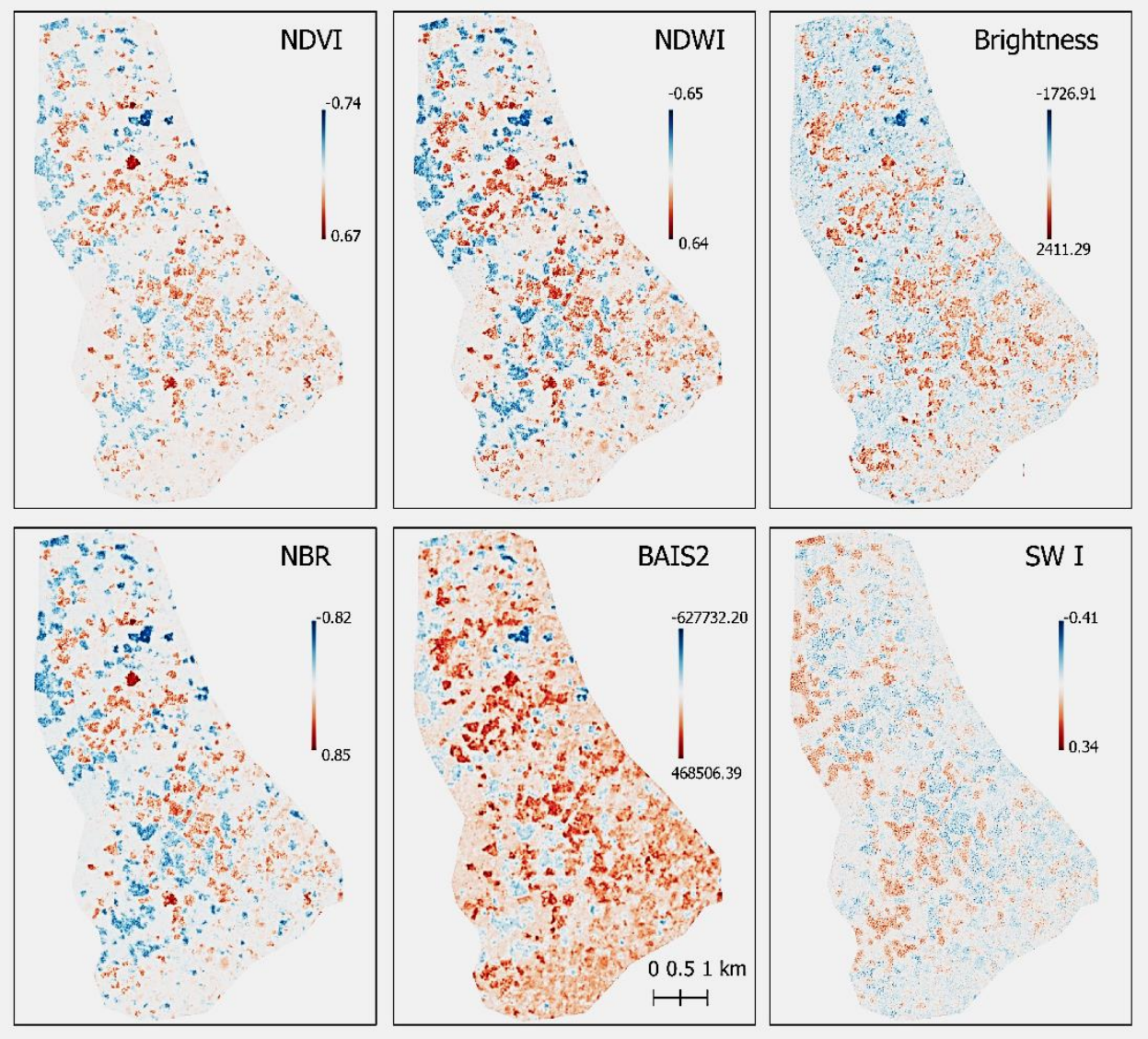


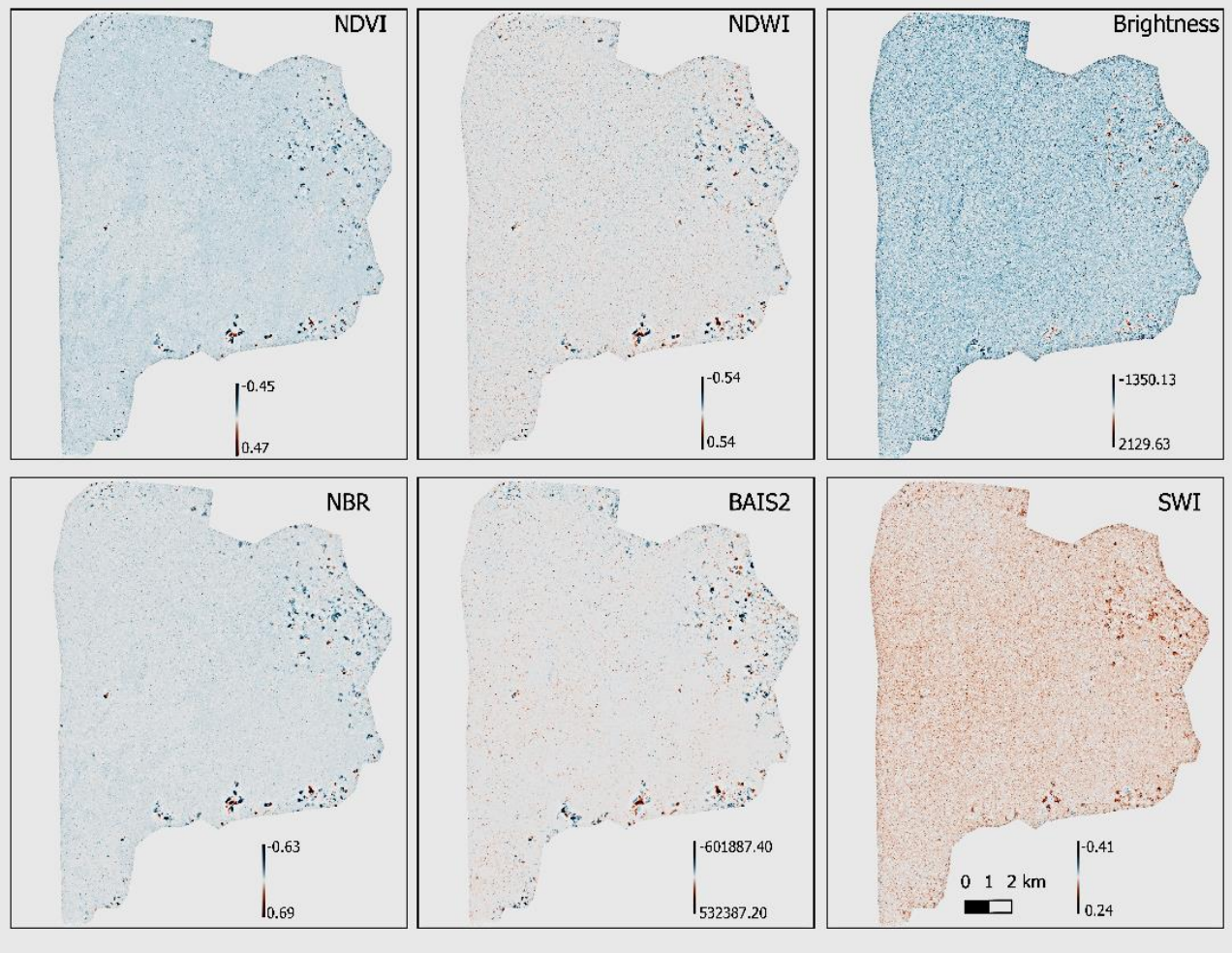
9.2 Example of S1 Metrics in LS-2 and LS-5



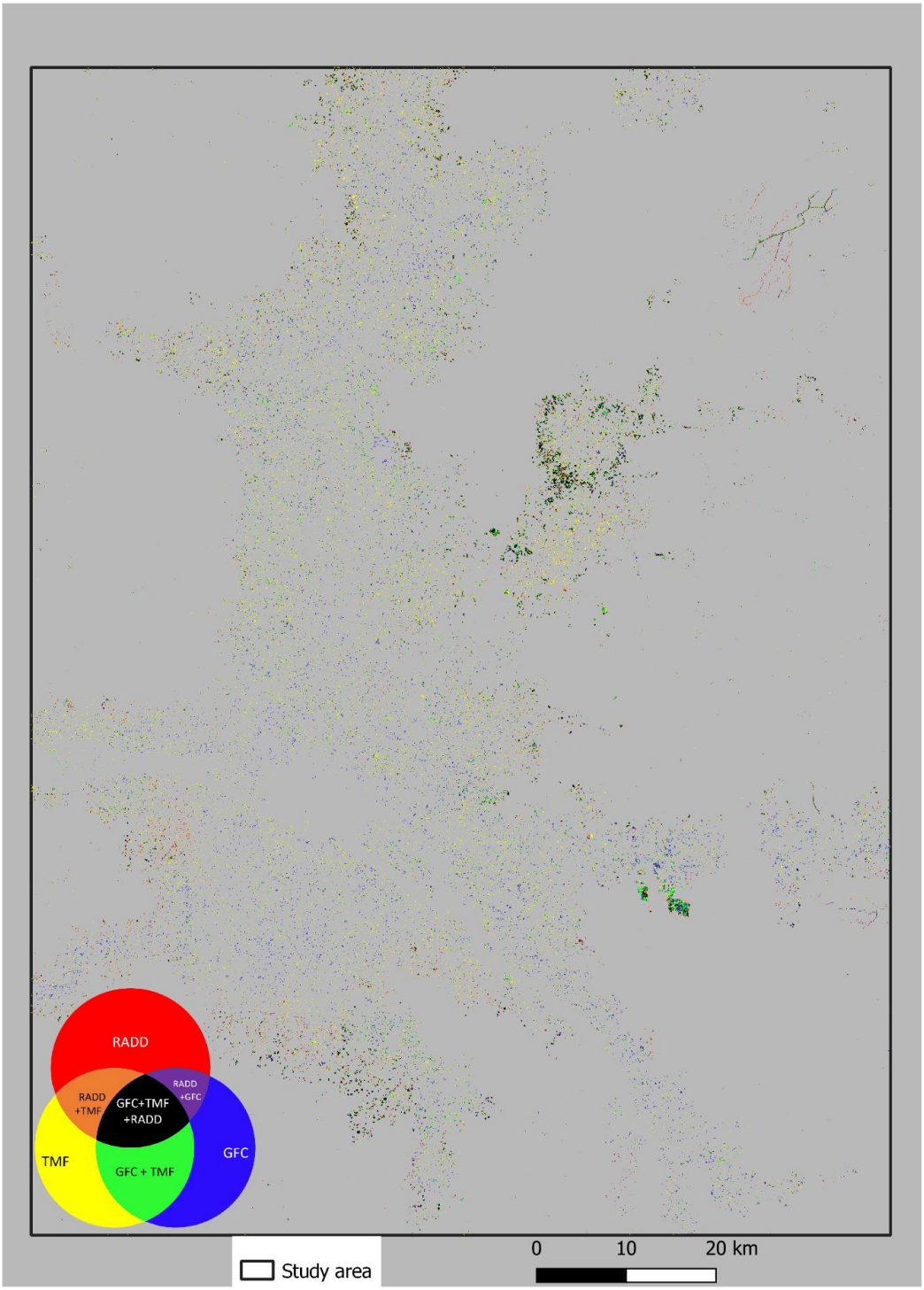

VV
VH
RVI
VV/VH
0 1 2 km

9.3 Example of Indices differences within 1 year in LS-2 and LS-5

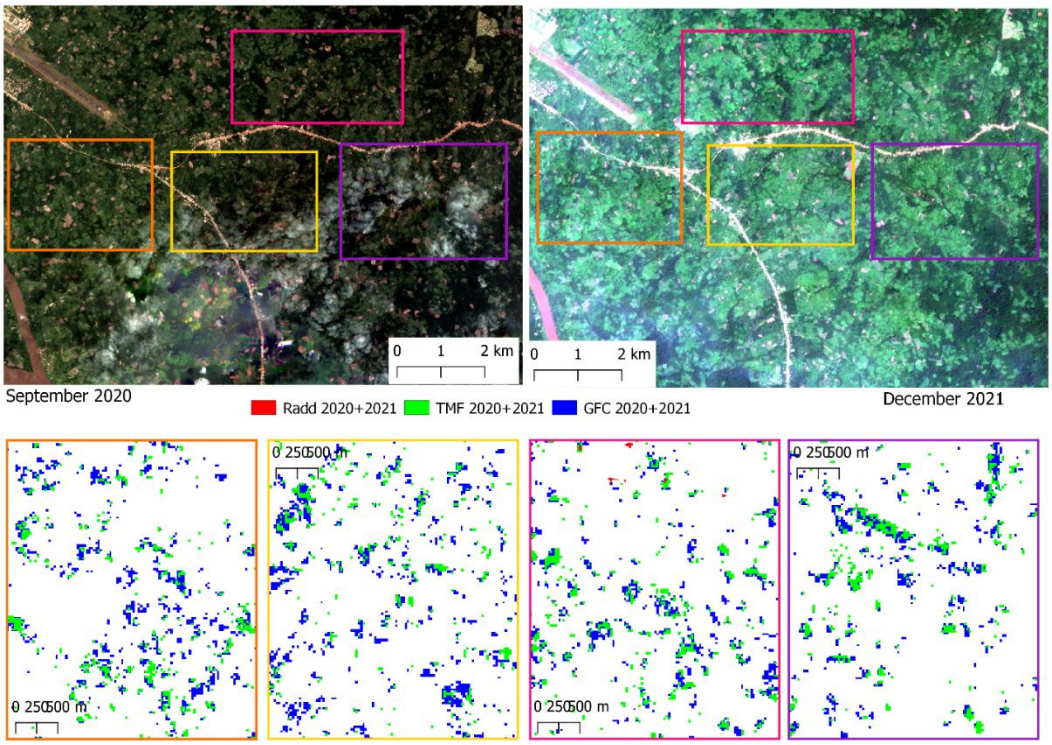



NDVI
-0.45
0.47
NDWI
-0.54
0.54
Brightness
-1350.13
2129.63
NBR
-0.63
0.69
BAIS2
-601887.40
532387.20
SWI
-0.41
0.24
0 1 2 km

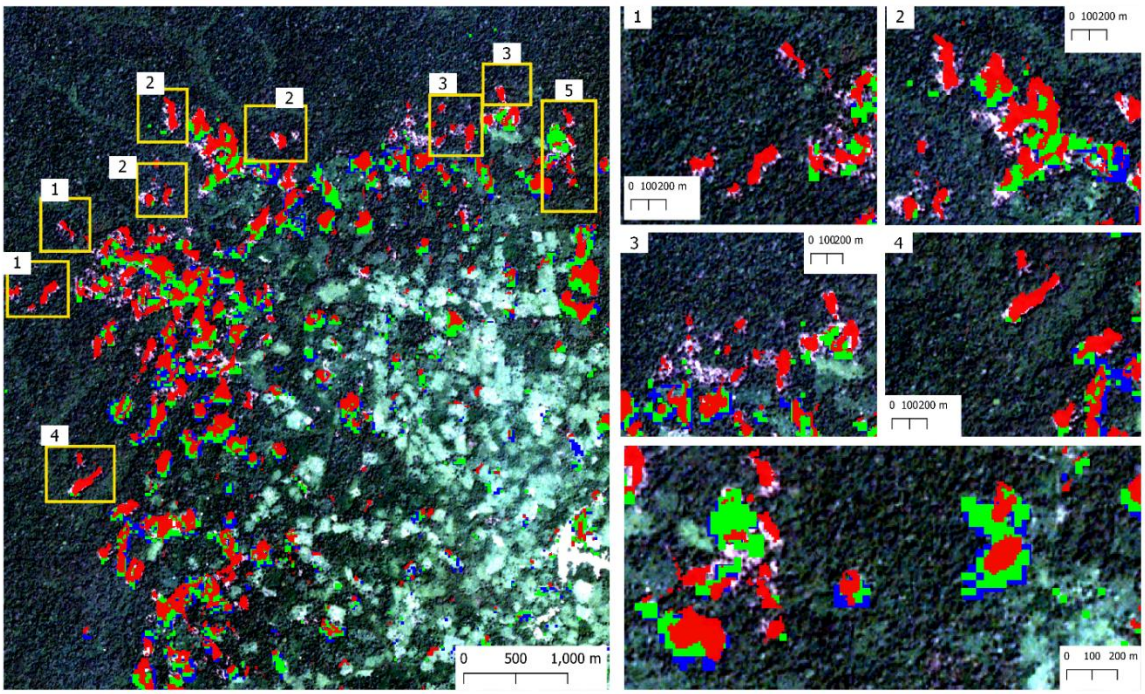
9.4 Comparison of optical and SAR-based deforestation products GFC+TMF+RADD for 2020



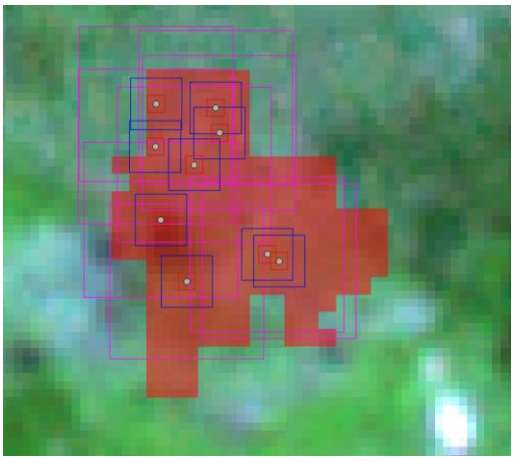
9.5 Optical products work better in mixed landscapes



9.6 SAR works better in forest dominated landscapes



9.7 Scenarios for LS-4 products precision analysis: Due to lower area coverage combined layer GFC+TMF+RADD, validation points are located closely



9.8 Classification confusion matrix in different landscapes

Predicted value	Reference value				
	LS1 2 classes	Forest	non-Forest	total	UA
	Forest	84	6	90	93.33
	non-Forest	17	34	51	66.67
	total	101	40	141	
	PA	83.17	85.00	OA	83.69
		Forest	non-Forest	total	UA
	Forest	82	6	88	93.18
	non-Forest	15	25	40	62.50
	total	97	31	128	
	PA	84.54	80.65	OA	83.59

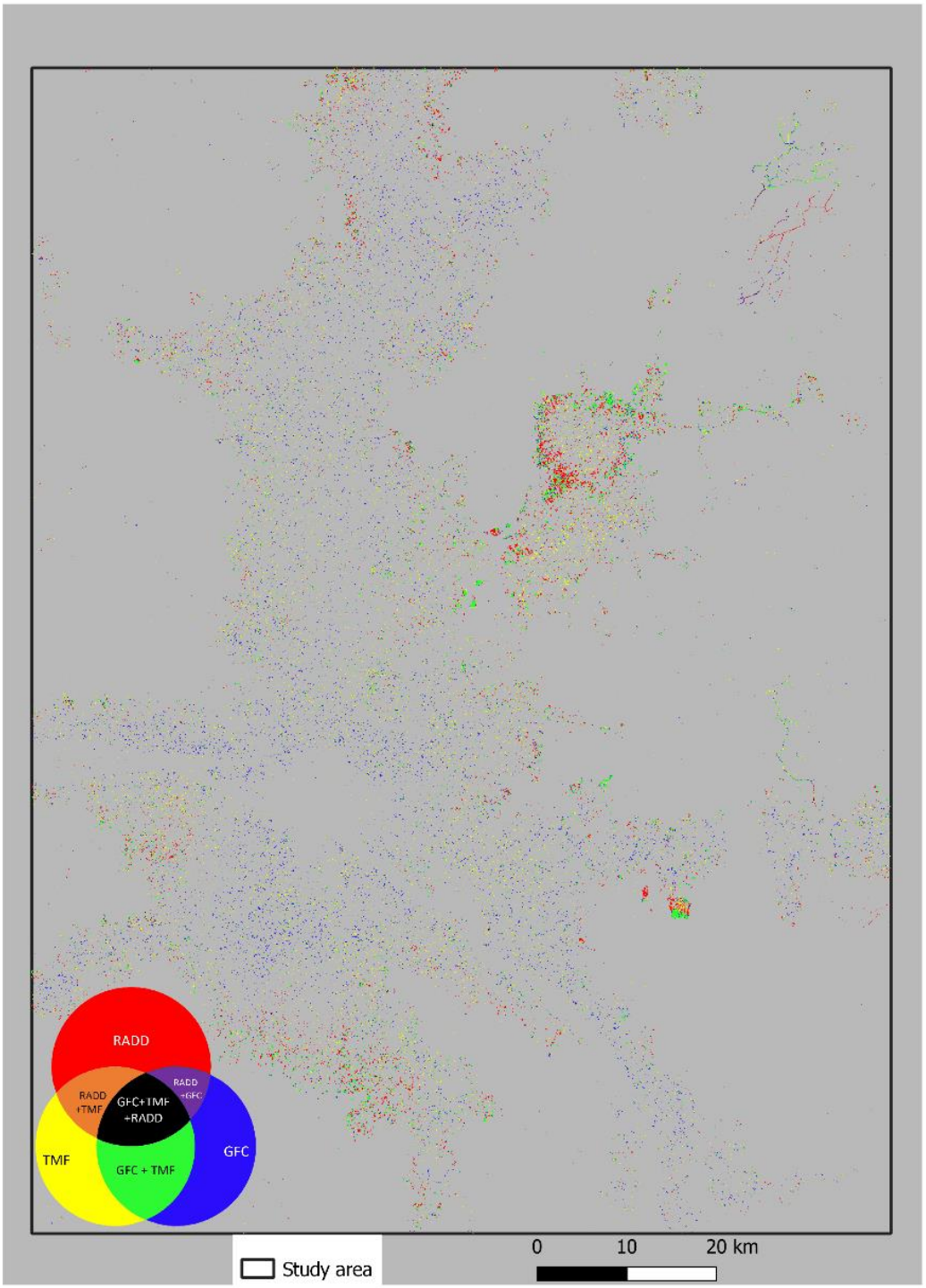
Predicted value	Reference value				
	LS4 2 classes	Forest	non-Forest	total	UA
	Forest	67	13	80	83.75
	non-Forest	12	23	35	65.71
	total	79	36	115	
	PA	84.81	63.89	OA	78.26
	LS4 3 classes	Forest	non-Forest	total	UA
	Forest	84	11	95	88.42
	non-Forest	10	33	43	76.74
	total	94	44	138	
	PA	89.36	75.00	OA	84.78

Predicted value	Reference value				
	LS5 S1	Forest	non-Forest	total	UA
	Forest	78	7	85	91.76
	non-Forest	18	24	42	57.14
	total	96	31	127	
	PA	81.25	77.42	OA	80.31
	LS5 S2	Forest	non-Forest	total	UA
	Forest	81	4	85	95.29
	non-Forest	7	35	42	83.33
	total	88	39	127	
	PA	92.05	89.74	OA	91.34
	LS5 S3	Forest	non-Forest	total	UA
	Forest	82	3	85	96.47
	non-Forest	6	36	42	85.71
	total	88	39	127	
	PA	93.18	92.31	OA	92.91

Reference value					
LS3 2 classes	Forest	non-Forest	total	UA	
Forest	80	2	82	97.56	
non-Forest	17	28	45	62.22	
total	97	30	127		
PA	82.47	93.33	OA	85.04	
LS3 3 classes	Forest	non-Forest	total	UA	
Forest	70	6	76	93.33	
non-Forest	14	27	41	66.67	
total	84	33	117		
PA	83.33	81.82	OA	82.91	

Predicted value	Reference value				
	LS2 S1	Forest	non-Forest	total	UA
	Forest	70	6	76	92.11
	non-Forest	23	16	39	41.03
	total	93	22	115	
	PA	75.27	72.73	OA	74.78
	LS2 S2	Forest	non-Forest	total	UA
	Forest	71	5	76	93.42
	non-Forest	10	29	39	74.36
	total	81	34	115	
	PA	87.05	85.29	OA	86.96
	LS2 S3	Forest	non-Forest	total	UA
	Forest	72	4	76	94.74
	non-Forest	10	29	39	74.36
	total	82	33	115	
	PA	87.80	87.88	OA	87.83

9.9 Comparison of optical and SAR-based deforestation products GFC 2020+TMF 2020+RADD for 2021



Mapping the Tropical Moist Forests Extent Based on Combination of Sentinel-1 and Sentinel-2 Satellites

Sourav Karmakar

Tropical moist forests are crucial for their biodiversity and ecological significance, but they are threatened by deforestation and forest degradation driven by increasing population growth, agricultural expansion, shifting cultivation, logging, and mining activities in an unsustainable manner. To effectively preserve these forests, up-to-date yearly forest extent maps are needed for research and development, carbon estimation and monitoring national forest status. However, existing remote sensing based forest cover maps are designed for wide areas, have coarse resolution and lacks details at landscape level.

In this study, we developed local forest extent maps for five different landscapes using high-resolution Sentinel-1 (S1) based optical and Sentinel-2 (S2) based SAR sensors. Two existing optical products GFC and TMF and one SAR-based product RADD are compared at five different landscapes. Their best combination describing the actual ground situation was chosen as training samples. One-year S1 time series of four SAR features and one-year difference of six multispectral S2 indices was used for forest extent classification. Classification based on only S2 was carried out in three landscapes dominated by heterogenous landcovers, where S1, S2 and their combination were used in two other landscapes dominated by forest area according to the analysis of existing products.

In the landscapes where both S1 and S2 are used, S1 achieved an overall accuracy of 75% to 80% while S2 achieved 87% to 91% of overall accuracy. Combining the two sensors improved overall accuracy by 1% to 2% compared to their single sensor products. Optical sensors detected forest, non-forest areas with higher confidence, but cloud-free images are scarce in tropical regions. On the other hand, although SAR performed with a lower accuracy, it can provide ground information regardless of all-weather condition. Thus, complementary use of Sentinel-1 and Sentinel-2 enhances the accuracy and temporal coverage in tropical forests. However, further investigation is needed to assess various data fusion techniques for combining optical and SAR data to fully leverage their potential benefit in tropical forests.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

Faculté des bioingénieurs

Croix du Sud, 2 bte L7.05.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/agro**