MONITORING FOREST COVER CHANGE USING RADAR DATA

A Case Study of Gombe Forest in Kigoma

MUFURUKI ADVERA

A Dissertation Submitted to the Department of Geospatial Sciences and Technology in Partially Fulfilment of the Requirements for the Award of Bachelor Science in Geoinformatics (BSc. GI) of Ardhi University

CERTIFICATION AND COPYRIGHT

The undersigned certify that they have read and hereby recommend for acceptance by the Ardhi University dissertation titled "**Monitoring Forest Cover Change Using Radar Data: A Case Study of Gombe Forest in Kigoma.**'' in partial fulfillment of the requirements for the award of degree of Bachelor of Science in Geoinformatics at Ardhi University.

………………………………………… ………………………………….

Dr. Anastazia Msusa (Supervisor)

Date………………………………

Mr. Iriael Mlay (Supervisor)

Date………………………………

DECLARATION AND COPYRIGHT

I, MUFURUKI ADVERA hereby declare that, the contents of this dissertation are the results of my own findings through my study and investigation, and to the best of my knowledge they have not been presented anywhere else as a dissertation for diploma, degree or any similar academic award in any institution of higher learning.

………………………………

MUFURUKI ADVERA

22694/T.2019

(Candidate)

ACKNOWLEDGEMENT

Accomplishment of this research project has been a result of cooperation and support from several people but mostly by the grace of the Almighty God.

I would like to express my sincere gratitude to my supervisors Dr. Msusa and Mr. Mlay for their encouragement, continuous support and challenges that helped to shape this dissertation.

I would like to extend this appreciation to the staff members of Geospatial Science and Technology Department (DGST) for their moral and academic support, my advantageous classmates at Ardhi University and all my close friends who always supported and gave me full attention in solving different problems concerning my dissertation.

Last but not least, my heartfelt gratitude goes to my family for their unconditional love, prayers and support, but also for selflessly encouraging me to explore directions in life and seek my own destiny. I owe them a lot.

DEDICATION

Glory and honor be to the Almighty God for his protection from the beginning of this dissertation up to the end. This dissertation is dedicated to my parents Mr. & Mrs. Mufuruki, my supervisors Dr. Msusa and Mr. Mlay. You are all of great importance to me. God bless you.

ABSTRACT

Forest cover change is a significant environmental issue that has received important attention in recent years. The changes in forest cover in many parts of the world lead to increase the accumulation of atmospheric carbon and thus accelerate the process of global warming. Monitoring forest cover change is crucial for assessing ecological changes and targeting conservation efforts. Optical remote sensing has been used to monitor forest cover changes over four decades but its application is limited because of the presence of cloud coverage on the images. Recently availability of several space borne synthetic aperture radar (SAR) missions has widened the scope of utilizing Radar images for monitoring of forest cover change. Radar data has emerged as a valuable tool for monitoring forest cover change due to its ability to penetrate through clouds, vegetation, and soil. This research aimed at investigating the use of Radar data to fill data gap caused by limitation of optical remote sensing technique in monitoring the changes of Gombe Forest cover in Kigoma. In this study, Google Earth Engine was used to analyse the data on how Forest cover changes as time goes, from data collection, processing and land cover generation. Land cover and graphs from 2017 to 2021 were generated to meet the study's objectives, using Optical (Sentinel 2) MSI, Level-1C images and dual polarized (VV and HV) Radar (Sentinel) 1 SAR GRD: C- band data. Five different categories of land covers (Permanent water bodies, Tree Cover, Shrubland, Grassland and Bare/ Sparse Vegetation) were recognized on sentinel 2 and both SAR scenes. Both the images were classified using Random Forest algorithms. The classification accuracy was computed from the randomly selected independent validation pixels where over 88.64% for optical and 74.46% for Radar data overall accuracies were obtained. The results of this study reveal that Optical data and Radar data can be used interchangeably in forest cover monitoring hence helps in understanding the applicability of SAR to map and quantify (Monitor) forest cover changes in cloud prone areas (Gombe Forest).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATION AND ACRONYMS.

ARD	Analysis Ready Data
ESA	European Space Agency
GEE	Google Earth Engine
GHG	Green House Gas
GIS	Geographical Information Systems
GRD	Ground Range Detection
MSI	Multispectral Instrument
QGIS	Quantum Geographic Information System
RADAR	Radio Detection and Ranging
REDD	National Oceanic and Atmospheric Administration
RS	Remote Sensing
SAR	Synthetic Aperture Radar
SDG	Sustainable Development Goals
SNAP	Sentinel Application Platform.
TIFF	Tagged Image Format
TOA	Top of Atmosphere
USGS	United States Geological Survey
VH	Vertical-Horizontal Polarization
VV	Vertical-Vertical Polarization

CHAPTER ONE

INTRODUCTION

1.1 Background

Forest is a major resource and play vital role in maintaining the ecological balance, environmental set up and in the Earth's climate system, in a number of different ways like producing and regulating the world's temperatures (Kurnar, 2011). Forests provide essential ecosystem services lead to human live-hoods. However, the distribution and condition of tropical forests are under significant pressure, causing reduction and risking biodiversity loss (Elikana, 2021).

Over utilization of forest resources has resulted in the depletion and changes in forest cover are the matter of global concern due to its ability of promoting role in carbon cycle. (Kurnar, 2011) Africa is home to some of the world's most magnificent tropical forests. With more than 60 million people dwelling within or near these forests, they are relied upon for many ecosystem systems, with live-hoods dependent on them for providing food, medicinal plants, fuel, fibres, and non-timber forest products. At the same time, they are also important for societal and cultural purposes (De Wasseige, 2013). Despite the importance of these forests, the spatial and temporal variation in tropical forest cover has raised the dynamic state and increased forest cover decline. However, throughout many parts of Africa, forest cover change remains poorly understood due to field-based monitoring challenges and a lack of remote sensing studies. Government policies have often failed to prevent illegal forestry activities due to the absence of defined tenure, which has also increased over-utilisation. This is also linked to the lack of national capacity for monitoring and reporting deforestation, especially in many sub-Saharan Africa countries and hence remains a barrier. (Van Passel, 2020).

The situation is more complicated in developing countries such as the United Republic of Tanzania, where many households are highly dependent on forest resources. Despite the value of forests, changes in use patterns pose a noticeable threat to forest resource sustainability on both socio-economic and ecological functioning. This has led to the increased scarcity of forest resources, further aggravated by the continuing high deforestation rate and associated with a change in climatic conditions experienced in many parts of Tanzania (Hansen, 2013). Therefore, the need for data and information on the state of Tanzania's forest resources is of increasing importance. Yet, Tanzania's forest resource status and trends are mostly unknown, with current data being fragmented and outdated (Vesa, 2010).

Tanzania is currently undergoing significant forest cover changes but monitoring is limited, in part due to a lack of remote sensing knowledge tools and methods. For instance, Forest loss identified for 2019 was found to be 157,204 hectares, with an overall accuracy of 82%. These forest losses within Tanzania have already triggered ecological problems and alterations in ecosystem. (Elikana, 2021). The lack of institutional capacity has largely constrained data reliability on Tanzanian forest resources, with inadequate national-wide forest monitoring coverage using an Earth Observation-based system (Anderson, 2017).

According to (Elikana, 2021) who studied Forest Monitoring System for Tanzania, Gombe Forest being part of Tanzanian forests in Kigoma region, concluded that Monitoring of Forest is still the challenge as seasonality changes and persistent cloud cover in Tanzania create low data availability and excessive gaps (missing data).

Using freely accessible satellite data and advanced remote sensing methods can provide a cost-effective and timely approach to achieving systematic wall-to-wall information for Tanzania's forests. This information is required to support national policy processes aimed at improving sustainable forest management at the same time addressing the issues of Reduced Emissions from Deforestation and Forest Degradation (REDD+) and Green House Gas (GHG) as international reporting obligations. Also, the 2030 Agenda for Sustainable Development Goals (SDGs) for enhancing life on land and its targets through combating deforestation (Anderson, 2017).

However, monitoring of forest cover change is crucial for assessing ecological changes and targeting conservation efforts, the most biodiverse forests on the planet are also the most difficult to monitor remotely due to their frequent cloud cover. Optical sensors measure reflected solar light, thus only function in the day time and can't penetrate cloud. The use of Optical remote sensing technique in monitoring forest cover change like the provision of satellite-based data (Landsat data series) has come a long way since the 1970 and it has been used in monitoring forest cover change but are most difficult and still the challenge as seasonality changes and persistent cloud cover in Tanzania create low data availability and excessive gaps (missing data) in the Landsat archive (Elikana, 2021)**.**

Active remote sensing, Radar (Radio Detection and Ranging) is the remote sensing technique that based on the use of radio waves Example Synthetic Aperture Radar (SAR) imagery like Sentinel 1. Radar data provides 24-hour all weather Earth observation and has ability to penetrate atmospheric condition (cloud free) providing near real time visibility in cloud covered area, both day and night thus due to the weather capability of the radar instruments as the potential application of remote sensing, can be used in environmental monitoring example forest cover change monitoring.

In order to adapt to forest cover change, sufficient and cost-effective method should be provided to obtain wall-to-wall information on the forest extent and associated changes of Gombe. To begin to resolve the problem, we develop and implement an effective and efficient approach in monitoring forest changes of Gombe Forest in Kigoma Region.

Therefore, this study intends to determine the use Radar data to fill the data gap caused by limitation of optical remote sensing technique in monitoring the changes in forest cover and enhancing conservation programs and support efforts to save the Gombe Forest.

1.2 Statement of the Research Problem

Monitoring forest cover change is crucial for assessing ecological changes and targeting conservation efforts. Optical remote sensing data has been used for decades (more than 40 years) but its application is limited due to the presence of clouds on the images leading to data gaps. The presence of data gaps leads to lack of some important information on the quantification of spatial and temporal changes in forest cover which is an essential component of forest monitoring programs. Since monitoring of Forest cover changes requires systematically collected data, this study tend to investigate the use of radar data to fill the data gap.

1.3 Objectives

1.3.1 Main objective

The main objective of this study is to investigate the use of Radar data to fill data gap in monitoring Gombe Forest cover change in Kigoma.

1.3.2 Specific objectives

The specific objectives to achieve the main objective of the study are:

i. To develop a long-term Forest cover of Gombe using Optical images and Radar Images from 2017 to 2021.
ii. To evaluate the performance of Radar data in Monitoring forest cover change of Gombe Forest.
iii. To develop a long-term Forest cover distribution map of the Gombe Forest from 2017 to 2021 using Radar images.

1.4 Research Questions

Specifically, this study will focus on the following research questions:

i. How has the forest cover of Gombe been changing from 2017 to 2021 for both optical and radar images?
ii. How effective is the use of radar data in detecting and monitoring changes in forest cover of Gombe Forest over time?
iii. What is the Forest cover distribution of Gombe Forest from 2017 to 2021 using Radar data?

1.5 Significance of the Research

The significance of this research includes:

i. The study provides a step by step (systematic) methodology for using radar data in monitoring of land cover changes in forested and cloud-prone regions like Gombe Forest in Kigoma with great promise for application to improved monitoring of other forest characterized by high cloud cover.
ii. The study provides essential information to guide policy formulation and implementation in protecting forests with better decision making in government programmes and other forest protection for instance, 2030 Agenda for Sustainable Development Goals (SDGs) over fighting deforestation.
iii. The study demonstrates the potential application of open-free software, freely available satellite data, and advanced remote sensing techniques (understanding the application of radar data to map and quantify forest cover changes) to provide a cost-effective method to obtain wall-to-wall information on the forest extent and associated changes of Gombe. It will bridge the information gap and knowledge concerning the use of radar data for generating forest information about status, extent and types. Therefore, the information is vital for developing practical, long-term plans in monitoring to enhance conservation programs system and support efforts to save the last leftovers of Kigoma forests and Tanzania's socio-economic growth at large.

1.6 Beneficiaries of the Research

i. The beneficiaries of this research include:
ii. TANAPA (Tanzania National Parks) as Gombe forest is within Gombe stream National Park.

iii. Decision makers, planners and environmental scientists in Kigoma region and Tanzania at large will be able to come up with proper adaptation measures to cope with forest cover change after quantification of spatial and temporal land cover change in the forest. (sustainable development goals)
iv. The Government will be able to plan and prepare strategies to enhance conservation programs and support efforts to save the last leftovers of Kigoma forests.
v. Forest cover information is needed to support the national forest policy to sustainably manage, conserve, restore and utilize forests and associated resources for Tanzania's socio-economic growth.

1.7 Dissertation Organization

This dissertation consists of five chapters, explaining in details what the research deals with, how it was conducted, methods used and results obtained in investigating the use of Radar data to fill the gap caused by optical data in monitoring forest cover change.

Chapter 1 explains the background of the study, which gave rise to the problem. The objectives, research questions, significance and beneficiaries of the research. Chapter 2 gives the review of the study including different concepts related to monitoring forest cover change using Radar data. Chapter 3 covers all the techniques, methods and procedures used to achieve the results of the study. Chapter 4 provides the analysis and discussion of the results. It shows the results obtained in this study and the results obtained by previous studies and Chapter 5 contains the conclusion and recommendations for the study.

CHAPTER TWO

LITERATURE REVIEW

2.1. Overview

This chapter reviews various issues, theories and researches done in the study, with the purpose of adding knowledge and familiarizes the researchers with relevant information about the problem being studied. In this chapter, the major concern is made on the various definition and views related to the research.

2.2. Description of Forest cover changes

2.2.1 Forest

Forests are one of the most important ecosystems on Earth. They provide vital ecological, economic, and social benefits, including oxygen production, carbon sequestration, biodiversity conservation, and wood and non-wood forest products (Shimada et al, 2015). Forests in Tanzania play a significant role in the daily livelihoods. They are an important source of energy for cooking, building timber, traditional medicine, tourism, fodder, water catchments, shelter for wildlife and estuaries for fish breeding areas. Also, these forests have high biodiversity, containing over 10,000 plant species, hundreds of which are nationally endemic, 724 species of flora and fauna identified as threatened in the IUCN Red List, and 276 species of flora and fauna classified as endangered (IUCN, 2013)

2.2.2 Forest cover change

Forest cover refers to the extent or area of land that is covered by trees or woody vegetation. Forest cover change, refers to the process of changes in the extent, structure, and composition of forest ecosystems over time (Shimada et al, 2015). Forest cover change can occur due to various factors such as human activities, natural disasters, and climate change.

2.2.3 Monitoring forest cover change

Monitoring forest cover change is essential for sustainable forest management and conservation. It enables the assessment of the state of forest ecosystems, identification of areas under threat, and development of appropriate strategies for forest restoration and management (Saatch et al, 2011).

2.3. Remote Sensing.

Remote sensing is the science and art of acquiring information about the Earth's surface without actually being in contact with it (Campbell, 2002). This is done by sensing and recording reflected or emitted energy and processing, analysing, and applying that information. Remote sensing using is an effective tool for monitoring forest cover change. Methods for monitoring forest cover change using Radar data involves (Shimada et al, 2015);

Synthetic Aperture Radar (SAR) data is a remote sensing technique that uses radar signals to create high-resolution images of the Earth's surface. SAR can penetrate deeper into the forest canopy and provide information on the structure and composition of forest ecosystems, including the height, biomass, and density of trees.

Interferometric SAR (InSAR) is a technique that uses two or more radar images to create a digital elevation model of the Earth's surface. InSAR can provide information on the topography and surface deformation of forest ecosystems, which can be used to detect changes in the forest cover.

2.4. Sensor system.

The sensor systems are simply the eyes of the satellites that view and record the scene. Sensors are the special instruments mounted on the platforms (satellite) usually having sophisticated lenses with filter coatings, to focus the area to be observed at a specific region of electromagnetic spectrum (Campbell, 2002). Solar radiation is the main source of electromagnetic radiation and is a combination of several wavelengths such as gamma ray, x-ray, visible, infrared, and thermal and microwaves. Sensor systems mainly operate in the visible, infrared, thermal and microwave regions of EMR. Based on the source of EMR sensor systems can be broadly classified as passive or active systems (Joseph, 2005):

2.4.1. Passive sensors.

Are sensors that can only be used to detect energy when the naturally occurring energy is available. For all reflected energy, this can only take place during the time when the sun is illuminating the Earth. There is no reflected energy available from the sun at night. Energy that is naturally emitted (i.e. thermal infrared) can be detected day or night, as long as the amount of energy is large enough to be recorded (Jensen, 2000).

2.4.2. Active sensors.

These sensors provide their own energy source for illumination. The sensor emits radiation, which is directed toward the target to be investigated. The radiation reflected from that target is detected and measured by the sensor. Active sensors have the ability to obtain measurements anytime, regardless of the time of day or season. However, active systems require the generation of a large amount of energy to adequately illuminate targets. Some examples of active sensors are a laser flour sensor and synthetic aperture radar (SAR) (Jensen, 2000).

2.5 Optical remote sensing concept

Optical remote sensing is the process of acquiring information about the Earth's surface using sensors that capture data in the visible, near-infrared, and shortwave-infrared regions of the electromagnetic spectrum. The sensors collect reflected solar radiation from the Earth's surface and atmosphere, and the data is used to create images or maps of the surface (Jensen, 2007).

Optical remote sensing is a powerful tool for studying the Earth's surface and environment. It is widely used in various fields, such as agriculture, forestry, geology, oceanography, urban planning, and environmental management. Some of the key applications of optical remote sensing include (Jensen, 2000):

i. Land cover and land use mapping to classify different land cover and land use types, such as forests, croplands, water bodies, and urban areas. This information is used for monitoring changes in land use and land cover over time, which is critical for environmental management and conservation.
ii. Vegetation and Forest monitoring to monitor growth and health of vegetations, including monitoring vegetation stress due to drought, pests, or diseases. This information is useful for agricultural planning, forestry management, and monitoring of natural habitats.
iii. Oceanography to study the properties of the ocean, including ocean temperature, colour and chlorophyll concentration. This information is used for monitoring ocean productivity, fisheries, and ocean currents.
iv. Atmosphere to study atmospheric properties such as aerosols, clouds, and gases. This information is used for weather forecasting, climate modelling, and air quality monitoring.

2.5.1 Sentinel-2

Sentinel-2, launched as part of the European Commission’s Copernicus program on June 23, 2015, was designed specifically to deliver a wealth of data and imagery. The satellite is equipped with an opto-electronic multispectral sensor for surveying with a sentinel-2 resolution of 10 to 60 m in the visible, near infrared (VNIR), and short-wave infrared (SWIR) spectral zones, including 13 spectral channels, which ensures the capture of differences in vegetation state, including temporal changes, and also minimizes impact on the quality of atmospheric photography. The orbit is an average height of 785 km and the presence of two satellites in the mission allow repeated surveys every 5 days at the equator and every 2-3 days at middle latitudes (ESA, 2015).

The Sentinel-2 data provides GMES (Global Monitoring for Environment and Security) program, jointly implemented by the EC (European Commission) and ESA (European Space Agency) services related, for example, to land management, agricultural production and forestry, and monitoring of natural disasters and humanitarian operations (ESA, 2015).

2.5.2 Sentinel 2 image pre-processing

Image pre-processing is the preparation of image for subsequent analysis, correction of deficiencies example clouds and noises in the image, and removal of flaws. When an image is being pre-processed, some features are enhanced. This qualifies the image for further processing (Sonka et al., 2008).

2.6 Radar remote sensing concept.

Radar (Radio Detection and Ranging) is the most common type of active microwave remote sensing. Radar system sends out pulses of microwave energy toward a target and detects the energy that is reflected back. A major advancement in radar remote sensing has been the improvement in azimuth resolution through the development of synthetic aperture radar (SAR) systems (Jensen, 2000). Synthetic aperture radar (SAR) is the sides looking airborne radar system combine radar and signal processing to form high-resolution backscatter images and have the following advantages:

- Active microwave energy penetrates clouds and can be an all-weather remote sensing system.
- Synoptic views of large areas, for mapping at 1: 25,000 to 1: 400,000; cloud-covered countries may be imaged.
- Coverage can be obtained at user-specified times, even at night.
- May penetrate vegetation, sand, and surface layers of snow.
- Enables resolution to be independent of distance to the object, with the size of a resolution cell being as small as 1 x 1 m.
- Images can be produced from different types of polarized energy.

Polarization refers to the orientation of electromagnetic waves in space. In remote sensing, polarization is used to enhance the information content of the signal received by the sensor

(Lee et al, 2009). Based on this, four possible combinations of transmit and receive polarization exists:

- HH - for horizontal transmit and horizontal receive
- VV - for vertical transmit and vertical receive
- HV - for horizontal transmit and vertical receive, and
- VH - for vertical transmit and horizontal receive.

The following are the types of polarimetric radar data (Lee et al, 2009).

i. Single polarization refers to radar data that measures only one component of the electromagnetic wave, either the horizontal or vertical component. This type of data provides limited information about the target being observed.
ii. Dual polarization refers to radar data that measures both the horizontal and vertical components of the electromagnetic wave. This provides more detailed information about the target, such as its orientation, structure, and composition.

2.6.1 Sentinel-1

The Sentinel-1A and Sentinel-1B satellites formed the first part of the Copernicus space mission, launched by the European Space Agency in 2014 and 2016 respectively. They offer a 6-day repeat cycle at the equator and a spatial resolution down to 5m (Haq, 2012).The Sentinel-1 is equipped with twin polar orbiting satellites designed to provide a spatial data for environment and security warranting, global economic and business growth. The satellites are to operate day-and-night and perform a synthetic aperture with radar imaging. Sentinel-1 bands allow to get imagery in all weather conditions.

There are four exclusive acquisition modes produced using satellite bands, unique to the Sentinel-1 (Haq, 2012):

- Strip map (SM)
- Interferometric wide swath (IW)
- Extra wide swath (EW)
- Wave (WV)

This research has employed Interferometric Wide Swath (IW)

Interferometric Wide Swath (IW): IW mode, short for interferometric wide swath mode is the next one to consider. It combines swath width of 250 km with a moderate resolution of 5 m by 20 m, so appears to be the most used over land as helps to spot a target on the ground.

Data Products

Sentinel 1 products involves the following products;

- Level-0 Products
- Level-1 Products
- Level-2 Products

This research has employed Level-1 ground range Detected geo-referenced products (GRD

Level-1 ground range Detected geo-referenced products (GRD) is the type of Level-1 product using focused data which projected to ground range, then detected and multi-looked. All data is projected to ground range using an Earth ellipsoid model corrected with specified terrain height which maintains the original satellite path direction and includes complete georeferencing information (Haq, 2012). The result has nearly square spatial resolution and square pixel spacing with reduced speckle.

2.6.2 Sentinel 1 Image pre-processing.

Referred to as image restoration and rectification, are intended to correct for sensor and platform specific radiometric and geometric distortions of data. Radiometric corrections may be necessary due to variations in scene illumination and viewing geometry, atmospheric conditions, and sensor noise and response. Each of these vary depending on the specific sensor and platform used to acquire the data and the conditions during data acquisition (Russ, 1995). The reason for image pre-processing is due to the occurrence of the following distortions;

i. Surface roughness – This measured in cm where strongly influences the strength of the returning signal. A surface is rough if the modified Rayleigh criteria give $h > \lambda/(4.4\sin\gamma)$ where h is the height in cm of surface objects and γ is the depression angle. Rough surfaces return a strong signal. The surface is smooth if $h < \lambda/(25\sin\gamma)$ and acts as a specular reflector, returning a weak or non-existent signal.

ii. Image foreshortening- All terrain that has a slope inclined toward the radar will appear compressed or foreshortened relative to slopes inclined away from the radar.

iii. Image layover - an extreme case of image foreshortening, in which for instance the top of a mountain can appear closer than its base.

iv. Shadowing – refer to the situation whereby portions of the ground are in the beam's shadow and totally obscured.

v. Image speckle – This caused by constructive and destructive interference, and hence random bright and dark areas in a radar image. The speckle can be reduced by processing separate portions of an aperture and recombining these portions so that interference does not occur.

2.7. Image Classification.

Image classification is the process of categorizing and labelling group of pixels or vectors within an image based on specific rules. It extracts information from the multiband raster image, For example in land cover classification from the satellite images. Digital image classification uses the spectral information represented by the digital numbers in one or more spectral bands, and attempts to classify each individual pixel based on this spectral information. This type of classification is termed spectral pattern recognition. In either case, the objective is to assign all pixels in the image to particular classes or themes e.g., water bodies, forest, vegetation, grassland, etc. The resulting classified image is comprised of a mosaic of pixels, each of which belong to a particular theme, and is essentially a thematic "map" of the original image (Lillesand, 1994). There are two methods of image classification, which are:

i. Supervised classification
ii. Unsupervised classification

2.7.1. Supervised Classification.

With Supervised classification a user can choose sample pixels from an image that are characteristic of particular classes, and then instruct the image processing software to utilize these training sites as references for the categorization of all other pixels in the image. It depends on the prior knowledge of the location and identity of land cover class that are in the image such that it is based on the external knowledge of the area of interest (Mather, 2004).

There are different kinds of classification algorithms characterized by employing sophisticated mathematical and statistical methods to generate predictions about the likelihood of data input being classified in a given way (Classifier, n.d.) which are;

- Logistic Regression
- Naïve Bayes
- Stochastic Gradient Descent
- K-Nearest Neighbours
- Decision Tree
- Support Vector Machine

- Random Forest

This research has employed Random Forest classification algorithm.

Random Forest (RF) is a machine learning algorithm used for classification and regression tasks. It works by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or the mean prediction (regression) of the individual trees. The algorithm randomly selects a subset of the features for each tree to be grown, which introduces additional randomness and helps to prevent overfitting. Furthermore, it uses bootstrapped samples of the training data to build each decision tree, further enhancing the variability of the trees in the forest. The final prediction of the random forest is based on the collective votes (classification) of the individual decision trees. (Breiman, 2001). This algorithm classifies each pixel individually, resulting in a final classification by majority voting. In other words, each decision tree is generated using a different subset of training samples in each interaction, thus building multiple singular or disconnected trees, which may not contain all classes in each tree (Gareth et al., 2013).

Random forest is more stable in the presence of outliers and in very high dimensional parameter spaces than other machine learning, because it follows specific rules for tree growing, tree combination, self-testing and post-processing (Sarica et al., 2017)

2.7.2. Unsupervised classification

Unsupervised classification occurs when the software analyses an image and groups pixels with similar properties without the user defining training samples for each land cover class as it does not depend on the prior knowledge. Besides, the computer will perform classification automatically (Mather, 2004). This method is suitable in evaluating areas where you have no or a little knowledge of the site and it is advantageous because it has unbiased geographical assessment of pixels.

2.8 Accuracy Assessment

Accuracy assessment is an important part of any classification assignment as it compares the classified image to another data that is accurately taken to present land cover classes on the ground (ground truth data). The system checks from the classification results, on whether if they comply with the ground to which they represent (Congalton, 2001). This is achieved through comparison of signature files used for the classification process and the ones used to check of the accuracy of the image classification output.

This assessment is critical for a number of reasons, including the requirement to know how well you are doing and learn from your mistakes, the ability to compare techniques quantitatively, and the capacity to apply the information from your spatial data analysis in various decision-making processes. The image classification accuracy's is evaluated using an error matrix, from which user accuracy, producer accuracy, and overall accuracy are produced (Congalton, 2001).

2.9. Land cover.

Land cover refers to the surface cover on the ground, whether vegetation, urban infrastructure, water, bare soil or other. Identifying, delineating and mapping land cover is important for global monitoring studies, resource management, and planning activities. Identification of land cover establishes the baseline from which monitoring activities can be performed, and provides the ground cover information for baseline SAR images (Jensen, 1986).

Land-cover defines the physical condition and characteristics of the biotic component at the local, regional, national, or continental scale. It also shows how human needs and actions have modified the environment to give it uses different from those of its original condition and aptitude (Chu, 2020; Hassan et al., 2016).

2.10 Output Validation

Validation refers to the process of assessing the accuracy and reliability of data or information. It involves comparing the data or information against a known standard or reference to determine whether it is correct or accurate (Congalton et al, 2008). In remote sensing, validation is an important step in ensuring that the results obtained from the analysis of remotely sensed data are accurate and reliable.

Validation in remote sensing involves comparing the results of the analysis of the remotely sensed data with a reference data set that is considered to be accurate. The reference data set can be obtained from ground-based measurements, aerial photographs, or other sources. The comparison can be done using statistical techniques such as accuracy assessment or error matrices (Congalton et al, 2008).

- Accuracy assessment involves comparing the results obtained from the analysis of the remotely sensed data with the reference data set to determine the level of agreement or disagreement between the two. This is usually done by overlaying the two data sets and counting the number of pixels that match or do not match. The results are then expressed as an accuracy percentage (Lillesand et al, 2014).

- Error matrices are another way of assessing the accuracy of remotely sensed data. They involve comparing the results obtained from the analysis of the remotely sensed data with the reference data set and categorizing the errors into different types such as omission, commission, and misclassification (Foody, 2010).

2.11 Land Cover Mapping

The use of remote sensing analysis and interpretation techniques facilitate land cover mapping by using remotely sensed data such as satellite images. Land cover mapping involves generating maps showing various land covers (Chu, 2020; Hassan et al., 2016). The land cover maps are essential in resource management, monitoring studies and planning activities and they provide the base for thematic mapping.

2.12 Google Earth Engine.

The Google Earth Engine (GEE) is a web portal that offers access to software and algorithms for processing time-series satellite imagery and vector data from all over the world (Gorelick et al., 2017). The data repository contains approximately over 40 years' satellite images for the whole world, including a sizable amount of daily and sub-daily data as well as repeat data for several locations for the entire duration of two weeks. The data available are from multiple satellites that provide data, including the entire Landsat series, Moderate Resolution Imaging Spectrometer (MODIS), National Oceanographic and Atmospheric Administration Advanced very high-resolution radiometer (NOAA AVHRR), Sentinel 1, 2, and 3, Advanced Land Observing Satellite (ALOS), and other (Kumar & Mutanga, 2018).

GEE makes easy to access high performance computing resources for processing very large geospatial datasets, where by the accessibility is controlled by Internet-accessible application programming interface (API) and an associated web-based interactive development environment (IDE) (Gorelick et al., 2017). Users can create and run customized algorithms using the programming interface, and analysis is parallelized such that many processors are used for each computation, greatly accelerating the process. This makes it much easier than with desktop computing to undertake global-scale analyses (Kumar & Mutanga, 2018).

GEE uses a code editor to write scripts for performing different geospatial works. Code editor is the web-based IDE for Earth Engine Java script API, it has features which are designed to make developing complex geospatial workflows fast and easy. The code editor has the following elements (Platform – Google Earth Engine, n.d.);

- JavaScript code editor.
- Map display for visualizing geospatial datasets.
- API reference documentation (Doc tab).
- Git-based Script (Script tab).
- Console output (Console tab).
- Task Manager (Task tab) to handle long-running queries.
- Interactive map query (Inspector tab).
- Search of the data archive or saved scripts.
- Geometry drawing tools.

CHAPTER THREE

METHODOLOGY

3.1 Overview

This section describes the study area (area of interest), presents the overall workflow to achieve the research objectives, it shows the methods ranging from data collection to results. It includes the data to be used, its acquisition, data pre-processing, processing, data analysis methods and list of software used in obtaining the results.

3.2 Description of the study area.

The study area is Gombe Forest in Kigoma region as it is shown in Figure3.1. It is located in between Kigoma rural and the shores of Lake Tanganyika at the North-West corner of Tanzania. The region is situated between latitude: 4^{0}45’00” S and 4^{0}40’00” S and between longitude: 29^{0}30’00” E and 29^{0}40’00” E. The forest is within Gombe Stream National Park, 16 km north of Kigoma, the capital of Kigoma Region Gombe Forest occupies a total area of 14 square miles of which 35.69 square kilometers making it the smallest national park in Tanzania.

Kigoma region is on plateau that slopes from Northeast at about 1,750 meters down to 800 meters at the shore of Lake Tanganyika at the North-West corner of Tanzania. In Kigoma the wet seasons oppressive and overcast, the dry season is partially cloudy, and it is warm year around. Over the course of the year, the temperature typically varies from 17ºC to 30ºC and rarely below 15ºC or above 32ºC.

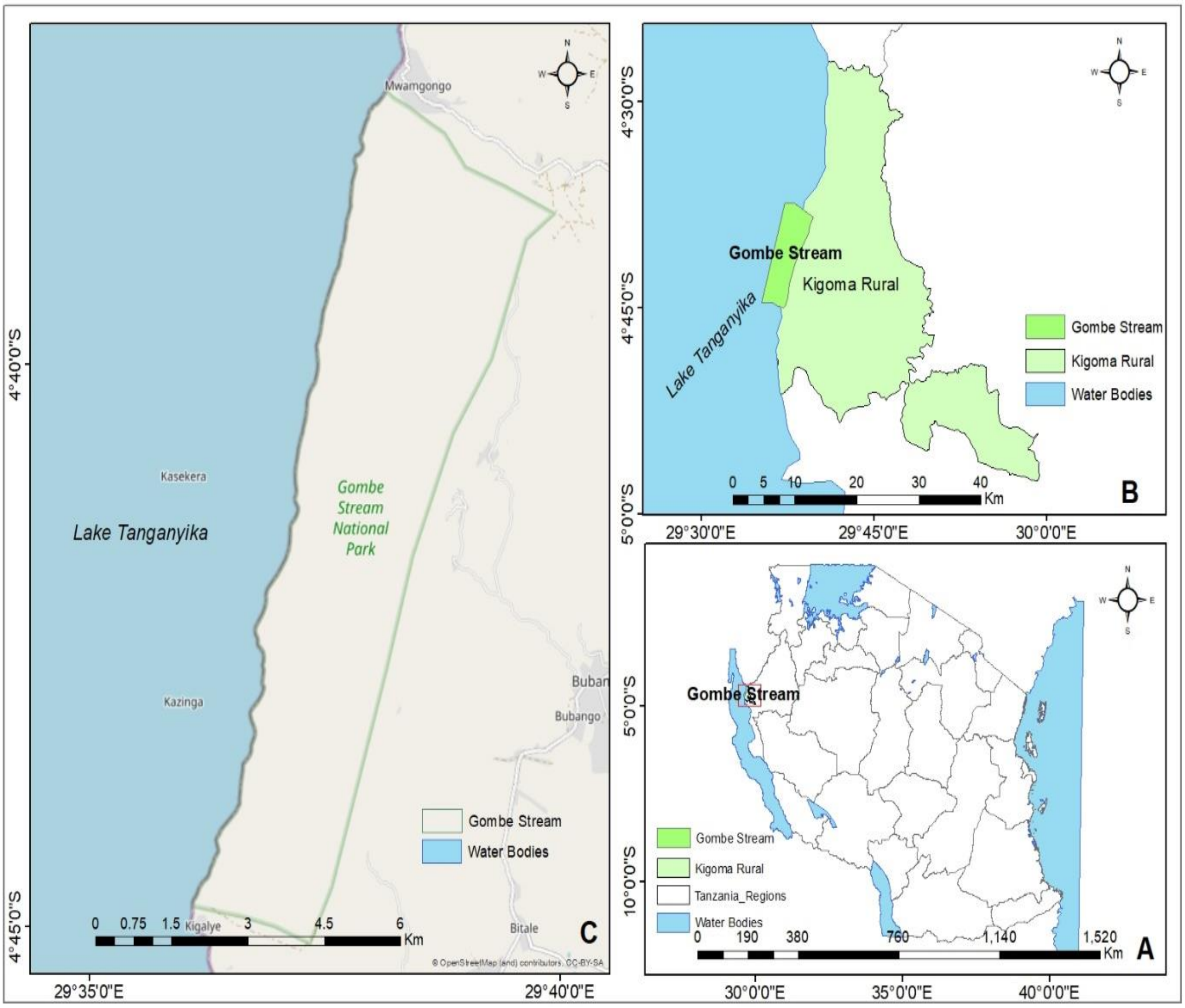

Figure 3.1: Location Map of the study area

A-Shows the Area of interest (Gombe Forest) in wider perspective, **B**- Shows Gombe Forest in District level (within Kigoma Rural District) and **C**-Shows Gombe Forest.

3.3 Methodological flow chart.

A flowchart is used to explain the separate steps of a process used to achieve the research objectives in sequential order, it shows the overall steps ranging from data acquisition to the end results. The step-by-step process to be used in this study is summarized in Figure 3.2

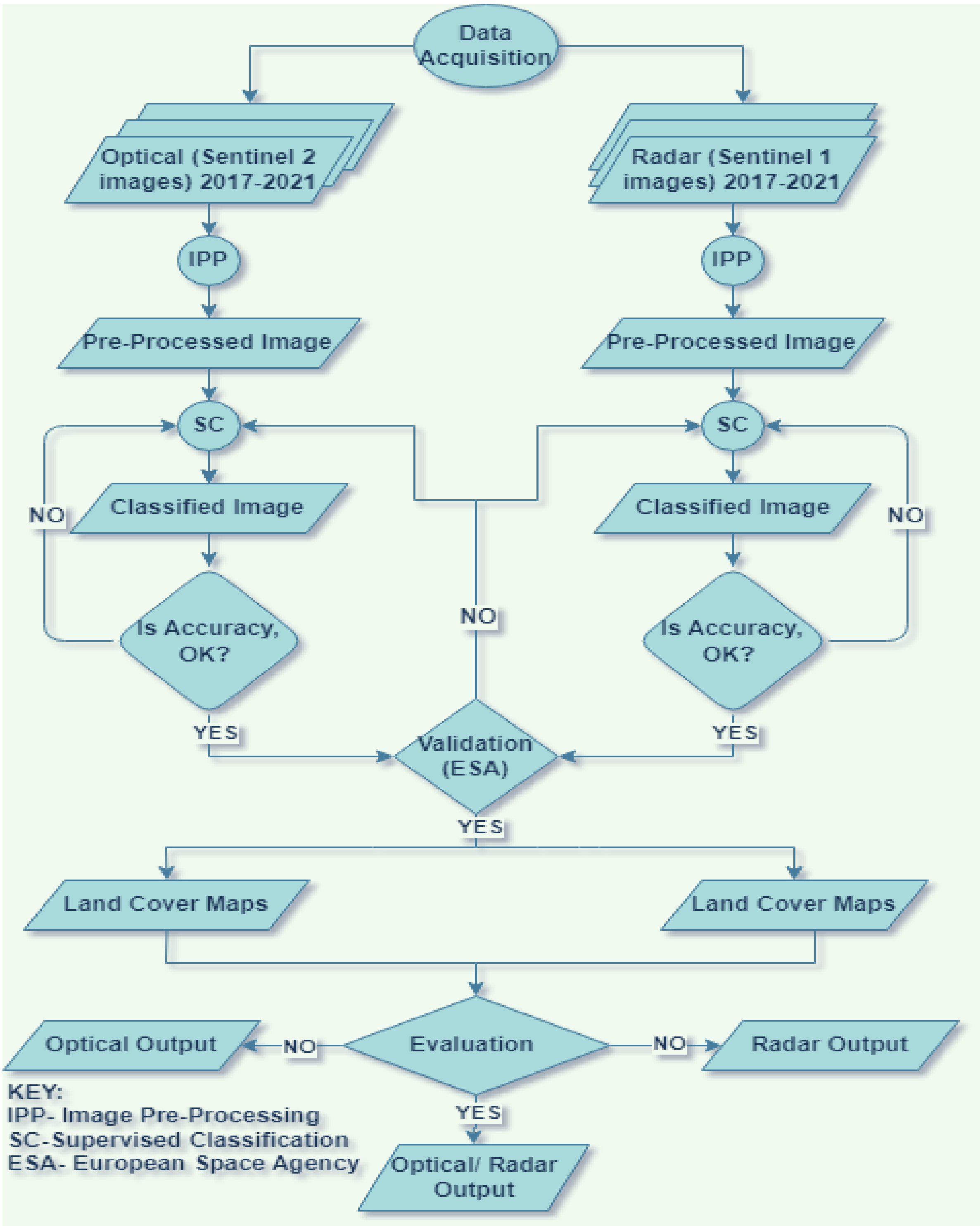

Figure 3.2: Methodological flowchart

3.4 Data Acquisition

In this study, a collection of images from two sensors were used. One is Optical sensor, Multispectral Instrument (MSI) while the other is Synthetic Aperture Radar (SAR) operating in the C-band. The Sentinel-2 satellite provide Optical images and Sentinel-1 satellite constellation provides SAR images as shown in Table 3.1.

Table 3.1: Summary of the research data characteristics.

NO	DATA	SOURCE	Epoch	Format
1	Optical (Sentinel 2) images	ESA/Sentinel-2	2017-2021	TIFF
2	Radar (Sentinel 1) images	ESA/Sentinel-1	2017-2021	TIFF
3	Administrative boundary of the area of interest	DIVA GIS(http://www.divagis.og	-	Shapefile
4	High Resolution ESA World cover for Validation.	ESA	2021	TIFF

3.4.1 Sentinel 2 data acquisition

Sentinel 2 MSI level 1c products were used, which are a Top-Of-Atmosphere (TOA) reflectance product acquired on the GEE platform over the study area for the years 2017,2019 to 2021 (Table 3.2). GEE stores a large number of geospatial datasets as it loads the entire data collection for the whole planet. Therefore, some additional searching criteria in the GEE code editor functions were performed like Filtering a collection basing on;

i. **Date**: Period/ time interested, here dry season (July) was selected over wet season to avoid much effect of clouds in the images.
ii. **Bounds(geometry):** Filtering region of interest, here the shapefile of the study area was used (Gombe Forest)
iii. **Metadata**: Selecting % of cloud cover, here less than 30% of cloud cover was used.

Table 3.2: Optical (Sentinel 2) data characteristics

S/No	Data	Date of acquisitions
1.	Sentinel-2	July 11, 2017
2.	Sentinel-2	July 7, 2019
3.	Sentinel-2	July 9, 2021

3.4.2 Sentinel-1 data acquisition

SAR data of the Sentinel-1 SAR Ground Range Detected (GRD) product from the Google Earth Engine (GEE) platform were used, and the imaging time was 2017,2019 to 2021. (Table 3.3) and since GEE stores a large number of geospatial datasets as it loads the entire data collection for the whole planet, some additional searching criteria in the GEE code editor functions were performed like Filtering a collection basing on;

i. **Date**: Period/ time interested, here dry season (July) was selected over wet season in order to match with optical (Sentinel 2)

ii. **Bounds(geometry):** Filtering region of interest, here the shapefile of the study area was used (Gombe Forest)

iii. Metadata specification:

- **Data product:** Level-1 Ground Range Detected High Resolution (GRDH) modes, tthese scenes are projected on a regular 10 m × 10 m grid, with a reference spatial resolution of 20 m × 22 m (range and azimuth directions, respectively).
- **Transmitter receiver polarization**: Dual-polarization (VV and VH) high resolution products, which can provide more information about the targets being observed. In particular, the combination of VV and VH polarizations helps to distinguish between different types of vegetation and other land cover types.
- **Instrument (Imaging) Mode**: Interferometric wide swath (IW) mode with a spatial resolution of 20 m × 5 m.
- **Orbit properties pass (acquisition orbit)**: Descending orbit were considered.
- **Resolution mode:** 10-m

Table 3.3: Radar (Sentinel 1) data characteristics

S/No	Data	Date of acquisitions	Orbit properties	Polarization	Beam mode
1.	Sentinel-1A	July 09, 2017	Descending	VV+VH	IW
2.	Sentinel-1A	July 11, 2019	Descending	VV+VH	IW
3.	Sentinel-1A	July 12, 2021	Descending	VV+VH	IW

3.4.3 Acquisition of Reference data for Validation.

ESA World cover v200 resolution product was used in validation, these products were downloaded from the Google Earth Engine (GEE) platform and the imaging time was 2021, a freely accessible global land cover product at 10m resolution based on both Sentinel-1 and Sentinel-2 data, containing 11 land cover classes developed by the United Nations (UN) Food and Agriculture Organization (FAO) and independently validated with a global overall accuracy of 76.7% as summarized in Table 3.4 in appendix 1.

3.5 Software utilized

In this research, the software used was Google Earth Engine (GEE), Arc Geographic and Information System (ArcGIS 10.1) and Microsoft Excel software.

- Google Earth Engine was used in the data acquisition, pre-processing, processing of both Optical (Sentinel 2) images and Radar (Sentinel 1) images such processes include the image classification. All orbital data were acquired and processed using the GEE cloud computing platform, which runs on the portal: https://earthengine.google.com which allows handling and processing large amounts of data, with high computing power.
- Arc Geographic and Information System (ArcGIS 10.1) was used for mapping (preparing Land cover maps).
- Microsoft Excel was used in preparing tables and graphs/charts used to quantify the results obtained from Land cover.

3.6 Data preparation and pre-processing

3.6.1 Sentinel 2 data preparation and pre-processing

As a source, MSI level 1c products were used, which are a Top-Of-Atmosphere (TOA) reflectance product, which is radiometrically and geometrically corrected, including orthorectification using Sentinel-2 Toolbox (ESA, 2021). Optical satellite imagery is an important Earth observation data source, yet when clouds are present, they provide limited utility for land surface applications. Therefore, after accessing or loading sentinel 2A into GEE Code Editor and visualization some additional pre-processing steps were followed

i. Making a cloud-free images mosaic of the study area (cloud masking).
   This pre-processing was performed to remove clouds, cirrus clouds, and shadows, using a cloud scoring algorithm to mask polluted pixels as described in ESA (2021). Sentinel 2MSI level 1C has QA60 bitmask band (16-bit) with cloud mask information which was used to mask cloud from the images, the information was extracted as binary values where Bit 10 represents Opaque clouds (0 for No opaque clouds and 1 for Opaque clouds present), and Bit 11 represents Cirrus clouds (0 for No cirrus clouds and 1 for Cirrus clouds present).
ii. Creating a cloudless satellite image mosaic.
   Composite cloudless image mosaic using the cloud-masked images was created in GEE using per-band “median” statistics.

iii. Image sub setting for the study area

The subset of area of interest was created by importing a “Shapefile” with my study area information (Gombe Forest) and then the study area was masked.

iv. Creation of RGB Combination from Sentinel 2 images (Composite).

Sentinel 2 image composite was created using RGB combination of True color in order to aid visualization of features in an image when performing classification where Band 4,3,2 were used. For this research Band4 represented Red, Band3 represented Green and Band 2 represented Blue.

3.6.2 Sentinel-1 data preparation and pre-processing

After loading an image into GEE code editor and visualization, the Sentinel-1 data were pre-processed using the Sentinel-1 Toolbox software (version 1.1.1). Sentinel-1 GRD images available in the Google Earth Engine collection (Google Earth Engine 2020) have already undergone pre-processing using the European Space Agency’s Sentinel-1 Toolbox. GEE pre-processing includes the application of the orbit files, thermal noise and GRD border noise removal, radiometric calibration to sigma naught, and range-Doppler terrain correction.

Although SAR data were unaffected by rain and cloud cover, noise has a significant influence on data quality. Therefore, additional pre-processing steps was applied to further data enhancement, which includes removing additional GRD border noise occasionally caused by heavy convective rain cells, in addition to applying radiometric terrain correction and finally adaptive speckle filtering. The final output was geocoded and topographically normalized gamma-naught VV- and VH-polarized backscatter images at 10 m pixel spacing as shown Figure 3.3.

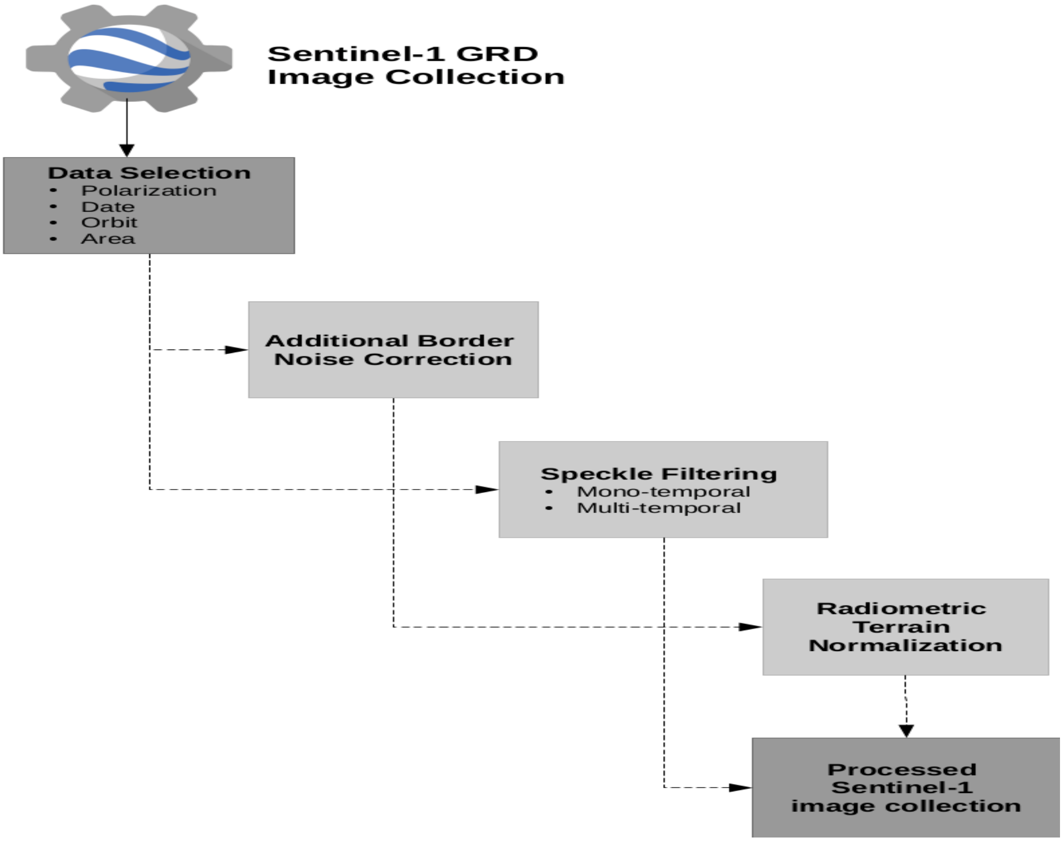

Figure 3.3: The preparation framework to create Sentinel-1 ARD in GEE.
Source: https://github.com/adugnag/gee_s1_ard.

The dark grey box represents the compulsory steps and the light grey ones represent the elective pre-processing steps.

Then, Image sub setting for the study area followed where region of interest was created by importing a “Shapefile” with my study area information (Gombe Forest) and then the study area was masked.

Finally, Creation of RGB Combination from Sentinel 1 images (Composite). Sentinel 1 image composite was created using RGB combination of different bands to aid visualization of features in an image during classification where Band VV, VH, VVVH_ratio was used. For this research Band VV represented Red, Band VH represented Green and Band VVVH_ratio represented Blue.

3.7 Data processing

In this research, Data processing involved Image classification of both sentinel 2 and Sentinel 1 images for year 2017,2019 and 2021.

3.7.1 Image classification

This processing step followed after performing pre-processing procedures and creating the image composites for both Sentinel 2 and Sentinel 1 images. This was performed in order to determine the land cover of the area of interest, and for that case Supervised classification was used in this research. The following are the steps/procedures used in performing classification;

i. Land cover class selection

For image classification, the researchers' knowledge of the area under study (Gombe Forest), was used as well as ESA World cover high-resolution images available from the Google Earth Engine for the identification of land cover as described early in Table 3.4. It resulted to five different categories of land covers (Permanent water bodies, Tree Cover, Shrubland, Grassland and Bare/ Sparse Vegetation) on sentinel 2 and both SAR scenes as shown in Table 3.5 (Suhet, 2015).

Table 3.5: Description of the Land cover classes used for image classification

ID	**Land cover class**	**Palette**	**Color**
10	Tree cover	#006400	
20	Shrubland	#ffbb22	
30	Grassland	#ffff4c	
60	Bare /sparse vegetation	#b4b4b4	
80	Permanent water bodies	#0064c8	

ii. Training samples selection

A selection criterion was applied in training areas under the following conditions: (a) samples within the area of interest, (b) basing on the ESA world cover classes, palette (color) as is the data that is to be used for independent classification validation where for about 1000 samples were selected where the training areas, selected by polygonal vectors, were uniformly distributed over the collected images to represent each class to be classified.

To avoid overfitting, the data were pseudo-randomly separated in the proportions observing the best model fit in the proportion 70 -30% where 70% sample was used for training while 30% for validation /measuring classification reliability.

iii. Selection of Classification algorithm

Both the images (Sentinel 2 and Sentinel 1) were classified in Google Earth Engine (GEE) software using a supervised training method by applying the non-parametric machine-learning Random Forest algorithms, which is a widely used high-precision classifier. Its robustness lies in the fact that it can combine multiple decision-tree results through a selection of random subsets within the training set. The Sentinel-2 RF model used the bands B2, B3, B4, B5, B6, B7, B8, B8A, B11, B12 while Sentinel-1 RF model used the bands VV, VH, and quotient band of VV and VH as inputs.

iv. Classification Accuracy Evaluation

The accuracy of the classified images was computed from the independent sample, 300 (30%) samples were randomly selected from high Google Earth imagery images for validation purpose. In this study, the confusion matrix was used to evaluate classification accuracy, Overall Accuracy (OA), Kappa Index, Commission Error (CE), Omission Error (OE), User Accuracy (UA), Producer Accuracy (PA). The overall accuracy referred to the percentage of sampled pixels correctly classified, while the class accuracy referred to the percentage of sampled pixels correctly classified in a specific class. An error of omission indicated that a class had not fully included features of that class, while an error of commission indicated a class incorrectly included features belonging to a different class. These metrics were used to evaluate the error of both Sentinel-1 and Sentinel-2 classification images

3.6 Output/Result Validation

The validation of the result was done by comparing the Optical (Sentinel 2) and Radar (Sentinel 1) data with ESA World Cover (v200) reference data for year 2021 in terms of land cover and area coverage. This process was done after classification accuracy evaluation/ accuracy assessment and after obtaining acceptable overall accuracies (above 70%) and Kappa values above 0.4 for both Optical and Radar outputs (Congalton, 1991). Then, Area coverage using both maps and statistical data (table and graphs) were used in result comparison to quantify the forest spatial and temporal status of Gombe forest.

CHAPTER FOUR

RESULTS ANALYSIS AND DISCUSSION

4.1. Overview

In this chapter, the results obtained through the implementation research methodology and discussion are presented according to the intended objectives of this research which was to investigate the use of Radar data to fill data gap in monitoring Gombe Forest cover change in Kigoma. It involves data presentation, analysis, interpretation of the results and products obtained from the methodology, aimed to narrate the findings and also to provide the direction on the discussion section of the research. The results obtained in this research are Land cover maps, Charts and graphs representing Forest cover changes of Gombe Forest from 2017 to 2021 for both Optical (Sentinel 2) and Radar (Sentinel 1).

4.2 Result of Accuracy Assessment

Table 4.1 and Table 4.2 in Appendix 3 provides the confusion error matrix (validation error matrix) and Kappa statistics that were used to determine the classification accuracy of the land cover for 2017, 2019, and 2021 for both Optical (Sentinel 2) and Radar (Sentinel 1) product.

For Optical products, the Total accuracy for 2017, 2019, and 2021 land cover was 91.67%, 93.94%, and 89.7% respectively and the corresponding Kappa statistics were 0.89, 0.92, and 0.85. The strongest agreement according to the Kappa numbers was for the year 2019 followed by the years 2017 and 2021.

For Radar products, the Total accuracy for 2017, 2019, and 2021 land cover was 74.46%, 77.17%, and 81.52% respectively and the corresponding Kappa statistics were 0.63, 0.67, and 0.73. The strongest agreement according to the Kappa numbers was for the year 2021 followed by the years 2019 and 2017.

4.3 Result Importance

During Optical (Sentinel 2) classification, Band 4 (B4) resulted in a higher record of importance variable being important than other bands with records above 95 for year 2017, greater than 105 for year 2019 and dropped to 80 for year 2021 followed by Band 12 (B12) while the lowest being Band 6 (B6) for year 2017 and 2019 and Band 7 (B7) for year 2021. It suggests that, in the case of Optical (Sentinel 2) discrimination between the evaluated classes is influenced by B4 rather than other Bands as shown in Figure 4.1, Figure 4.2 and Figure 4.3.

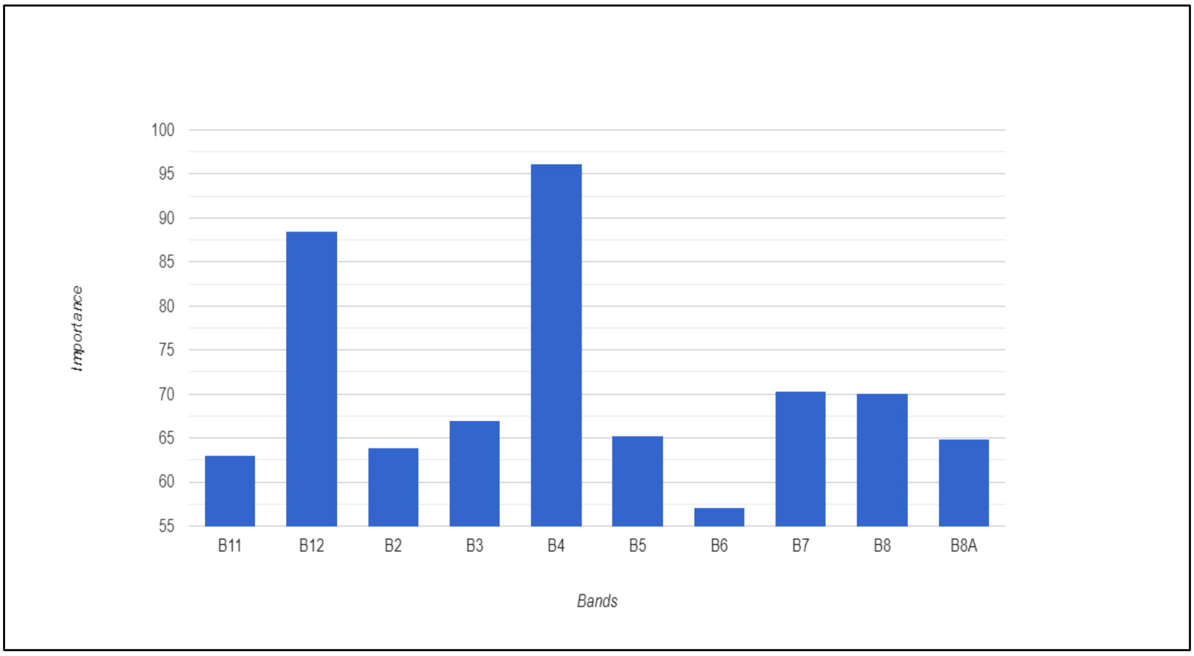

Figure 4.1: Ranking showing the importance of the Ten variables (Bands) used in the Random Forest algorithm with Optical (Sentinel-2) for year 2017.

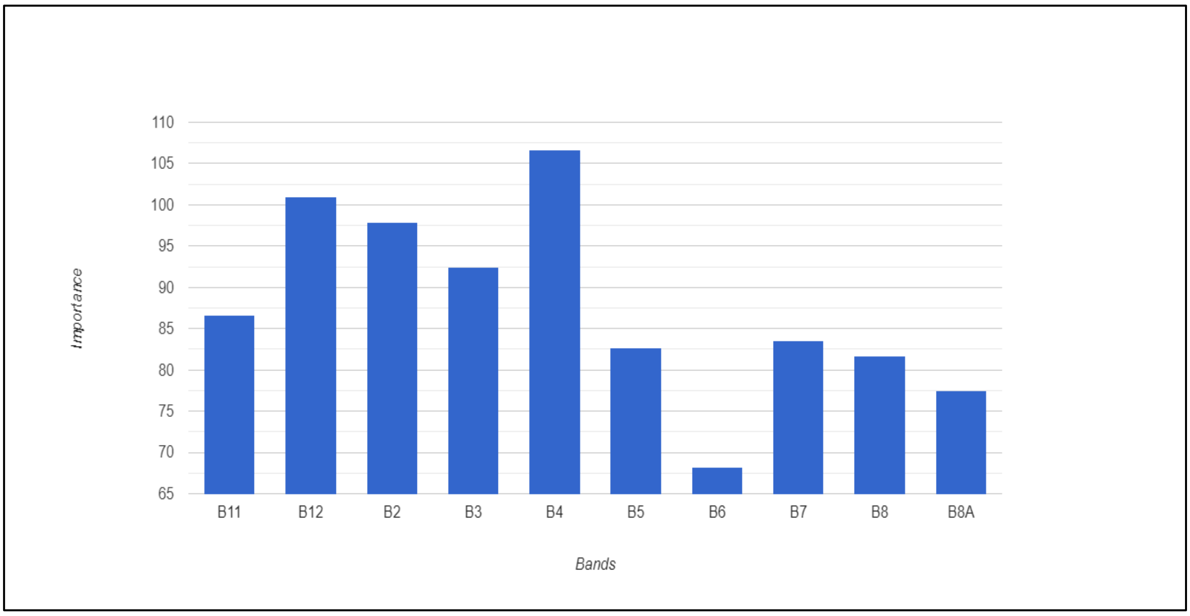

Figure 4.2: Ranking showing the importance of the Ten variables (Bands) used in the Random Forest algorithm with Optical (Sentinel-2) for year 2019.

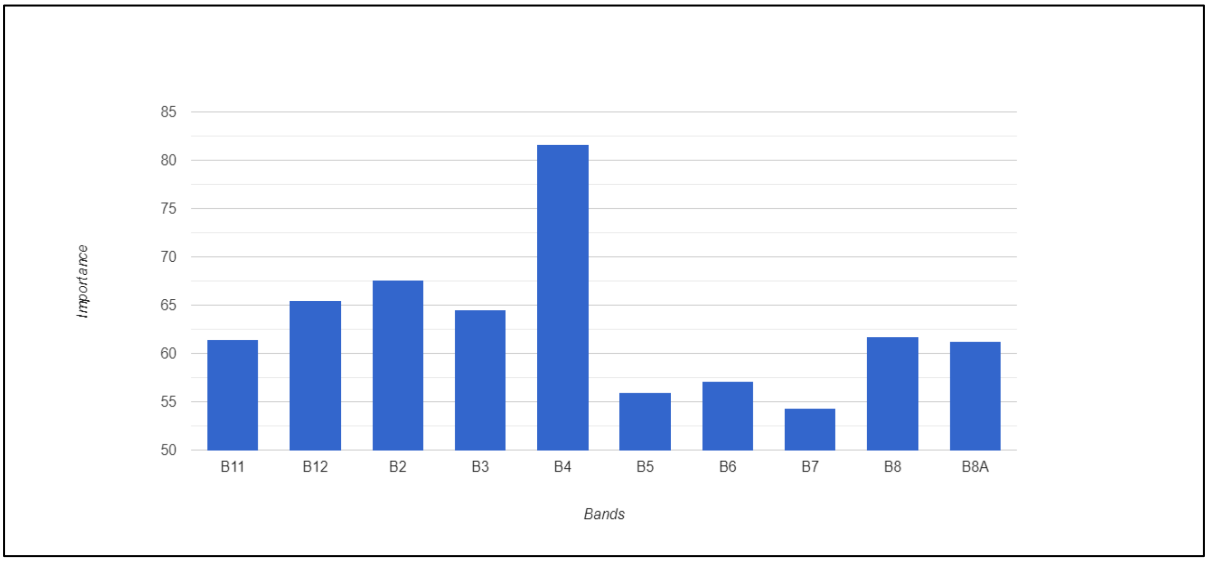

Figure 4.3: Ranking showing the importance of the Ten variables (Bands) used in the Random Forest algorithm with Optical (Sentinel-2) for year 2021.

During sentinel 1 classification, VH Band resulted in a higher record of importance variable for year 2017 and 2019 and VV Band for year 2021. This variable is used to rank the classification from the algorithm (Behnamian et al., 2017). In this same scenario, the (VV, VH) Band was more important than the angle range, with records greater than 600 for all years. It suggests that, in the case of radar, discrimination between the classes is influenced by backscatter VH then VV rather than by angle as shown in Figure 4.4, Figure 4.5 and Figure 4.6.

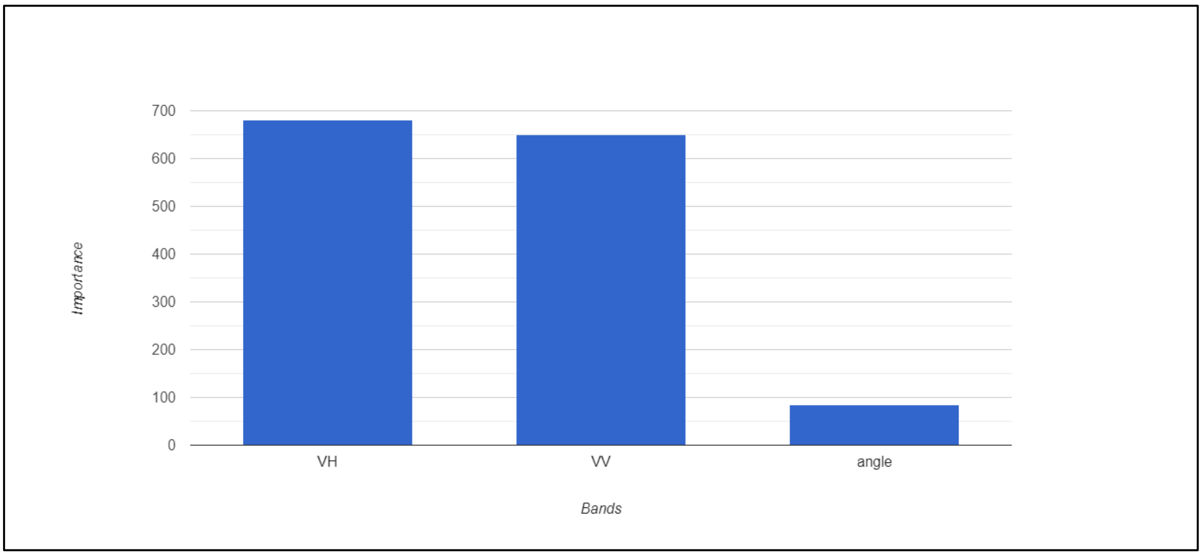

Figure 4.4: Ranking showing the importance of the three variables (Bands) used in the Random Forest algorithm with Radar (Sentinel-1) for year 2017.

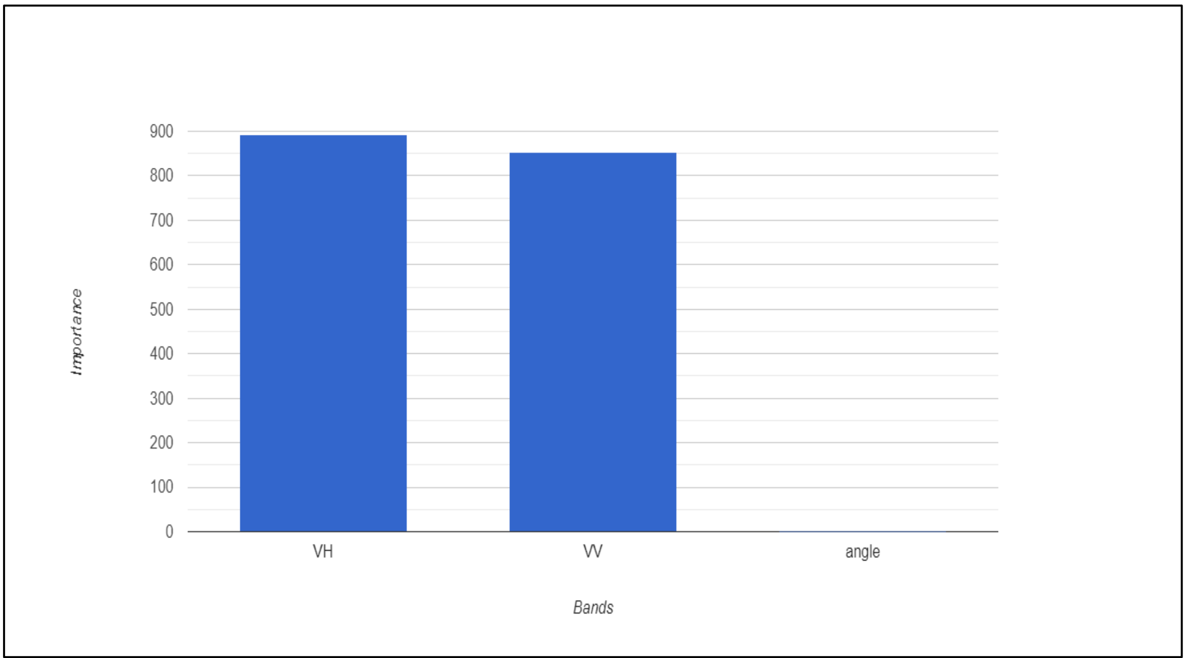

Figure 4.5: Ranking showing the importance of the three variables (Bands) used in the Random Forest algorithm with Radar (Sentinel-1) for year 2019.

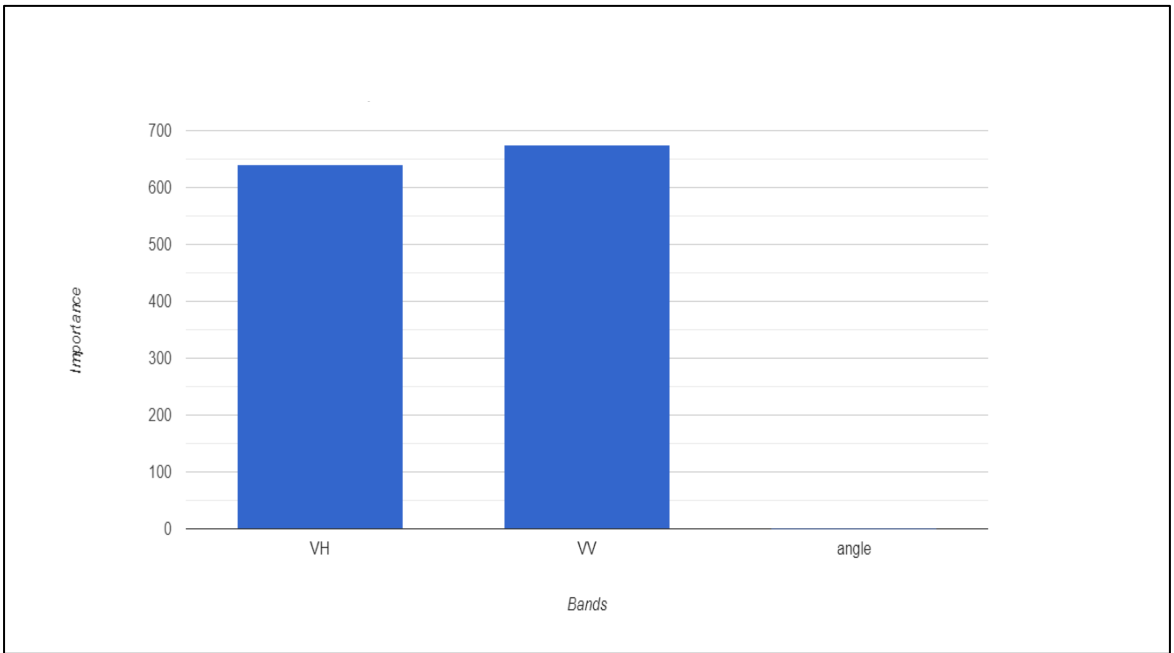

Figure 4.6: Ranking showing the importance of the three variables (Bands) used in the Random Forest algorithm with Radar (Sentinel-1) for year 2021.

4.4 Result of classification

This part involves developing a Forest cover change of Gombe Forest using both Optical and Radar images from 2017 to 2021 to meet objective (i) of this study. This has been presented visually using maps (Land cover maps) and statistically using Tables and Charts. During classification of both Optical (Sentinel 2) and Radar (Sentinel 1) products, five land cover were classified which were Permanent water bodies, Tree cover, Shrubland, Grassland and Bare/sparse Vegetation for year 2017, 2019 and 2021 as shown in Figure 4.7 to Figure 4.10 and Table 4.3 to Table 4.4 below.

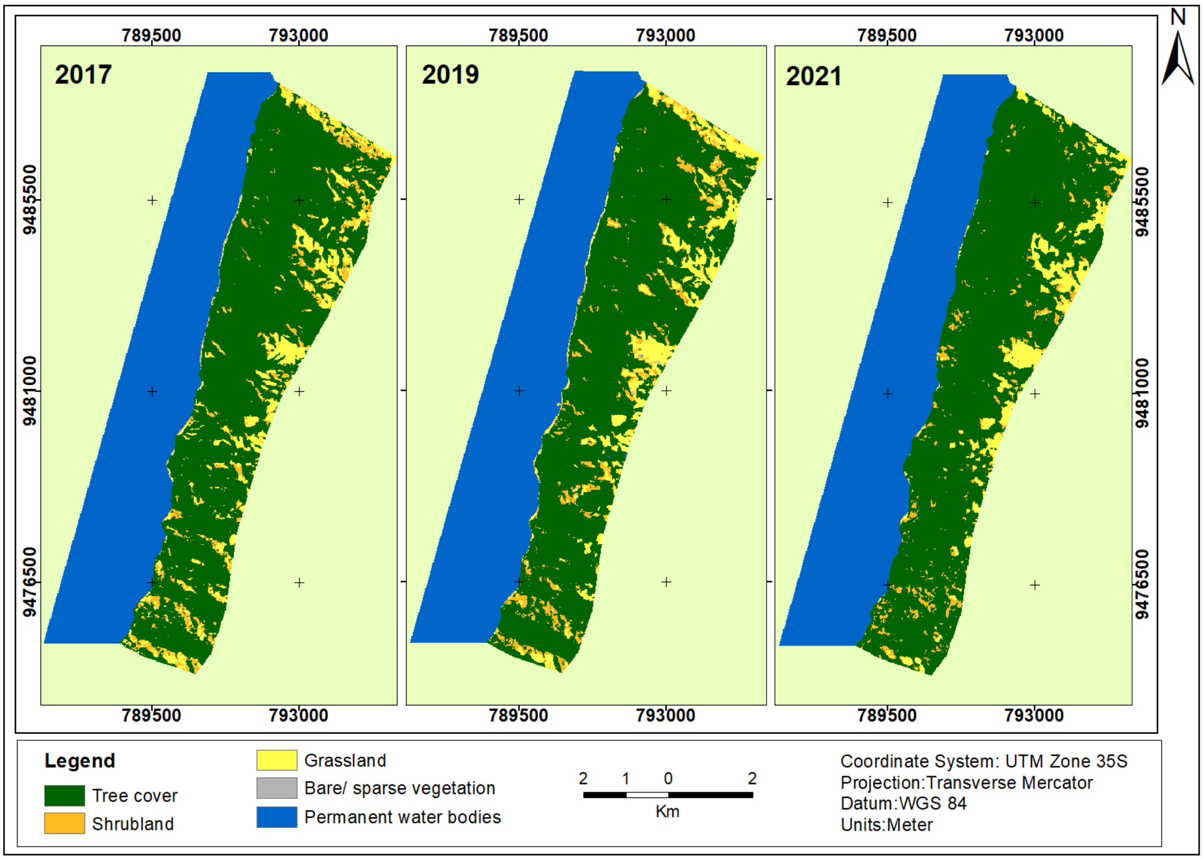

Figure 4.7: Land cover maps from 2017 to 2021 obtained with Optical products.

Table 4.3: Area coverage obtained with Optical products

Class Name	Area (Ha) 2017	Area (Ha) 2019	Area (Ha) 2021
Permanent water bodies	2403.80	2405.88	2404.08
Tree cover	2698.40	2688.80	2770.92
Shrubland	245.56	265.60	182.28
Grassland	262.76	251.04	260.16
Bare/sparse vegetation	13.84	13.04	6.92
Total	**5624.36**	**5624.36**	**5624.36**

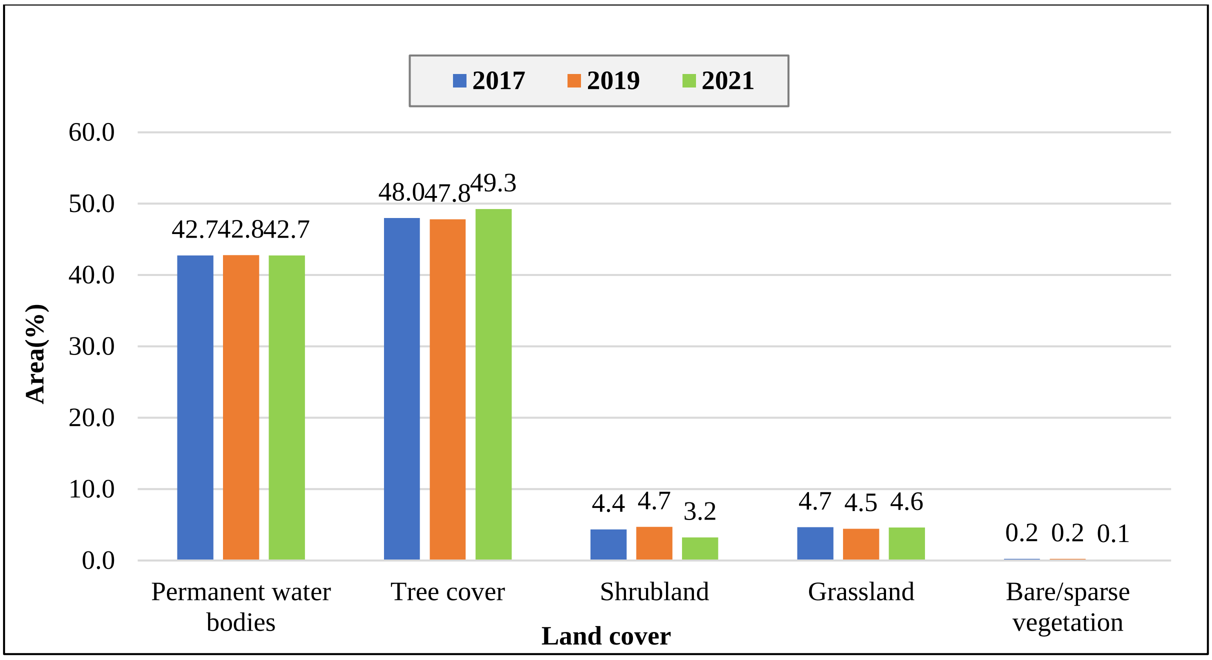

Figure 4.8: Land cover change distribution chart obtained with Optical products

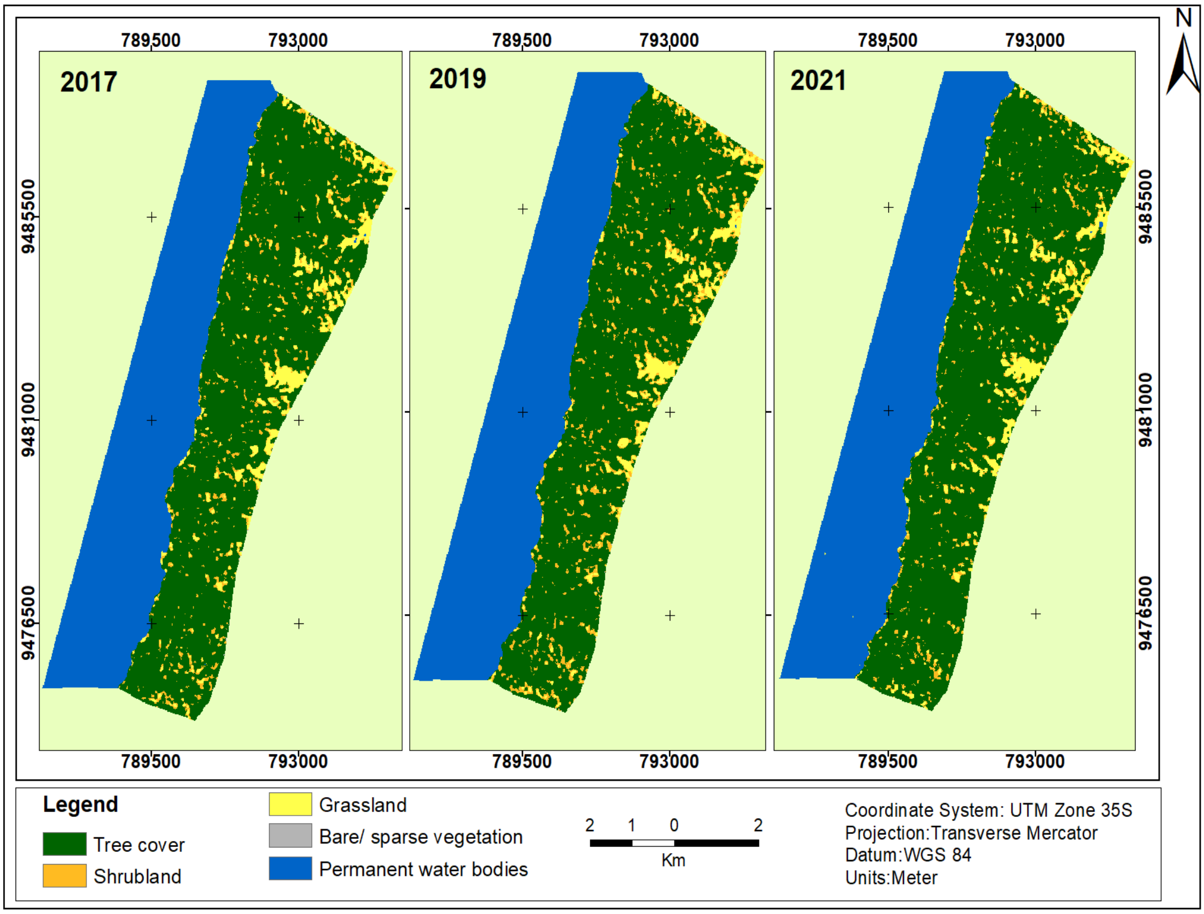

Figure 4.9: Land cover maps from 2017 to 2021 obtained with Radar products.

Table 4.4: Area coverage obtained with Radar products

Class Name	**Area (Ha) 2017**	**Area (Ha) 2019**	**Area (Ha) 2021**
Permanent water bodies	2404.24	2407.16	2404.68
Tree cover	2761.44	2724.12	2799.88
Shrubland	225.16	296.56	197.60
Grassland	231.12	194.44	220.56
Bare/sparse vegetation	2.40	2.08	1.64
Total	**5624.36**	**5624.36**	**5624.36**

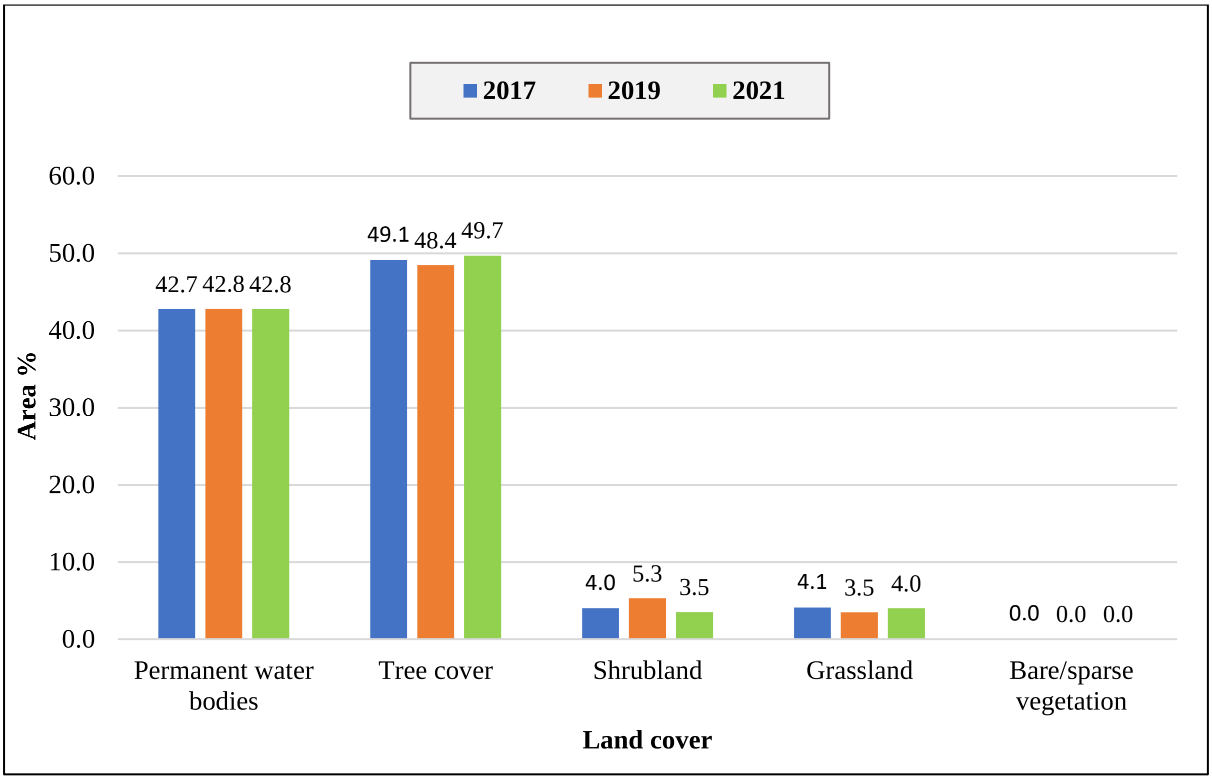

Figure 4.10: Land cover change distribution chart obtained with Radar products

4.5 Result Validation

This part involved the comparison of both Optical (Sentinel 2) and Radar (Sentinel 1) land cover with independent ESA World cover as Reference data for validation for year 2021 where Optical and Radar land cover agreed to ESA products as presented in visually using maps in Figure 4.11 and statistically using table in Table 4.5 in appendix 3 and chart in Figure 4.12.

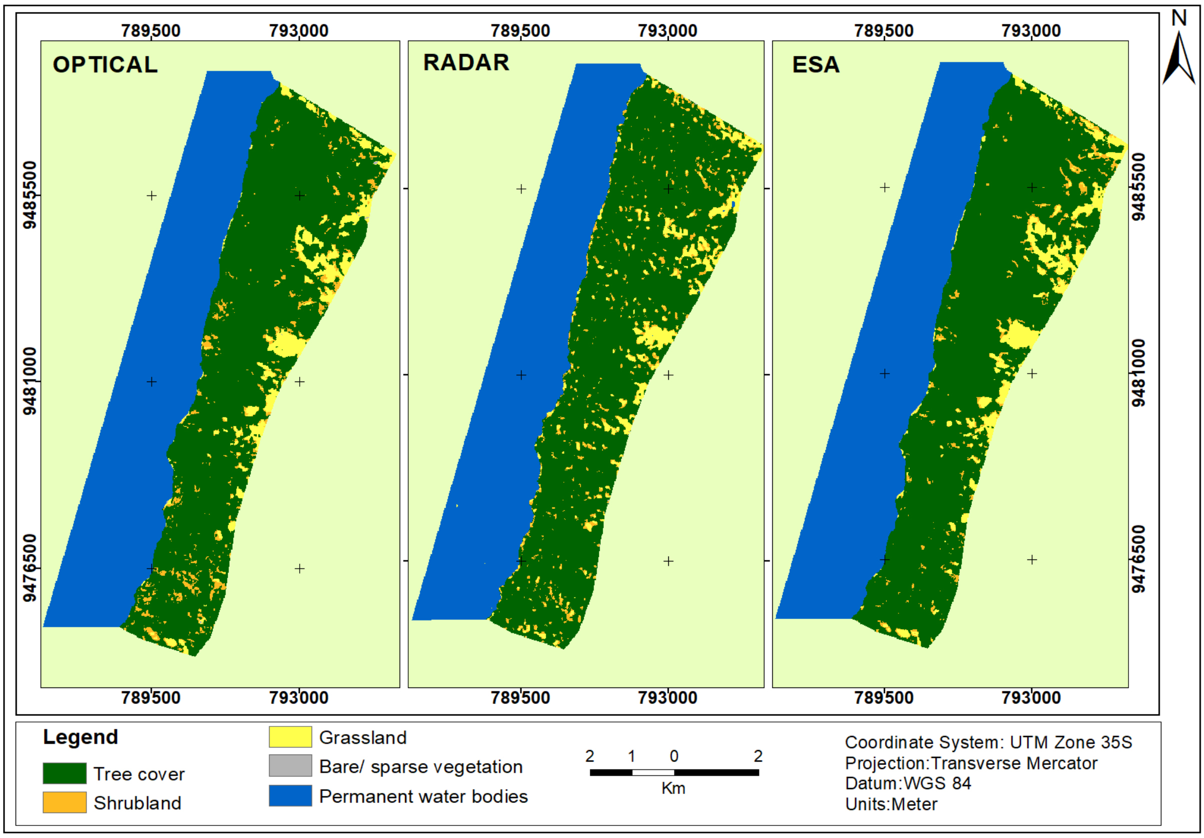

Figure 4.11: Land cover maps of 2021 obtained showing result validation

Table 4.5: Area coverage in 2021 showing result validation.

Class Name	**OPTICAL Area (Ha)**	**RADAR Area (Ha)**	**ESA Area (Ha)**
Permanent water bodies	2404.08	2404.68	2429.60
Tree cover	2770.92	2799.88	2787.00
Shrubland	182.28	197.60	130.24
Grassland	260.16	220.56	275.28
Bare/sparse vegetation	6.92	1.64	2.24
Total	**5624.36**	**5624.36**	**5624.36**

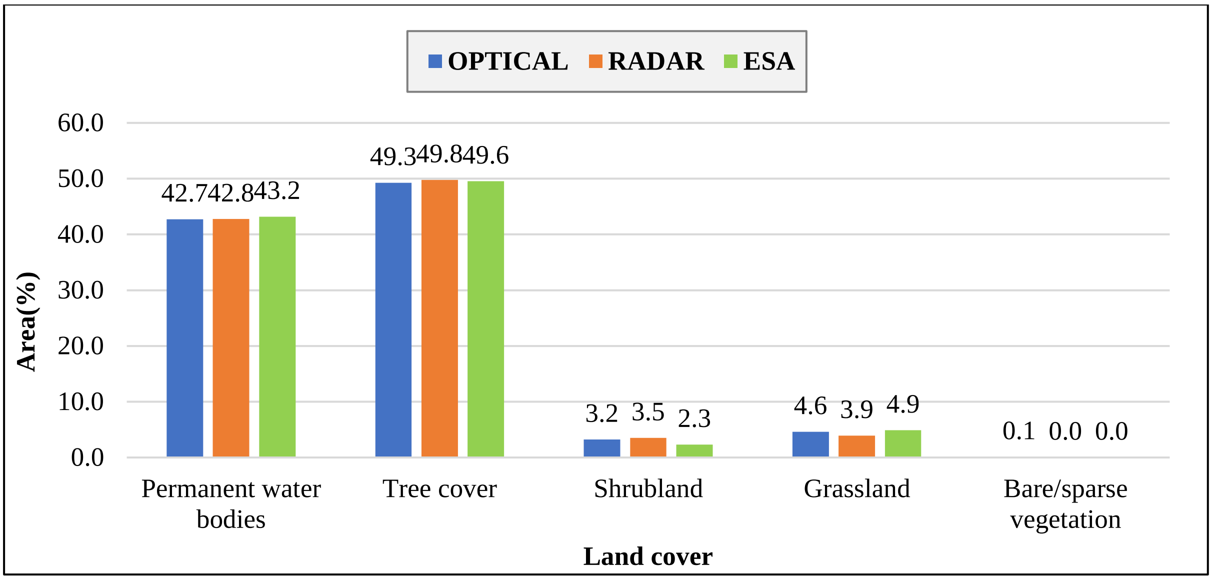

Figure 4.12: Land cover change distribution chart of 2021 showing result validation

4.6 Result Evaluation

This section involves evaluating the performance of Radar data in monitoring forest cover change of Gombe Forest as to meet objective (ii) of this study. Here Optical products are compared with Radar output visually using maps (Land cover maps) and statistically using Tables and Charts as shown in Figure 4.13 to 4.18 and Table 4.6 Table 4.8.

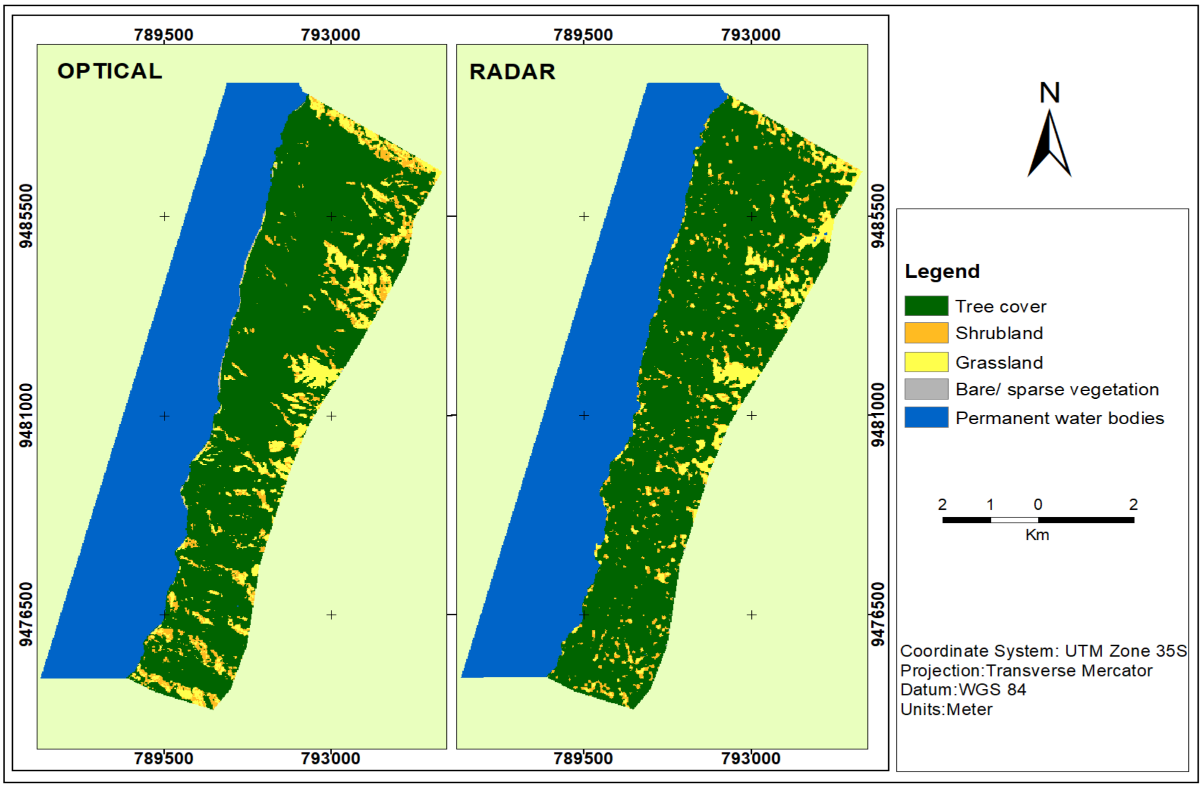

Figure 4.13: Land cover map of 2017

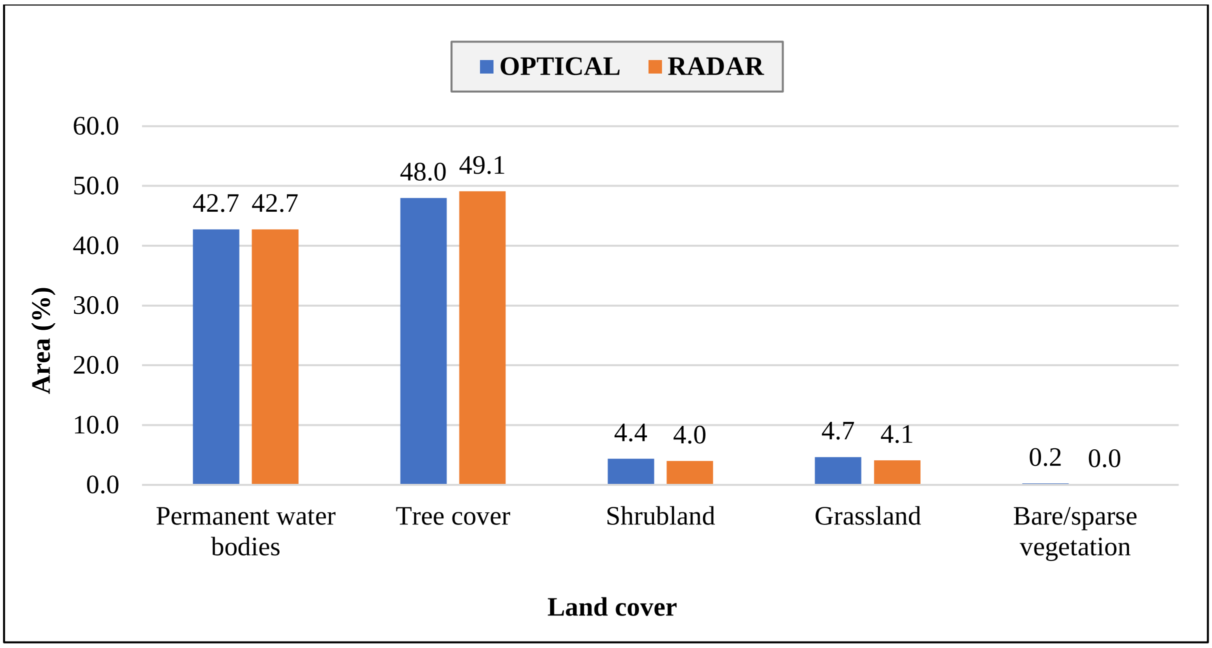

Figure 4.14: Land cover change distribution chart of 2017

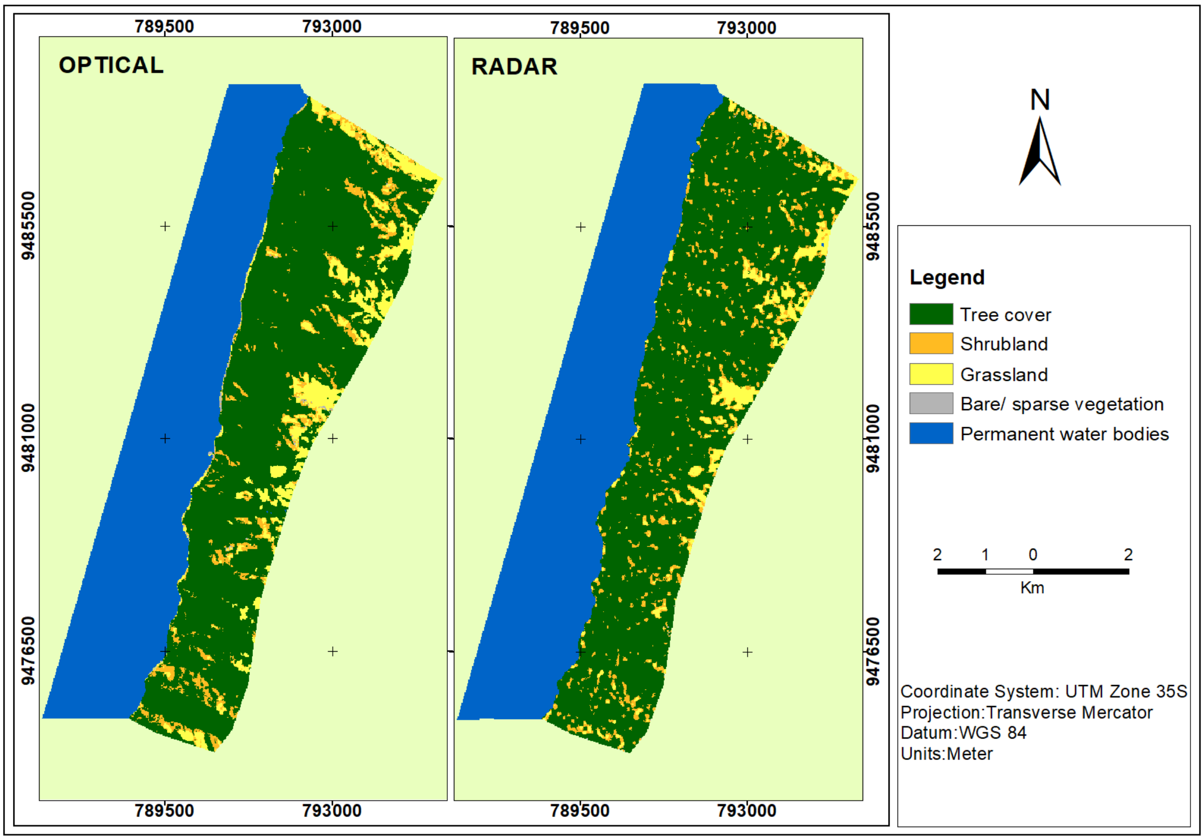

Figure 4.15: Land cover map of 2019

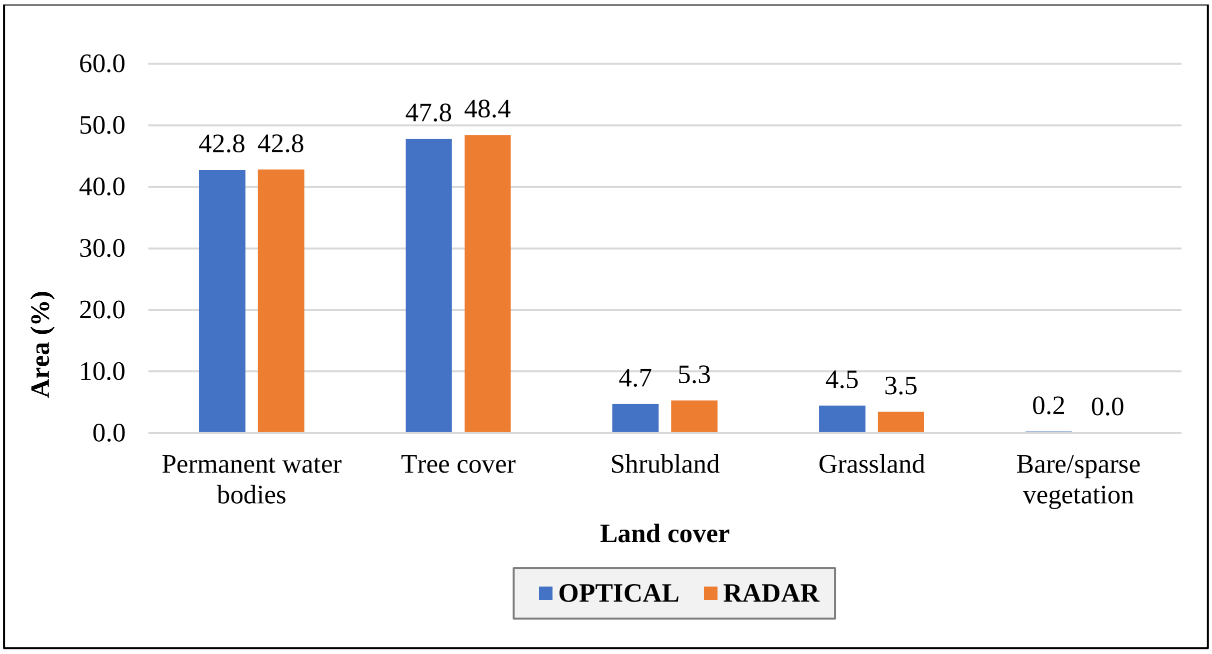

Figure 4.16: Land cover change distribution chart of 2019

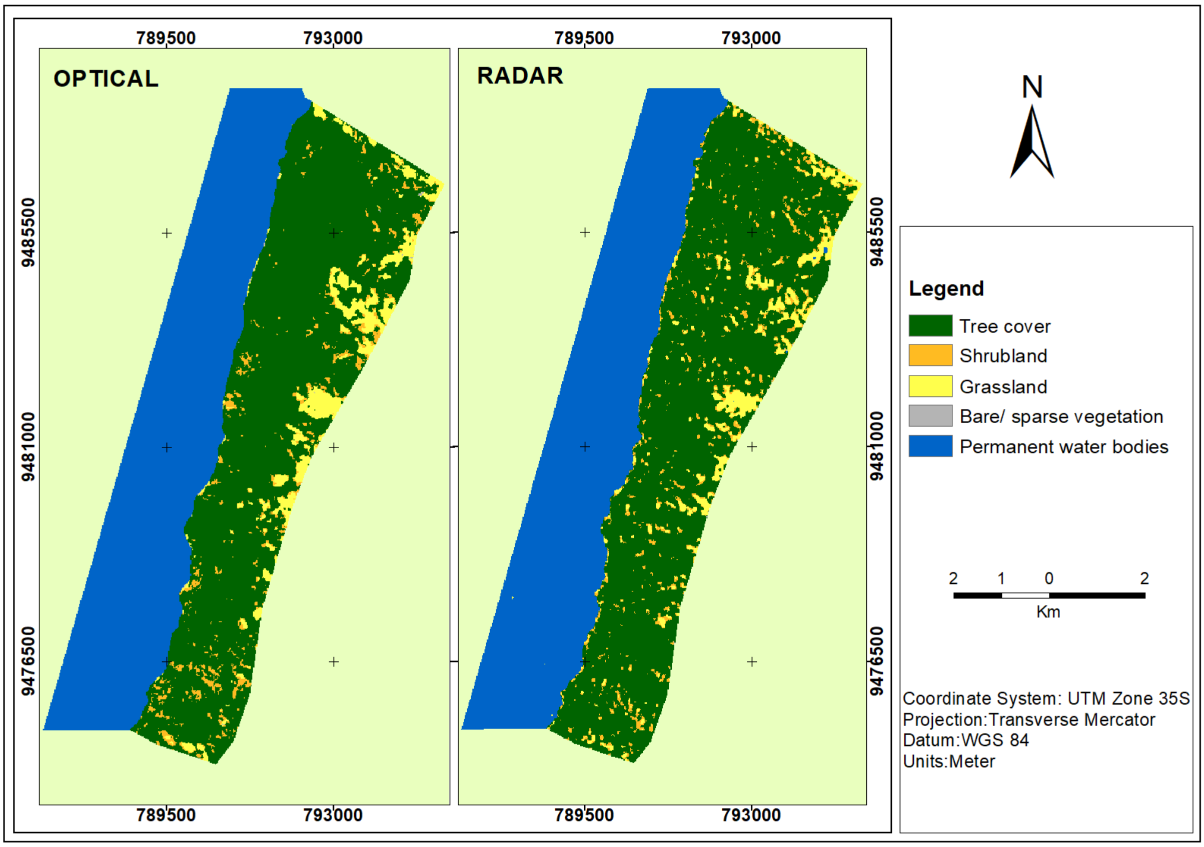

Figure 4.17: Land cover map of 2021

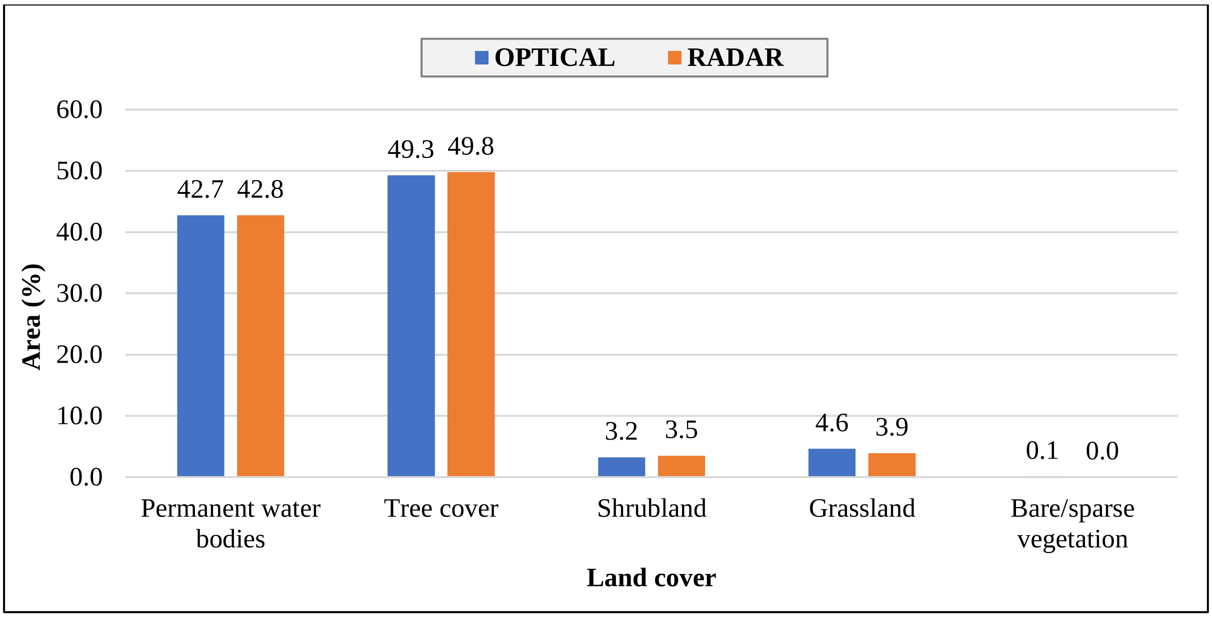

Figure 4.18: Land cover change distribution chart of 2021

4.7 Discussions

4.7.1 Discussion on Accuracyy Assessment Results

Results of this study indicate that (Optical) Sentinel-2 products possess a higher overall performance in delineating land cover types than the (Radar) Sentinel-1products where over 88.64% for optical and 74.46% for Radar data overall accuracies were obtained. This result corroborated with the finding report by (Meneghini, 2019), Meneghini stated that the results indicate that Sentinel-2 multispectral bands possess a higher overall performance in delineating land cover types than the Sentinel-1 backscatter bands.

4.7.2 Discussion on Land cover products

This study focused on to investigating the use of Radar data to fill data gap in monitoring Forest cover change of Gombe Forest in Kigoma. Our results show that Optical and Radar remote sensing data presented similar image characteristics in all year (from 2017 to 2021). This result agrees with (Meneghini, 2019) on an evaluation of Sentinel-1 and Sentinel-2 for Land Cover Classification. Meneghini concluded that while Sentinel-2 demonstrated the capability to consistently capture land cover, there is potential for single date Sentinel-1 backscatter image to act as ancillary information in Sentinel-2 scenes affected by clouds or for increasing separability across classes of mixed multispectral qualities but distinct surficial roughness, such as bare ground versus sparsely vegetation areas. The details of interpretation are presented in Table 4.9, Table 4.10 and Figure 4.19.

Table 4.9: LC area coverage, status, and changes between 2017, 2019, and 2021 obtained with Optical products.

LC Types	Area						Change (Gain/Loss)					
	2017		2019		2021		2017-2019		2019-2021		2017-2021	
	Ha	%	Ha	%	Ha	%	Ha	%	Ha	%	Ha	%
PWB	2403.8	42.74	2405.9	42.8	2404.1	42.74	2.08	0.04	-1.80	-0.03	0.28	0.01
TC	2698.4	47.98	2688.8	47.8	2770.9	49.27	-9.60	-0.17	82.12	2.21	72.52	1.29
SL	245.56	4.366	265.6	4.72	182.28	3.24	20.04	0.36	-83.32	-1.37	-63.28	-1.13
GL	262.76	4.672	251.04	4.46	260.16	4.63	-11.72	-0.21	9.12	-0.71	-2.60	-0.05
B/SV	13.84	0.246	13.04	0.23	6.92	0.12	-0.80	-0.01	-6.12	-0.10	-6.92	-0.12
Total	**5624.4**	**100**	**5624.4**	**100**	**5624.4**	**100**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**

Here, **PWB**- Permanent water bodies, **TC**- Tree cover, **SL**- Shrubland, **GL**-Grassland, **B/SV**-Bare/sparse vegetation, **LC**-Land Cover.

Table 4.10: LC area coverage, status, and changes between 2017, 2019, and 2021 obtained with Radar products.

LC Types	Area						Change (Gain/Loss)					
	2017		2019		2021		2017-2019		2019-2021		2017-2021	
	Ha	%	Ha	%	Ha	%	Ha	%	Ha	%	Ha	%
PWB	2404.2	42.75	2407.2	42.8	2404.7	42.8	2.92	0.05	-2.48	-0.04	0.44	0.01
TC	2761.4	49.10	2724.1	48.4	2799.9	49.7	-37.32	-0.66	75.76	1.28	38.44	0.61
SL	225.2	4.00	296.6	5.27	197.6	3.51	71.40	1.27	-98.96	-1.76	-27.56	-0.49
GL	231.1	4.11	194.4	3.45	220.6	3.99	-36.68	-0.66	26.12	0.54	-10.56	-0.12
B/SV	2.4	0.04	2.1	0.04	1.6	0.03	-0.32	0.00	-0.44	-0.01	-0.76	-0.01
Total	**5624.4**	**100**	**5624.4**	**100**	**5624.4**	**100**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**

Here, **PWB**- Permanent water bodies, **TC**- Tree cover, **SL**- Shrubland, **GL**-Grassland, **B/SV**-Bare/sparse vegetation, **LC**-Land Cover.

Figure 4.19 shows the trend of land cover change from 2017 to 2021 in terms of percentage area coverage where both Optical and Radar showed similar characteristics in land cover changes with some few differences. Here Permanent water bodies, Grassland and Bare/sparse vegetation had almost same value of land cover change being almost 0% while Shrubland and Tree cover being almost -1% and 1% respectively for both optical and radar products.

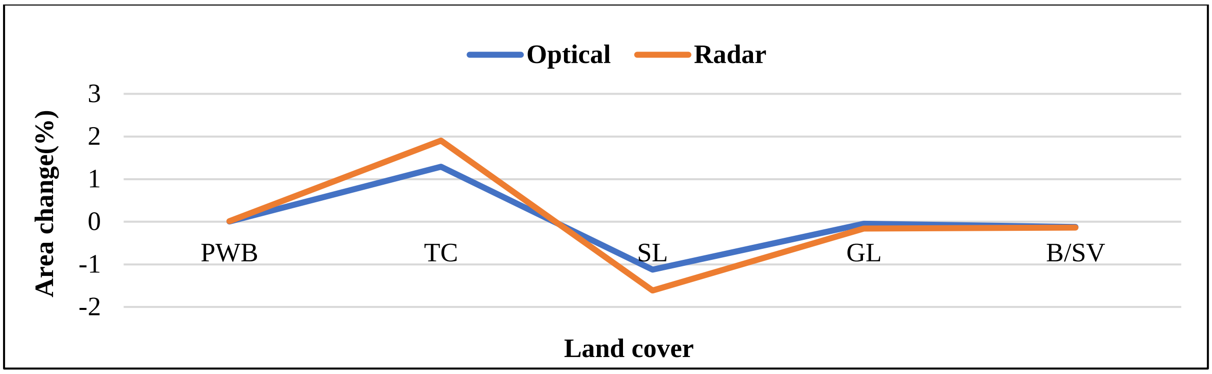

Figure 4.19: Land cover change trend graph from 2017-2021

Here, **PWB**- Permanent water bodies, **TC**- Tree cover, **SL**- Shrubland, **GL**-Grassland, **B/SV**-Bare/sparse vegetation.

Also (Figure 4.20), Radar data from active sensor mapped the surface water that was omitted by Optical data from passive sensor where by this result agrees with (Meneghini, 2019) an evaluation of Sentinel-1 and Sentinel-2 for Land Cover Classification.

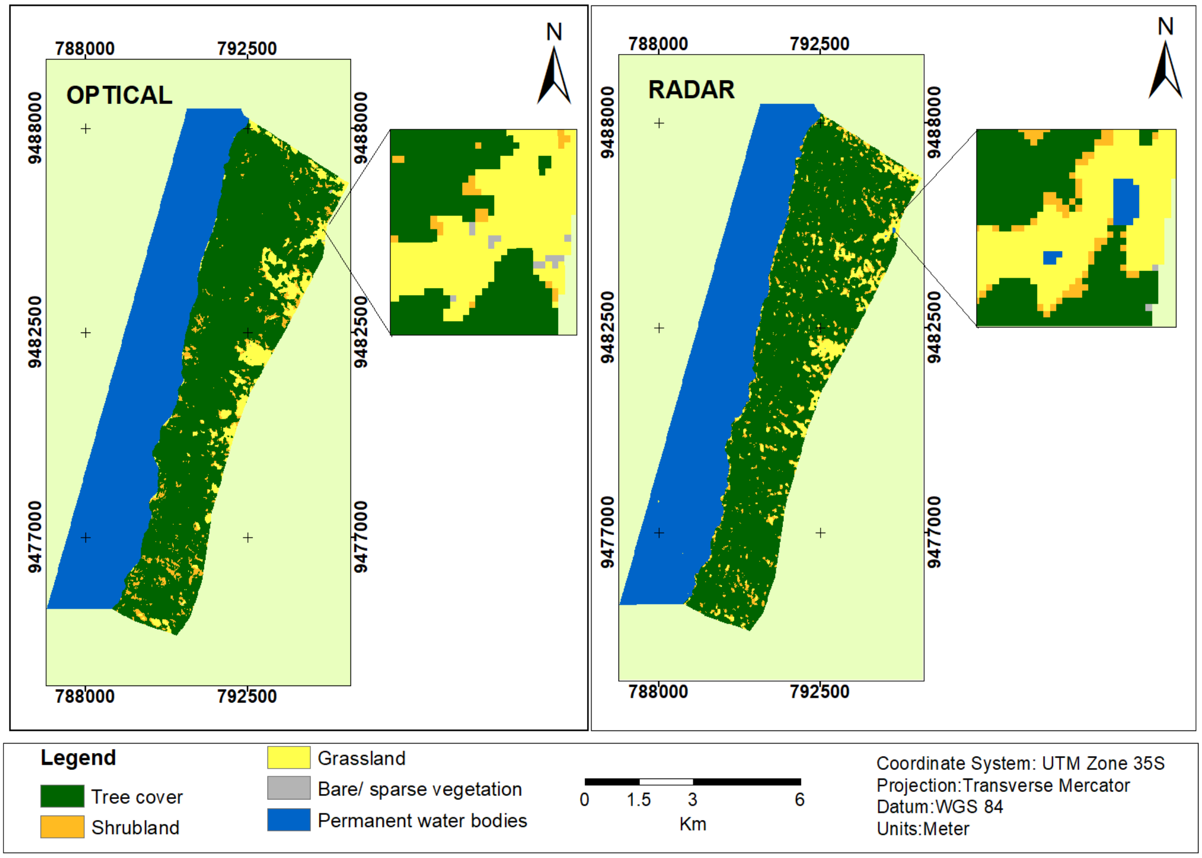

Figure 4.20: Land cover map showing water detection in 2021.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATION

5.1 Conclusion

The study investigated the use of Radar data to fill data gap caused by Optical data limitations in monitoring Forest cover change of Gombe Forest in Kigoma. The classification of both Optical and Radar images for three different time period i.e. 2017, 2019 and 2021 were performed to produce the Land cover maps whereas over 88.64% for Optical and 74.46% for Radar data overall accuracies were obtained and both Optical and Radar results agreed to ESA World cover products used as the Reference data for result validation. These Land cover maps have shown a clear forest cover where Radar maps showed the similarities with Optical maps in area coverage and land cover change from one year to the other in all land cover classes thus can be used interchangeably. For instance, Tree cover has declined from 49.1% in 2017 to 48.4% in 2019 and then increased to 49.7% in 2021 from Radar products being not far from Optical products where Tree cover has declined from 48.0% in 2017 to 47.8% in 2019 and then increased to 49.3% in 2021.

5.2 Recommendation

This study in view of the findings and conclusion recommends the following;

- Further studies should be conducted on the same forest to analyze the changes in forest cover using different approach (integrating/ Fusion of Optical data and Radar data) which will further describe the changes and assist in improving adaptation strategies in Forest cover change. This study used optical and radar data independently, the GEE facilitates the combination of optical and SAR data. Such a combination would thus be another strategy to further improve the results and performance compared to using any of the two alone of this study.

REFERENCES

Behnamian A, Millard K, Banks SN, White L, Richardson M, Pasher J. (2017). *A systematic approach for variable selection with random forests.* IEEE Geoscience and Remote Sensing Letters.

Campbell, J. (1987). *Introduction to Remote Sensing.* New York. : The Guilford Press.

Campbell, J. (2002). *Introduction to Remote Sensing* (3rd ed.). New york: Guilford Press.

Campbell, J. B., & Wynne, R. H. (2011). *Introduction to remote sensing*. Guilford Press.

Cheng, J., & Greiner, R. (2002). Learning Bayesian Networks from data. *An information theory based approach. Artificial Intelligence, 137*, 43-90.

Chu, D. (2020). *Remote Sensing of Land Use and Land Cover in Mountain Region*. Springer Singapore.

Classifier. (n.d.). C3 AI. Retrieved June 30, 2022, from https://c3.ai/glossary/data-science/classifier/

Congalton, R. G. (1991). *Assessing the Accuracy of Remotely Sensed Data:* Principles and Practices.

Congalton, R. G. (2001). Accuracy assessment and validation of remotely sensed and other spatial information. *International Journal of Wildland Fire*, *10*(4), 321.

Congalton, R. G., & Green, K. (2008). *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, (2nd ed.). CRC Press.

De Wasseige, C., Flynn, J., Louppe, D., Hiol Hiol, F., Mayaux, P. T.-S., Weyrich: Neufchâteau, B. 2., & Scholar.(2013). *The forests of the Congo Basin-State of the forest.* p. 2.–2. (n.d.).

Elikana John, P. B. (2021). *A Forest Monitoring System for Tanzania*. Remote Sensing.

Foody, G. M. (2010). Accuracy assessment: A review of the issues, its terminology and progress in achieving consensus. *Remote Sensing of Environment*, 114(1), 128-137.

Gareth J, Witten D, Hastie T, Tibshirani R (2013). An Introduction to Statistical Learning: with *Applications in R. 4 Classification.* Springer-Verlag, New York, p127-168.

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

Haq, M. (2012). *Techniques of Remote Sensing and GIS for flood monitoring and damage assessment:* Pakistan.

Hassan Z, Shabbir R, Ahmad SS, Malik AH, Aziz N, Butt A, Erum S. (2016). *Dynamics of land use and land cover change (LULCC) using geospatial techniques*: Islamabad Pakistan. SpringerPlus.

IUCN, I. U. (2013). *Red List of Threatened Species.*

Jensen, J. R. (1986). *Introductory Digital Image Processing.* New Jersey: Prentice-Hall.

Jensen, J. R. (2000). *Introductory Digital Image Processing.* New Jersey: Prentice-Hall.

Jensen, J. R. (2007). Remote sensing of the environment: *An earth resource perspective.* Pearson Education.

John E, Bunting P, Hardy A, Silayo DS, Masunga E. (2021). *A Forest Monitoring System for Tanzania.* Remote Sensing.

Joseph, G. (2005). *Fundamental of Remote Sensing* (2nd ed.). India: University of Press Pvt. Ltd.

Kumar, L., & Mutanga, O. (2018). *Google Earth Engine Applications Since Inception:* Usage, Trends, and Potential. Remote Sensing, 10(10), 1509.

Kurnar, D. (2011). Monitoring forest cover changes using remote sensing and GIS: a global prosective. *Reseach Journal of Environmental Sciences*, 105.

Lee, J.-S., & Pottier, E. (2009). *Polarimetric Radar Imaging*: *From Basics to Applications* (2nd ed.). CRC Press.

Lillesand, T. A. (1994). *Remote Sensing and Image Interpretation.* New York: John Wiley and Sons Inc.

Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (2014). *Remote Sensing and Image Interpretation,* Seventh Edition. Wiley.

Lillesand, T.M, Kiefer, R.W., and Chipman, J. (2008). *Remote Sensing and images interpretation.*

Mather, P. (2004). *Computer Processing of Remotely Sensed Images.* New york: Guilford Press.

Meneghini, Aaron, (2019). *An Evaluation of Sentinel-1 and Sentinel-2 for Land Cover Classification.* International Development, Community and Environment (IDCE). 235.

Platform – Google Earth Engine. (n.d.).

Rahman, M. M. (2010). *Mapping tropical forest cover and deforestation using synthetic arpeture radar (SAR) images*. Applied Geomatics, 113-121.

Russ, J. C. (1995). *The Image Processing Handbook* (2nd ed.). CRC Press.

Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T., Salas, W., ... & Houghton, R. A. (2011). *Benchmark map of forest carbon stocks in tropical regions across three continents.* Proceedings of the National Academy of Sciences, 108(24).

Sabins, F. F. (1986). *Remote Sensing. Principles and interpretation* (2nd ed.). New York: Oxford: W. H. Freeman & Co.

Sarica, A., Cerasa, A., & Quattrone, A. (2017). Random Forest Algorithm for the Classification of Neuroimaging Data in Alzheimer's Disease: A Systematic Review. *Frontiers in Aging Neuroscience*, *9*, 329.

Shimada et al. (2015). Forest monitoring using synthetic aperture radar. *Journal of Forestry Research*, 791-798.

Sonka, M., Hlavac, V., & Boyle, R. (2008). *Image processing, analysis and machine vision (3. Ed.)*

Suhet, B. Hoersch. (2015). *Sentinel-2 User Handbook*. European Space Agency (ESA).

Taiwo, O. A. (2010). Types of Machine Learning Algorithms. (Y. Zhang, Ed.) *New Advances in Machine Learning*, 3-31.

Van Passel, J.; De Keersmaecker, W.; Somers, B. (2020) *Monitoring woody cover dynamics in tropical dry forest ecosystems using sentinel-2 satellite imagery*.

Vesa, L.; Malimbwi, R.E.; Tomppo, E.; Zahabu, E.; Maliondo, S.; Chamuya, N.; Nsokko, E.; Otieno, J.; Dalsgaard, S. (2010). *A National Forestry Resources Monitoring and Assessment of Tanzania* (NAFORMA).

APPENDENCES

1: Table of ESA World cover class definition and characteristics

2: Tables of Accuracy Assessment results

3: Tables of Result Evaluation

APPENDIX 1

Table 3.4: Map layer and definition of the classes

Map code	**Land cover class**	**Descriptions**	**Palette/ color code**
10	Tree cover	This class includes any geographic area dominated by trees with a cover of 10% or more. Areas planted with trees for afforestation purposes and plantations. This class also includes tree covered areas seasonally or permanently flooded with fresh water except for mangroves.	
20	Shrubland	This class includes any geographic area dominated by natural shrubs having a cover of 10% or more. Shrubs are defined as woody perennial plants with persistent and woody stems and without any defined main stem being less than 5 m tall. Trees can be present in scattered form if their cover is less than 10%. Herbaceous plants can also be present at any density. The shrub foliage can be either evergreen or deciduous.	
30	Grassland	This class includes any geographic area dominated by natural herbaceous plants (Plants without persistent stem or shoots above ground and lacking definite firm structure): Woody plants (trees and/or shrubs) can be present assuming their cover is less than 10%. It may also contain uncultivated cropland areas (without harvest/ bare soil period) in the reference year	
40	Cropland	Land covered with annual cropland that is sowed/planted and harvestable at least once within the 12 months after the sowing/planting date. The annual cropland produces an herbaceous cover and is sometimes combined with some tree or woody vegetation. Note that perennial woody crops will be classified as the appropriate tree cover or shrub land cover type. Greenhouses are considered as built-up.	

50	Built-up	Land covered by buildings, roads and other man-made structures such as railroads. Buildings include both residential and industrial building. Urban green (parks, sport facilities) is not included in this class. Waste dump deposits and extraction sites are considered as bare.	
60	Bare /sparse vegetation	Lands with exposed soil, sand, or rocks and never has more than 10 % vegetated cover during any time of the year.	
70	Snow and Ice	This class includes any geographic area covered by snow or glaciers persistently.	
80	Permanent water bodies	This class includes any geographic area covered for most of the year (more than 9 months) by water bodies: lakes, reservoirs, and rivers. Can be either fresh or salt-water bodies. In some cases, the water can be frozen for part of the year (less than 9 months).	
90	Herbaceous wetland	Land dominated by natural herbaceous vegetation (cover of 10% or more) that is permanently or regularly flooded by fresh, brackish or salt water. It excludes unvegetated sediment (see 60), swamp forests (classified as tree cover) and mangroves see 95)	
95	Mangroves	Taxonomically diverse, salt-tolerant tree and other plant species which thrive in intertidal zones of sheltered tropical shores, "over wash" islands, and estuaries.	
100	Moss and lichen	Land covered with lichens and/or mosses. Lichens are composite organisms formed from the symbiotic association of fungi and algae. Mosses contain photo-autotrophic land plants without true leaves, stems, roots but with leaf-and stem like organs.	

APPENDIX 2

Table 4.1: Accuracy of Land Cover maps from 2017 to 2021 obtained with Optical product

	LC	**PWB**	**TC**	**SL**	**GL**	**B/SV**	**UA(%)**	**K**
	PWB	27	0	0	0	0	100	
	TC	0	41	2	0	0	97.62	
2017	**SL**	0	1	16	3	0	80	
	GL	0	0	2	33	0	84.62	
	B/SV	0	0	0	3	4	100	
	PA(%)	100	95.34	80	94	57.14	**91.67**	**0.89**
	LC	**PWB**	**TC**	**SL**	**GL**	**B/SV**	**UA(%)**	**K**
	PWB	27	0	0	0	0	100	
	TC	0	42	1	0	0	95.45	
2019	**SL**	0	2	13	5	0	92.86	
	GL	0	0	0	35	0	87.5	
	B/SV	0	0	0	0	7	100	
	PA(%)	100	97.67	65	100	100	**93.94**	**0.92**
	LC	**PWB**	**TC**	**SL**	**GL**	**B/SV**	**UA(%)**	**K**
	PWB	27	0	0	0	0	100	
	TC	0	41	2	0	0	93.18	
2021	**SL**	0	2	18	0	0	64.29	
	GL	0	1	8	26	0	92.86	
	B/SV	0	0	0	2	5	100	
	PA(%)	100	95.35	90	74.28	71	**88.64**	**0.85**

Here, **PWB**- Permanent water bodies, **TC**- Tree cover, **SL**- Shrubland, **GL**-Grassland, **B/SV**-Bare/sparse vegetation, **UA**- User’s accuracy, **PA**- Producer’s accuracy, **K**- Kappa statistics.

Table 4.2: Accuracy of Land Cover maps from 2017 to 2021 obtained with Radar product.

	LC	PWB	TC	SL	GL	B/SV	UA(%)	K
	PWB	34	0	0	0	0	100	
	TC	0	54	1	6	0	71.05	
2017	**SL**	0	7	0	2	1	0	
	GL	0	14	8	48	1	77.42	
	B/SV	0	1	0	6	1	33.33	
	PA(%)	100	88.52	0	67.61	12.5	**74.46**	**0.63**
	LC	**PWB**	**TC**	**SL**	**GL**	**B/SV**	**UA(%)**	**K**
	PWB	33	0	0	1	0	100	
	TC	0	51	5	3	2	95.45	
2019	**SL**	0	9	1	0	0	92.86	
	GL	0	9	1	56	5	87.5	
	B/SV	0	1	0	6	1	100	
	PA(%)	100	97.67	65	100	100	**77.17**	**0.67**
	LC	**PWB**	**TC**	**SL**	**GL**	**B/SV**	**UA(%)**	**K**
	PWB	34	0	0	0	0	94.44	
	TC	0	53	3	5	0	76.81	
2021	**SL**	0	6	3	1	0	37.5	
	GL	0	9	2	60	0	84.51	
	B/SV	2	1	0	5	0	0	
	PA(%)	100	86.89	30	84.51	0	**81.52**	**0.73**

Here, **PWB**- Permanent water bodies, **TC**- Tree cover, **SL**- Shrubland, **GL**-Grassland, **B/SV**- Bare/ sparse vegetation, **UA-** User's accuracy, **PA-** Producer's accuracy, **K-** Kappa statistics.

APPENDIX 3

Table 4.6: Area coverage in 2017

Class Name	OPTICAL Area (Ha)	RADAR Area (Ha)
Permanent water bodies	2403.80	2404.24
Tree cover	2698.40	2761.44
Shrubland	245.56	225.16
Grassland	262.76	231.12
Bare/sparse vegetation	13.84	2.40
Total	**5624.36**	**5624.36**

Table 4.7: Area coverage in 2019

Class Name	OPTICAL Area (Ha)	RADAR Area (Ha)
Permanent water bodies	2405.88	2407.16
Tree cover	2688.80	2724.12
Shrubland	265.60	296.56
Grassland	251.04	194.44
Bare/sparse vegetation	13.04	2.08
Total	**5624.36**	**5624.36**

Table 4.8: Area coverage in 2021

Class Name	OPTICAL Area (Ha)	RADAR Area (Ha)
Permanent water bodies	2404.08	2404.68
Tree cover	2770.92	2799.88
Shrubland	182.28	197.60
Grassland	260.16	220.56
Bare/sparse vegetation	6.92	1.64
Total	**5624.36**	**5624.36**