



JIMMA UNIVERSITY
COLLEGE OF SOCIAL SCIENCES AND HUMANITIES
DEPARTMENT OF GEOGRAPHY AND ENVIRONMENTAL
STUDIES

SATELLITE BASED TIMESERIES FOREST CARBON STOCK ESTIMATION AND MODELING UNDER DIFFERENT SCENARIOS. A CASE OF SHIWSHIWE FOREST, SOUTH-WEST ETHIOPIA.

By

OYTISA ONATO OLADO

A Thesis Submitted to School of Graduate Studies, Jimma University, College of Social Sciences and Humanities, Department of Geography and Environmental Studies in Partial Fulfillment of the Requirements for the Degree Master of Science in Geographic Information Systems and Remote Sensing.

July, 2021

Jimma, Ethiopia

JIMMA UNIVERSITY
COLLEGE OF SOCIAL SCIENCES AND HUMANITIES
DEPARTMENT OF GEOGRAPHY AND ENVIRONMENTAL STUDIES

SATELLITE BASED TIMESERIES FOREST CARBON STOCK ESTIMATION AND
MODELING UNDER DIFFERENT SCENARIOS. A CASE OF SHIWSHIWE
FOREST, SOUTH-WEST ETHIOPIA.

Advisor: Dr. Ajay Babu (Associate Prof.)

Co-Advisor: Mr. Ashenif Melese (MSc.)

A Thesis Submitted to School of Graduate Studies, Jimma University, College of Social
Sciences and Humanities, Department of Geography and Environmental Studies in
Partial Fulfillment of the Requirements for the Degree Master of Science in Geographic
Information Systems and Remote Sensing

July, 2021

Jimma, Ethiopia

JIMMA UNIVERSITY
COLLEGE OF SOCIAL SCIENCES AND HUMANITIES
DEPARTMENT OF GEOGRAPHY AND ENVIRONMENTAL STUDIES
GEOGRAPHIC INFORMATION SYSTEMS AND REMOTE SENSING

SATELLITE BASED TIMESERIES FOREST CARBON STOCK ESTIMATION AND
MODELING UNDER DIFFERENT SCENARIOS. A CASE OF SHIWSHIWE
FOREST, SOUTH-WEST ETHIOPIA.

A Thesis Submitted to School of Graduate Studies, Jimma University, College of Social
Sciences and Humanities, Department of Geography and Environmental Studies in
Partial Fulfillment of the Requirements for the Degree Master of Science in Geographic
Information Systems and Remote Sensing

Graduate Thesis Ownership Agreement

This thesis is a property of Jimma University, an institution that awarded **MSc** Degree to the graduate student and funded its research cost fully or partly. The research work was accomplished under the close support and supervision of the assigned University's academic staff. It is therefore, strictly forbidden to publish, modify, or communicate to or put at the disposal of third party the entire document or any part thereof without the common consent of the research supervisor(s) and the graduate student. Disregarding this agreement would lead to accountability according to the Jimma University's Research and Publication Misconduct Policy **Article 1.7** of the University's Document for "Guidelines and Procedures for Research, March 2013".

Name of the Graduate Student	Signature	Date
<u>OYTISA ONATO OLADO</u>	_____	_____
Name (s) of the Research Supervisor (s)	Signature	Date
Main Advisor: _____	_____	_____
Co-Advisor: _____	_____	_____
Name of Chairperson	Signature	Date
_____	_____	_____
Name of Internal Examiner	Signature	Date
_____	_____	_____
Name of External Examiner	Signature	Date
_____	_____	_____

Title of the Thesis:

Degree Awarded: **MSc**

ACKNOWLEDGMENTS

Above all, I would like to thank the Almighty God for His existence in every journey of my life. At the bottom of my heart, I would like to express my sincere and deepest gratitude to my advisors Dr. Ajay Babu (Associate Prof.) and co-advisor Mr. Ashenif Melese (MSc) for their unlimited and continuous support, guidance, and comments throughout my work.

Secondly, I would like to thank Ms. Tishager Zelalem and my family for their endless support in every path of my life.

To this extent, my sincere hearted gratitude also goes to Prof. Tsegaye Bekele, Dr. Kefelegn Getahun, Dr. Kefyalew Sahle, Dr. Marcelo Rezende, and Dr. Florence Landsberg for their inexpressible input into my academic-life and their kin encouragement to carry my thesis work.

To this end, I would like to thank Hawassa University for providing me sponsorship to study my MSc in GIS and Remote Sensing at Jimma University. My deepest thanks also go to ‘Alle Special Woreda Administration Office’ for providing me logistics during field campaigns and data collection.

ABSTRACT

Accurate quantification of forest above-ground live carbon stock is crucial for sustaining forests as a climate change mitigation mechanism. However, the major challenge is to find an accurate technique for timely processing and deriving biomass information estimates at the extent of big data. This study is thus mainly aimed in estimating and modeling forest aboveground biomass and carbon stock using field inventory data and automated scenario based analysis of satellite images in Google Earth Engine cloud computing environment. The carbon stock estimation was computed at three stages viz: at plot level, at full-hectare basis and using satellite images. The plot level biomass was estimated based on field inventory of 20 circular plots which then extrapolated to determine the biomass of the study area. The resulting biomass was then converted into carbon based on IPCC guidelines. For satellite based biomass and carbon stock estimation, 10m surface reflectance bands and respective vegetation indices were geocomputed for the years 2016 to 2020. These were later correlated with derived satellite variables to validate the AGC derived from sentinel-2 and SAR images of 2019. The coefficient of determination (R^2) between observed and the predicted AGC was then used to validate the estimated result. The sentinel 2 NDVI showed the strongest correlation ($r = 0.9$) with AGC in the study area. Sentinel 1 variables revealed moderate correlation ($r = 0.08$ to 0.7) with the on-situ AGC. To develop AGB predictive model, NDVI, NDWI, EVI and Entropy were selected based on their correlation coefficient and variable importance. The model has a coefficient of determination value of 0.86. Forest above ground carbon stock map was produced by the developed model and masked by the forest cover produced from fusion image composited from B2, B3, B4, B8, NDVI, EVI, NDWI, VH, elevation and slope band. The per-pixel AGC for the study area ranged from 0.538 – 1.6153, 0.353 -1.518, 0.319-1.480, 0.517 – 1.614 tons for the years 2016 to 2020 respectively. The total plot level per pixel AGC was found between 6 and 7.4 tons. The fusion of Sentinel-2 variables with sentinel-1 GRD images, GLCM and ALOS DSM scenario performed better in estimating AGB and carbon stock compared to use of sentinel-1 and Sentinel-2 alone with model accuracy of 98%. Overall, integrating field data with multisensor remote sensing method increases the accuracy of modeling and estimating forest AGC stock.

Key Words: Above Ground Carbon, GEE, GLCM, Image Fusion, SAR, Sentinel-2.

TABLE OF CONTENTS

DECLARATION	I
ACKNOWLEDGMENTS	II
ABSTRACT.....	III
TABLE OF CONTENTS.....	IV
LIST OF TABLES	VII
LIST OF FIGURES	VIII
ACRONYMS	IX
CHAPTER ONE	1
INTRODUCTION	1
1.1 Background	1
1.2 Statement of the Problem	7
1.3 Objectives of the Study	8
1.3.1 General Objective	8
1.3.2 Specific Objectives	8
1.4 Research Questions	9
1.5 Significance of the Study	9
1.6 Scope of the Study.....	10
1.7 Limitation of the Study	10
1.8 Organization of the Paper.....	11
CHAPTER TWO	12
LITERATURE REVIEW	12
2.1 Forests Resources in Climate Change Perspective	12
2.2 Policy Directions and Forestry Sector in Ethiopia	13
2.3 Forest Inventory and Biomass Estimation	14
2.4 Forest Carbon Stock Estimation and Accounting Tools	17
2.5 Potential Remote Sensing Contributions to Carbon Stock Modeling	18
2.6 Multisource Data Fusion for AGC Stock Monitoring.....	22
CHAPTER THREE	23
METHODS AND MATERIALS.....	23
3.1 Description of the Study Area.....	23
3.1.1 Location	23

3.1.2 Climate Condition.....	23
3.1.3 The livelihood and Landcover of the Area.....	25
3.2 Materials.....	26
3.2.1 Data Type and Source.....	26
3.2.2 Software and Tools.....	27
3.2.3 Field Materials.....	28
3.3 Methods.....	28
3.3.1 Research Design.....	28
3.3.2 Study variables.....	29
3.3.3 Sampling Procedures and Sample Size.....	30
3.3.4 Google Earth Engine (GEE) Data Processing.....	33
3.3.5 Spectral Indices and Transformations.....	34
3.3.6 Variable Selection for Forest AGC Prediction.....	34
3.3.7 Prototype Model Development.....	35
3.4 Data Processing and Analysis.....	35
3.4.1 Tree level Above Ground Biomass Estimation.....	35
3.4.2 Plot Based Carbon Stock Estimation.....	37
3.4.3 Satellite Based AGC Mapping prototype Development Scenarios.....	38
3.4.4 Predictive Model Development Scenario Analysis.....	57
CHAPTER FOUR.....	60
RESULTS AND DISCUSSION.....	60
4.1 Total Aboveground Biomass and Carbon Stock Estimation.....	60
4.2 GEE Timeseries Prototype of Forest AGC Stock Development.....	62
4.3 The aboveground carbon stock predictive model development.....	74
4.3.1 Correlation of Sentinel-2 satellite variables with Observed AGC.....	74
4.3.2 Correlation of Sentinel-1 satellite variables with Observed AGC.....	74
4.3.3 Correlation Summary of the Independent Variables with Observed AGC.....	76
4.3.4 Selection of the Best Variable to develop predictive equation.....	77
4.3.5 Comparison Vegetation Indices and Backscatter.....	79
4.3.6 Aboveground Carbon Stock Estimation Based on Best Variable.....	80
4.3.7 Correlation between Observed and predicted AGC.....	86
CHAPTER FIVE.....	88

CONCLUSION AND RECOMMENDATION.....	88
5.1 Conclusion.....	88
5.2 Recommendation.....	90
REFERENCES	92
APPENDICES	97
ANNEX 1 Field Photos.....	97
ANNEX 2 Field Based AGB and AGC Stock Calculations	104
ANNEX 3 Satellite variables extracted for field plots.....	109
ANNEX 4 Summary of AGC and Satellite derivatives regression analysis.....	110
ANNEX 5 Regression Analysis Equations for Predictive Model Development	113
ANNEX 6 Sentinel-1 GRD products spatial resolution.....	115
ANNEX 7 Sentinel-2 Spatial Resolution Spectral Bands.....	116
ANNEX 8 Prototype Modules Codes	117

LIST OF TABLES

Table 1: Forest carbon pools.....	14
Table 2: Main methods for forest AGB estimation	15
Table 3: Wavelengths most frequently used in SAR	20
Table 4: Average climate and vegetation periods (source: Researcher).....	25
Table 5: Sentinel-2(A and B sensors) bands in GEE.....	26
Table 6: Summary of data type and sources	27
Table 7: Software and application tools.....	27
Table 8: Field survey equipment.....	28
Table 9: Vegetation index variables used in GEE	34
Table 10: Tree level AGB calculations.....	37
Table 11: Sentinel-1 image preprocessing steps in GEE.....	45
Table 12: Bands of the multispectral fusion image	50
Table 13: Binary classification scheme for forest masking	54
Table 14: Random forest classification Scenarios	68
Table 15: Accuracy assessment in GEE for scenario-3	72
Table 16: Summary of regression statistics between AGC and vegetation indices.....	74
Table 17: Summary of regression statistics between AGC and Sentinel-1 variables.....	75
Table 18: Results of multiple regression	78
Table 19: Field survey and plot (tree) level AGB estimations	104
Table 20: Plot level field AGC stock estimations.....	109
Table 21: Satellite variables used as input in regression analysis	109
Table 22: AGC and NDVI regression analysis.....	110
Table 23: AGC and SAVI and NDWI regression analysis.....	111
Table 24: Results of GLCM Sum-Average texture and AGC regression analysis.....	112
Table 25: Results of GLCM Entropy texture and AGC regression analysis	112
Table 26: Results of GLCM correlation texture and AGC regression analysis.....	113
Table 27: Results of GLCM Variance texture and AGC regression analysis.....	113
Table 28: Modes and Spatial resolution of Sentinel-1 GRD Products	115
Table 29: GEE Sentinel-2 image computational time, memory and algorithms	124

LIST OF FIGURES

Figure 1: Forest carbon pools and fluxes (adapted from Vickers B. et al. 2012)	2
Figure 2: Functions of the national monitoring system (source: Bekele et al., 2019)	4
Figure 3: Share of GHG emissions of the different economic sectors in Ethiopia.....	14
Figure 4: Radar backscattering in forests.....	20
Figure 5: Sentinel-1 data acquisition modes.....	21
Figure 6: Location map of study area (source: Researcher)	23
Figure 7: Timeseries temperature distribution overtime in ⁰ c	24
Figure 8: Seasons and climate of the study area (source: Researcher)	24
Figure 9: LULC Map of the study area (source: Researcher).....	25
Figure 10: Research design framework	29
Figure 11: Circular plot of 10m radius used in the field.....	30
Figure 12: Field inventory plots.....	31
Figure 13: Tree caliper held at right angles to the tree trunk for DBH measure.	31
Figure 14: Measurement of tree diameter at breast height (DBH)	32
Figure 15: Example 2016 cloudy Sentinel-2 median raw image.....	39
Figure 16: Sentinel-1 products naming convention (Source: ESA, 2020).....	44
Figure 17: Example of RandomForest prediction.....	52
Figure 18: Input training points for the RF classifier	53
Figure 19: Sentinel-2 image processing and Random Forest classification	56
Figure 20. Overall methodological flow chart.....	59
Figure 21: DBH distribution across plots	60
Figure 22: Plot level AGB in tonnes per hectare	61
Figure 23: Plot level AGC stock (kg C/ha).....	61
Figure 24: Processing chain of the JavaScript prototype.....	64
Figure 25: ANN-based Sentinel-2 band selection and aggregation GUI.....	65
Figure 26: Comparison of raw Sentinel2 bands (A) with Fusion image bands (B).....	66
Figure 27: Timeseries of multispectral pixel level data fusion result.....	67
Figure 28: Graph of Variable Importance.....	68
Figure 29: Scenario-1 and Scenario-2 Forest cover Maps.....	69
Figure 30: Scenario-3 based forest cover maps	70
Figure 31: R ² of independent variables based on simple linear regression	76
Figure 32: The correlation between AGC and NDVI and the predictive model	77
Figure 33: Aboveground Carbon Maps	81
Figure 34: Plot level Predicted AGC	82
Figure 35: Peak NPP from 2016 to 2020 (kgC/m ²)	85
Figure 36: Sentinel-2 10m spatial resolution bands.....	116
Figure 37: Sentinel-2 20m spatial resolution bands.....	116
Figure 38: Sentinel-2 60m spatial resolution bands.....	117

ACRONYMS

ACA	Agricultural Commercialization Agency
AFOLU	Agriculture Forestry and
AGB	Above Ground Biomass
AGC	Above Ground Carbon
AI	Artificial Intelligence
ALOS	Advanced Land Observing
ANN	Artificial Neural Network
AOI	Area of Interest
API	Application Programming Interface
BAU	Business As Usual
BGB	Below Ground Biomass
BOA	Bottom of Atmosphere
CEO	Collect Earth Online
CHIRPS	Climate Hazard InfraRed
CPU	Central Processing Unit
CRGE	Climate Resilient Green
DBH	Distance at Breast
DEM	Digital Elevation Model
ESA	European Space Agency
FAO	Food and Agricultural
FDRE	Federal Democratic Republic
FRA	Forest Resource Assessment
GDAL	Geospatial Data Abstraction
GEE	Google Earth Engine
GHG	Green House Gases
GIS	Geographic Information Systems
GRASS	Geographic Resource Support
GRD	Ground Range Detected
INDCs	Intended Nationally Determined
IPCC	Intergovernmental Panel on
INDC	Intended Nationally Determined
IW	Interferometric Wide
LIDAR	Light Detection and
LULC	Landuse and Land
ML	Machine learning (ML)
MMUs	Minimum Mapping Units
MRV	Measurement Reporting and
MSI	Multispectral Instrument
NASA	National Aeronautics and

NDC	Nationally Determined
NDVI	Normalized Difference Vegetation
PALSAR	Phased Array type
RADAR	Radio Detection And
REDD+	Reducing Emissions from
RMS	Root Mean Square
SAR	Synthetic Aperture Radar
SNNP	Southern Nations Nationalities
SRTM	Shuttle Radar Terrain
TOA	Top of Atmospheric
UAV	Unmanned Aerial Vehicle
UNFCCC	United Nations Framework
WBISPP	Woody Biomass Inventory
WAPOR	Water Productivity Database

CHAPTER ONE

INTRODUCTION

1.1 Background

Climate change is one of the world's most pressing challenges these times. Anthropogenic activities such as deforestation and the expansion of agricultural land become the main factors of climate change. These activities have led to the release of numerous pollutant gases into the air such as carbon dioxide (CO₂), methane (CH₄), and various organic gases (Adnan, Yusoff, et al., 2017). The global temperature has increased by around 1°C since pre-industrial times as the emissions of these greenhouse gases' concentration (GHG) rise (Reiche, 2015).

The changing climate as a result of these gases has thus a range of potential ecological, physical and health impacts. These impacts include but not limited to extreme weather events such as floods, droughts, storms, and heatwaves; sea-level rise; altered crop growth; and disrupted water systems. A process that contributes to the removal of CO₂ from the atmosphere and storage for a certain period is known as a carbon sink. This carbon storage occurs mainly in oceans, soils, and forests, where organisms capture carbon and release oxygen into the atmosphere. To enhance the carbon storage ecosystems and make the climate at its normal condition globally, there has been thus an increasing effort by nations. This basically includes responding to the effects of the changing climate within the extent of the United Nations Framework Convention on Climate Change (UNFCCC) and protect forest resources (UNFCCC, 2016, SERVIR, 2019).

Protecting forests has thus gained concern as a mitigation mechanism to combat climate change. This is because forests represent one of the most important ecosystems in our planet and a fundamental element to fight climate change. They provide fundamental ecosystem services such as absorbing, storing and converting carbon dioxide (CO₂) into oxygen. Moreover, forests absorb and store about 30% of current levels of carbon emissions from fossil fuels and industry into their carbon pools(GFOI, 2013).

The main forest carbon pools within and around vegetation that have the potential to store much more carbon are Above-ground biomass (stems, bark, and leaves), Below-ground biomass (roots of all sizes), Deadwood (organic matter in deadwood), Litter or dead organic matter in litter and soil organic carbon (SOC). These pools are interconnected to one another through several important fluxes such as photosynthesis as shown in Figure 1 below. These fluxes keep a balance between CO₂ captured by forests through photosynthesis and the CO₂ released through respiration and forest fires, representing a positive balance of capture and storage (*The Carbon Cycle*, 2011).

However, the AGB in natural forest trees often represent the greatest fraction of total biomass of a forested area and respond more rapidly and significantly as a result of land-use change than other carbon pools. Consequently, the majority of carbon accounting efforts are focused on tree AGB as the carbon stored in the living biomass of trees is typically the largest pool and the most directly impacted by deforestation and degradation. The estimate of forest AGB carbon is thus a critical step in quantifying carbon stocks and fluxes (*Figure 1*) from tropical forests (Barredo et al., 2016).

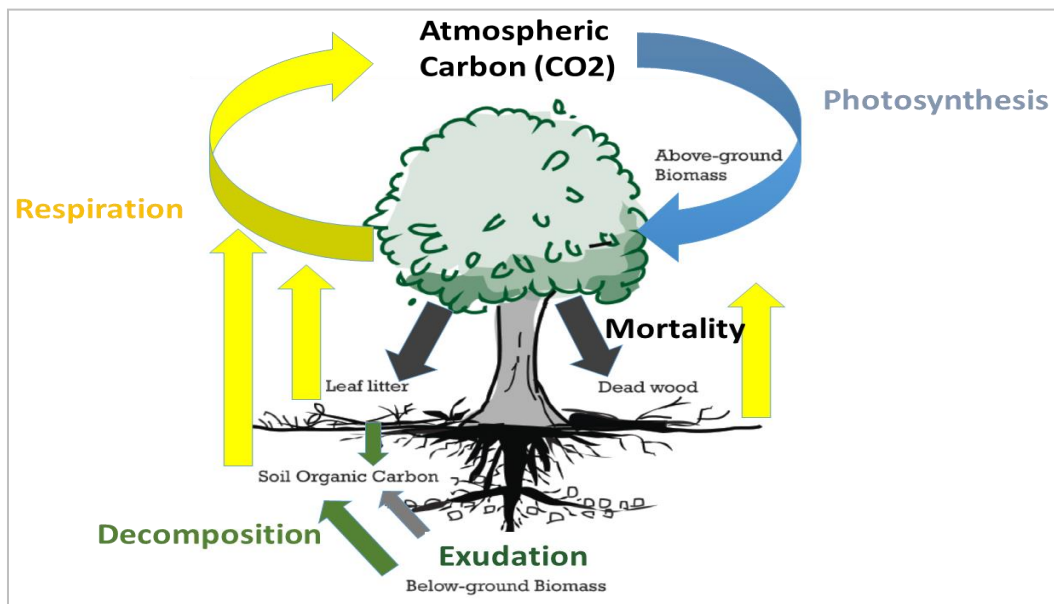


Figure 1: Forest carbon pools and fluxes (adapted from Vickers B. et al. 2012)

Moreover, forests mitigate climate change through carbon sequestration, contribute to the balance of oxygen, carbon dioxide, and humidity in the air supplying 75% of freshwater worldwide. Forests are the most biologically diverse ecosystems on land, home to more

than 80% of the terrestrial species of animals, plants, and insects. This makes forests especially, tropical forests as vital in the global carbon balance. Moreover, forests can effectively contribute to CO₂ capture and sequestration over different periods through protection, conservation, and forest management in light of the high existing degradation and deforestation rates (Plugge et al., 2016).

Conversely, the rapid shrinking of the forests creates a major threat and increase the effects of climate change. This ends the net flow of carbon from the atmosphere into the forest. In forests that are left standing, logging, wood fuel extraction, fires and grazing typically reduce carbon stocks faster than they can naturally recover. Deforestation and forest degradation, located primarily in tropical regions, have thus reported accounting for 12–20% of global anthropogenic greenhouse gas (GHG) emissions in the 1990s and early 2000s. These processes also impact the future potential of forests to remove additional carbon from the atmosphere (Nunes et al., 2020).

An earlier study in the tropics by (Chave et al., 2005) has investigated that tropical forests are important in global climate change mitigation as they contain as much carbon in their vegetation and soils as temperate and boreal forests combined. The Intergovernmental Panel on Climate Change (IPCC) estimates that forestry or, rather, deforestation, forest degradation and other changes in forests-contributes 17.4% of global annual greenhouse gas (GHG) emissions mainly in tropical developing countries. This is 5.8 gigatons (Gt) of carbon dioxide equivalent (CO₂-e) per year (Brack, 2019).

Reducing Emissions from Deforestation and Forest Degradation in developing countries (REDD) (UNFCCC,2013) is one of the key global initiatives under the UNFCCC. REDD was extended to REDD+ in 2014 by adding the carbon sink potential as well as conservation and sustainable forest management issues. It aims at mitigating climate change by rewarding tropical developing countries for reducing emissions by avoiding deforestation and forest degradation while also slowing the drivers of landuse change (Thapa, et al., 2015).

However, successful implementation of such mechanisms to reliably estimate GHG emissions require consistent and accurate detection of tropical forest cover change (Reiche, 2015) and the corresponding carbon stock of the land that is cleared (Saatchi et

al., 2011). Ethiopia under the Paris agreement (Nations, 2016) is therefore inquired to design and implement an operational monitoring, reporting, and verification (MRV) system (see Figure 2: Functions of the national monitoring system) to report changes in forest and forest carbon stock for receiving rewarding compensations (Reiche, 2015). This has been clearly stated in the Ethiopian national CRGE strategy, in particular, the green economy strategy as a basis for determining Ethiopia's proposed contribution to GHG emission reductions by sector.

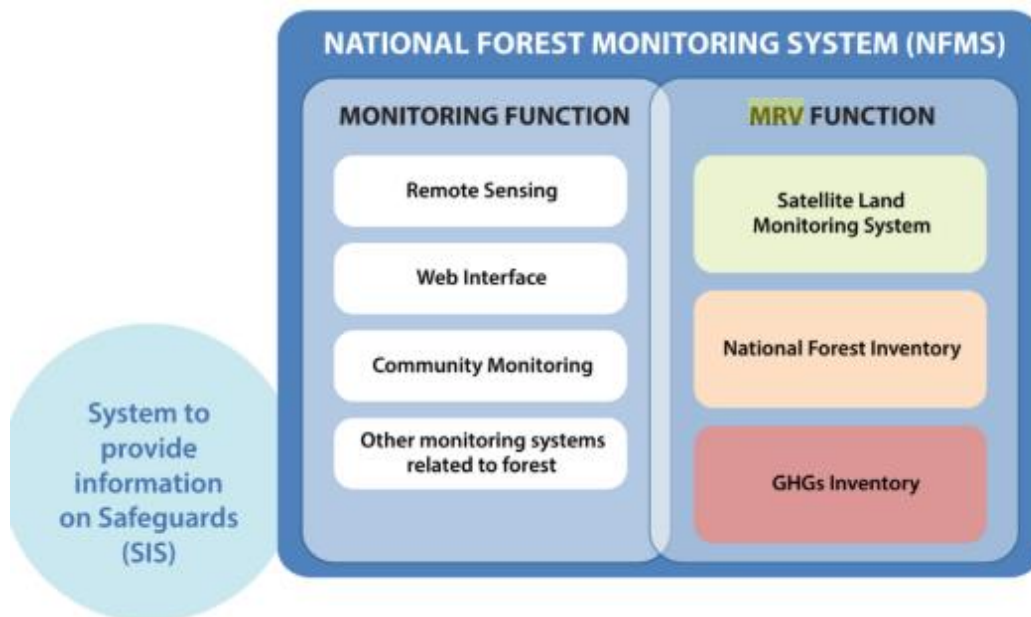


Figure 2: Functions of the national monitoring system (source: Bekele et al., 2019)

To contribute and Ethiopia to complement this agreement, managing forests is crucial to respond to the rising climate change issue. Besides this, it enables to generate carbon credit in line with conserving biodiversity and improving the livelihoods of the communities (Abera et al., 2017). Assessing the forest carbon stock is thus one of the important steps to start with sustainable landuse planning concerning low carbon emission. This is because in developing countries for example, about 38 % of the primary energy consumption comes from forest biomass. In Ethiopia, biomass supplies 93 % of total household energy consumption (Shiferaw et al. 2010).

To successfully implement mitigating policies and take advantage of the REDD+ program of the UNFCCC, a well-authenticated estimates of forest carbon stocks is

needed. Though efforts to measure and map biomass stocks nationwide is growing, obtaining high-quality information on biomass carbon stocks is one of the most evasive tasks associated with REDD+ MRV in Ethiopia. These efforts on natural forest restoration and the development of plantations are expected to result in a reduction of forest degradation and deforestation. For this reason, different methodologies are being piloted to assess and monitor the state of forests such as the use of dense time-series image analysis at project scale (MEFCCC, 2017).

To develop a robust methodology for monitoring forest AGB carbon and reduce carbon emissions, successful attempts at the local level such as forest inventory, wall-to-wall and sample-based remote sensing are thus vital. Hence these efforts may be transferred into a cost-effective accounting mechanism at the national level for its future inclusion in the forest reference level (FRL).

However, the quantification and estimation of spatially explicit services for carbon stock potential may differ based on climate, management applications, ecosystem, species, and the local communities (Ndalowa, 2014). The amount and spatial distribution of carbon stored in forests is still uncertain. This is due to the difficulties in measuring AGB and carbon stock at the field scale that complement the national requirements. So far, three kinds of methods have been proposed to estimate the forest AGB, including national forest inventory (NFI)-based methods, process model-based methods, and remotely sensed data-based methods.

The field-based inventory methods can observe accurate parameters such as tree height and diameter at breast height (DBH) for AGB estimation. Though destructive sampling methods were initially used to estimate living tree biomass, several studies (Tesfaye et al., 2016) have used allometric equations as an alternative method for biomass estimation (Adnan et al., 2017; Reiche, 2015). The most important AGB a tree predictor parameters as reported by Chave et al., (2005) in decreasing order of importance are tree trunk diameter, wood specific gravity, total height, and forest type (dry, moist, or wet) with an overestimate bias of 0.5- 6.5% when errors were averaged across all stands. However, these methods mostly focus on some of the representative wood rather than all forest

biomass and are always labor-intensive and time-consuming. This has called to the use of remote sensing satellite data for forest biomass estimations.

Remote sensing thus offers an economic, efficient and robust mode of AGB monitoring and activity-data at different scales than field-based inventory methods (Adnan et al., 2017). The use of band ratio vegetation indices derived from remote sensing satellite images such as enhanced vegetation index (EVI), normalized difference vegetation index (NDVI), and simple ratio vegetation index (SRVI) is essential to calculate the biomass density. Thus, the advancement of remote sensing technologies and methods to monitor forest cover across large extent landscapes these times has become a substantial priority.

Yet, more continuous information has been available since the start of the more recent space programs, like the Sentinel program of the European Space Agency (ESA) to build observations based on time series of images (Udali, 2019). This helps for monitoring the state of forest carbon pools (Dick OB, 2015) in combating forest disturbances which is a big issue in the study area threatening the tree species diversity and value of the forests to the community. Using large data sets, excellent progress has been made in detecting changes in AFOLU in the new big-data era (Wulder, et al., 2018). The best-known existing algorithms typically use most often calibrated optical satellites such as NASA's Landsat imagery to compare bands or build robust indices that are differenced or otherwise interpreted through time for stability and change (Fortin et al., 2020).

Though such optical satellites are incontrovertibly powerful, they cannot easily detect and sense forest areas especially in cloudy weather and night times. This could likely be made even more beneficial through image fusion which is less considered in previous studies. To capture current and future changes in forest carbon stocks, satellite remote sensing at medium spatial resolution (10 - 30 m) supported by field observations is the appropriate tool for most tropical countries (Betru et al., 2019a; De Sy, et al., 2012).

Therefore, the rapid and accurate estimation and monitoring of AGB over various scales of space and time are crucial for greatly reducing the uncertainty in carbon stock assessments, and for informing strategic forest management plans. This further enhances the operationalization of an integrated policy and strategy to increase adaptability to the

adverse impacts of climate change and foster climate resilience and low GHG emissions development in a manner that does not threaten food production.

This paper is thus aimed at the use of Google Earth Engine (GEE) satellite timeseries images to extract forest carbon stocks. It basically uses data from multi-temporal Sentinel-2 MSI and Sentinel-1 backscatter data and assesses its applicability in estimating AGC timeseries maps under different scenarios. The GEE makes available not only near real-time data access but also a very powerful application programming interface (API) for processing and visualizing the data. The study was further intended to use generalized allometric equations developed for a similar biophysical environment as an alternative method for AGB and carbon stock estimation. The results of the analysis were validated with a field survey and compared to global a reference forest mask derived from global forest mosaics rescaled to local scale.

1.2 Statement of the Problem

Ethiopia is facing rapid deforestation and degradation of land resources and experiencing the effects of climate change (MEFCC, 2018). Illegal harvesting of important timber species such as *Cordia africana*, *Hagenia abyssinica*, and *Afrocarpus falcatus* has degraded the quality and altered the structure and the species composition of the forest over time (Abere et al., 2017). In general, large forested areas are cleared out every year due to different driving forces. This is thriving issue in the study area in particular where there is no information on the forest structure.

These problems call for enhanced and continuous efforts to improve forest management information to enhance their productivity and production including for biodiversity and carbon stock. In addition, the forest carbon markets such as REDD+ require spatially explicit patterns of forest carbon stock timeseries values with high accuracy. Such markets need to identify where to be focused to enhance the value of the forest to the community (Thapa et al., 2015). However, the tools to extract such timeseries information for biomass estimation remain scarce and existing generalized models experience high uncertainty to represent biomass carbon stock in the actual forests of Ethiopia and the study area in particular where all these are lacked.

The use of optical satellite-based monitoring of forest biomass and carbon stock has been applied in similar studies but this also remains problematic due to cloud cover, seasonality, and the limited spatio-temporal resolution (Su et al., 2020). Earth observation data analyses also become more difficult and more expensive with smaller minimum mapping units (MMUs) that increase mapping efforts and usually decrease change mapping accuracy. The use of texture parameters is indicated as prominent for land cover and vegetation classification in previous research, but image texture and its relation with forest carbon stock have not yet been fully investigated. Moreover, Scientific research investigating the relation between specifically the Sentinel-1 backscatter and VIs was not found.

Besides this, the sheer size and a large number of open satellite images make developing prototypes for monitoring the abrupt changes in a forest carbon stock status rather cumbersome at larger extent. To get from an idea to a prototype product oftentimes end up spending more time on downloading and pre-processing than the actual prototyping and geocomputation (Canty et al., 2019). The study thus took the advantages of the combined use of sentinel-1 and Sentinel-2 in GEE to fill the information gap in timeseries carbon stock monitoring.

1.3 Objectives of the Study

1.3.1 General Objective

- The aim of this research is to estimate and model satellite-based timeseries forest carbon stock under different scenarios between 2016 and 2020 in Shiwshiwe forest.

1.3.2 Specific Objectives

- To estimate the total aboveground carbon stock of the forest areas
- To develop a Google Earth Engine prototype model of timeseries forest aboveground carbon stock
- To develop a predictive model of aboveground carbon stock estimation

1.4 Research Questions

1. What is the total potential forest aboveground live tree biomass carbon stock in the study area?
2. Can fusion of multisource sentinel-1 and sentinel-2 products be integrated to enhance the prototype models for forest carbon stock monitoring at a different frequency or with an optical data scenario?
3. How can be predictive aboveground carbon model developed for improved aboveground carbon stock estimation?

1.5 Significance of the Study

Geocomputational modeling of the amount of tree carbon stock is important for the design and implementation of effective forest management plans and policies. The study can contribute to continuous timeseries monitoring of biomass & carbon stocks at an efficient time and optimized accuracy. It will further enhance the knowledge gap in estimating timeseries above ground carbon stock taking advantage of opensource geospatial tools and RADAR techniques. The automated satellite image time series analysis carried in this study as monitoring approach is crucial to detect and analyze the recurring changes.

This provides the ability to constantly monitor and accumulate miscellaneous information from the available petabytes of data. The use of GEE in this regard reduces the processing of bigdata to derive patterns, determine trends, and identify deviations in vegetation status. The methodology adopted in this study can also be applied by researchers to assess land cover changes, bioenergy mapping and ecosystem services monitoring at larger extent. It will also allow planners, policy makers and practitioners with enhanced capacity to timely measures for integrated and reliable decisions for maximized community benefit.

The study will also provides new insights to develop national-level technologies for the management of natural resources such as portals that provides various online GIS to its users without having any specific GIS software installed in the client's machine. This is

very important especially for resource limited sectors to share and disseminate information to end users. The prototype developed can be easily shared to established observational network and remote sensing-based spatial databases for modeling and periodic assessment of carbon stock and balance at different tiers.

Thus, in general, reliable and continuous information on forest carbon content is crucial especially for the REDD+ readiness program (MRV) to mitigate climate change impacts. With this in the center, it goes in line with the advancement in the earth observation and geospatial technologies and current government policy and directions towards mitigating climate change impacts while the sustainable benefit of the ecosystem to the society is attained.

1.6 Scope of the Study

The study focuses on estimating and modeling forest above-ground biomass carbon using Sentinel-1 SAR GRD backscatters and Sentinel-2 optical MSI data in Shiwshiwe forest, Alle Special district, South-West of Ethiopia. It was carried between January 01, 2020 to December 30, 2020 by analyzing satellite data using opensource methods as a means of forest monitoring and assessment through developing a prototype model in GEE and a predictive model of timeseries of forest AGB carbon applying multisource satellite fusion of sentinel-1 SAR and Sentinel-2 as well as live tree inventory.

It was also restricted to satellite-based timeseries forest aboveground carbon stock estimation and prediction under business as usual (BAU): using field inventory only and optical satellites only scenarios and combined multisensor fusion scenarios excluding other biomass carbon pools and land cover type. The models developed were specific to extent of the study area because of the difference in information gained from satellite images of different landforms and forest types but the methodology developed through the research can be applied elsewhere with similar environmental setup.

1.7 Limitation of the Study

To come up with the best results, this study used forest inventory and sentinel-1 and sentinel-2 data as well as a literature review to automate the analysis method in GEE.

However, the limitation of this study concerns with absence of unavailability of long wavelength SAR imagery such as L band for free and laser scanners for detailed forest above ground biomass and carbon stock estimation.

The study solely focuses on developing a point based timeseries prototype and predictive model and did not considered the area based timeseries loss and gain of forests over the period. There is also no continuously inventoried field data for AGC monitoring in the area to predict future potential changes in aboveground forest carbon stock. Furthermore, there are no high resolution satellite imageries at the spatial resolution specified in this study for historical monitoring of the carbon stocks.

1.8 Organization of the Paper

The research report comprises five chapters. Chapter 1 – *Introduction*: introduces the background, statement of the problem, the objective of the study, the significance of the study, the conceptual framework of the research as well as the scope of the study. Chapter 2 – *Literature review*: reviews the concept of a forest, current and future information requirements for policy directions in the forestry sector, above-ground biomass, methods to estimate AGB, and application of synthetic aperture radar and optical sentinel-2 remote sensing data. Chapter 3 – *Materials and methods*: describes the characteristics of the study area, materials and methods to be used in the research to achieve the research objectives and intermediate result of analysis to be achieved. Chapter 4, *Results and Discussion*, presents the output of all the steps followed in chapter 3 from the selection of the datasets to the processing techniques and algorithms used for timeseries image analysis and classification of the satellite images. It discusses the result and development of the methodology. Finally, Chapter 5 presents *conclusions and recommendations*.

CHAPTER TWO

LITERATURE REVIEW

2.1 Forests Resources in Climate Change Perspective

Human beings have become a component in the earth's system, driving and accelerating global warming through the rapid release of greenhouse gases (GHGs) into the atmosphere (NAP, 2010). The composition of the atmosphere is being altered in such a way through increasing the concentration of greenhouse gases (GHGs) in the atmosphere by fossil fuel burning and by deforestation and land use changes. These are anthropogenic causes of climate change that has led to climate change discourse posing huge challenges to nations, organizations, enterprises, cities, communities and individuals. Developing countries will suffer most from adverse consequences of climate change, and the highly vulnerable regions and people are already being affected (Plugge et al., 2016).

The global climate change discourse recently has emphasized increased commitment and contribution inquiring individual partner countries to submit their national contribution regarding climate change mitigation and adaptation activities referred to as Nationally Determined Contributions (NDCs). In this context, there has been an increasing effort by nations to respond to climate change issues within the extent of the United Nations Framework Convention on Climate Change (UNFCCC) and protect their forest resources (Nations, 2016), (SERVIR, 2019).

Forest resources gained attention in this regard as forest carbon stocks have a key role in mitigation and adaptation with climate change. The forest carbon is mainly stored in the above-ground biomass (AGB) covering 70-90%. Forests absorb large amounts of atmospheric carbon dioxide (CO₂) in terrestrial land and soil. Forests play a crucial role in climate mitigation providing two-thirds of the cost-effective nature-based climate

mitigation needed to hold warming to below 2°C. Despite this, deforestation and degradation processes continue. The quality of natural forests is considered poor and forest areas are fragmented or degraded, with poor natural biodiversity (Ma et al., 2021).

Hence, modeling of forest resource and status are very important economically, ecologically and for sustainable forest management. The sustainable forest management as mechanism to mitigate climate change is affected by several factors linked to global change such as deforestation. Accurate and reliable information is therefore required to reduce the impact of climate change. This is crucial in the MRV system to provide spatially explicit estimates that adhere to IPCC principles of transparency, comparability, consistency, completeness and accuracy (Mitchell et al., 2017).

2.2 Policy Directions and Forestry Sector in Ethiopia

The efforts under the IPCC have led to multilateral negotiations and the adoption of intermingling policies and measures to circumvent the dangerous interference of anthropogenic activities in the climate system. The policies and measures are principally aimed at improving resilience and reducing climate change impact risks. The Ethiopian Policy and Strategy on the development, conservation and use of forests, adopted in 2006, states that the strategy is developed to bring sustainable development through community participation. The policy strategies pursued include a systematic control of forest resources from possible threats of theft and misuse. The preamble of the Forest Development, Conservation and Utilization Proclamation No 542/2007 states that the sustainable utilization of the country's forest resources is possible through ensuring the participation of, and benefit sharing by the concerned communities (FARM-AFRICA, 2010).

Under the business-as-usual scenario (BAU) however, the forestry sector will remain the second largest contributor to land-based emissions, due to forest loss from other sectors such as agriculture, livestock, and energy. Under this scenario, the pressure on forests and related increased rate of deforestation and forest degradation will ultimately deplete the natural resource base and negatively affect the performance of the economy in the long-

term. Under the current practices, greenhouse gas (GHG) emissions would more than double from 150 Mt CO₂e in 2010 to 400 Mt CO₂e in 2030 (CRGE, 2011).

Ratifying the Paris Agreement in 2017, the Ethiopian first NDC builds extensively on and is well aligned with the CRGE strategy. Ethiopia's submitted its NDC, containing mitigation actions and goals based on the national strategy to achieve a climate-resilient green economy (CRGE) (FDRE, 2015). In particular, the green economy strategy is set as a basis for determining Ethiopia's proposed contribution to GHG emission reductions by sector. The forestry sector contributes to half of all emissions next to the agricultural sector. It accounts for about 37% of national GHG emissions (Melaku, et al., 2018).

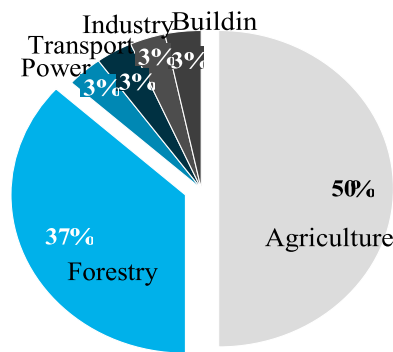


Figure 3: Share of GHG emissions of the different economic sectors in Ethiopia

2.3 Forest Inventory and Biomass Estimation

A forest is composed of five carbon pools stored in the living trees above and belowground, in dead matter including standing dead trees, down woody debris and litter, in non-tree understory vegetation and the soil organic matter. When trees are cut down there are three destinations for the stored carbon – dead wood, wood products or the atmosphere (GOFC-GOLD, 2003).

Table 1: Forest carbon pools

	Pools	Description
Living biomass	Above ground biomass (AGB)	All living biomass above the soil including stem, stump, branches, bark, seeds and foliage
	Below ground biomass (BGB)	All living biomass of live roots

Dead organic matter	Dead wood	Includes all non-living woody biomass not contained in the litter either standing or lying on the ground
	Litter	Includes all non-living biomass with a diameter less than a minimum diameter, lying dead, in various states of decompositions above the mineral or organic soil
Soil	Soil organic matter (SOM)	Includes organic carbon in mineral and organic soils (including peat) to a specific depth

Among the five components, the AGB accounts the largest carbon stored in tropical region and it is directly affected by deforestation and degradation (Kumar and Sharma, 2015). The change in the forest areas and corresponding changes in forest biomass due to management and regrowth greatly influence the transfer of carbon between the terrestrial forest ecosystem and the atmosphere. Assessment of the AGB helps to quantify the carbon stocks, which in turn enables to analyse the current status and project the near future changes.

Accordingly, several methods are being developed for AGB estimation with different demands and level of accuracy. The most accurate and straightforward method to quantify the forest biomass is establishing sampling sites on the field and harvesting all the trees, drying them (in particular, for carbon stock estimation) and weighting the biomass. This method is very expensive and has a destructive effect on the experimental site. It is precisely for that one location only, and impractical in other regions for country-level analysis (Chen et al., 2018; Huang et al., 2018).

Such methods cannot be therefore applied across a landscape. Subsequently, many investments were put into the development of models that can extend the field measurement results over larger areas.

Table 2: *Main methods for forest AGB estimation*

Method	Benefits	Limitations	Uncertainty
Biome averages	<ul style="list-style-type: none"> • Accuracy could increase due to immediate availability • Data refinements could increase accuracy and consistent globally 	<ul style="list-style-type: none"> • Fairly generalized • Data sources not properly sampled to describe large areas 	High
Forest inventory	<ul style="list-style-type: none"> • Generic relationships readily available 	<ul style="list-style-type: none"> • Generic relationships widely understood • Relatively field-labor is costly 	Low

Optical remote sensors	<ul style="list-style-type: none"> • Satellite data routinely available at the global scale • Globally consistent 	Limited modeling ability for tropical VIs saturates at low carbon stocks	High
Radar remote sensors	<ul style="list-style-type: none"> • Satellite data are not always free • New systems launched are expected to provide improved data • accurate for young or sparse forest 	<ul style="list-style-type: none"> • Less accurate in mature forests and mountainous terrain because signal saturates also increases errors • Can be expensive 	Medium

Ethiopia in this context adopted a new forest definition in February 2015 stating 'A forest is land spanning at least 0.5 ha covered by trees and bamboo), attaining a height of at least 2m and a canopy cover of at least 20% or trees with the potential to reach these thresholds in situ in due course" (MEFCC, 2015).

There exist different forest definitions (Schwieder et al., 2018). The Ethiopian forest definition differs from the definition used for international reporting to the Global Forest Resources Assessment and from the forest definition used in the National Forest Inventory which both applied the FAO forest definition with the thresholds of 10% canopy cover, a 0.5 ha area and a 5 m height. The forest definition included Ethiopia's dense woodlands which have a wider distribution through the country (see Figure 1). Commercial agriculture is expanding mainly on dense woodlands and Ethiopia desires to allow the FRL to create REDD+ incentives for the conservation of these important areas (FRL, 2016).

Forest inventory is an activity of data collection that supports extraction of the required information base on the forest resource within an area of interest. Estimation of aboveground biomass and carbon stocks in the carbon pools requires conducting forest carbon inventory based on sampling. A sampling design defines the number and spatial distribution of sample elements drawn from the sample frame which consists of the set of all elements of interest to which the estimation relates. (UNFCCC, 2015).

A full inventory of forest trees is thus practically not possible. Consequently, sampling has to have resorted to biomass and to carbon stock at a landscape level. Estimates based on data collected from the measured samples are then extrapolated to the entire population, the majority of which has not been measured. The method of selection of

sampling units from the sampling frame should meet the twin requirements of randomness (equally being selected) and coverage (the sample should be spread over the entire sampling frame). A commonly used sampling design in forest inventory that aims to meet these requirements cost-effectively is stratified systematic sampling with a random start using fixed area sample plots (UNFCCC, 2015).

2.4 Forest Carbon Stock Estimation and Accounting Tools

There exist multiple opensource to proprietary carbon accounting tools suitable for the AFOLU sector designed for different purposes. Many of these carbon models used to estimate forest carbon for reporting purposes use the same data that is used by forest management agencies. This includes forest inventory, land use land cover, ownership data, change and disturbance information, growth and yield estimates, biodiversity data, etc. These models have varying levels and types of ecological processes represented and range from simple statistical relationships, such as emission factors used in systems where few data are available to complex combinations of empirical models (equations), process models (deterministic equations), and assumptions that enable these models to be executed promptly (Boisvenue & White, 2019).

Physical process-based ecological models such as the BEPS model, for instance, have been used to simulate key physiological processes including both carbon and water cycles at national or regional scales. The ‘best’ tool (model) depends on the use to which it is being put, the result needed and the accuracy required for that result. Ecosystem models such as CENTURY, DayCent, or DNDC use complex functions to describe the processes of an ecosystem. They attempt to represent the dynamic processes that occur within an ecosystem which result in changes in stocks and fluxes of carbon and nitrogen.

However, the currently existing models do not provide sufficiently accurate or consistent carbon estimates across spatiotemporal scales. More often, these models are applicable at a specific project level and limited to get the advantage of the emerging remote sensing technologies and are not able to easily adapt to new data types or inputs.

2.5 Potential Remote Sensing Contributions to Carbon Stock Modeling

Increasing the frequency and extent of observations and developing inferences about forest attributes will in itself contribute to advancing our understanding of processes and their cross-scale interactions. Besides, remote sensing satellite sensed information has the potential for the scaling of various processes across spatiotemporal scales. Information about forest composition, structure, productivity, and disturbances are useful for the more practical applications of carbon modeling, and are also useful for tracking changes in forests, and therefore for studying and understanding the forest system. With a proliferation of remote sensing satellites offering an increasing number and variety of observations, the potential for virtual constellations expands, further increasing the observation capacity for forests (Boisvenue & White, 2019).

The rapid evolution of remotely sensed data products and methods provides at least two further opportunities: the use of machine learning algorithms, and the advancement of the cross-scale tracking of processes (Belgiu & Drăguț, 2016). Remote sensing cannot measure AGB directly, therefore, algorithms for estimating AGB has been developed by linking ground truth data to the potential variables derived from remotely sensed images (Gasparri et al. 2010; Hansen et al. 2013).

Expressing the relationship between AGB and remotely sensed data using a simple linear or non-linear regression model is often difficult yet. Many researchers have therefore used non-parametric algorithms such as boosted trees, local kernel-weighted methods, k-nearest neighbors (k-NN), artificial neural networks (ANN), random forest (RF), and support vector machine (SVM)) to estimate AGB and improve AGB carbon stock estimation (Amani et al., 2019; Apley & Zhu, 2019; Chen et al., 2018; Khudinyan, 2019; Santi et al., 2017).

In the previous state of data scarcity, machine learning techniques were of little use in forest modeling. However, remote sensing provides an increasingly overwhelming observational capacity, permitting the use of these powerful tools, and with them, possibilities of increasing the available information and associated understanding of our forested landscapes. Thus, there is an urgent need for more detailed observations to

support the practical modeling of forest carbon. The selection of, the sensor type, quality, and accuracy, as well as spatial and temporal resolution for mapping and monitoring of the forested landscapes and the algorithm to be used in the extraction of useful information remain important (Romijn, et al., 2015).

Thus, optical remote sensing products such as Landsat imagery for indirect forest AGB estimation has long historical data and has widely been used in forest biomass and carbon stock monitoring. But the penetration possibility of an optical sensor remains weak as it records the information reflected from the forest canopy and is dependent on the leaf structure, pigmentation and moisture. The heterogeneity of forest stand structures may also be the major reason for data saturation in optical imageries (Zhao et al., 2016). This results in high uncertainty especially, in dense tropical forests with a big amount of biomass (Betru et al., 2019b; Chen et al., 2018).

Earth observation technologies are also unceasingly emerging and protonated the development of new satellite and airborne sensors. This derived the emergence of new analysis and methods (de Sousa et al., 2020) with an increased resolution for better forest information extraction. Evolving satellite technologies include the use of light detection and ranging (LIDAR), Unmanned Aerial Vehicles (UAV) and Synthetic Aperture Radar (SAR) observations for forest characterization and biomass estimation. These techniques can help to overcome the challenge of cloud cover in tropical areas. New techniques for acquiring, processing and managing immense satellite remote sensing data include cloud-based databases and data processing platforms that offer space for large datasets and computational resources for processing (Filipponi, 2019) (GOFC-GOLD, 2014).

Sentinel-1 satellite mission is developed by ESA under the Copernicus initiative with a constellation of two satellites: Sentinel-1A and Sentinel-1B. The satellites are launched in April 3rd, 2014 and April 25th, 2016 respectively (ESA, 2013). Like sentinel-2, Sentinel-1 satellites orbit in near-polar, sun-synchronized orbit at 693 km altitude and in the same orbital plane (Geudtner et al., 2014). It collects data in four modes: Stripmap mode (SM), Interferometric Wide Swath mode (IW), Extra Wide Swath mode (EW) and Wave mode (WV) (Niculescu, et al., 2018).

Table 3: Wavelengths most frequently used in SAR

Band Designation*	Wavelength (λ), cm	Frequency (ν), GHz
Ka (0.86 cm)	0.8 – 1.1	40.0 – 26.5
K	1.1 – 1.7	26.5 – 18.0
Ku	1.7 – 2.4	18.0 – 12.5
X (3.0 cm, 3.2 cm)	2.4 – 3.8	12.5 – 8.0
C (6.0)	3.8 – 7.5	8.0 – 4.0
S	7.5 – 15.0	4.0 – 2.0
L (23.5 cm, 25 cm)	15.0 – 30.0	2.0 – 1.0
P (68 cm)	30.0 – 100.0	1.0 – 0.3

Unlike optical waves, Sentinel-1 is a radar sensor that uses radio waves that can penetrate clouds and detect the surface structure. The parameters to consider sentinel radar imagery for a land cover mapping study at wavelength level are polarization and incidence angle. At surface level, structure and dielectric properties which may cause layover or foreshortening in the images and a notable increase in backscatter are key parameters. The penetration capacity of the designated bands is thus a primary factor in wavelength selection for forest monitoring.

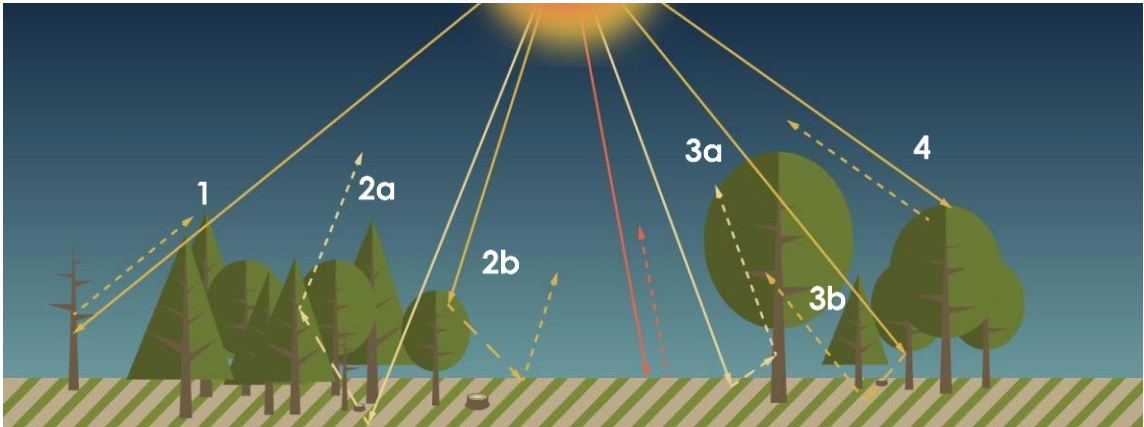


Figure 4: Radar Backscattering in Forests

The dominant backscattering sources in forests include (1) direct scattering from tree trunks, (2a) ground-crown scattering, (2b) crown-ground scattering, (3a) ground-trunk scattering, (3b) trunk-ground scattering, (4) crown volume scattering. Generally, the longer the wavelength, the greater the penetration into the target and the availability of the data also remains important.

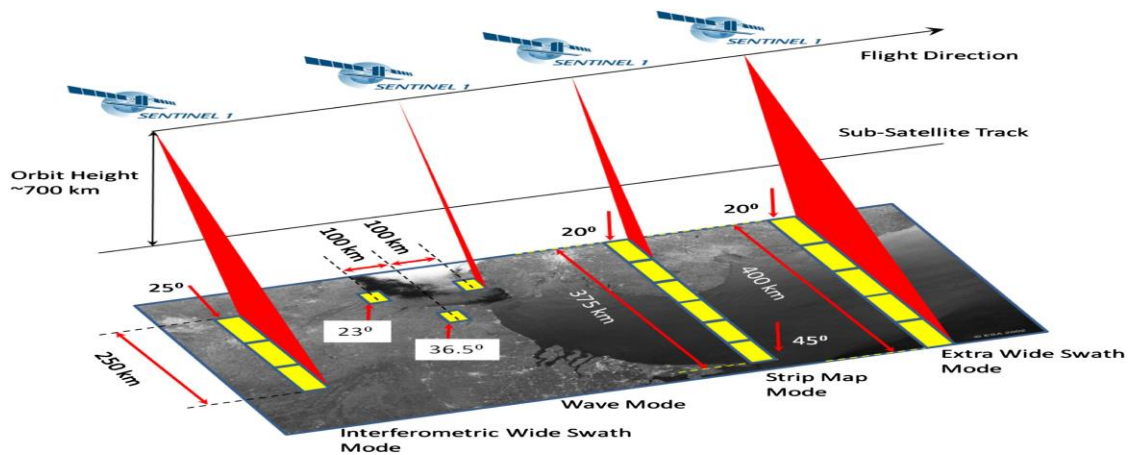


Figure 5: *Sentinel-1 data acquisition modes*

Interferometric Wide Swath (IW) allows taking measurements on a scan swath of 250 km with a resolution of 5m on 20m. This mode uses Terrain Observation with Progressive Scans SAR (TOPSAR) to create higher quality data. Polarimetric SAR sensors are capable to transmit and detect the vertical (V) and horizontal (H) components of the backscattered radiation. Hence, there are four possible polarization configurations viz: co-polarization- HH, VV, cross-polarization- HV, VH, and their combinations. Polarimetric SAR backscattered energy is directly dependent on the physical properties of the vegetation elements and influences on backscattering mechanism. The scale of the surface relative to the wavelength determine how rough or smooth they appear and how bright or dark they will appear on the image (Pohl C., 2017).

The advantage of using Sentinel-1 images is its short revisit time favorable for forest change detection applications. However, SAR with longer wavelengths (e.g., ALOS PALSAR products) is more suitable for forest applications since it is more penetrating than shorter wavelengths. To compensate for this limitation, more information needs to be derived aside from the provided VV and VH polarization of Sentinel-1 (Argamosa et al., 2018). Studies have shown that saturation remains a dominant issue when directly estimating AGC stock using SAR data in high biomass areas. A consideration of multi-temporal SAR data with multiple polarizations and the use of rule-based algorithms can help to mitigate the saturation problem and improve AGC stock estimations (Thapa et al., 2015).

2.6 Multisource Data Fusion for AGC Stock Monitoring

Finding an optimum combination of accuracy of measurements and the cost of the technology is often a major challenge in remote sensing projects for estimating forest biomass and carbon stock. Using SAR data, biomass estimation and change detection can be best achieved. Studies have shown that SAR data can detect half of the tree trunks removed by selective logging. Modeling biomass production or carbon sequestration, however, needs both optical and SAR data to be combined (Scott et al., 2009).

The synergistic use of SAR and optical remote sensing was applied in several studies for describing vegetation in diverse ecosystems, for example, forests, wetlands, agricultural areas, upland vegetation types and also to differentiate broad land cover classes (Braun & Hochschild, 2015; Christiansen et al., 2019; Lu et al., 2017; Mahdianpari et al., 2018; Martín, n.d.; Quang & Hoa, 2018; Tavares et al., 2019). Moreover, the synergy of Sentinel 1 SAR with optical data is important for it is much higher in Spatio-temporal resolution compared to other SAR data such as RADARSAT or ALOSPALSAR. The SAR side looking also makes it extremely sensitive to the relief, even under vegetation cover in tropical forests. However, the sensitivity of Sentinel-1 to forest timeseries has still not been assessed over larger undulating topography and in areas of difficult access and is urgent to use Sentinel-1 for estimating forest carbon monitoring.

A study by (Tsai, et al., 2018) indicated that a combination of terrain ancillary data, spectral vegetation indices, and simple illumination transformation could predict vegetation classification with accuracy above 70%. A more refined and accurate 30 m NDVI composite of the past 30 years was developed for the United States of America (USA) using the Google Earth Engine filling the missing data due to clouds by using a climate-driven modeling approach producing data at multiple scales (Robinson, et al., 2017)

CHAPTER THREE

METHODS AND MATERIALS

3.1 Description of the Study Area

3.1.1 Location

The research was conducted in Southern Nations, Nationalities and Peoples` Region (SNNPR) at Shiwshiwe forest of Alle Special Wereda. The study area is geographically located at the south-eastern escarpment of the Omo-Gibe watershed covering an area of 3349.22ha. It is 650km far southwest of Addis Ababa and 415kms south of regional town Hawassa. It extends from 5° 38' - 5° 50' N latitude to 37° 21' - 37° 40' E longitude.

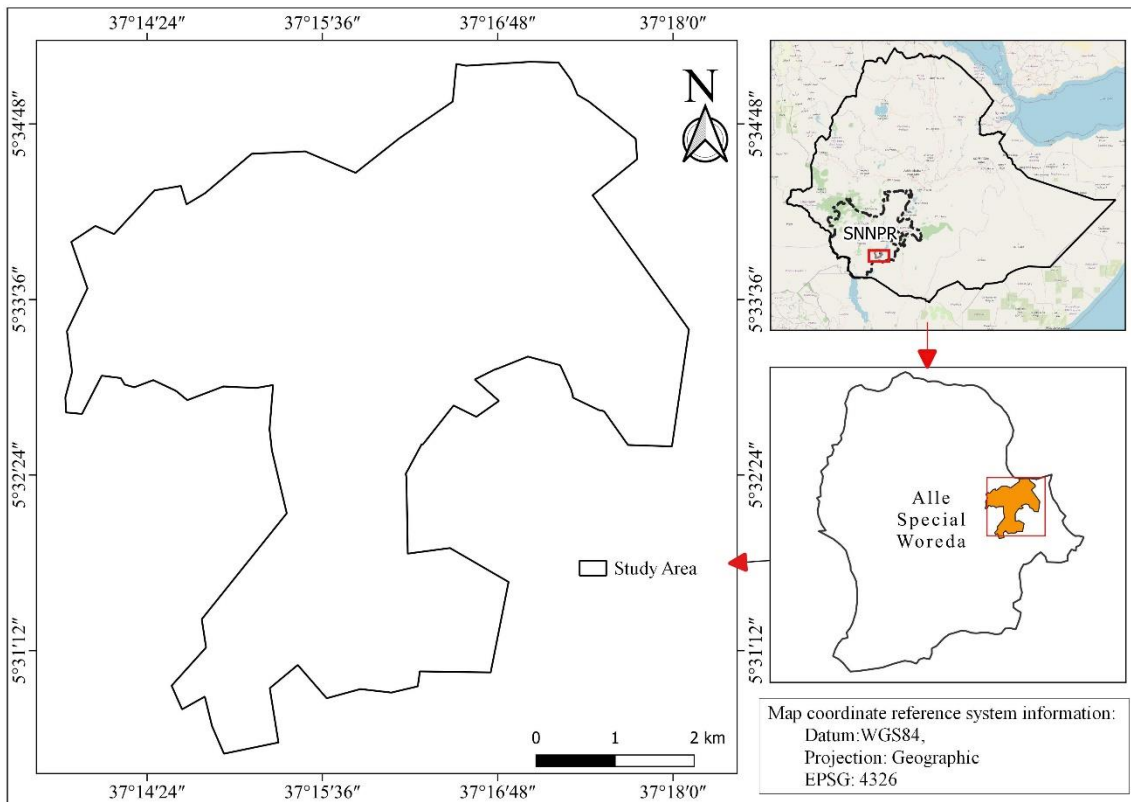


Figure 6: Location Map of Study Area (source: Researcher)

3.1.2 Climate Condition

The study area attains hot to warm humid lowlands mountains tepid to cool humid mid highland agro-ecological zones with altitudinal range of 1600m to 2600m. The rainfall

pattern of the study area is a bimodal type with minimum annual precipitation of 1200mm and a maximum of 2600mm with major peak in April and a small peak in October. The temperature of the study area ranges from 14°C to 28°C (FAO, 2010). The average temperature distribution over time calculated from GEE CHIRPS satellite image is depicted in Figure 7 below.

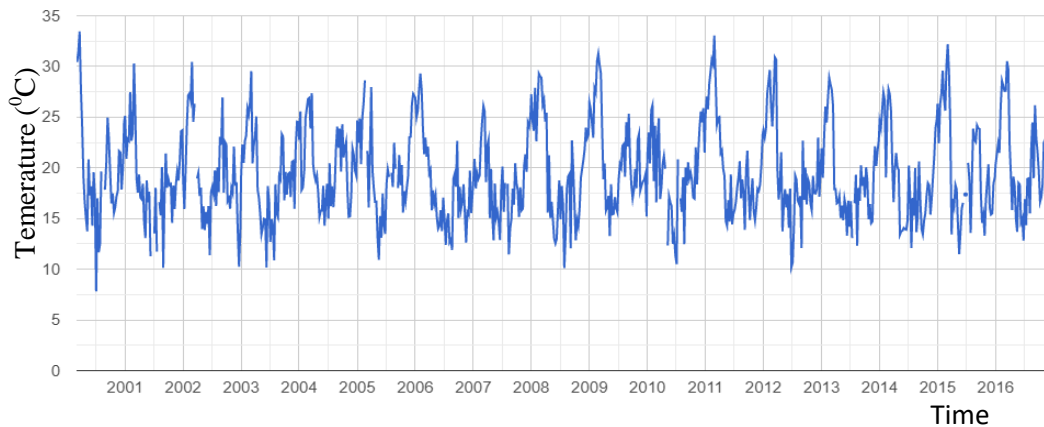


Figure 7: *Timeseries temperature distribution overtime in °C*

The local climatic conditions including growing seasons as computed from the local climate estimator of FAO from the FAO agroclimatic database with observations from nearly 30,000 stations worldwide have long-established. The seasons and climate of the study area is illustrated in Figure 8 below.

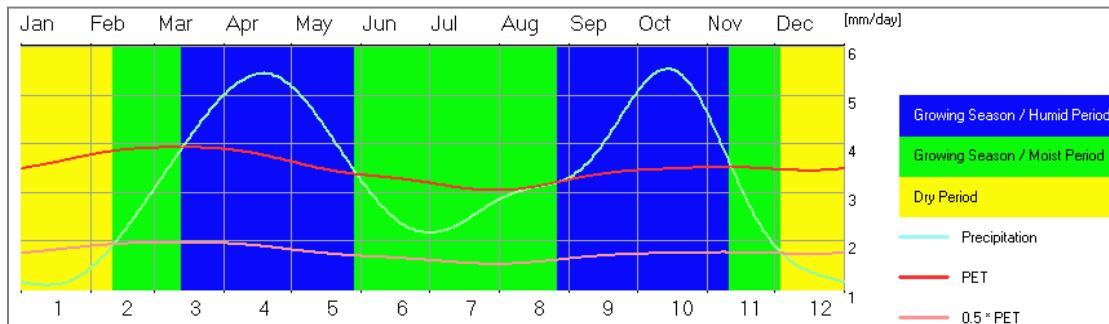


Figure 8: *Seasons and climate of the study area (source: Researcher)*

The dry season ranges from December to late-February with total length of 69 days. This season is important for the selection of optical satellite images due to lower precipitation exhibited. The details of the average vegetation periods seasons are presented in Table 4.

Table 4: Average climate and vegetation periods (source: Researcher)

Period	Nr.	Length	Begin	Begin	End	End	T_Mean	T_Min	T_Max	Precip.	PET	Wind Speed	Sun fract.	Vapor Pressure
		Days	Date	Day	Date	Day	°C	°C	°C	mm/day	mm/day	km/h	%	hPa
Dry	1	69	4 DEC	338	10 FEB	41	16.4	8.7	24.1	1.3	3.6	3.5	75.5	10.3
Moist+Hum.	1	296	11 FEB	42	3 DEC	337	16.6	10.5	22.8	3.7	3.5	4.4	54	12.1
Humid	1	77	13 MAR	72	28 MAY	149	17.5	11.4	23.5	4.8	3.7	5.5	55.1	13.5
Humid	2	76	27 AUG	239	10 NOV	314	16.7	10.8	22.7	4.5	3.4	3.8	53.6	12.2

3.1.3 The livelihood and Landcover of the Area

The northern and eastern part of the study area is characterized by a crop livelihood zone where sorghum, enset, maize & teff are dominant crops grown and mainly cattle & shoat livestock. The south and northeastern areas of the study area are dominated by agricultural livelihood zone with dominant maize, sorghum & teff growing (FAO, 2015). The area is also well known by its bamboo coverage which is main source of income and as a construction material. The Shiwshiwe area is covered with natural forest with dominant tree species of *Hagenia abyssinica*, *Arundinaria alpina* (Highland Bamboo), *Albizia schimperiana*, *Croton macrostachyus*, and *Trichilia dregeana*.

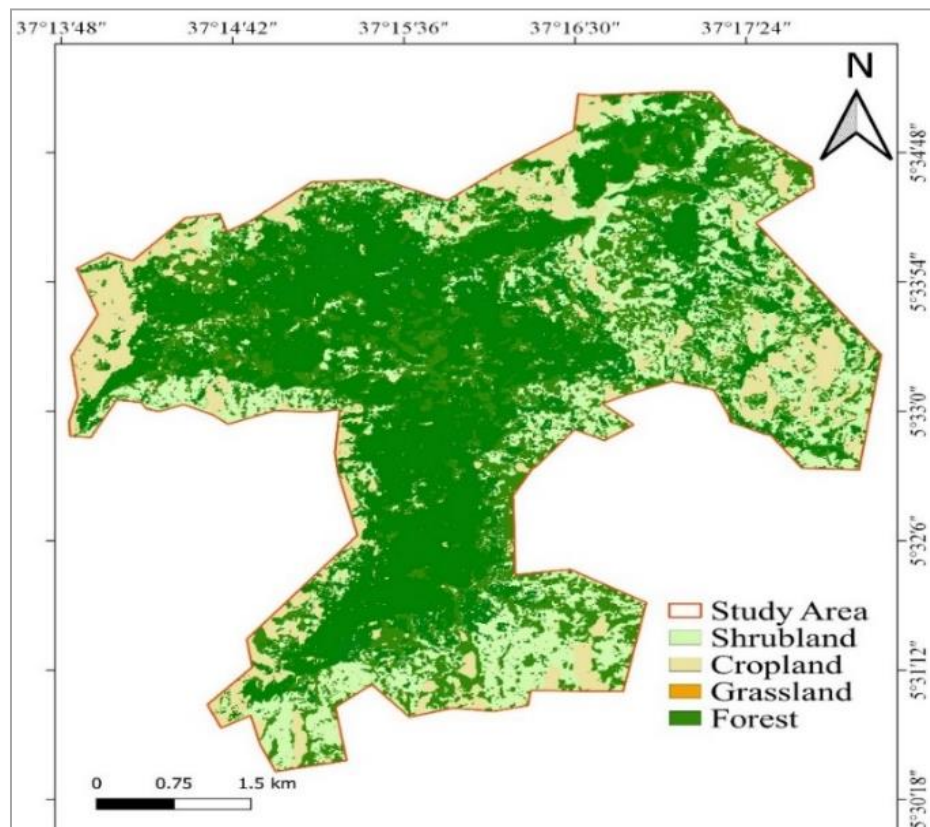


Figure 9: LULC Map of the study area (source: Researcher)

3.2 Materials

3.2.1 Data Type and Source

3.2.1.1 Optical Sentinel-2 MSI imagery Data

The optical Sentinel-2 Multispectral Instrument (MSI) imagery available in Google Earth Engine (GEE) for big data analysis were used for the extraction of vegetation indices and forest cover extent mapping. In this study, the sentinel-2 Level-1C 10m bands (B2, B3, B4 and B8) available in GEE were primarily used.

Table 5: *Sentinel-2(A and B sensors) Bands in GEE*

Name	Resolution	Wavelength (2A)	Wavelength (2B)	Description
B2	10 meters	496.6nm	492.1nm	Blue
B3	10 meters	560nm	559nm	Green
B4	10 meters	664.5nm	665nm	Red
B8	10 meters	835.1nm	833nm	NIR

3.2.1.2 Radar Sentinel-1 GRD imagery Data

The dual-polarization C-band Synthetic Aperture Radar (SAR) multi-looked and geocoded Level-1 Ground Range Detected (GRD) 10db products were used in this study. To gain information about the conditions and status of vegetation, the topographically corrected VV band and VH polarizations were retrieved from GEE with 10m spatial resolution and a temporal resolution of six days (see ANNEX 6 **Sentinel-1 GRD products spatial resolution**). The polarizations were used as two independent input variables for estimating forest biomass and carbon stocks.

3.2.1.3 Field Inventory Data

For the survey of attributes of the vegetation and of the site characteristics, a circular sample unit is often used according to FAO standards. The circular plots used in this study were laid at 10m radius using a meter tape to compromise with the resolution of the satellite images used (see Figure 11). The field data inventoried mainly comprises of sampling plot coordinate locations, photographs of the sample plots, tree species, DBH and H. Other ancillary information about the forest area such as the boundary of the forest reserve is obtained from the local office of the Environment Forest Climate Change Control Authority in Alle Special Woreda.

Plot data were collected in summer 2019 (September to December). Plot centers were located using handheld GPS Map 78 units augmented with Wide Area Augmentation Systems. All live trees with a DBH > 5cm, species, diameter, and height were measured at each plot. Smaller trees with DBH < 5cm were also tallied for total stem number, but not used in biomass calculations.

Table 6: *Summary of Data Type and Sources*

Data	Resolution	Purpose	Source
Sentinel-1 GRD	10m	Extract timeseries texture characteristics and backscatter information using grey level co-occurrence matrix and resampling the result	GEE
Sentinel-2 MSI	10m	Drive multispectral band, biophysical variable and vegetation indices, and fractional cover	GEE
DBH and H		For calculating field AGB and AGC	Field survey
ALOS DEM	25m	For topographic correction	GEE

3.2.2 Software and Tools

Open Source software can be used to undertake spatial analysis of datasets in the context of REDD+ which is released under a license that allows the software to be freely used, modified, and shared (Grinand et al., 2013). Hence for this study, an opensource and free software such as QGIS, Notepad++, Zotero and GEE applications were used.

Table 7: *Software and application tools*

Software	Purpose
GEE	For automation of satellite image processing using API.
Zotero	For reference citations management
Spreadsheet	For statistical analysis
QGIS	For extracting pixel value and producing maps

3.2.3 Field Materials

The materials used during field campaign for forest tree inventory in the study area are presented below in Table 8. For details see 0.

Table 8: *Field survey equipment*

Material	Purpose
Measuring tape, Caliper*	For forest tree measurement and outline the plot
Clinometer*	For tree height measurement
GPS Map 78*	To record tree and plot coordinates
Ipad Tablet Computer	To interactively record field data
Sleeping Bag, Axes, first aid kit	Camping and safety

* These field equipment were obtained from Alle Special Woreda Forest Environment and Climate Change Control office upon formal request.

3.3 Methods

3.3.1 Research Design

The modeling frameworks for the satellite based aboveground carbon stock estimation followed a quantitative study approach. With hypothecation of the problem statement, a field observation was carried. A systematic random sampling-based field inventory data, sentinel-1 SAR backscatters, grey level co-occurrence matrix (GLCM), sentinel-2 reflectance bands and sentinel-2 vegetation indices were exploited for the estimation and modeling of forest AGC.

GEE java scripts were primarily developed for processing the satellite images to drive the necessary parameters for AGC predictive model development on opensource platforms. Different pixel-based methods were used to answer the research questions and achieve the objectives. The general framework is as in Figure 11 below. The thesis report writeup followed the guideline for report writing amended in the Jimma University's Guidelines and Procedures for Research booklet, February 2013.

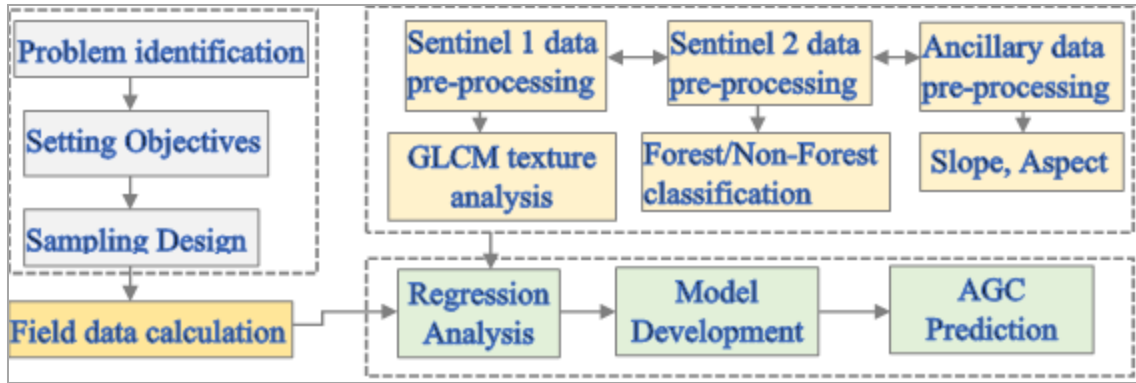


Figure 10: *Research Design Framework*

3.3.2 Study variables

Two types of variables are used in this paper. Those are dependent variables and independent variables.

A. Dependent Variable

The dependent variable of this study is a response variable that is output of the character that has been changed because of variations in the independent variables. This means that if the value of independent variable changes due to any factor, its effect may consequently observe on the output. The dependent variable of this study is aboveground carbon stock (AGC).

B. Independent Variable

The independent variables are factors or parameters those affect output of dependent variables. In this study the common factors affecting the performance of AGC estimation are:

- Tree parameters (Height, DBH and Wood-specific-density)
- Sentinel-1 backscatters (VV and VH polarizations)
- Sentinel-1 GLCM textures (Entropy, Angular Second Moment, Contrast, Correlation, Difference entropy, Difference variance, Difference entropy, Energy, Homogeneity, Variance, Sum average, Sum variance, Sum entropy, and Maximum correlation)
- Sentinel-2 Bands (B2, B3, B4, and B8)
- Sentinel-2 Indices (NDVI, NDWI, EVI, and SAVI)

3.3.3 Sampling Procedures and Sample Size

The sampling design protocol includes the specification of the sample size, sample locations, and the reference assessment units (i.e., pixels or image blocks). For the study, a systematic random sampling design was exploited. Hence, systematic sampling with a random starting point is generally more efficient than simple random sampling and is also more traceable. Sampling variability can be quantified with standard unbiased estimators in the form of statistical formulas (GOFC-GOLD, 2016).

Field forest inventory data were gathered from systematically designed Collect-Earth Online (CEO) plots as ground truth data. Collect Earth Online is an opensource crowdsourcing product for earth science analysis within the Open Foris software suite. It is a user-friendly, online Java-based tool that draws upon a selection of other software such as Google Earth and Saiku to facilitate data collection and analysis. The Collect Earth also allows the interpretation of high and medium spatial resolution imagery in Google Earth, Bing Maps and Google Earth Engine.

3.3.3.1 Plot Size selection

A circular plot of $10\text{m} \times 10\text{m}$ (see figure 10) was designed to collect sample data from the field. The circular-shaped plot type is chosen because it is easy to establish in the field. The sample size was selected for the study to increase the probability of selecting a median of a pixel in the imagery complementary to the spatial resolution of the images used.

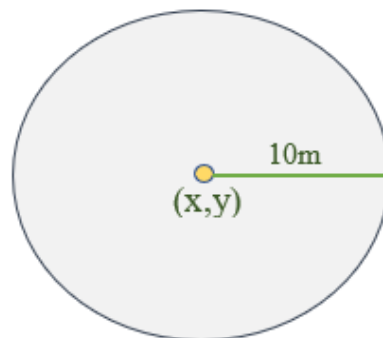


Figure 11: *Circular plot of 10m radius used in the field.*

3.3.3.2 Tree Sampling and inventory

A field campaign was carried out in Shiwshiwe forest to obtain a robust set of calibration and validation measurements at a convenient time. The data were collected using a systematic randomly taken at-start stratified plot. The plot coordinate location was taken using handheld GPS Map 78 by considering the different zone of the forest viz; transitional, buffer and core zones and the center coordinates of each plot was recorded.

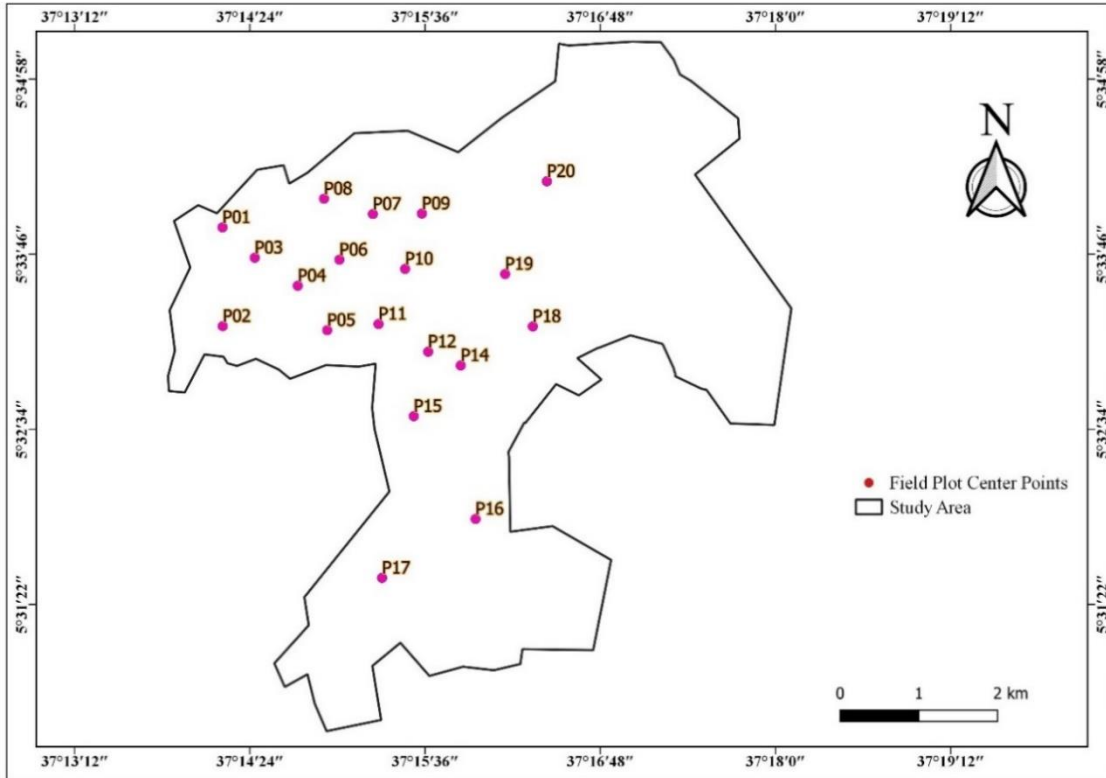


Figure 12: Field inventory plots

Trees were inventoried from circular plots with 10m radius () applying a systematic random sampling method for the estimation of above-ground biomass at the plot level along the topographical gradients.



Figure 13: Tree caliper held at right angles to the tree trunk for DBH measure.

The minimum tree diameter often used in dbh is 5 cm but it can vary depending on the expected

size of trees and agroecology. For this study, the minimum dbh is chosen as tree growth is faster in tropical areas and a minimum height of 5m based on FAO forest definition. All tree diameters in the plots were recorded at 1.3m aboveground level as distance at breast height (DBH) and at 30cm aboveground level as stump height (H) for bamboo species (Mac Dicken, 1997) except where trunk irregularities at that height occur (plank woods, tapping or other wounds) and necessitate measurement at a greater height. Furthermore, if tree branch is below the measurement height, an equivalent diameter is defined as $\text{SQRT}(D^2)$ on the basis of all D values.

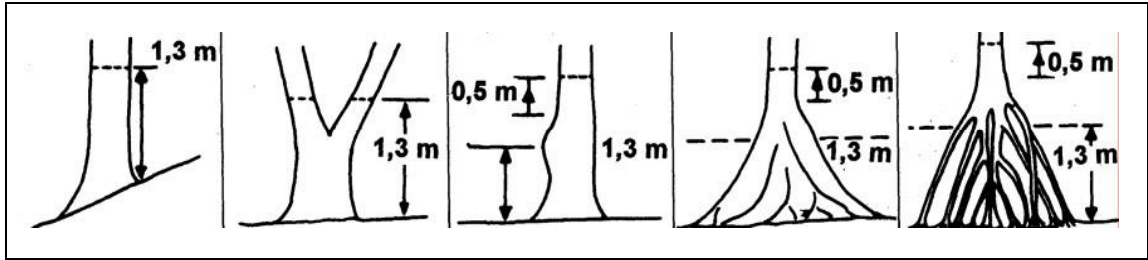


Figure 14: Measurement of tree diameter at breast height (DBH)

Moreover, the tree heights were measured using Clinometer. The Clinometer determines the angle to the tip of the tree based on 15m to 20m fixed distance to the target tree. The tree height was then recorded according to the clinometer scale. To measure the tree height with the clinometer in the field, the tip of the target tree was sighted and read the scale. Next, the bottom of the tree was then sighted and read the scale to measure the angle of elevation. The tree height is then obtained by adding the two measurements. The DBH and tree height (H) were used to compute field-based forest AGB using allometric equations.

In the plot, local names of trees species were recorded and later scientific names of each tree species were identified from a source as published by *Bekele T., (2007)*. These forest stand parameters were measured in the field through systematic random sampling. The plots represent a statistical subsample of the measured trees within each forest stratum, which can be scaled to represent population-level estimates using standard allometric techniques.

3.3.4 Google Earth Engine (GEE) Data Processing

The increasing availability of free time series data for large areas over the previous years has led to a shift from traditional bi-temporal change detection approaches, in which many changes are missing and the timing of changes is disregarded, to time series based methods (Reiche, 2015). This trend is expected to continue due to the availability of more open-access RS datasets and daily advancement in sensor, image processing, and computer vision technologies. Cloud computing platforms are efficient ways of storing, accessing, and analyzing datasets on very powerful servers, which virtualize supercomputers for the user. These systems provide infrastructure, platform, storage services, and software packages in a variety of ways for the customers (Amani et al., 2020).

Google Earth Engine (GEE) is thus a cloud-based automatic parallel processing and fast computational platform for processing of remote sensing data facilitating scientists to develop their own algorithms with less effort than before. It brings together the world's petabytes satellite imagery of scientific measurements. Earth Engine Code Editor is one of the web-based integrated development environments of GEE's tools which facilitates developing complex geospatial workflows. The Sentinel-1 SAR data and Sentinel-2 MSI image collections used in this research were retrieved and preprocessed in the Google Earth Engine cloud computing Code Editor platform. This platform efficiently runs computations using parallel computing, meaning that the workloads are distributed between many CPUs in the Google's data centers. The most cloud-free image collections were composited at annual and per seasonal levels to minimize cloud pixels exploiting the power of GEE. To remove the impacts of cloud cover on sentinel-2 spectral signatures, a cloud mask based on the 'QA60' band in Sentinel-2, which uses a bitmask to mark the opaque clouds (B10) and cirrus clouds (B11) (ESA, 2015) was used. Then both the clouds were masked out by selecting only the clear pixels that have a 'QA60' band equals to 0. The preprocessed data were then converted into top of atmospheric reflectance (TOA) and were mosaiced as ImageCollections. The sentinel 2 TOA images were further reduced using the median filter function.

3.3.5 Spectral Indices and Transformations

Landuse and landcover and Vegetation index algorithms such as Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Normalized Difference Water Index (NDWI) and Soil Adjusted Vegetation Index (SAVI) were extracted from sentinel-2 images for the years 2016 to 2020 in GEE using java API. Using a user-defined baseline period, they could monitor the degradation or improvement of vegetation and the impact of mitigation efforts.

Table 9: *Vegetation index variables used in GEE*

Name	Equation	Description
NDVI	$(B8-B4) / (B8+B4)$	Rouse Jr. et al., 1974
NDWI	$(B8 - B3) / (B8 + B3)$	Gao et.al. 1996
EVI	$2.5 * (B8 - B4) / (B8 + 6 * B4 - 7.5 * B2 + 1)$	Huete et al., 2002
SAVI	$((b8 - b4) / ((b8 + b4 + 0.5)) * 0.5)$	Huete, 1988

3.3.6 Variable Selection for Forest AGC Prediction

The actual path of AGB carbon estimation and analysis consists of many decision points to minimize uncertainty in the output result. The most important decision is the proper prediction model selection. Forest biomass prediction as in Chen et al (2018) is defined as a mapping process for estimation of biomass values at the location without observation based on the points values at nearby observations and/or considering other factors at the site using various methods categorized mainly as parametric (various statistical regression methods) and non-parametric (machine learning algorithms). In parametric models, the expression relating to dependent and independent variables is easy to interpret and calculate. Two of the most often used methods are stepwise regression and multivariate linear regression. To handle high biomass variability and saturations of remote sensing imagery, Berninger et al. (2018) assume multivariate linear regression to be superior to other models. Different literature reviewed however declares that the parametric regression prediction models are preferable and commonly used models compared with the non-parametric ones (Khudinyan, 2019).

In this study, a parametric regression was chosen and multi-year time series of SAR acquisitions were used to understand the backscattering behavior of different tree species.

Multiple SAR images of the same polarization grouped into a time series can reveal temporal signatures that monitor the phenology of the species as in (Udali, 2019). The field AGB and information gained from both sentinel-1 backscatters, GLCM textures, and sentinel-2 VIs were correlated statistically using regression analysis. The variable with a high coefficient of determination (r^2) and low root mean square error (RMS) became the final predictor of AGC. A prediction model was then developed by considering the coefficient of selected variables and constant values from a multilinear regression analysis.

3.3.7 Prototype Model Development

The most important task in satellite image processing and pattern recognition is the classification of image textures. Over the years, extensive researches have been made for the classification of texture images. Image textural features, primarily developed by Haralick et al. (1973), measure the spatial homogeneity of the backscattering and contain information such as forest structure. To determine the more relevant SAR variables to this research, different modules of image analysis were developed. The modules were coded in GEE environment using Java API. These involve separate scripts for optical image preprocessing, VIs extraction, Sentinel-1 backscatters preprocessing, DSM topographic data extractor, and multisensor image fusion modules.

Various approaches for fusing satellite-based time series data can be applied. However, attempts were made to combine time series at the signal level, normalizing optical and SAR signals to simulate similar trends in a fused time series. This attempt was carried to fill Sentinel-2 gaps with simulated VH and VV SAR backscatter assuming similar behaviors). The designated bands of the fusion image composite are B2, B3, B4, B8, NDVI, EVI, SAVI, MNDWI, VH, VV, elevation and slope band.

3.4 Data Processing and Analysis

3.4.1 Tree level Above Ground Biomass Estimation

The first objective of this research was to geocompute forest biomass components using the regression models estimated from the sample plots to improve our understanding of

the spatial distribution of each component for the whole site. To do so, tree species-specific regression models were used to predict the above-ground biomass (AGB).

For model training and statistical analysis, information was collected from 20 circular plots (Figure 10) in the Shiwshiwe forest. Forest aboveground biomass (AGB) is usually determined through a combination of forest inventories with allometric tree biomass regression equation models. This is usually based on plot inventories along the following three steps. 1) selection and application of an allometric biomass function for the estimation of individual tree biomass, 2) sum-up individual tree AGB to estimate plot AGB, and 3) computation of an across-plot average to full hectare-based conversion.

The plot-level forest variables were computed based on the individual live tree data, including basal area (m^2/ha). The DBH and height (H) as tree parameter measurements were loaded into excel database for the derivation of individual tree-based biomass components which was statistically linked with satellite data-derived proxies. To calculate plot-level biomass on a per hectare basis, all the recorded DBH of each tree in the plot were tallied by DBH to get biomass in kg per hectare. Species allometric equations for tropical forests developed from tree inventory data by (Chave et al., 2014) were selected for this research.

The allometric equation is selected as it best-fits the bioclimatic conditions of the area; incorporating all the tree species in Shiwshiwe forest. It has the highest R^2 compared to similar allometric equations and availability of wood-specific gravity, trunk diameter (DBH), and total tree height.

$$Tb = 0.0673 * (\rho * (DBH)^2 * h)^{0.976} \text{ ----- Equation 1}$$

Where Tb = the total biomass (kg), DBH = diameter at breast height (cm), ρ = wood density (g/cm^3) and h = height (m)

Based on the measurements of the diameter and height of tree stands on each sample plot, calculations of above ground biomass were carried out in Microsoft Excel using Equation 1. The DBH and height measurements were converted into the same unit by converting tree height in meters to centimeters. The values were then converted into tree-level

biomass in kilograms (kg) by multiplying by 1000 for all the 20 plots as illustrated for P01 in Table 10: Tree level AGB calculations.

The AGB for each tree in the plots was then converted into full hectare (kg/ha). This was done by asking how many plots of 10mX10m exist in a full hectare. This was carried by multiplying the plot level AGB in kg by 10000 and then divided by the area of the 20 plots.

Table 10: Tree level AGB calculations

Plot 01											
Tcode	Family	Genus	Species	dbh (cm)	POM	Height (m)	Height (cm)	Wood D. (g/cm3)	AGB (g)	AGB(kg)	AGB (kg/ha)
IT1	Rosaceae	Hagenia	abyssinica	36	1.3	10	1000	0.5905	37207.3012	37.2073012	186.036506
IT2	Poaceae	Arundinaria	alpina	7	1.3	5	500	0.63	824.1334446	0.824133445	4.12066722
IT3	Poaceae	Arundinaria	alpina	7	1.3	5.5	550	0.63	904.4754837	0.904475484	4.52237742
IT4	Euphorbiaceae	Croton	macrostachyus	54	1.3	17	1700	0.518	121269.5249	121.2695249	606.347625
IT5	Poaceae	Arundinaria	alpina	6.5	1.3	8.5	850	0.63	1196.991824	1.196991824	5.98495912
IT6	Poaceae	Arundinaria	alpina	6	1.3	7	700	0.63	847.1071907	0.847107191	4.23553595
IT7	Poaceae	Arundinaria	alpina	6	1.3	8	800	0.63	965.024876	0.965024876	4.82512438
IT8	Poaceae	Arundinaria	alpina	5	1.3	5	500	0.63	427.322355	0.427322355	2.13661178
IT9	Poaceae	Arundinaria	alpina	5	1.3	7	700	0.63	593.4396721	0.593439672	2.96719836
IT10	Poaceae	Arundinaria	alpina	9	1.3	8	800	0.63	2129.455879	2.129455879	10.6472794
IT11	Poaceae	Arundinaria	alpina	8	1.3	9	900	0.63	1898.207934	1.898207934	9.49103967
IT12	Poaceae	Arundinaria	alpina	7	1.3	8	800	0.63	1303.823012	1.303823012	6.51911506
IT13	Poaceae	Arundinaria	alpina	7.5	1.3	9	900	0.63	1673.523117	1.673523117	8.36761559
IT14	Fabaceae	Albizia	grandibracteata	20	1.3	14	1400	0.534	14870.54363	14.87054363	74.3527181
IT15	Rubiaceae	Rothmannia	urcelliformis	12.1	1.3	9	900	0.642	4336.176822	4.336176822	21.6808841
IT16	Boraginaceae	Cordia	africana	27	1.6	13	1300	0.482	22485.35523	22.48535523	112.426776
P01AGB=									212932.4066	212.9324066	1064.662

The plot level above ground biomass of the entire stands per hectare was then calculated using the following equation:

$$AGBt/ha = \frac{\sum_{i=1}^n (TBpi)}{A} * 10000 \text{ ----- Equation 2}$$

Where AGBt/ha: Total Biomass (t/ha)

TBpi: Tree Biomass (t)

A: Plot size (m²)

n: number of trees

Tree-based biomass estimates were summed by plot to determine total plot-level biomass components.

3.4.2 Plot Based Carbon Stock Estimation

The plot level full hectare (kg/ha) biomass summaries were converted into above ground carbon stock using default IPCC conversion factor of 0.47 as in Equation 3.

$$AGC_{(kg/ha)} = AGB_{kg/ha} * 0.47 \text{ ----- Equation 3}$$

Where $AGC_{(kg/ha)}$ = Aboveground carbon in "kg/ha"

AGB = Aboveground biomass in "kg/ha"

3.4.3 Satellite Based AGC Mapping prototype Development Scenarios

The biomass estimation based on optical satellite data is revealed by the correlation between multispectral images' spectral responses such as EVI and biomass. The following steps were followed to develop the satellite-based AGB and AGC mapping prototype tool in GEE. The prototype is designed at modular level each working on a specific task and named accordingly. The modules are described in the sub-topics below which when integrated yields the prototype for aboveground carbon stock mapping.

3.4.3.1 Optical Sentinel-2 data Preprocessing Prototype

Sentinel-2 constellation is a European Space Agency's (ESA) Copernicus Program earth observation mission which systematically acquires optical imagery over the earth's surface at a high spatial resolution (10m to 60m). The other important objective of the Sentinel-2 mission is its complementation with other global satellite programs such as the Landsat, SPOT, and other satellite programs for ensuring continuity in monitoring the dynamics on Earth's surface. Sentinel-2 data have been used for different applications such as agricultural, forest monitoring, urban development, wetland monitoring.

Further important examples of the Sentinel-2 application include the development of CCI-S2 prototype Land Cover map of Africa at 20m resolution, the Copernicus Land cover services high spatial resolution maps, and the new pan-European high spatial resolution land cover/use maps (Phiri et al., 2020). For this study, the prototype development stages in the GEE code editor range from instantiating JavaScript to call for required Sentinel-2 bands from petabytes of images to final export of the final parameters for carbon stock estimation. The optical sentinel-2 data preprocessing prototype involves the following submodules.

A. Sentinel-2 Image Acquisition Query Module

The atmospherically corrected Sentinel-2 surface reflectance data were used in this study. The images were retrieved by filtering the collection to images overlapping the study area and those within the study time range. This filtering involves defining sensor type, the desired time frame and the AOI for which cloud-free images were to be created by *ee.ImageCollection.filterDate()* function.

B. ImageCollection, Cloud removal and Median pixel value Retrieval Module

To use the image collection in the analysis, cloud masking was applied to remove cloudy pixels from the analysis. It is fairly straightforward where functions are mapped (looped) to remove cloud and add NDVI bands and then filter it down to a date range and location. For this purpose, Sentinel-2 QA60 quality assurance band was used to conduct cloud masking. This study further presents an automatic workflow to aggregate Sentinel-2 cloud-free images, by removing cloud areas for user defined areas of interest and periods as a multitemporal image aggregation approach.

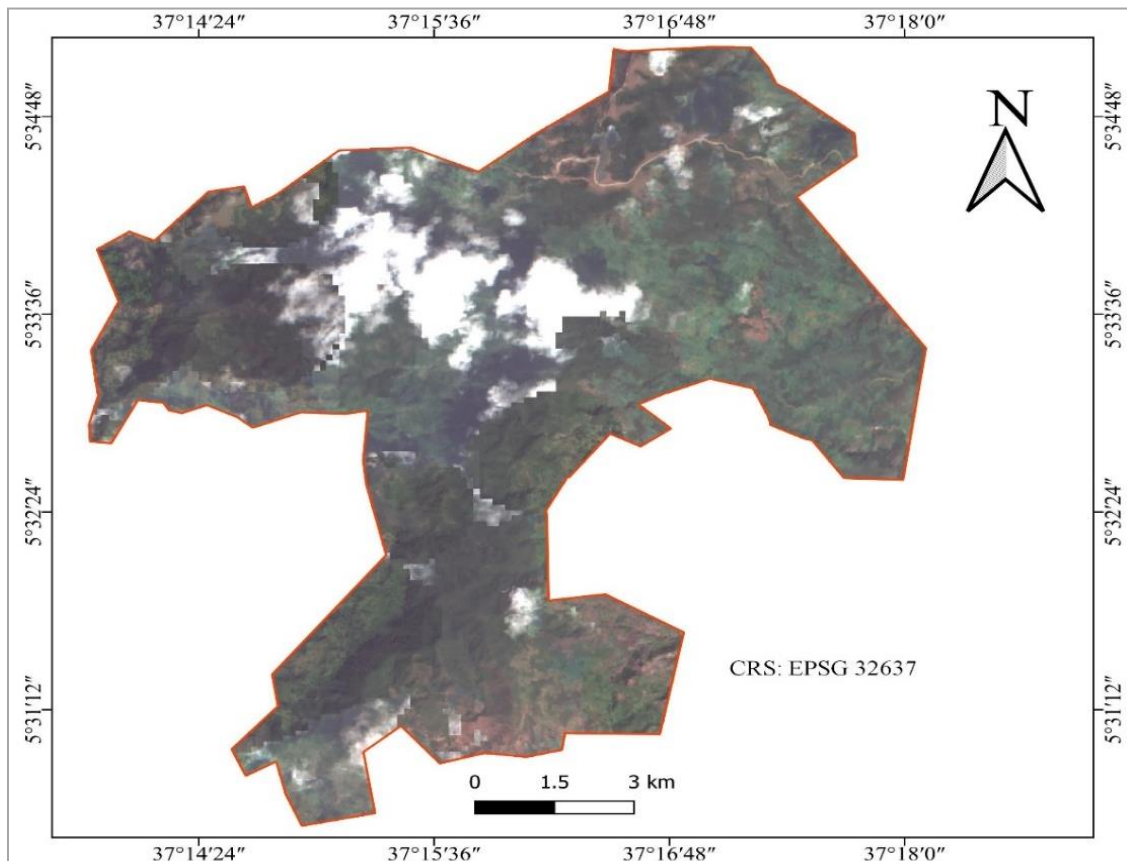


Figure 15: Example 2016 cloudy Sentinel-2 median raw image

For this purpose, the 'QA60' band from the sentinel image collections was selected by using the 'maskS2clouds' JavaScript function over the image collections of the driest months of each year.

```
function maskS2clouds(image) {var qa = image.select('QA60');
```

Then 'cirrusBitMask' function was used to detect and mask clouds and cloud shadows using Bits 10 and 11 as they are clouds and cirrus, respectively. Both the flags were set to zero, indicating clear conditions.

```
1 //call DefineStudyArea code
2 var studyArea = require('users/oytisa/FinalTheisisMSc:DefineStudyArea');
3 var aoi = ee.FeatureCollection("users/oytisa/shiwshiwe_bound");
4 // Expression to create cloud-free Sentinel.
5 function maskS2clouds(image) {
6   var qa = image.select('QA60');
7   // Bits 10 and 11 are clouds and cirrus, respectively.
8   var cloudBitMask = 1 << 10;
9   var cirrusBitMask = 1 << 11;
10  // Both flags should be set to zero, indicating clear conditions.
11  var mask = qa.bitwiseAnd(cloudBitMask).eq(0)
12    .and(qa.bitwiseAnd(cirrusBitMask).eq(0));
13  return image.updateMask(mask);
14 }
```

The masked and scaled data were then returned without the QA bands in the Google Earth Engine code editor environment using the 'return' function as in the script below.

```
return image.updateMask(mask).select("B.*")
    .copyProperties(image,
    ["system:time_start"]);
```

In the next step, the cloud-masked Sentinel-2 data were loaded and converted into TOA reflectance data and reduce the collection into one image. The timeseries daily 10m bands available each year over the DOY's were then averaged using the 'median' function to create a median ImageCollection band. Then the median imageCollections were reduced to multispectral reduced images that were later be used for random-forest classification and further analysis. With the loaded ImageCollections of Sentinel-2 imagery, a true color composite was set using bands B4, B3 and B2 that is adequate and pertinent for the training data sampling.

```

var image = collection.median();
var image = image.clip(boundary).uint16();

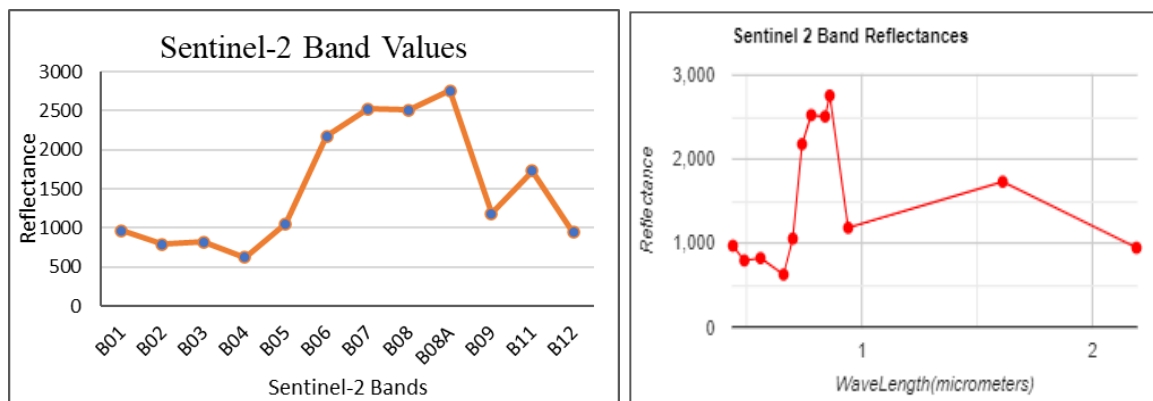
```

C. Spectral Band Selection Module

The sentinel-2 Red, Green and Blue bands were selected and examined to make the images discernible. Yet, it is hard to differentiate grass from vegetation, and once the RGB bands are replaced with Red, Near-Infrared, and Shortwave-Infrared bands in the ImageCollections, features previously ambiguous or indiscernible became clear and prominent to extract the information in need for the model development.

The distinguishing of a given land cover from other LULC categories like water, vegetation or bareland areas depends on the spectral reflectance pattern. This is basically due to their absorption and reflection of different kinds of light. These different combinations of reflectance values can be used to create spectral ‘signatures’ for different land-cover types.

The bands were selected using artificial neural network (ANN) and GEE hence this is proved to provide the most suitable bands for further image classification. The band selection was examined for each date while asynchronously updating the image information.



To get results faster, the scaling of the bands' selection code was improved which allows improving the availability of computing resources for all users was improved. This was carried using the ‘reduceRegion()’ and ‘scale’ function. This was able to avoid scaling error over processes such as image stretching and vegetation indices enhancing the image computation time. Earth Engine has limits in place to prevent too many such aggregations

from being run concurrently. Therefore, the "Too many concurrent aggregations" error was triggered by a reduction within a map function while creating the timeseries chart.

3.4.3.2 Vegetation Indices Extraction as proxies

A vegetation index derived from satellite image can be an indicator to describe vegetation greenness, density and health. The selection of vegetation indices to perform the AGC spatial modeling was based on the performance in previous studies and the same procedure was followed as for sentinel-2 band selection using ANN. Three vegetation Indices were calculated from summer Sentinel-2 surface reflectance optical imagery as independent variables in this study. The timeseries vegetation indices generated through here were used as input data in later data fusion prototype. The vegetation indices extraction module involves the scripts to derive NDVI, EVI, SAVI and NDWI.

Normalized Difference Vegetation Index (NDVI)

The is an index of plant “greenness” or photosynthetic activity. It is the most commonly used vegetation ratio index which was calculated using the NIR (B8) and R(B4) of sentinel-2 images. But it is known to be insensitive to small changes in areas with dense vegetation with higher biomass as the vegetation index values approximate to 1. This saturates the above-ground biomass and carbon stock estimation hence; the value of persistent cloud shadows also gains 1 because the red band reflectance values are close to zero.

$$NDVI = \frac{NIR-R}{NIR+R} \text{----- Equation 4}$$

Enhanced Vegetation Index (EVI)

EVI is designed to enhance the saturation problem of NDVI to high biomass vegetation signal with improved sensitivity. It separates the influences of atmospheric scattering and canopy background signal using suitable correction factors. The preprocessed Sentinel-2 data of GEE was employed to derive EVI map using Equation 5 below.

$$EVI = 2.5 * \frac{NIR - R}{(NIR + 6 * R - 7.5 * Blue + 1)} \text{----- Equation 5}$$

where NIR is B8, Red is B5, Blue is B2,

2.5 is gain factor,

1 is the canopy background adjustment, and

6 and 7.5 are coefficient factors to correct the influence of aerosol content.

Normalized Difference Water Index (NDWI)

In this study, the original NDWI index proposed by Mahdianpari *et.al* (2018) is used to obtain accurate results at the 10m target resolution. Moreover, NDWI is calculated using the green and NIR bands of Sentinel-2 data both of which are at a 10 m spatial resolution.

$$NDWI = \frac{NIR-G}{NIR+G} \text{----- Equation 6}$$

Where NIR = Near Infrared (Band8)

G = Green (Band3)

Soil Adjusted Vegetation Index (SAVI)

The SAVI vegetation index is closely related to tree standing biomass and is calculated using the NIR and Red bands of sentinel-2 as in Equation 7 below.

$$SAVI = \frac{NIR-R}{NIR+R+L} * (1 + L) \text{----- Equation 7}$$

Where L = 0.5 as canopy adjustment factor to reduce soil brightness,

NIR = Band8 and

R = Band4

3.4.3.3 Sentinel-1 RADAR data preprocessing

The open availability of some of the previously restricted SAR datasets such as ALOS, PALSAR and Sentinel-1 is increasing. Sentinel-1 radar backscatter patterns are used across many domains of humanitarian aid, crisis response, agriculture, maritime activity and forest monitoring (Filipponi, 2019).

The Google Earth Engine Sentinel-1 C-band image collections of Level-1 Ground Range Detected (GRD) scenes were filtered using the ‘.filter’ function for the study area. The filtering was carefully made based on the naming convention of Sentinel-1 products with extension of “.SAFE”.

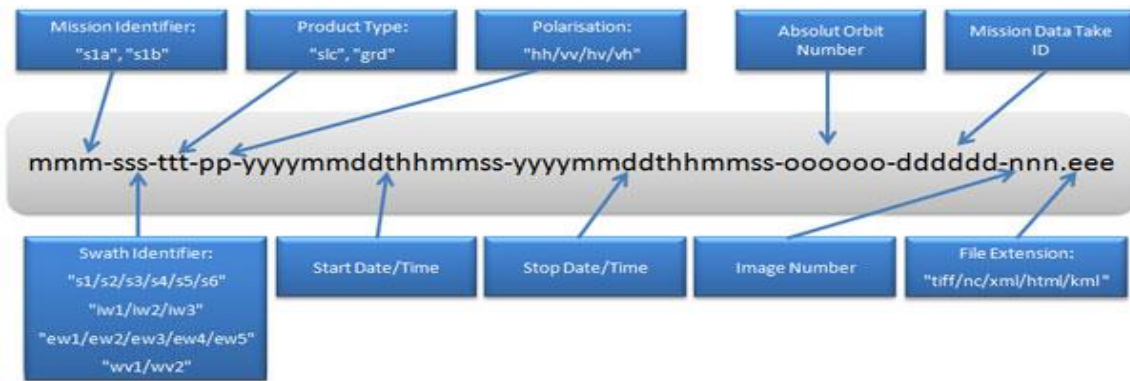


Figure 16: Sentinel-1 products naming convention (Source: ESA, 2020)

To create a homogeneous subset of Sentinel-1 data, the following metadata property fields were used. These include the *transmitterReceiverPolarisation*, *instrumentMode*, *orbitProperties_pass*, *resolution_meters*, *relativeOrbitNumber_start* and *resolution* (as 'M' for medium) or 'H' for high) functions.

```

5 // selectig Sentinel-1 Backscatter polarizations based on sensor peroperties
6 var collection = ee.ImageCollection('COPERNICUS/S1_GRD')
7     // .filter(ee.Filter.listContains('transmitterReceiverPolarisation', 'W'))
8     // .filter(ee.Filter.listContains('transmitterReceiverPolarisation', 'VH'))
9     .filter(ee.Filter.eq('instrumentMode', 'IW'))
10    .filter(ee.Filter.eq('resolution', 'H'))
11    .filter(ee.Filter.eq('resolution_meters', 10))
12    // .filter(ee.Filter.eq('orbitProperties_pass', 'ASCENDING'))
13    .filter(ee.Filter.eq('relativeOrbitNumber_start', 152))
14    .filter(ee.Filter.date('2020-01-01', '2020-12-10'))
15    .sort('system:time_start')
16    .filterBounds(geometry)
17    .map(function(image) {
18        var edge = image.lt(-30.0);
19        var maskedImage = image.mask().and(edge.not());
20        return image.updateMask(maskedImage);
21    });
22 // Set map center and scale
23 Map.centerObject(table, 11);

```

The above code snippet filters all the possible bands as per the required parameters overlapping the study area. For the year 2016 for instance, 13 image collections were filtered. The number of all available image scenes and their respective bands were displayed using the 'print(collection)' code snippet. It prints the name of the scene with full information as printed below.

COPERNICUS/S1_GRD/S1A_IW_GRDH_1SDV_20161029T031817_20161029T031853_013699_015F9A_49E4 (3 bands)”.

The naming indicates the image provider, satellite name, sensor type, mode of acquisition, image type, date of acquisition and calibration, and image serial number and the bands included.

The import of Sentinel-1 data from GEE was followed similar approach used to import sentinel-2 images before merging the layers for classification from these two collections. This is because Sentinel-1 images are microwave data and require specially designed algorithms unlike optical Sentinel-2 images to get imagery orthorectified and calibrated. For this purpose, the following preprocessing steps were adopted in Google Earth Engine on the selected ImageCollections from 2016 to 2020 to derive the backscatter coefficient in each pixel (see Table 11 below). The IW mode GRD radiometric DN values of the VH and VV polarizations were converted to backscatter coefficient (σ°) in decibels (dB). The backscatter coefficient represents the target radar cross-section or backscattering area per unit ground area. Since it can vary by several orders of magnitude, it is converted to dB using Equation 8 below in the corresponding JavaScript.

$$10 \cdot \log_{10}(\sigma^\circ) \text{ ----- Equation 8}$$

It measures whether the radiated incident microwave radiation terrain scatters preferentially away from the SAR sensor dB < 0 or towards the SAR sensor dB > 0. This scattering behavior depends on the physical characteristics of the terrain, primarily the geometry of the terrain elements and their electromagnetic characteristics. Therefore, before the data analysis, a geometric correction image enhancement technique was applied including slant range to ground range conversion, multi looking and azimuth compression.

Table 11: *Sentinel-1 image preprocessing steps in GEE*

Step	Preprocessing steps	Description
1	Apply orbit file	Updates orbit metadata with a restituted orbit file.
2	GRD border noise removal	Removes low intensity noise and invalid data on scene edges (since January 12, 2018)
3	Thermal noise removal	Removes additive noise in sub-swaths to help

		reduce discontinuities between sub-swaths for scenes in multi-swath acquisition modes. This operation, however, cannot be applied to images produced before July 2015.
4	Radiometric calibration	Computes backscatter intensity using sensor calibration parameters in the GRD metadata
5	Terrain correction (orthorectification)	Converts data from ground range geometry, which does not take terrain into account, to unitless backscatter coefficient σ° using the ALOS DEM

The VV and VH dual polarization Sentinel-1 collections in IW mode were composited after initiating these steps for display and most importantly, for fusion with Sentinel-2 data to be used as a final classification composite for forest masking.

3.4.3.4 GLCM Extraction from Sentinel-1 Backscatter Polarizations

Texture is the frequency of tonal change on an image produced by the aggregation of unit features that may be too small to be discerned individually within the image. An example would be the smooth texture of green wet-grass against the rough texture of green tree crowns on an aerial photo. Image texture measures are often used for the reduction of spatial heterogeneity of forest stand structures, especially when forest sites have complex forest stand structures. Grey level co-occurrence matrix (GLCM) based texture measures are the most common approach for textural image extractions. The GLCM image textural features, developed by (Haralick, et al., 1973) measure the spatial homogeneity of the backscattering. The spatial information in the form of these texture features can be useful for image classification and produce new images such as forest structure.

Google Earth Engine has thus multiple ways to geocompute these satellite image textures. One of these for instance is entropy, which provides a measure of the disorder within a space. The Entropy texture measures the randomness of intensity distribution yielding low values for smooth images than for a coarse image. In forested landscapes, these variations in image textures can result from changes in stem crown closure, density, or species type (Franklin et al., 2001) enhancing the accuracy of AGB carbon stock estimation. This dependence of texture on density and size of tree crown is obvious on high-resolution images. This means that if a pixel falls on a tree, its neighbor may also fall on the same tree, resulting in a low local variance. As the image resolution increases

to comparable level to the dominant tree crown size, local variance increases, and this should be particularly true in tropical forests where the stands are heterogeneous and the species diversity is high.

Parameters	Description
Contrast	Calculates the level of local variation within a window (Yuan et al., 1991)
Correlation	Measures the grey-level linear dependency within an image (Kayitakire et al., 2006)
Dissimilarity	Is a measure of the local variation (Rubner et al., 2001).
Homogeneity	Measures the smoothness of image texture (Tuttle et al., 2006)
Mean	Average grey-level in the small neighborhood (Materka and Strzelecki, 1998)
Second Momentum	Second moment is an indicator of local homogeneity (Yuan et al., 1991)
Variance	Variability of the spectral response of pixels (Materka and Strzelecki, 1998).
Entropy	A statistical measure of uncertainty (Yuan et al., 1991)

In this research, to decrease the number of input variables for the regression analysis, the entropy, average, correlation, and variance textures with different window sizes were extracted from the time-series stacks of sentinel-1 VV and VH polarizations. GLCM analysis script was developed and executed in GEE with the input variables from time-series fusion image for each VV and VH polarizations. The detailed steps followed to calculate the GLCM textures are:

1. Defining study area and rectangular geometry extent encompassing the study area

```
//1. Defining study area and
//rectangular geometry extent
var geometry = ee.Geometry.Polygon(
  [[[37.22466865063448, 5.590110130946576],
    [37.22466865063448, 5.507415341802643],
    [37.303632884032915, 5.507415341802643],
    [37.303632884032915, 5.590110130946576]]], null, false);
var AOI = ee.FeatureCollection('users/oytisa/shiwshiwe_bound');
```

2. selecting Sentinel-1 polarizations and apply processing
3. Removing ugly edges
4. Convert sigma naught to gamma naught using equation
5. Refine Lee speckle filter

```

// img must be in natural units, i.e. not in dB!
// Set up 3x3 kernels
// Use a sample of the 3x3 windows inside a 7x7 windows to determine gradients and
directions
var sample_weights = ee.List([[0,0,0,0,0,0,0], [0,1,0,1,0,1,0],[0,0,0,0,0,0,0],
[0,1,0,1,0,1,0], [0,0,0,0,0,0,0], [0,1,0,1,0,1,0],[0,0,0,0,0,0,0]]);
// Calculate mean and variance for the sampled windows and store as 9 bands
var sample_mean = mean3.neighborhoodToBands(sample_kernel);
var sample_var = variance3.neighborhoodToBands(sample_kernel);
// Determine the 4 gradients for the sampled windows and find the maximum gradient
amongst gradient bands
// Determine the 8 directions and "collapse" the stack into a single band image (due to
masking, each pixel has just one value (1-8) in its directional band, and is otherwise
masked)
// Calculate localNoiseVariance
// Set up the 7*7 kernels for directional statistics
// Create stacks for mean and variance using the original kernels. Mask with relevant
direction.
// "collapse" the stack into a single band image (due to masking, each pixel has just one
value in its directional band, and is otherwise masked)
// A finally generate the filtered value
////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////////
6. // Combine 3 steps of speckle filtering into one simple function
function preProcessing(image) {
  var collection = toDB(RefinedLee(toNatural(image)));
  return collection;
}
7. // Convert images values to Natural-->Apply Refined Lee filter-->convert to DB
*****Time-series-2017-Sentinel1-VH *****
// Time-series creating, stack averaging. Converting to gamma naught
var stack1 = ee.ImageCollection(collection.VH.filterDate('2016-01-01', '2016-02-01'));

```

```

var stack1_gamma = stack1.map(toGamma0);
var stack1_Natural = stack1_gamma.map(toNatural);
var stack1_DB = toDB(stack1_Natural.map(RefinedLee).reduce('mean'));
//Map.addLayer(stack1_DB.clip(table), { min: -25, max: 5}, 'stack1_avg', 0);
var stack1_DB = stack1_DB.rename('stack1_VH')
8. // For GLCM texture analysis, build a stack for all the images in the collection
var VH_12_gamma = stack12_Natural.reduce('mean').rename('VH12');
print(VH_10_gamma, 'VH12')
////////////////////////////////////
9. //***** GLCM texture analysis *****
//var textureMeasuresVH = ['VH_asm', 'VH_contrast', 'VH_corr', 'VH_var', 'VH_idm',
'VH_savg',
//          'VH_ent', 'VH_sent', 'VH_dvar', 'VH_dent', 'VH_imcorr1'];
var textureMeasuresVH10 = ['VH10_savg', 'VH10_ent', 'VH10_corr', 'VH10_var'];
var textureMeasuresVH12 = ['VH12_savg', 'VH12_ent', 'VH12_corr', 'VH12_var'];
var          glcmVH_10          =
VH_10_gamma.unitScale(0,1).multiply(255).toByte().glcmTexture({ size:
4}).select(textureMeasuresVH10);
var          glcmVH_12          =
VH_12_gamma.unitScale(0,1).multiply(255).toByte().glcmTexture({ size:
4}).select(textureMeasuresVH12);
var VH_final =
(stack10_DB.addBands(stack12_DB).addBands(glcmVH_10).addBands(glcmVH_12)).cl
ip(table)
10. Print the final GLCM textures into consl and visualize in Code Editor
print(VH_final, 'glcm')
Map.addLayer(VH_final/* .clip(table)*/, {}, 'glcmVH');
//Map.addLayer(VH_final, {bands: 'VH_ent'}, 'VH_ent');

```

3.4.3.5 DSM Data Preprocessing

Topographic variation causes differences in reflectance for similar features with different terrain positions. To account for these effects, ALOS DSM was used. The ALOS Global

Digital Surface Model in GEE was resampled to equivalent resolution of sentinel images using the nearest neighbor method. The DSM was used to generate elevation and slope as forest ancillary data at 10m spatial resolution. The DSM was also used to apply the doppler terrain correction over the SAR GRD data.

3.4.3.6 Optical and Radar Data Fusion

Data availability and performance issues due to sensor failure or clouds and cloud shadows can be alleviated by filling the gaps in a multispectral time series with SAR data. The computing power of these new platforms, enables the application of complex algorithms to large amounts of data. These new satellite image information extraction methods are exploited in current studies to fill the gaps generated by the lack of high-resolution data for certain dates or by automated cloud detection and masking procedures. The gap-filling methods such as best-pixel selection, data fusion, data interpolation , or pixel unmixing have all been applied to timeseries data for the landuse landcover mapping of large areas (Carrasco et al., 2019). The combined use of Sentinel, Sentinel-2, and their derivatives as well as ALOS DSM enabled to create reliable aggregated measurements over a period of time with increased likelihood of obtaining cloud-free images.

In this study, the optical and radar datasets from 2016 to 2020 were aggregated temporarily at pixel level over the whole year into one image for each respective time. The timeseries aggregated images were median images of all the available cloud masked images of each DOY. The aggregated metrics from the different sensor type and the bands used in the image fusion process are illustrated in *Table 12* below.

Table 12: *Bands of the multispectral fusion image*

Sensor	Type	Bands	Metric
Sentinel-2	MSI	B2	median
	MSI	B3	median
	MSI	B4	median
	MSI	B8	median
Sentinel 2 Indices	MSI	NDVI	median
	MSI	EVI	median
	MSI	NDWI	median
	MSI	SAVI	median

Sentinel-1	GRD	VV	backscatter
	GRD	VH	backscatter
	GLCM	entropy	texture
ALOS	DSM	slope	topography
	DSM	elevation	topography

3.4.3.7 Machine Learning using Random Forests

Machine learning (ML) methods, also called artificial intelligence (AI) methods, use algorithmic models to analyse time series data. These methods treat the way in which the data were generated or relationships between the variables as unknown (Breiman, 2001). There exist many ML algorithms that perform different tasks. Some of the algorithms that are relevant to EO data are grouped according to four main analytic aims: classification, clustering, regression and dimension reduction.

In this study, the pixel-based Random Forest classification algorithm within GEE was used to model the land cover classes through supervised machine learning. The use of GEE automation in this case, enables the user to access information on any time scales at any location either input by country or by geometrical selection of a region. The RF was chosen as it is known to be resistant to noise and overfitting issues and its successful handling of high dimensional data attaining satisfactory results (Ghorbanian et al., 2020). The algorithm has several adjustable parameters that directly affect the final mapping accuracy of the image classification result. The number of trees and the number of variables in each node are the most critical parameters for RF-based satellite image classification. The algorithm constructs a decision tree for each sample based on the predictors (bands from the synergized bands). The trees then vote for each pixel to detect forest vs. non forest and then assigns the most supported value to each pixel. The example below illustrates tallies of Six **1**s and three **0**s which results a prediction of 1.

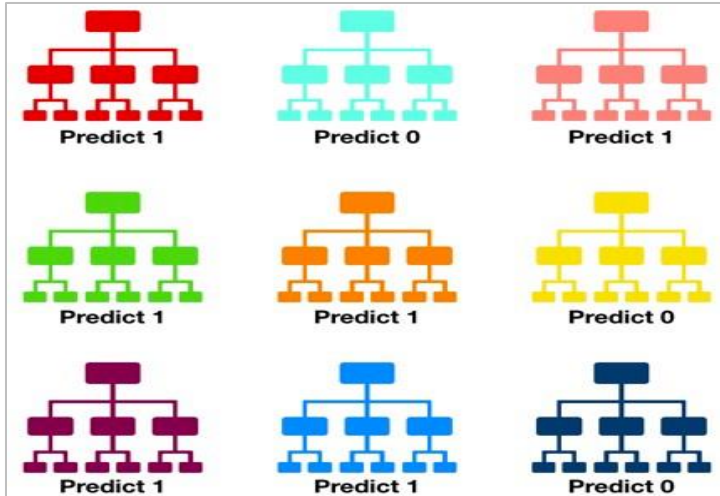


Figure 17: *Example of RandomForest prediction*

Input training points for the RF classifier were developed from high resolution imagery through Google earth time slider using visual inspection and field data collection. For the shapefile containing the field data, separate attribute columns were created for the study image time frames of 2016, 2017, 2018, 2019, and 2020. Respective attributes of each year were field with the land type visually interpreted from the high resolution Google Earth images.

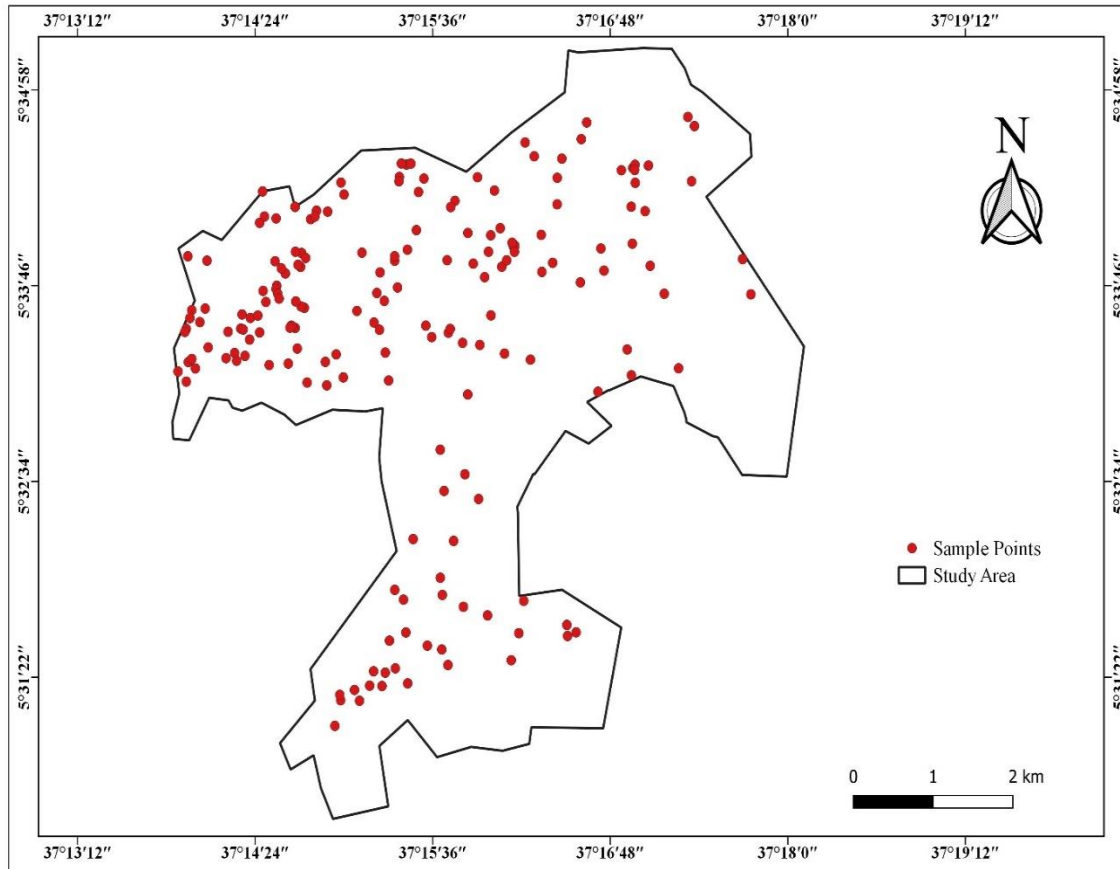


Figure 18: *Input training points for the RF classifier*

These training points were imported as geometry assets into GEE assets and were converted into FeatureCollection using the ‘merge’ function using the java script below.

```
//merge all geometry imports for classification
var tr_aoi = closedforest .merge(medforest)
                    .merge(openforest)
                    .merge(cropland)
                    .merge(grassland);
```

The input training FeatureCollection was then split into training and validation fractions at 70 to 30 percentage ratios to create feature collection with spectral signatures from composite with 'label' property. The training signatures were further used to train the classifier algorithm to assign unknown pixels into the LULC categories of interest for the study periods. A total of 160 training points, minimum of 20 for each LULC were used as input into randomForest classifier algorithm. The training FeatureCollection asset of point geometries was assigned a numeric value attribute indicating the associated land

cover class and a year that corresponds to the land cover label. The land cover attributes or labels were created as numeric so that it cannot be a string (such as “openforest”) or a numeric string. The binary classification scheme for training sample development and forest masking scheme for forest cover assessment is described in Table 13.

Table 13: *Binary classification scheme for forest masking*

Macro	Class	Description
Forest	Closed Forest	Formations where trees in the various storeys and the undergrowth cover a high proportion with tree canopy density of 70% and above. (adapted from FRA 2000 and ISFR, 2019))
	moderately dense forest	All lands with tree cover of canopy density between 40% and 70% above (India State of Forest Report (ISFR), 2019)
	Open Forest	Formations with discontinuous tree layer but with a coverage of at least 10 percent and less than 40 percent. With tree canopy density of 40% covered by primary or secondary, and more but less than 70%. (adapted from FRA 2000)
	Bamboo Forest	Forest on which more than 75% of the crown cover consists of tree species other than coniferous or broadleaved species (e.g., tree-form species of the bamboo, palm and fern families)
Non-Forest	Cropland	Arable and fallow land that grows annual crops (e.g., wheat, maize, sorghum, teff, cotton) or perennial crops (e.g., sugar cane, coffee, enset) on the small scale or commercial level by rain fed or irrigation schemes (adapted from FDRE, 2017).
	Grassland	Land covered with the natural growth of graminea and herbaceous vegetation or a land sown with introduced grass and leguminous for the grazing of livestock. This class also includes grasslands with scattered trees, also known as open woodlands (adapted from FDRE, 2017).

To use and instantiate these training points to be coded in the JavaScript API, a dictionary was created to explain the class meanings and run the model easily.

```
//Create a dictionary explaining the class meanings
var classes = [
  {'label':1,'description':'closedforest'},
  {'label':2,'description':'medforest'},
  {'label':3,'description':'openforest'},
  {'label':4,'description':'cropland'},
  {'label':5,'description':'grassland'}
];
```

The dictionary labeled feature collections were then imported from the asset, which is already in the required format. For the forest masking analysis, Sentinel-2 MSI Surface Reflectance imagery archives available on GEE were used for the years 2016 and 2020. GEE's cloud screening algorithm based on quality assessment bands (QA) was applied to remove cloud and cloud shadow contaminated pixels for each of the 10m resolution scenes covering the study area. Creating monthly and annual 10-m cloud-free imagery composition is a challenging task because of the constant clouds across the forested area most of the time. For this purpose, the surface reflectance bands of sentinel-2, sentinel-1 VV and VH backscattering polarizations and ALOS DEM were concatenated to replace the cloud masked area pixels as a patch. Zone 48 North.

A binary classification (Class/Non-Class) strategy was adopted where each land cover class (Table 13) was mapped individually. Non-forested areas were then masked out from the classified image. This was done by assigning "NULL" to all other landcover classes except the forested areas classified as open-forest, moderately dense forest, and closed-forest.

The results of the binary classification were printed into the GEE console for visualization of their properties and then were exported to Google Drive in GeoTIFF format with a scale of 10m resolution. The forested areas were extracted to be further used to mask the optical-microwave fusion data before estimating the AGB and AGC. The process applied to accomplish this task is presented in Figure 19 below.

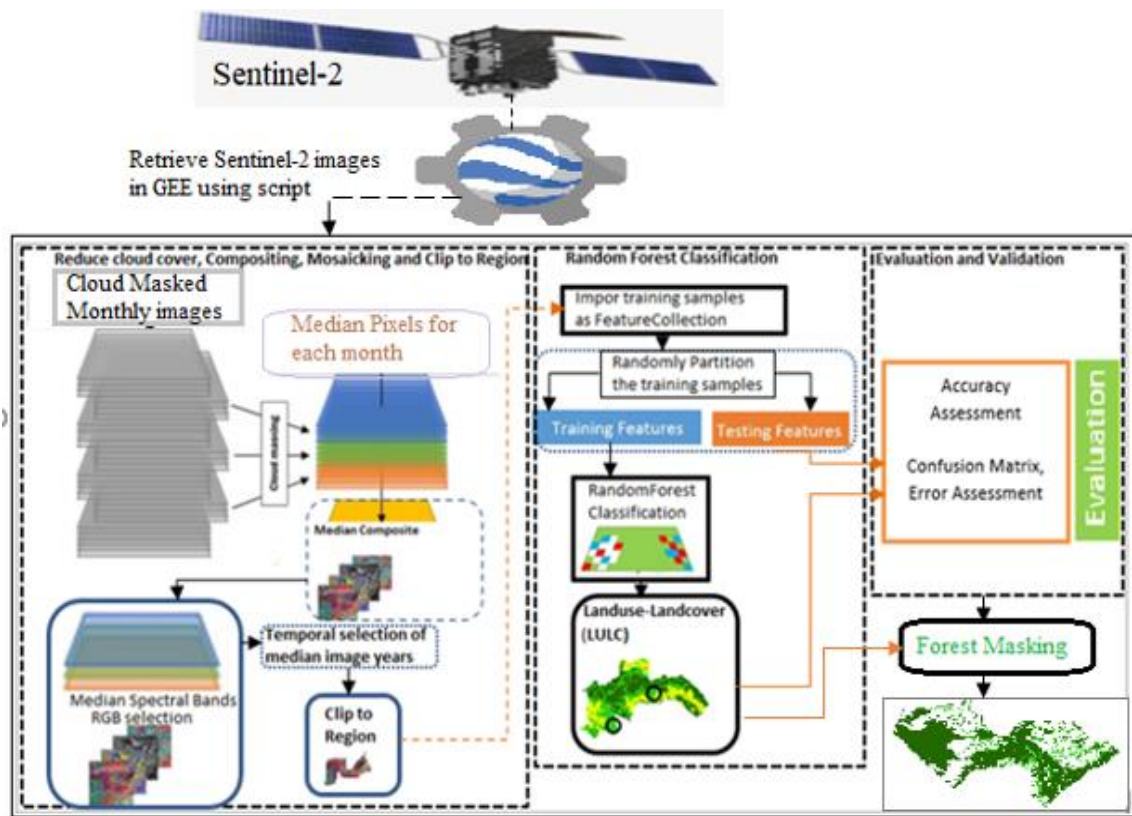


Figure 19: Sentinel-2 image processing and Random Forest classification

The ‘number of trees’ was set to 100 and the 10m resolution Sentinel-2 spectral bands, Sentinel-1 sigma naught backscatters and spectral indices described in Table 3 were used as predicting variables in the RF model depending on the category of each land cover class. The final classification was a binary output of forest and non-forest.

Accuracy assessment in GEE

Regardless of classified visualization, accuracy assessment is a quantification of results. This was done using the GEE code editor by partitioning the set of known sample data values for forested areas and non-forested areas to ratio 70 training to 30 validation. The validation samples (30% split of original training data) served as a testing and was used to deliver a confusion matrix. This confusion matrix is made by comparing the ‘label’ property with the newly created ‘classification’ property. From the confusion matrix, GEE can calculate overall accuracy, producer’s accuracy (omission error), consumer’s/user’s accuracy (commission error), and kappa – coefficient of agreement

and order value. These results were printed into the GEE console and saved as a CSV file into Google Drive.

3.4.4 Predictive Model Development Scenario Analysis

The estimation or prediction of a response (output) variable based on a set of covariates or explanatory (input) variables can be achieved by applying a regression. A time series regression is used to predict a future response based on the response history from relevant predictor values. The regression model is developed or trained based on set of input variables for which the response is known. The model can then be used to predict responses for test datasets (with input variables but no response variable), identifying which input variables are most important in providing good estimates or predictions, or inspecting the relationships between the input variables. Two common methods that are used for this purpose are parametric time series models and nonparametric convolutional neural networks. The parametric time series models capture temporal dynamics such as moving average while nonparametric convolutional neural networks which are an extension of static neural networks, are adapted to describe data over time.

In this study, the relationship between dependent variable and one or more independent variables was quantified using a regression analysis. For this purpose, the spectral value of the optical and SAR bands and their derivatives were extracted using the zonal statistics tool at the 20 plot locations with an average of three interconnected pixels. To develop the relationship between the Sentinel-2 bands, VIs, GLCM, and backscatter values and the in situ AGC, the pixels centering the plot location were considered to predict the quantities of carbon stored from satellite data alone. In this study, multiple linear regression was applied using the Sentinel satellites data as independent variables and the field AGC as a dependent variable.

Then, a carbon stock prediction equation was developed establishing a regression line between each independent variable and field-based carbon stocks. To select the best predictor variable, the coefficient of determination (R^2) value in the regression analysis was used as a comparison parameter. The R^2 is chosen to understand the relationship

between the independent and dependent variables as it shows the percentage of the variability explained by the model.

The best-fitting variables' values were then used to create the biomass and the carbon storage potential of the Shiwshiwe forest in QGIS (Annex 5). The AGB prediction accuracy was validated and tested using the R^2 values. The continuous carbon stock map for the study area was then mapped by multiplying the predicted AGB map using the IPCC carbon fraction factor of 0.47.

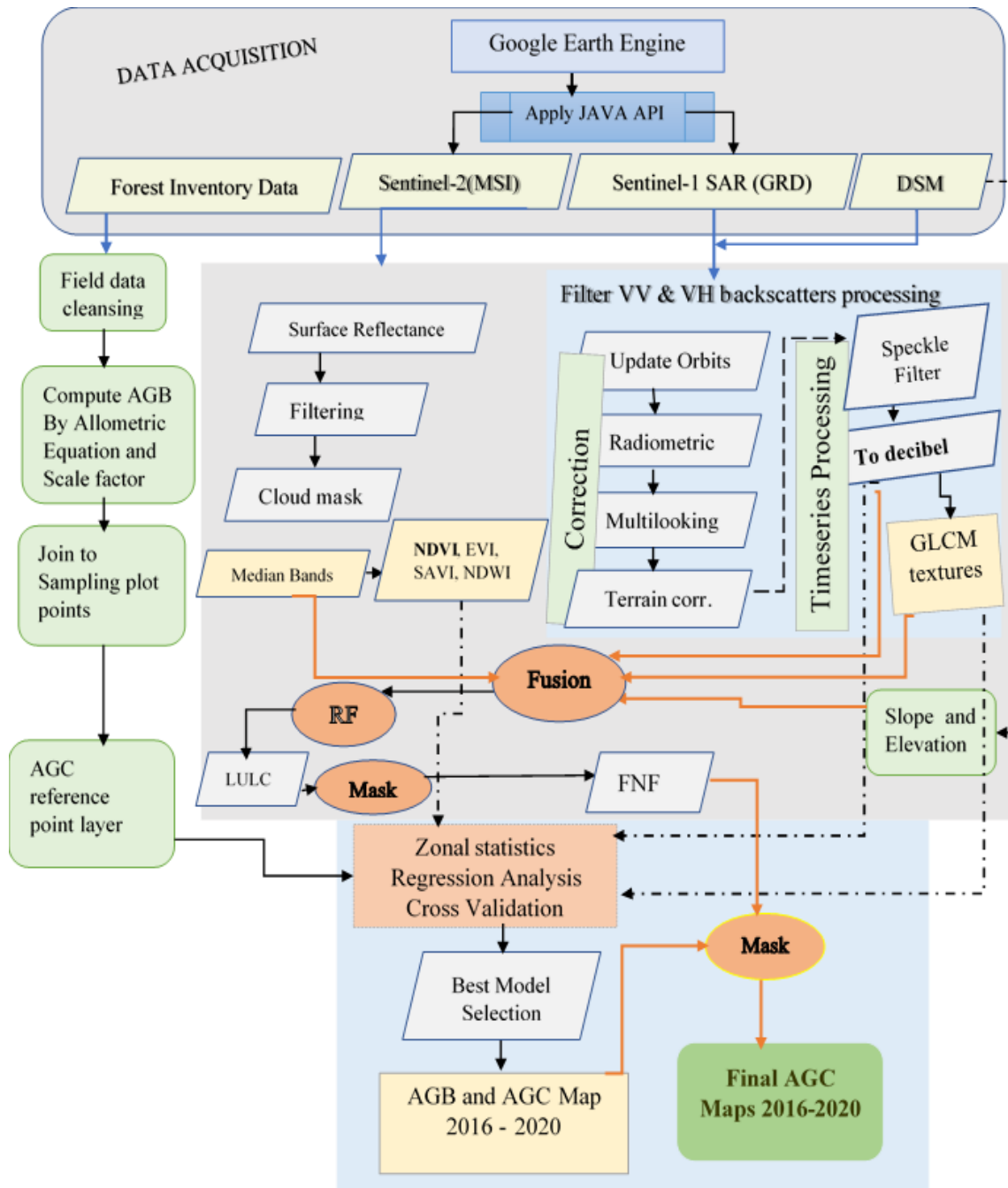


Figure 20. Overall Methodological Flow Chart

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Total Aboveground Biomass and Carbon Stock Estimation

The field forest aboveground carbon stock from live tree biomass was estimated based on field inventory carried out to collect plot level data in 20 sample plots in 2019. The field-biomass and carbon stock estimation were inferred from a total of 175 tree species sampled from 20 circular plots of 0.1ha plot area. The tree DBH measured varied between 5cm and 70 cm (see Figure 21). The measured tree height varied between 5m to 17.5m across the plots inventoried. The dominant species found in the study area were *Hagenia abyssinica*, *Arundinaria alpina* (Highland Bamboo), *Albizia schimperiana*, *Croton macrostachyus* and *Trichilia dregeana*.

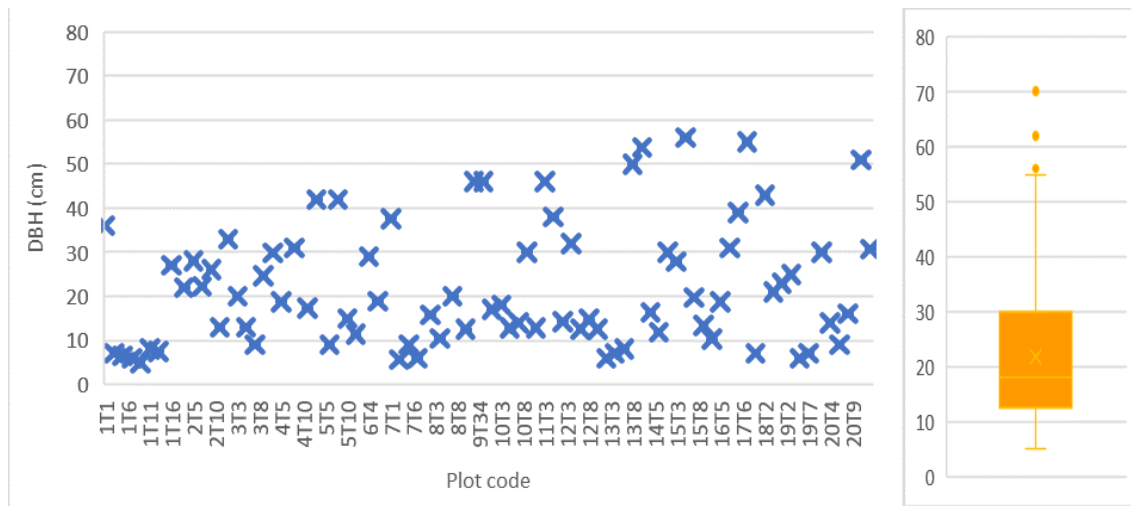


Figure 21: DBH distribution across plots

The AGB calculated based on the measured DBH(cm), H(cm), and species-specific wood-density was found between 0.42 t/ha (in plot P07) to 1.7 t/ha (in plot P17).

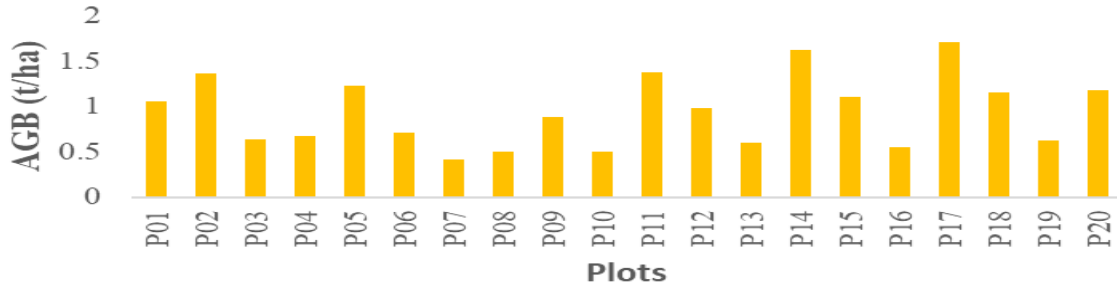


Figure 22: Plot level AGB in tonnes per hectare

The above ground biomass of the study area was found 18.91 t/ha or 18918.52 kg/ha. The above ground carbon was calculated by multiplying the AGB in t/ha by the conversion factor of 0.47 as dry-mass value of carbon. The results of the analysis of the average forest carbon stock in the above ground biomass (AGB) is plotted in Figure 22 below.

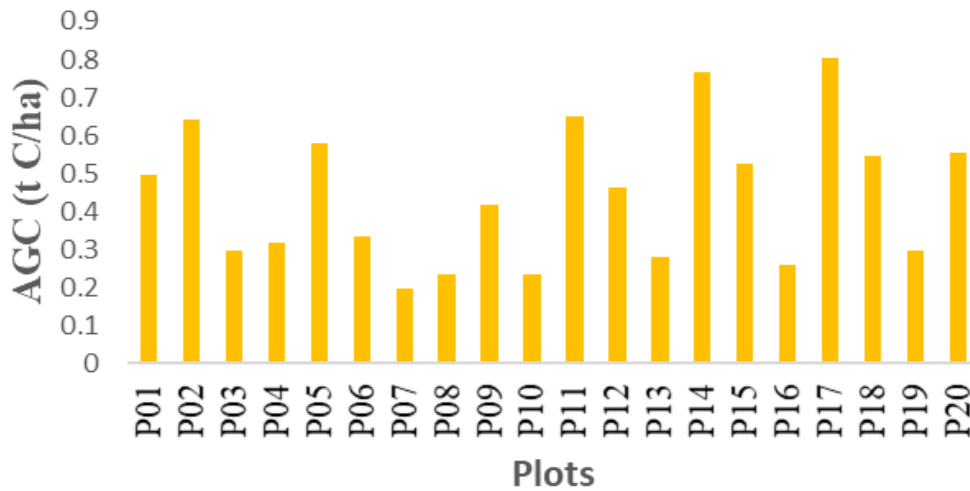


Figure 23: Plot level AGC stock (kg C/ha)

The measured tree height varied between 5m to 17.5m across the plots inventoried. The resulting calculated total AGC stock (kg C/ha) across the plots in the study area was found 8.8917 t C/ha. This value is comparable with the range of mapped carbon values found in the literature. Depending on the regional focus and the methods used, Ribeiro et al. (2011), and da Rocha et.al (2010) as cited in (Schwieder et al., 2018) obtained a range between 3.3 and 32.5 t/ha (mean 8.5 t/ha) or 5.0 -15.9 t/ha (mean 9.7 t/ha) respectively.

It is evident from the figures in the results section that the plot level and mean values of AGB density reflected variability within the forest areas. The variance of the amount of biomass and carbon stored in the forest is based on several factors including forest types, age, management practices, uncertainty in plot size and level of human and natural disturbances as indicated by (Chave et al., 2005; Dick OB, 2015; Saatchi et al., 2011).

4.2 GEE Timeseries Prototype of Forest AGC Stock Development

The prototype development process was carried in the planetary-scale image analysis GEE platform. This involved developing a JavaScript for processing the whole chain of filtering and making medians of satellite images, preprocessing, image fusion, random forest classification, accuracy assessment, forest area masking, to exporting the resulting outputs to google drive for a download (for details look at 0). The resulting prototype is a GEE graphical user interface (GUI) with underlining codes with the following main components.

1. Expression to define study area and create year list variables
2. Search all available sentinel-2 image pixels and filter based on the year list with <1% cloud
3. Commands to mask clouds and create cloud free sentinel-2 band stacks (ImageCollections)
4. Commands to extract vegetation indices
5. Commands to create mosaics of sentinel-1 backscatters
6. Commands to drive GLCM textures
7. Commands to create fusion of the different sensor images
8. Commands for binary classification of the fused image and mask forest areas

```

//call DefineStudyArea code
var studyArea = require('users/oytisa/FinalTheisisMSc:DefineStudyArea');
var aoi = ee.FeatureCollection("users/oytisa/shiwshiwe_bound");
// Expression to create cloudless Sentinel-2.
function maskS2clouds(image) {
  //----- create year list variables -----
  var year_list = ['2015', '2016', '2017', '2018', '2019', '2020'];
  var month_list = ['01', '02', '03', '04',
                   '05', '06', '07', '08',
                   '09', '10', '11', '12', 'year_median'];

  var bands = ['B1', 'B2', 'B3', 'B4',
              'B5', 'B6', 'B7', 'B8',
              'B10', 'B11', 'B12', 'B8A'];
}

```

The result of the commands for searching all available Sentinel-2 images was examined using the `.size()` function over the searched image collections in the dry season. This revealed 26 scenes for 2016 and when applied the `.updateMaskcloud` function to mask clouds, 6 scenes with cloud percentage less than one (<1%) was found. Similarly, for the year 2017, the search resulted 98 scenes and 12 cloud masked scenes. For the year 2018, 11 cloud unfiltered and 10 cloud masked images were obtained. The cloud unfiltered and cloud masked images for 2019 were 26 and respectively. For the year 2020, 10 cloudy scenes and 5 cloudless scenes were found.

The prototype has also the following processing chains for masking forest areas of the study site. The JavaScript prototype uses the filtered median satellite image data to add polygons of training data. With the next chain, input training data and predictors are defined via buffering field samples and drawing training polygons and merging them. This includes creating a variable called image to select the bands of interest and masking to the study area and assembling samples for the model using a class variable that holds labels from each sample geometry. The assemblage includes making each sample the same size as the satellite image pixel size. The next chains hold running the model, assessing accuracy, refining model, and final model result display.

To test the actual accuracy of the model rather than the fit, 50 random sampling points per class were created and run as an independent accuracy assessment using the `“stratifiedSample”` function in GEE. Each random point was buffered by a 10m-radius as shown in the code snippet below.

```

var modelActualAccuracy = classed.stratifiedSample({
  numPts:200,
  label: 'classification',
  region:oytAOI,
  scale: 10,
  geometries:true });
var ptsBuff = function(feature) {
  var pts = feature.get('classification');
  return feature.buffer(10).set('classification', pts);
};
var stratPoints = ptsBuff.map(ptsBuff)

```

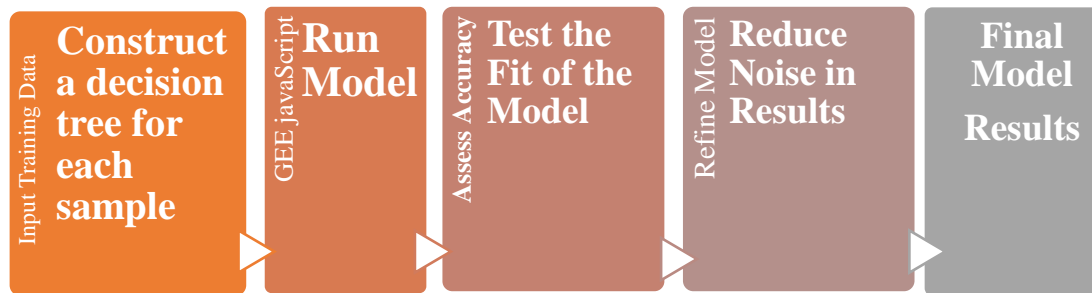


Figure 24: Processing chain of the JavaScript prototype

The resulting prototype model is a machine learning processing chain which is a “black box,” where we can only observe the inputs and outputs. It is a JavaScript in the GEE platform that can be used as a tool for retrieval of information over any landscape that will aid in biomass and carbon stock estimation from satellite images. The time series analysis prototype allowed to understand the potential forest biomass and carbon stock over time in the study area. It also allowed for timely processing of the big-data. The processing time of the whole chains of the algorithms applied over the sentinel-2 image for instance was found 67.138 seconds (see Annex9).

Moreover, the Artificial Neural Network (ANN) prototype submodule developed was able to select Sentinel-2 image bands and the derived vegetation indices (VIs) in GEE for the years 2016 to 2020 for satellite-based biomass and carbon stock estimation. This was carried by examining the spectral reflectance and surface reflectance curves of the bands. The result of the script development was a graphical user interface that depicts the variability of the raw DN and reflectance values of Sentinel-2 bands across respective wavelengths. The resulting GUI (Figure 25: ANN-based Sentinel-2 band selection and

aggregation GUI) has a thumbnail for visualizing the raw satellite image and their ID, study area enhanced layer viewer, and charts to examine the variance of the DN and reflectance values of Sentinel-2 images.

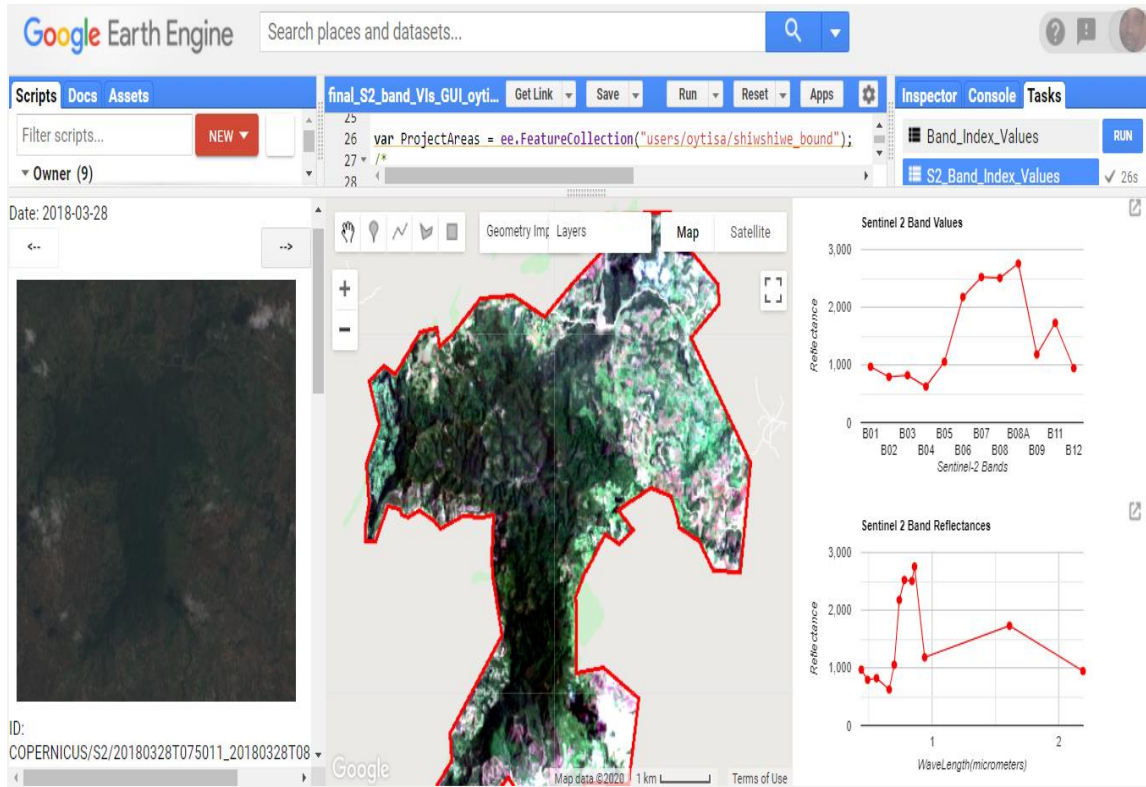


Figure 25: ANN-based Sentinel-2 band selection and aggregation GUI

As a result, the summer Sentinel-2 surface reflectance band values showed high availability across respective wavelengths when examined using the *Map.addLayer* function in the code editor. Therefore, the resulting spectrally variable band combinations from the bands B2 (blue), B3 (Green), B4 (Red), and B8 (NIR) were used to create median composite images of each band per year.

These median bands per year were used in the geocomputation of NDVI, EVI, SAVI, and NDWI indices. The corresponding visualization parameters were then later used in the further processing of the satellite images such as multispectral image fusion, random-forest classifications, and predictive model development.

Results of Multispectral Image Fusion Module

The fusion of the median sentinel-2 bands with Sentinel-1 and ALOS DSM has increased the separability of the image bands. This allowed to visualize the image with a better appearance and increased the visual classification of the images. The histogram of the stacked raw bands (A) and the fused median band histograms (B) in depict the removal of redundant information from the bands and increased separability for better information extraction.

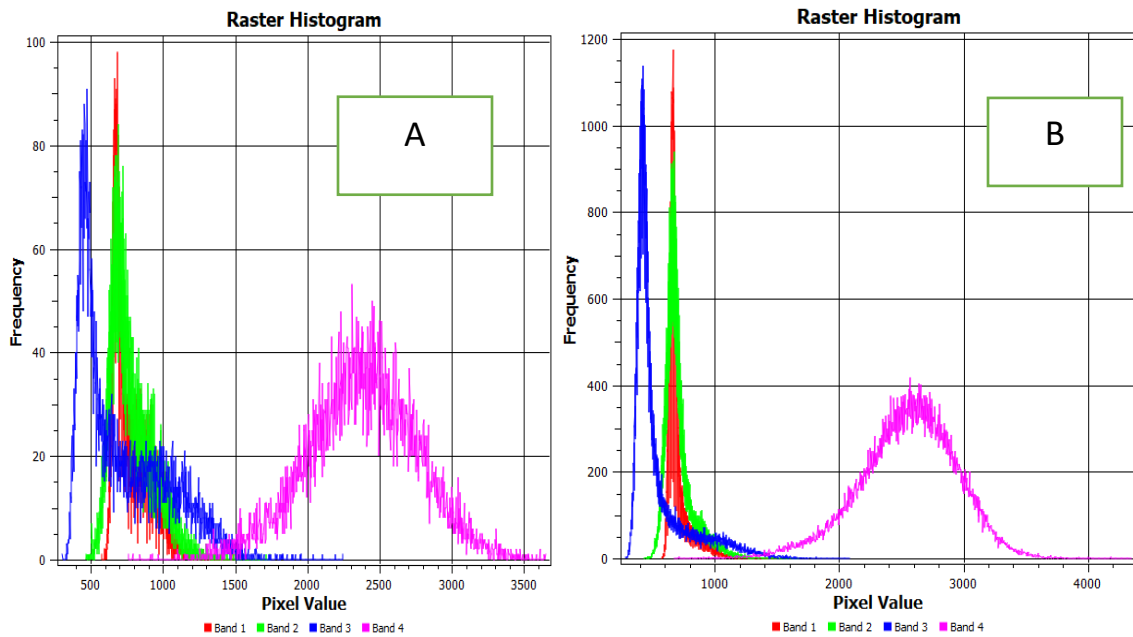


Figure 26: Comparison of raw Sentinel2 bands (A) with Fusion image bands (B)

The fusion of the temporal multispectral bands from Sentinel-2, Sentinel-1, and ALOS DSM was carried at a pixel level in GEE automation environment. This was carried over the median images of the sentinel images filtered at a yearly basis using the JavaScript prototype developed. The resulting image is a natural look multi-sensor image that is easy to interpret as shown in figure 27 below.

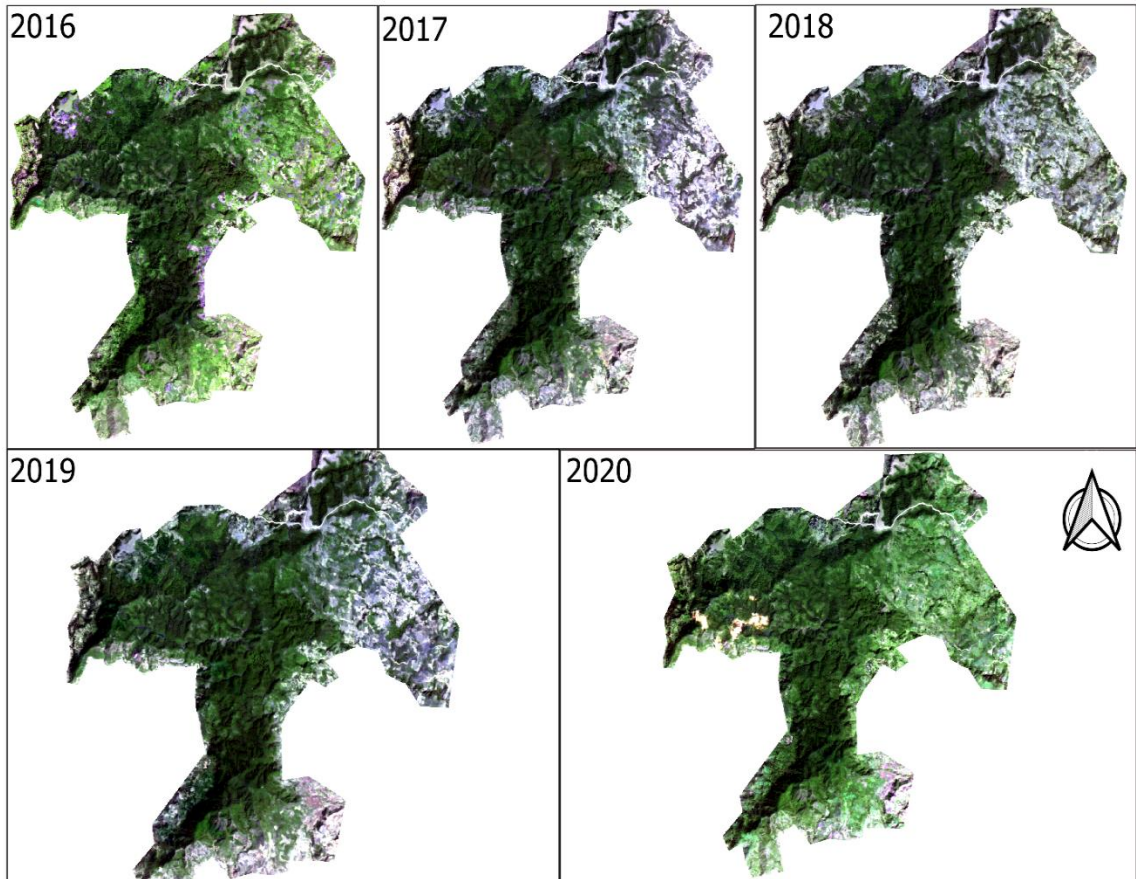


Figure 27: *Timeseries of multispectral pixel level data fusion result*

This study revealed that the variables used in the fusion of the Sentinel-2 reflectance bands, vegetation indices, Sentinel-1 backscatters, GLCM textures and topographical parameters varied by their attribute importance when run the following code of the RF model of the prototype.

```

s1...  Get Link  Save  Run  Reset  Apps  ⚙️
485  //RANDOM FOREST TREES AND VARIABLE IMPORTANCE
486  var dict = classifier.explain();
487  print('Explain:',dict);
488  var variable_importance = ee.Feature(null,
489  |         ee.Dictionary(dict).get('importance'));
490  var chart =
491  ui.Chart.feature.byProperty(variable_importance)
492  .setChartType('ColumnChart')
493  .setOptions({
494  title: 'Random Forest Variable Importance',
495  legend: {position: 'none'},
496  hAxis: {title: 'Bands'},
497  vAxis: {title: 'Importance'}
498  });
499
500  print(chart);

```


The code snippet above yielded the importance of each random forest algorithm trees as in the graph below.

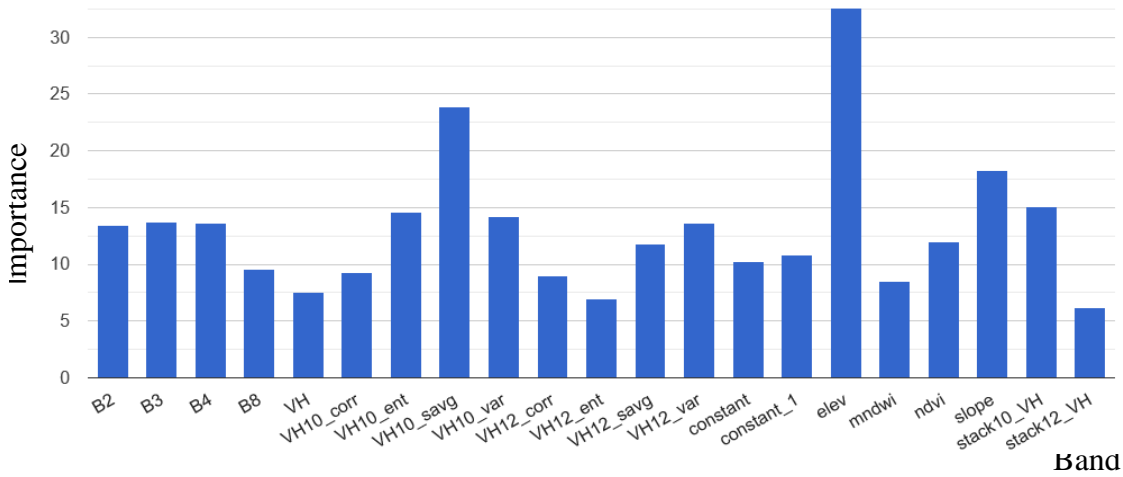


Figure 28: Graph of Variable Importance

VV_Entropy and VH_SumAV were the most important variables among the S1 variables for explaining the observed spatial patterns of forest AGC in the linear regression, and this was also shown in the RF models. The texture features were helpful to note that the texture characteristics of the Sentinel SAR backscatters with a larger window size achieved a greater potential for mapping forest AGB, based on the correlation analysis. It can be understood in this study that direct use of the backscatter coefficients from the SAR C band alone might not be appropriate. The nonexistence of penetrability could affect biomass carbon stock information extraction.

Result of Random Forest Classification and Accuracy Assessment in GEE

Forest cover area mapping is important to extract reliable input data for biomass estimation and a step-wise approach to national FRL development. The forest cover mapping for the area was based on a comparative analysis of three scenarios. For each scenario, a JavaScript code was developed to run the Random Forest classification and respective accuracy assessment in GEE to further mask the forest areas.

Table 14: Random Forest Classification Scenarios

Scenario	Description
Scenario 1	using only optical bands (B2, B3, B4, B8)
Scenario 2	using only SAR backscatters (VH and VV)
Scenario 3	using multispectral fusion (Optical bands, SAR backscatters, VIs and ALOS DSM)

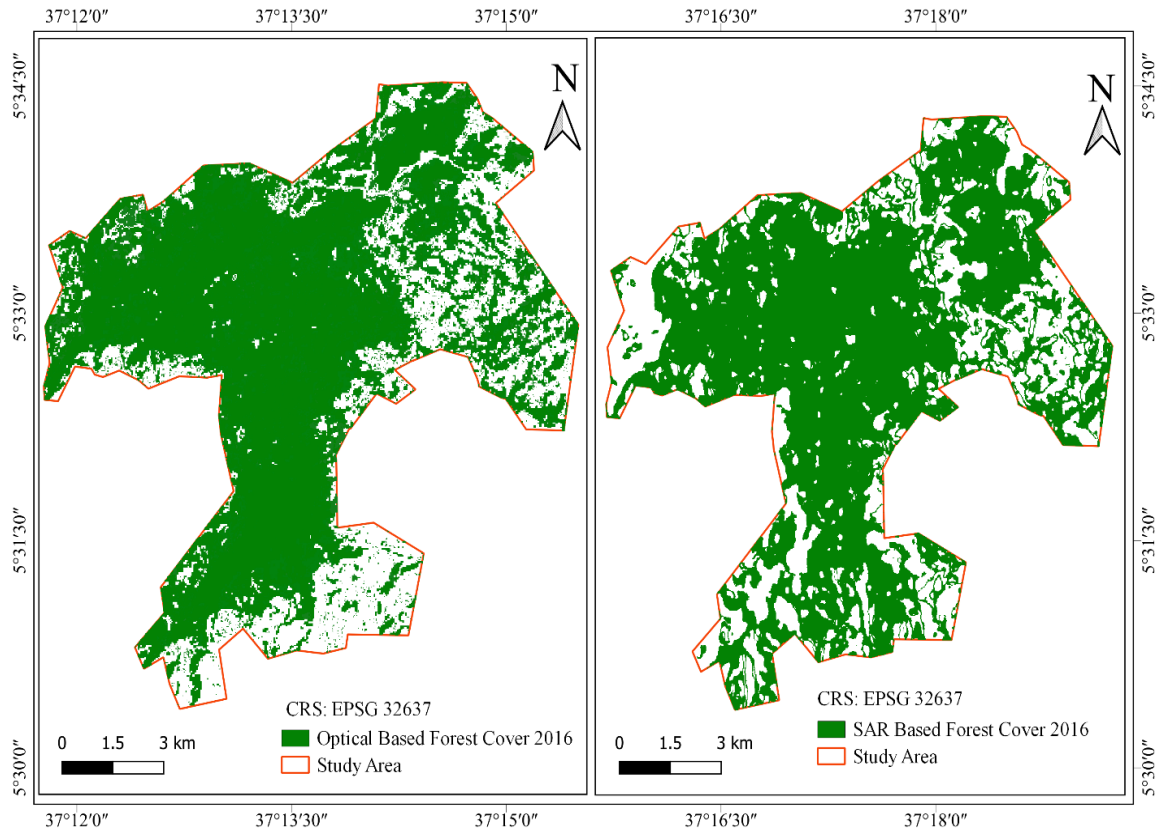


Figure 29: *Scenario-1 and Scenario-2 Forest cover Maps*

As a result, scenario-3 yielded a detailed timeseries of high accuracy binary classes of forest and non-forest images (*Figure 30: Scenario-3 based forest cover maps*). The output of the prototype developed based on this scenario has increased the details of the forest cover map and a highest accuracy of 95% was attained.

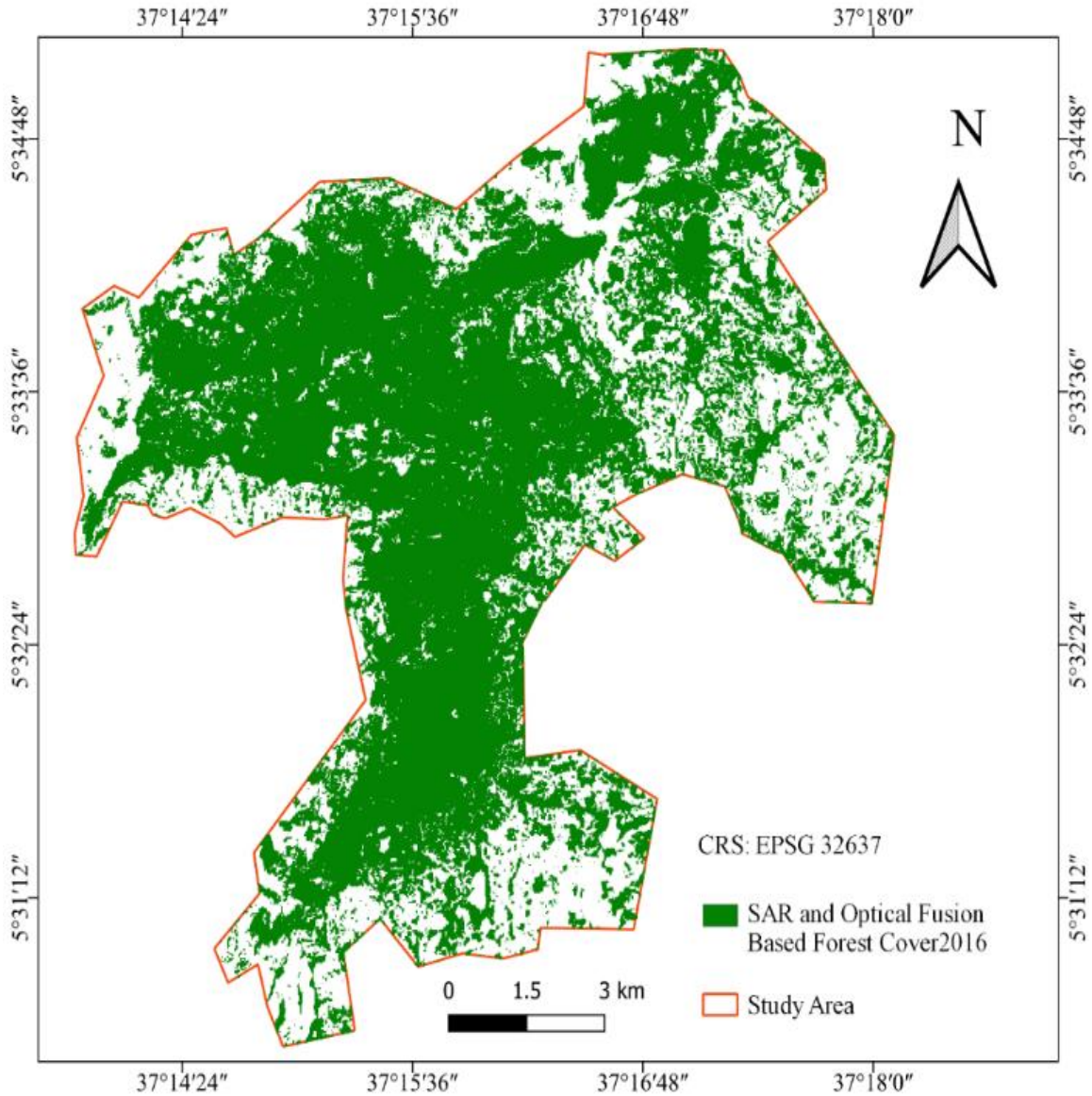


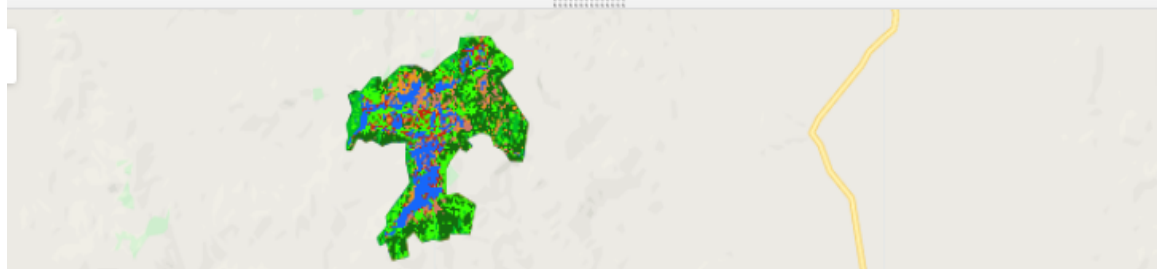
Figure 30: Scenario-3 based forest cover maps

The final forest cover map of the study area was thus produced using the combined composite of 10m sentinel-2 bands, Sentinel-1-backscatters, NDVI, NDWI, and ALOS DSM. The resulting JavaScript code and run-output for the compositing and random forest as well as GLCM texture extraction is shown below.

```

SAR_scenario_lulc_v2 *
117
118 //Train the classifier
119 var classifiers2 = ee.Classifier.randomForest().train({
120   features: s2training, classProperty: 'aoi2016', inputProperties: s2bands
121 });
122 //Run the Classifier
123 var classifieds2 = composite.select(s2bands).classify(classifiers2);
124
125 //Display the Classification
126 Map.addLayer(classifieds2.clip(oytAOI),
127 {min: 1, max: 7, palette: ['1667fa', 'c9270d', 'cf7b68', 'ee9a1c', '146d0e', '04bc
128
129 // Create a confusion matrix representing resubstitution accuracy.
130 print('RF-S2 error matrix: ', classifiers2.confusionMatrix());
131 print('RF-S2 accuracy: ', classifiers2.confusionMatrix().accuracy());
132
133 //Define both optical and SAR fusion to train
134 var opt_sar = ee.Image.cat(composite, SARVV_filtered, SARVH_filtered);
135 var bands_opt_sar = ['VH', 'VV', 'B2', 'B3', 'B4', 'B8', 'NDVI']; var training_opt_s
136   collection: mergeSamples, properties: ['aoi2016'], scale: 10
137

```



Accuracy Assessment of the Scenarios Module

Classification accuracy assessment verifies the accuracy of the classified image with reference data that is typically generated either through interpretation of imagery with higher spatial resolution or ground data collection (Lillesand et al. 2015). A complete and meaningful accuracy assessment framework should consider various components of map accuracy beyond traditional pixel-based approaches to paint a complete picture (Klotz et al., 2016). The prototype model accuracy assessment was based on the evaluation of the *ee.Classifier.smileRandomForest()* classifier with array of 40 number of trees. This was done using 75 to 25 ratios of training and validation samples for each scenario within the variant number of trees while processing the *smileRandomForest algorithm*. However, the results of Scenario-3 are presented here as a showcase as it yielded 0.98 overall accuracy with kappa statistics of 0.97. To overcome the computational timing out or “memory limit is being exceeded” error, the accuracy assessment results of confusion

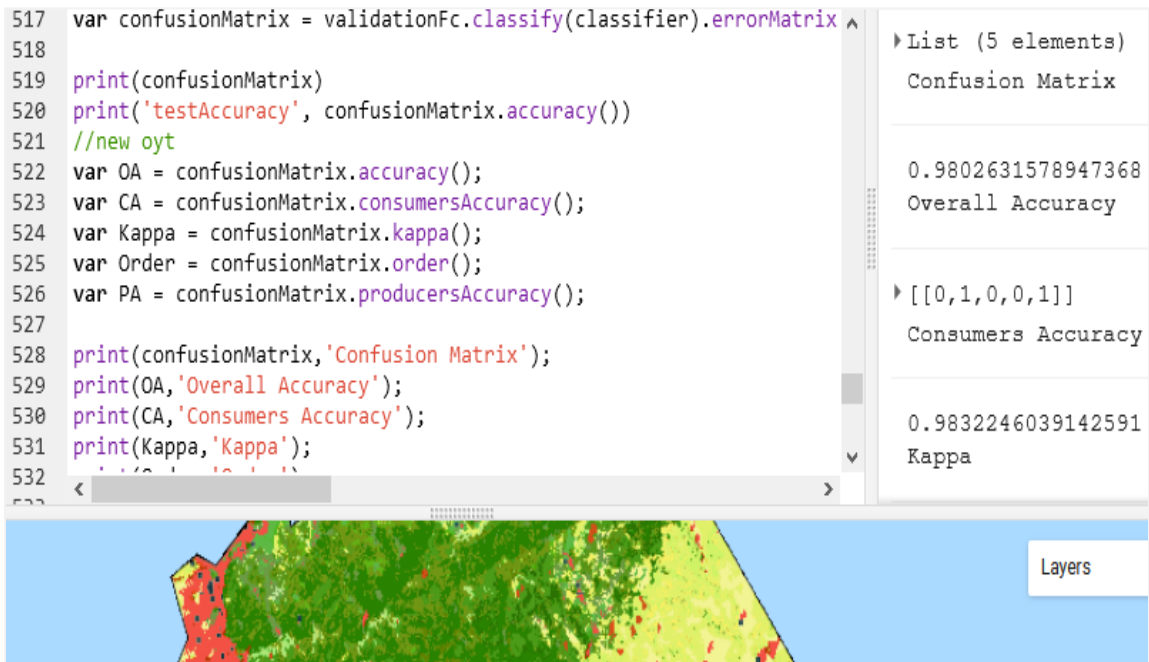
matrix and accuracy were exported as a CSV. This needed to create a 'feature collection' to be able to use the Export tools in GEE.

The pre-classified imagery was used for the train accuracy assessment that examines the accuracy of the classifier applying the `confusionMatrix().accuracy()` function. The confusion matrix was computed using `validationFc.classify(classifier).errorMatrix` expression. The results were then printed into the GEE console as depicted in the code below.

```

517 var confusionMatrix = validationFc.classify(classifier).errorMatrix
518
519 print(confusionMatrix)
520 print('testAccuracy', confusionMatrix.accuracy())
521 //new oyt
522 var OA = confusionMatrix.accuracy();
523 var CA = confusionMatrix.consumersAccuracy();
524 var Kappa = confusionMatrix.kappa();
525 var Order = confusionMatrix.order();
526 var PA = confusionMatrix.producersAccuracy();
527
528 print(confusionMatrix, 'Confusion Matrix');
529 print(OA, 'Overall Accuracy');
530 print(CA, 'Consumers Accuracy');
531 print(Kappa, 'Kappa');
532

```



The final summary of the different types of accuracy assessment method in GEE environment is summarized for the scenario-3 in below.

Table 15: Accuracy Assessment in GEE for Scenario-3

Accuracy Assessment type	Resulting Accuracy
Train Accuracy	0.999
Test Accuracy	0.985
Overall Accuracy	0.98
Consumers Accuracy (Kappa)	0.983
Producers Accuracy	0.964

The fusion of the remote sensing sensors image data were used to drive the forest cover of the area in GEE and the resulting LULC was compared to existing global landcover

such as the Global Copernicus Land Service (Buchhorn, Smets, et al., 2020) to examine the potential reliability of the prototype model. The combination between the satellite data helped to distinguish complex vegetation arrangements and improved forest cover estimates. It is hence an important output as suggested by (Isbaex & Coelho, 2021) that the fusion of vegetation indices with image texture measurements contribute more to improvements in classification accuracy than when they are used separately. This is clearly shown from the result from scenario-3 (see Table 15).

The prototype developed here as per the objective of the study, maximized the possibility to choose a specific period and create image mosaics, with the best cloudless pixel, solve terrain effects problems, and identify any changes in forest ecosystem in the world through classified images. Besides this, it can use several input parameters such as slope, aspect, radar textures and vegetation indices from multisensor bands at the same time when image classification is performed.

Therefore, this prototype provides a new baseline forest cover AGC product at 10 m resolution that is developed and validated in almost near-real time and at the same time maximizes the impact and uptake for the end users. The Sentinel-2 satellites experience high revisit time of five days. Sentinel-1 satellites penetrate areas that remain cloudy for a long time, while they also provide complementary information on the structural characteristics of the observed forest land cover. The research prototype model is thus capable to update the AGC map almost in real time when the two satellites data are jointly used with some ancillary data such as ALOS DEM. This research work is also a tremendous step forward towards the synergetic use of Sentinel-1 and Sentinel-2 data for worldwide land cover (Buchhorn, Lesiv, et al., 2020) and emerging AGC monitoring initiatives.

The statistical evaluation of the final AGC map product was 95% overall accuracy. Several studies in the biomass estimate indicate errors that can range from 5 to 30% (MEFCC, 2018) which arise from the different spatial and temporal resolutions, sensor type, scale, field data errors, and uncertainty. The study lies in the recommended accuracy for planning and management decision of forest which is recommended that in

forest research, accuracy reaches values greater than 90% for a regional scale and 80% for a national or global scale according to (Isbaex & Coelho, 2021; MEFCC et al., 2018)

4.3 The aboveground carbon stock predictive model development.

4.3.1 Correlation of Sentinel-2 satellite variables with Observed AGC

The sentinel products for predictive variable selection were resampled to the plot center points using the “Sample raster values” tool in QGIS. This algorithm creates a new vector layer with the same attributes of the input layer and the raster values corresponding to the point location. This means that if the raster layer has more than one band, all the band values are sampled. The sampled values from respective satellite products such as NDVI for the year 2019 were then used as input values in the correlation against the field AGC values inventoried. In this study, simple linear regression was first applied to understand the predictive power of each of the independent variables on the dependent variable once a causal relationship has been confirmed. Results obtained for the different Sentinel-2 based vegetation indexes are shown in Table 16 below.

Table 16: *Summary of Regression Statistics between AGC and Vegetation Indices*

<i>Regression Statistics</i>	NDVI	NDWI	EVI	SAVI
Multiple R	0.90	0.30	0.87	0.76
R Square	0.81	0.09	0.76	0.58
Adjusted R Square	0.80	0.04	0.75	0.55
Standard Error	0.08	0.18	0.09	0.12
Observations	20	20	20	20

4.3.2 Correlation of Sentinel-1 satellite variables with Observed AGC

A. Filtering GRD backscatter polarizations and GLCM textures

The availability of Sentinel-1 GRD SAR collections at full ‘spatial resolution’ via GEE potentially saves everyone’s time and hassle. The grey level co-occurrence matrix or GLCM textures are arrays of values that determines the varying combinations of pixel brightness values (grey levels) present in an image (Beyer, 2017). The GRD VV and VH backscatter polarizations and GLCM textures were extracted using the Java API in GEE. The final terrain corrected values are converted to decibels via log scaling ($10 \cdot \log_{10}(x)$) and quantized to 16-bits.

To determine the spatial co-occurrence of the pixel grey levels, 10 textures were computed for each sigma naught VV ($\sigma^{\circ}VV$) and sigma naught VH ($\sigma^{\circ}VH$) polarizations. These matrices of textures were angular second moment, contrast, dissimilarity, energy, entropy, correlation, mean, variance, homogeneity, and max. The textures were first composited at monthly level and then were stacked as median for each respective study years (2016-2020). However, to decrease the number of input variables for the regression analysis, entropy, Sum average, correlation, and variance textures with 3x3 and 7x7 moving window sizes were extracted from the time-series stacks of sentinel-1 VV and VH polarizations computed through the GEE. Only the 7x7 textures were used in this study. The variables were selected based on calculation of random-forest variable importance. The following code snippet of the prototype shows the procedure to extract the grey level cooccurrence matrix textures.

```

| /////// GLCM Texture Analysis *****
| ///////////////////////////////////////////////////////////////////
var textureMeasuresVH = ['VH_asm', 'VH_contrast', 'VH_corr', 'VH_var', 'VH_idm', 'VH_savg',
| 'VH_ent', 'VH_sent', 'VH_dvar', 'VH_dent', 'VH_imcorr1'];

var textureMeasuresVH10 = ['VH10_savg', 'VH10_ent', 'VH10_corr', 'VH10_var'];
var textureMeasuresVH12 = ['VH12_savg', 'VH12_ent', 'VH12_corr', 'VH12_var'];

var glcmVH_10 = VH_10_gamma.unitScale(0,1).multiply(255).toByte().glcmTexture({size: 4}).select(
var glcmVH_12 = VH_12_gamma.unitScale(0,1).multiply(255).toByte().glcmTexture({size: 4}).select(

var VH_final =(stack10_DB.addBands(stack12_DB).addBands(glcmVH_10).addBands(glcmVH_12)).clip(oyt
print(VH_final, 'glcm')
Map.addLayer(VH_final/*.clip(oytAOI)*/, {}, 'glcmVH');
//Map.addLayer(VH_final, {bands: 'VH_ent'}, 'VH_ent');

//***** Data export *****
// Export a cloud-optimized GeoTIFF to oytisa@gmail.com (Google Drive).
Export.image.toDrive({
  image: VH_final.reproject('EPSG:4326', null, 10).clip(oytAOI),
  description: 'glcmVH10m_f2019',
  folder: 'gee_oyt',

```

B. Correlation of Sentinel-1 Variables with Observed AGC

Table 17: Summary of Regression Statistics between AGC and Sentinel-1 variables

Regression Statistics	VH	VV	Entropy	SAvg	Correlation	Variance
Multiple R	0.84	0.64	0.78	0.45	0.46	0.28
R Square	0.70	0.40	0.61	0.20	0.21	0.08
Adjusted R Square	0.68	0.37	0.59	0.16	0.17	0.03
Standard Error	0.10	0.15	0.12	0.17	0.17	0.18
Observations	20	20	20	20	20	20

The correlation result found among AGC and the VV and VH backscatters were comparable with (Huang et al 2018). VH is shown to be more useful for the AGC estimation than the VV polarization backscatters due to less influence by soil moisture. Similar studies in the tropical forests have been made on the use of shorter wavelengths (e.g., C-band, 3.75 cm to 7.5 cm) by examining the features of GLCM textures that were derived from $\sigma^{\circ}VH$. Therefore, forest live biomass carbon stock is sensitive to $\sigma^{\circ}VH$ due to the occurring depolarization.

The correlation ($r^2 = 0.7$) observed in this research also agree with the findings of Thiel (2016) where the r^2 is 0.65. However there exist no comparable research in the study area. The grey level co-occurrence matrix texture (GLCM) analysis responds significant but relatively moderate relationship to AGC in the study area.

In general, a good relationship between grey level co-occurrence matrix and field above ground biomass was also observed with correlation coefficient ($r^2 =$ between 0.08 - 0.61). The $\sigma^{\circ}VH$ Entropy texture performed better with the highest correlation value of 0.61 and has contributed to the accuracy of the carbon stock mapped.

4.3.3 Correlation Summary of the Independent Variables with Observed AGC

To the linear regression predictive model, the R^2 values were compared to assess the variability explained by each of the independent parameters as backscatter in glcm textures and selected Sentinel-2 VIs. The model validation result of the input independent variables in the regression model are depicted in Figure 31 below.

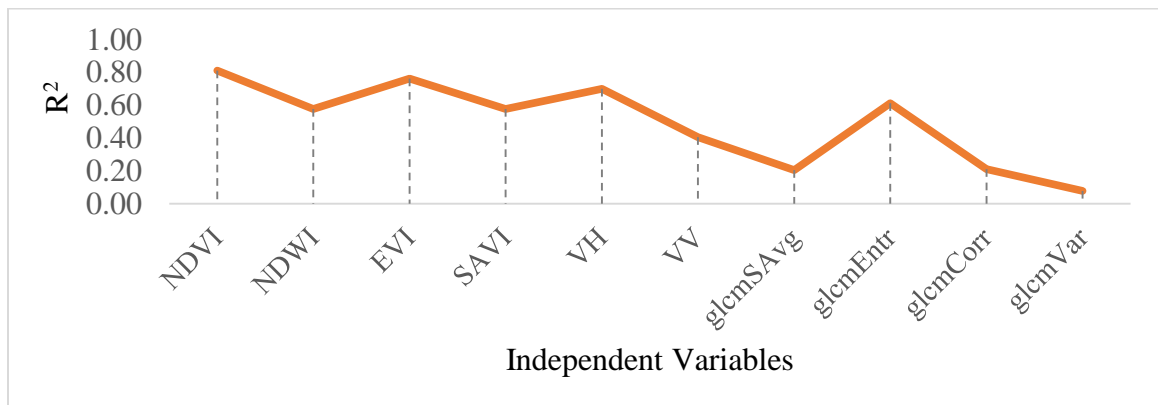


Figure 31: R^2 of independent variables based on simple linear regression

The correlation result depicted that the highest coefficient of determination value NDVI ($R^2 = 0.81$) and a comparatively lower R^2 for VH_variance texture (0.08). The Sentinel-1 gray level cooccurrence matrix textures, in general, showed less correlation with the above-ground biomass carbon stock but are very significant in the geocomputation of the AGC. The entropy texture resulted in a higher correlation with the plot biomass compared to the rest GLCM textures (see Table17).

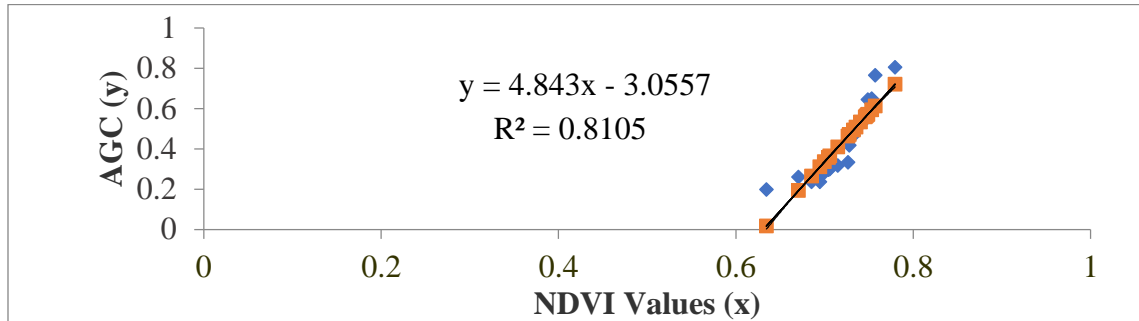


Figure 32: *The correlation between AGC and NDVI and the predictive model*

4.3.4 Selection of the Best Variable to develop predictive equation

The straight-line relationships between the dependent variable and the independent variables (predictors) were studied using multiple regression analysis as in Equation 9.

$$Y = mx_1 + mx_2 + mx_3 + \dots + mX_n + b \dots \dots \dots \text{Equation 9}$$

Where Y= the dependent variable

m= slope of the regression

$x_1, x_2, x_3, \dots, x_n$ = independent variables of the regression

b = constant

The resulting summary of the multiple regression analysis is depicted below.

Table 18: Results of multiple regression

<i>Regression Statistics</i>						
Multiple R	0.968960289					
R Square	0.938884043					
Adjusted R Square	0.870977423					
Standard Error	0.066197009					
Observations	20					
<i>ANOVA</i>						
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>	
Regression	10	0.60587	0.06059	13.8261	0.000267	
Residual	9	0.03944	0.00438			
Total	19	0.6453				
	<i>Coefficients</i>	<i>Std Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	0.300550796	1.31343	0.22883	0.82412	-2.67063	3.27173
NDVI2019	2.238629598	2.41582	0.92666	0.37829	-3.22633	7.70359
NDWI2019	2.418137101	0.78207	3.09196	0.01289	0.648967	4.18731
EVI2019	0.299892967	1.07071	0.28009	0.78574	-2.12223	2.72201
SAVI2019	-0.167091382	0.80442	-0.2077	0.84007	-1.98682	1.65264
gcm_Savg	-0.001713668	0.00193	-0.8892	0.39706	-0.00607	0.00265
gcm_ent	0.091130511	0.13454	0.67737	0.5152	-0.21321	0.39547
gcm_corr	-0.250478126	0.19597	-1.2782	0.23317	-0.69379	0.19283
gcm_var	0.000174455	0.00051	0.34046	0.74132	-0.00098	0.00133
VV	-0.015658079	0.00984	-1.5913	0.14601	-0.03792	0.0066
VH	-0.003401657	0.01365	-0.2492	0.80878	-0.03428	0.02747

When fitting a multiple linear regression model, it is likely to eliminate independent variables that are not important in predicting the dependent variable Y from the final equation. The objective in finding the “best equation” is to find the simplest model that adequately fits the data. The prediction equation is the equation with all of the independent variables in the equation but with the least number of variables that still explain a percentage of variance in the dependent variable that is comparable to the percentage explained with all the variables in the equation. The reason to develop the best Regression equation is not necessarily the equation that explains most of the variance in Y (the highest R²) but the one that fulfills the following best model selection techniques as published in <http://web.csulb.edu/~msaintg/ppa696/696regmx.htm>.

- This equation will be the one with all the variables included.
- The best equation should also be simple and interpretable. (i.e., contain a small number of variables).
- Simple (interpretable) & Reliable - opposing criteria.
- The best equation is a compromise between these two.

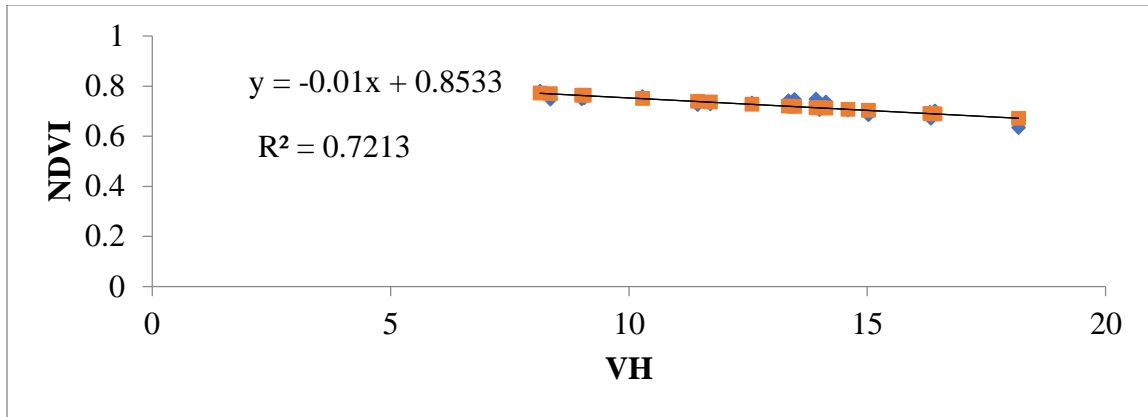
The predictive equation for AGC estimation was thus finally developed using variables with positive relationship. The equation developed Equation 10 below explains 93.4% a significant amount of the variation in the dependent variable of the AGC estimation with the selected variables.

$$AGC_p = (2.238 * NDVI) + (2.418 * NDWI) - (0.299 * EVI) + (0.09 * Entropy) - 0.3) * 0.47 \dots\dots\dots \text{Equation 10}$$

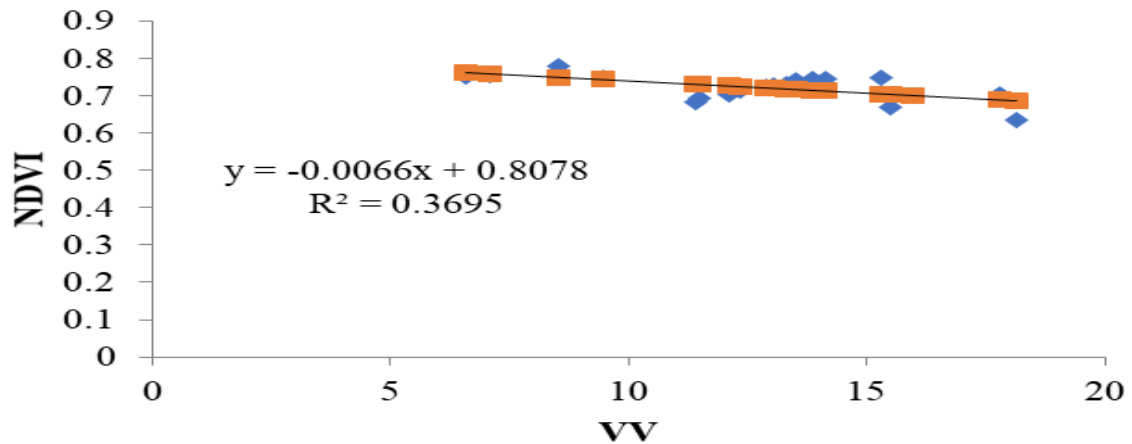
Based on the resulting R^2 value of the multiple regression, the NDVI, NDWI, EVI of sentinel-2 and Entropy texture of sentinel-1 were selected to compute the above-ground carbon stock of the Shiwshiwe forest as there is a significant linear correlation between AGC and the selected independent variables. The R^2 value of ~0.9388 indicates that the model accounts for about 94% of the variance in the dependent variable. It here interpreted as R^2 value closer to zero indicates no linear relationship, while a value close to one indicates a perfect linear fit.

4.3.5 Comparison Vegetation Indices and Backscatter

To investigate the relationship between Sentinel-2 VIs and Sentinel-1 polarization backscatters, in estimation of the plot-based AGC, Sentinel-1 Backscatters were examined by plotting the NDVI values against the VV and VH backscatter values. The NDVI vegetation index is selected as it has strong linear relationship with and it alone describes 81% of the variations in AGC. The resulting scatterplots trend-lines were fitted and the possible linear, exponential, or polynomial relation was examined. As a result, NDVI has positive linear relationship with VH.



The relationship between the NDVI and VV backscatter is also positive but the weaker than the relationship with VH.



Overall, this indicates that the backscatters can be used in vegetation mapping and monitoring their status. However, direct use of these backscatters might not be good estimators of AGB and AGC as depicted in Figure 31.

4.3.6 Aboveground Carbon Stock Estimation Based on Best Variable

The AGC map of the Shiwshiwe forest was geocomputed based on the equation derived from the multiple linear regression analysis (see Table 18) between AGC and the selected four-independent variables using Equation 10. This results annual timeseries AGC maps for the years 2016 to 2020. The resulting map shows the estimated tree biomass carbon content in a raster form for the study area. The map was then masked to the forest areas derived from randomforest classification for each respective year and then multiplied by the IPCC default conversion factor of 0.47. The final predictive equation developed for

the AGC estimation was used to produce continuous raster maps. As a result, the higher AGC was exhibited by the natural forest areas than other landcovers.

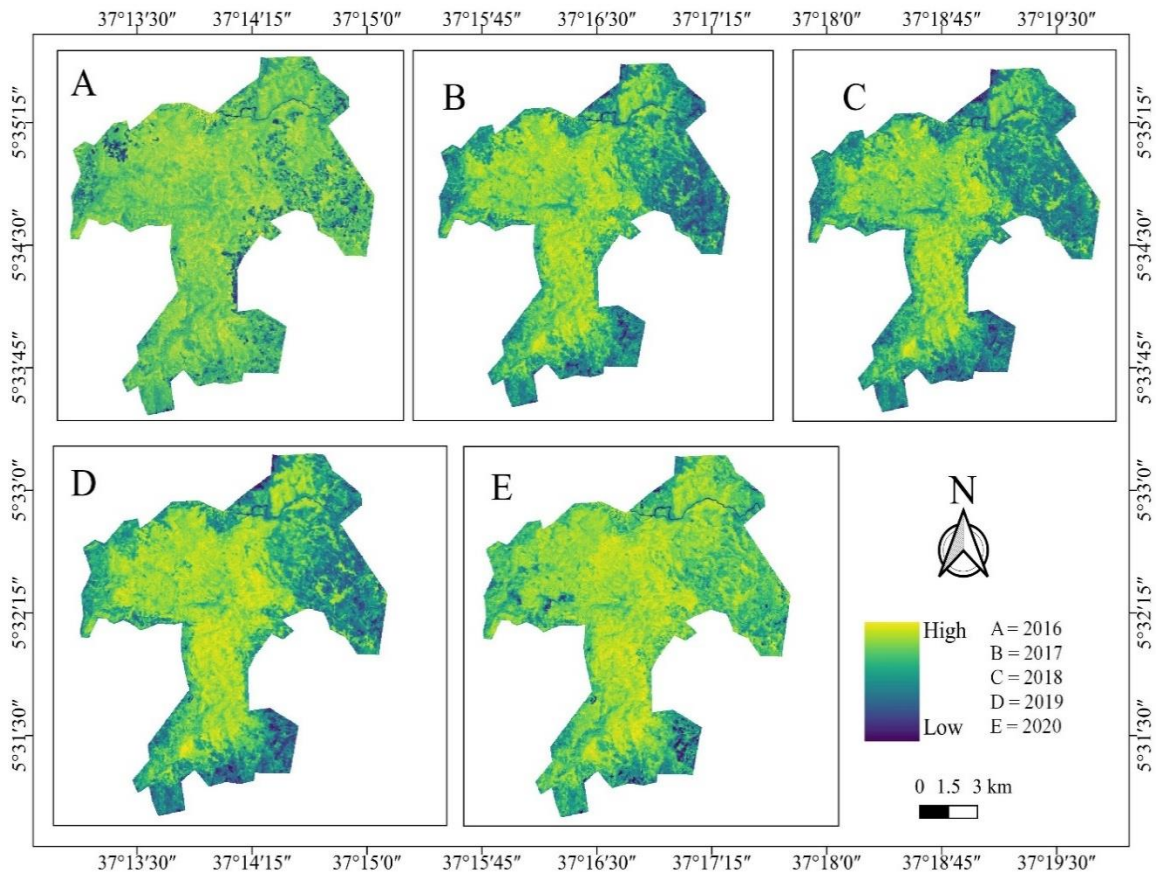


Figure 33: *Aboveground Carbon Maps*

The per-pixel AGC for the study area ranged from 0.538 – 1.6153, 0.353 -1.518, 0.319-1.480, 0.517 – 1.614 tons for the years 2016 to 2020 respectively. The total plot level per pixel AGC was found between 6 and 7.4 tons. Plot 17 generally showed lower predicted AGC across the study years than the other plots. It is clear from the map of AGC that the amount of AGC is lower in parts of the area which are close to road and other landuse types, especially near the agricultural lands. The inaccessible forest areas and dense forests showed the largest carbon stocks with the amount higher than 1.5 ton per pixel (176.25 t/ha).

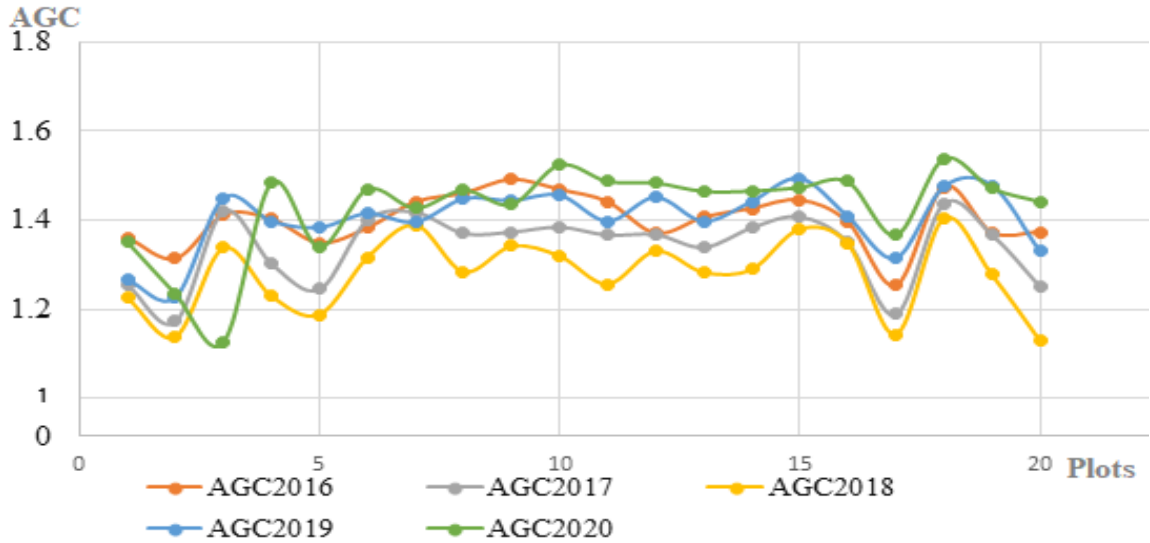
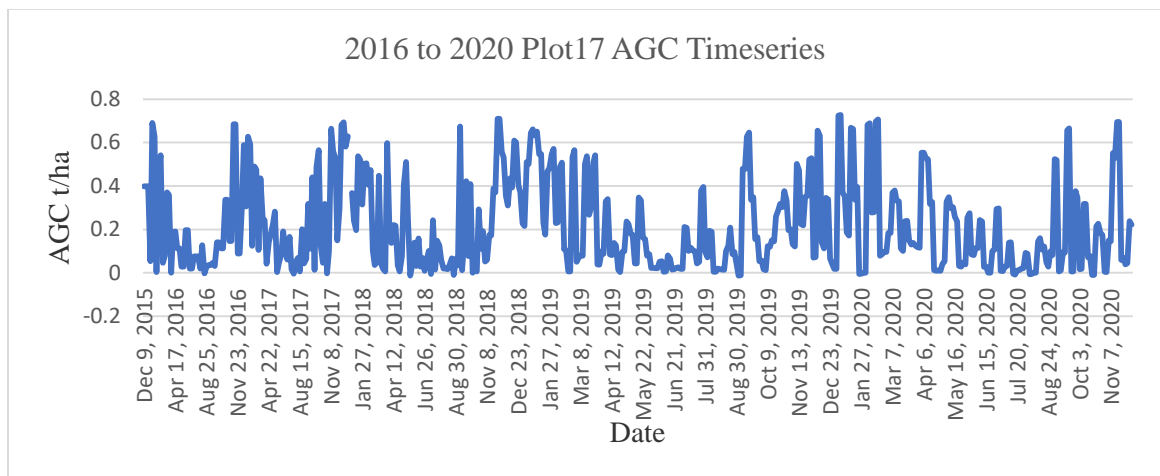


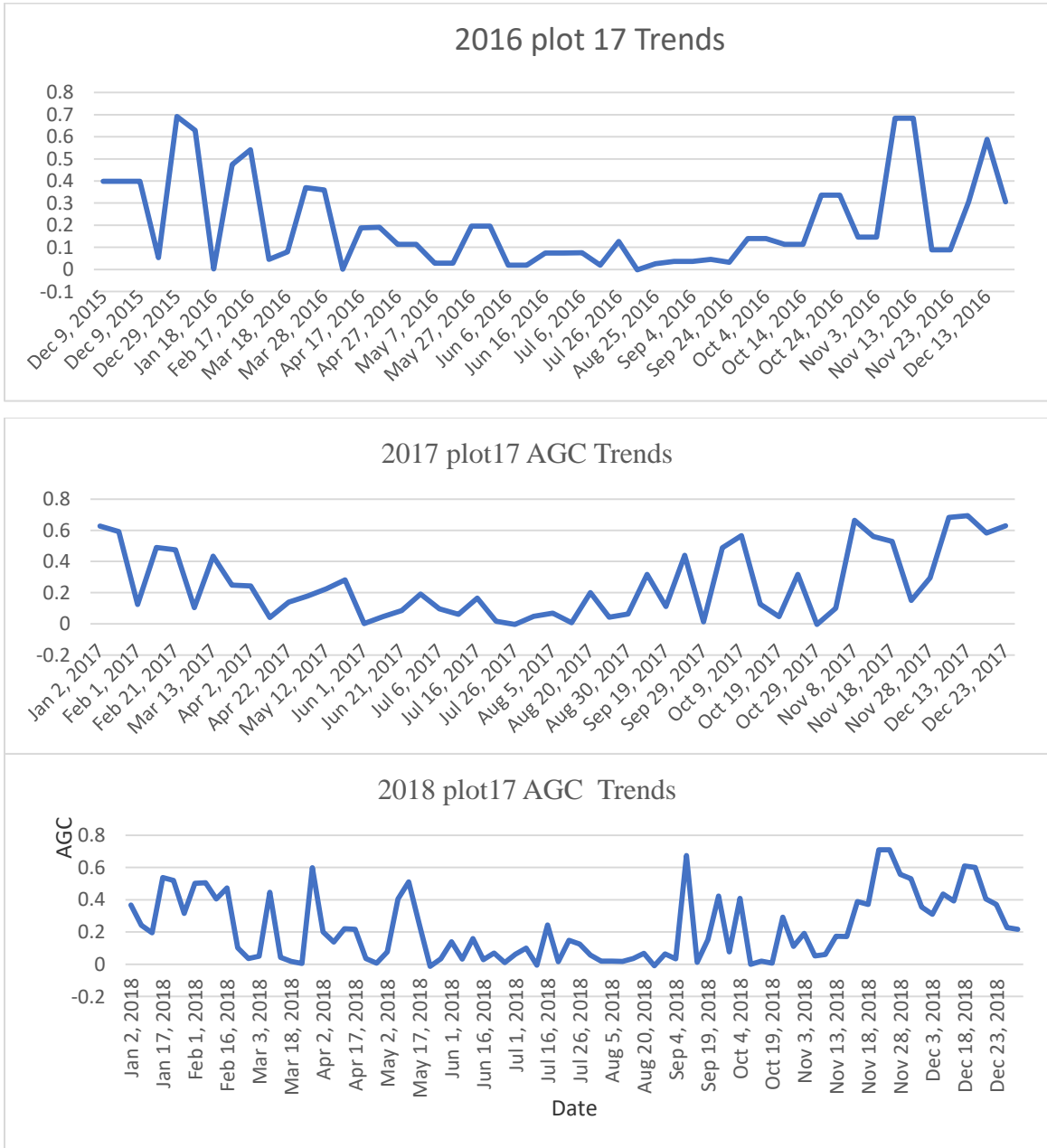
Figure 34: Plot level Predicted AGC

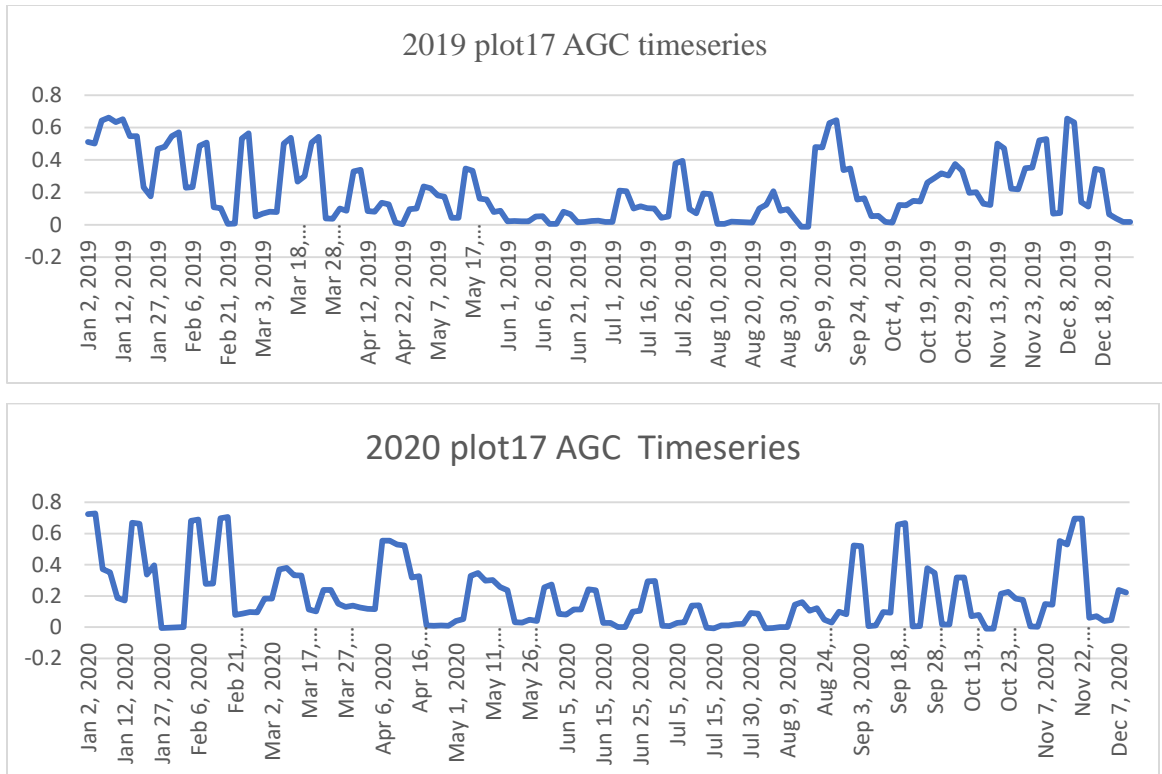
The estimation of the AGC map was based on a JavaScript developed for the predictive equation in GEE. The JavaScript allowed automating the annual timeseries AGC coverage mapping for the study area at a time less than a minute (see 8.2). The satellite based timeseries used in this study is a point-based timeseries. The plot centres were used as points for which the timeseries values for AGC is extracted from the maps produced based on the predictive equation. The total plot level per pixel AGC was 28.043, 26.804, 25.621, 28.076, and 28.551 for the years 2016, 2017, 2018, 2019 and 2020 respectively.

The timeseries AGC derived from Equation 10 across the study area at a single plot for instance is shown below.



The yearly satellite based point timeseries of AGB converted to AGC per plot can be separately generated using the prototype developed as shown below for plot17 for instance.





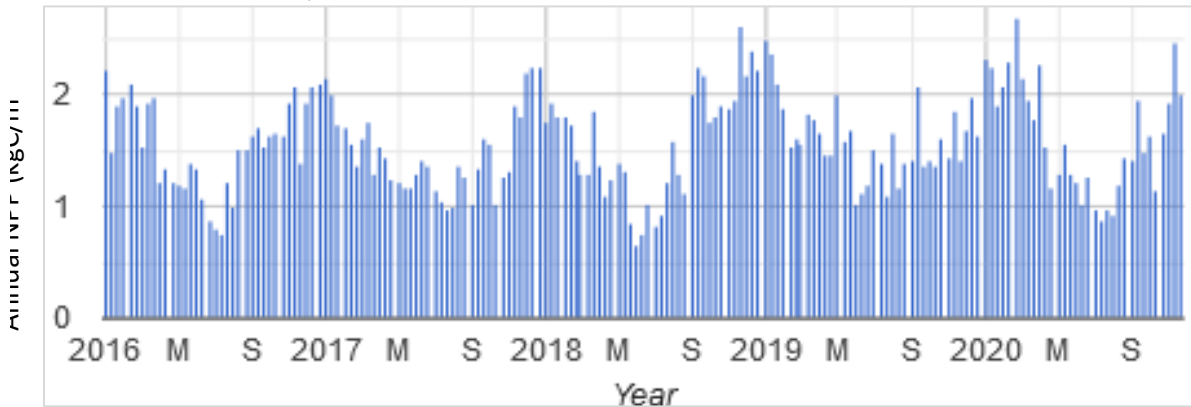
The above graphs depict the power of the prototype to derive timeseries above ground variability of the biomass carbon gC/m^2 per each plot across the study area. The aggregated annual above ground timeseries carbon stock was also derived using the FAO WAPOR data as Net primary production (NPP). NPP is a fundamental characteristic of an ecosystem, expressing the conversion of carbon dioxide into biomass driven by photosynthesis. The pixel value represents the mean daily NPP for that specific year from 2016 to 2020. This evaluates the performance of the forest areas.

```
var app = function (img)
{
var npp = img.select('L1_NPP_D').multiply(0.00047).clip(oytAOI);
return ee.Image(npp.copyProperties(img)).set('system:time_start',
};
```

The methodology for deriving the NPP was based on the framework of FAO that incorporates biome-specific light-use efficiencies (LUEs). The following data is used to calculate NPP: Daily: incoming solar radiation and temperature data (Tmin/Tmax); Dekadal: fAPAR and soil moisture stress; Seasonal: Land Cover. The NPP data deployed in GEE was used to compare the field-based carbon stock to globally available dataset. The resulting timeseries values are comparable. However, the global NPP values are

derived at low resolution of 250m and hence exaggerates the per pixel carbon stock values. The peak mean daily per pixel NPP in kgC/m^2 observed was:

- Jan 1, 2016 2.223
- Jan 1, 2017 2.224
- Nov 21, 2018 2.599
- Jan 1, 2019 2.477
- Feb 21, 2020 2.687



The peak NPP from 2016 to 2020 carbon stored in the forest per square meter showed a linear increment in carbon stock. This is feasibly due to increased management and awareness creation by the respective government authority in the society as reported by the community during reconnaissance survey and literature reviews.

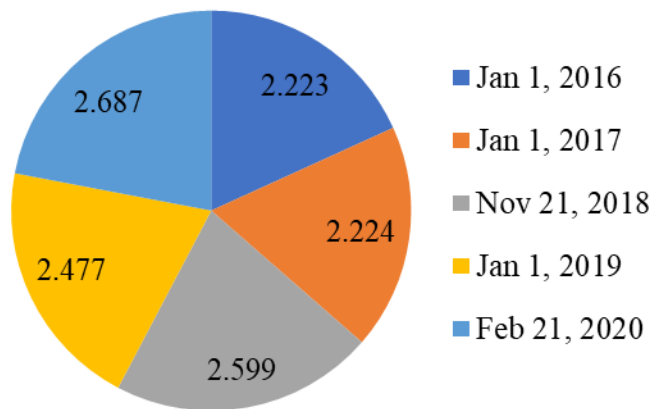


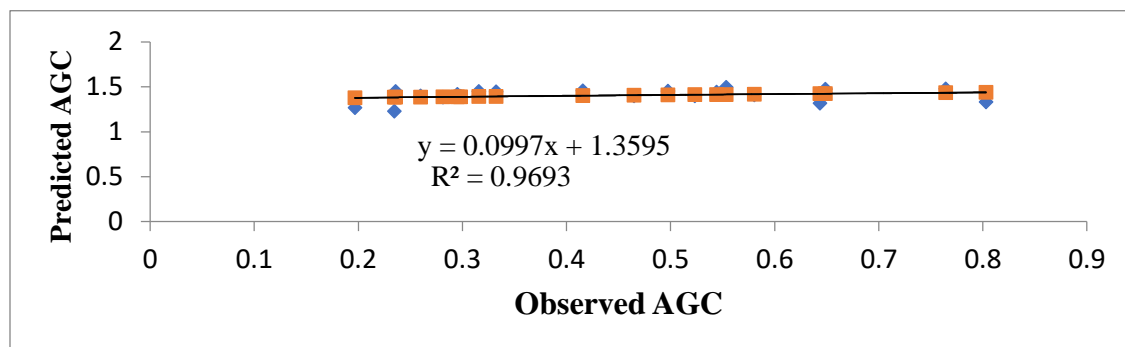
Figure 35: Peak NPP from 2016 to 2020 (kgC/m^2)

The net primary productivity as a function of the carbon stock due to the conversion of carbon dioxide into biomass driven by photosynthesis per pixel showed the highest peak in the year 2020 with 2. 687 kgC/m^2 .

The aboveground carbon stock variability over the study area from 2016 to 2017 generally showed higher separability than from 2018 to 2020. The geocomputation of the plot level point based timeseries can run through entering the plot center coordinate of the plot in GEE. The GEE prototype code developed plots the timeseries charts within 0.67seconds and postulates the seasonal trends in AGC. For sample visualization, plot17 timeseries of the predicted aboveground carbon stock is presented in the following figures. Thus, the prototype developed here in this research allowed to automatically monitor the state of AGC both at short-term (yearly basis) and at long-term (half a decade) that fits the REDD+ monitoring mechanism as adopted in the Ethiopian CRGE documents.

4.3.7 Correlation between Observed and predicted AGC

The prediction model was validated through observing the correlation between the observed and predicted AGC. An exponential trend was found with R^2 of 0.96.



The biomass estimates based on LULC classification with Sentinel images is still recent, but promising. For reliable information extraction, the current study used the fusion of radar (Sentinel 1) and optical data from (Sentinel 2), as well as incorporating models and algorithms, vegetation indices, textures, biophysical variables, and forest inventory data. The best variable selected was based on examining its correlation coefficient with field aboveground carbon values as used by (Chave et al., 2014; Dick OB, 2015; Tesfaye et al., 2016) and using the variable importance built-in module within the SmileRandomForest algorithm as used by (Apley & Zhu, 2019). The later method was applied on the fusion of images and the model performed the whole chain of processing within 0.6 to 5 seconds in GEE.

The predictive equation model was developed with SmileRandomForest algorithm. It can be noticed that most textures derived from $\sigma^{\circ}VV$ polarizations are correlated with each other. This finding is in line with a study by (Ghosh & Behera, 2018). Establishing a linear relationship using these features will not provide new information and could lead to over fitting the model. However, the capability of the RF algorithm to randomly select features assigned for each node takes care of this information redundancy.

Although the regression model performances vary between the sentinel satellite products and the observed AGC, the values are still comparable to results from studies dealing with the mapping of carbon in similar ecosystems. For example, González-Roglich and Swenson (2016) used spatially explicit predictions based on field samples ranging around a mean of 27 t/ha, had a mean prediction error of 9.6 t/ha at 60 m spatial resolution. Karlson et al. (2015) reported model performances with an r^2 of 0.57 when mapping aboveground biomass in Sudano-Sahelian woodlands using multi-temporal (7 observations) Landsat OLI products.

The performance of the model was also evaluated using net primary productivity to estimate the function and importance of the forests in mitigating climate change in the area. As a result, the NPP values linearly indicated the increased carbon storage in the forests in the year 2020 with the highest peak in 21st of February. This is indicator for good forest management in the area though it depends on climatic variables such as rainfall which would alter the value of carbon intake by forests.

Overall, the prediction model had exponential trend with r^2 of 0.96. This prediction power of the model was estimated through observing the correlation between the observed and predicted AGC. Given the lack of data on biomass, the current study provided valuable estimates of AGB and AGC storage and fills the data gap in an area under-represented by existing literature.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATION

5.1 Conclusion

This research is carried to estimate and model the timeseries forest aboveground carbon stock using GEE Java API. In this paper, a methodological prototype and predictive model for forest aboveground carbon stock monitoring using the fusion of Sentinel-2, ALOS DSM, and Sentinel-1 SAR data time series is introduced. To evaluate spectrally derived VIs and fusing Synthetic Aperture Radar (SAR) imagery for improving forest area masking and AGC monitoring, the freely available advanced Sentinel-1 radar (VV and VH backscatters) and Sentinel-2 optical imagery were utilized. This study contributes to the evaluation of economically viable, efficient prototype model development for scalable monitoring of AGC. The predictive equation developed for the AGC estimation can also contribute to the mapping of landscape-level forest aboveground biomass and carbon stock in the study area. Hence, the R^2 value of ~ 0.9388 indicates that the model accounts for about 94% of the variance in the dependent variable.

The results portrayed in the 'Results' section illustrate the general feasibility of the proposed approach for the study area where there exists recurrent cloud cover over time. In these cases, often already the least cloudy image provides a clean and artifact-free solution. The mean pixel value, which was used as the plot value of plot biomass carbon stock for the 20 plots prevented intrusion to the surrounding area where the tree canopy cover is different.

The conventional timeseries Sentinel-2 optical imagery was used as a benchmark for comparison with results achieved using Sentinel-2 VIs and fused use of Sentinel-2 VIs, ALOS DSM, Sentinel-1 SAR backscatters, and GLCM textures. The three scenarios as case studies showed the different accuracies that can be achieved based on different inputs and the relative importance of the different types of sensors in the generation of the AGB maps. The fusion of the satellite images and their derivatives increased overall classification accuracies when compared to traditional optical imagery results, while the

fusion of Sentinel-2 VIs and Sentinel-1 backscatters decreased the overall accuracies. Hence, this study demonstrated that fusion of multitemporal Sentinel satellite images have the potential to increase the monitoring and mapping of forest AGC stock.

The study provides that the use of automated prototyping and cloud-based infrastructures such as GEE overcomes the analysis of the current difficult and expensive remote sensing data with smaller MMUs. The use of such methods decreases the mapping efforts and usually increase mapping accuracy. The use of texture parameters is also indicated as prominent for forest cover and vegetation classification. Moreover, the use of C-band satellite backscatter signal based information with optical sentinel-2 bands and vegetation index values tackles the cloud cover problem.

Thus, the prototype model developed for processing the fusion of the Sentinel satellite, with optical and radar capabilities, holds great promise in the delineation of forests for enhanced timeseries monitoring of forest biomass and carbon stock. Hence the tasks of land cover classification and activity-data mapping for REDD+ and MRV missions can take advantage of fusion of the two data types leading generally to increased forest carbon stock mapping and estimation accuracy (Bagan, *et al*, 2012).

A study by Chen et al. (2018) indicated that texture characteristics of Sentinel-1 and the Sentinel-2 derivative VIs variables were the most relative and important predictors for explaining the observed variability of AGC. Although SAR C band and optical multispectral techniques have few advantages for detecting the sensibility of forest AGC compared to SAR P band or LiDAR, the high spatial resolution of sentinel satellite series with wider coverage is indeed useful information for applications in forest AGB estimation.

It is noticed that the outputs of this study have promising results for multi-temporal monitoring for long-term AGC mapping-based multispectral image analysis. The possibility to apply this approach to different types of sensors while generating complementary outputs is among its several advantages. These methods may prove important in supporting management and policies addressing carbon stocks in Ethiopia and elsewhere. Overall, the comparison assessment of this study provided a reference for the selection of combinations of predictors and algorithms for detailed forest AGB

estimation and modeling. This study therefore, drives that the detailed forest AGC estimation can be enhanced by the use of allometric equations and synergetic use of multispectral satellite timeseries data. Thus, it can be considered vital in alleviating climate change mitigation through supporting sustainable forest resource management.

5.2 Recommendation

- To take advantage of the massive open availability of satellite images, it is recommended to develop cloud based prototypes for monitoring the abrupt changes in forests that will help scientists and researchers to easily get from an idea to the actual prototyping and geocomputation towards mitigating climate change.
- The geospatial industry in Ethiopia is highly dependent on desktop software resources and existing geospatial data are locked within various sectors which leads to resource and time constraints and incompatible data. Therefore, the major stakeholder's have to turn their face towards effective cross-sectoral data sharing that will contribute to national MRV systems and forest policy amendments.
- Coding with JavaScript or Python APIs in GEE requires good bundle of knowledge on programming language which is not a primary objective of most of the Institutes offering GIS and Remote Sensing courses in Ethiopia and this needs to be addressed very soon integrating advanced programming languages in their curricula following "fit-for-purpose approach" that will meet the needs of society today and that can be incrementally improved over time.
- Though the result of the analysis for is crucial for making decisions in carbon stock modeling, intensive socio-economic data and increased number of plots need to increase its efficiency and accuracy.
- There is an urgent need to build cost-effective and sustainable MRV systems for flexibly designing a monitoring plan that can adapt remote sensing data to different forest definitions that holds consistent forest classifications for all REDD+ activities for critically integrating different types of information.

- The EFCCC and the regional government should promote carbon trading for additional financial incentive to the local community who are depending on the forest resource based on accurate spatial estimation of their forest carbon stocks.

It has been indicated that the forest sector in Ethiopia is the second largest contributor of GHG emissions in the country after agriculture (CRGE, 2011). Under the BAU scenario, the pressure on forests however will ultimately deplete the natural resource base and negatively affect the performance of the economy in the long-term. The use of complex GIS, satellite imagery interpretation software and the treatment of large amount of data in this regard, can only be effective if coupled with new algorithms and cloud-based optimizations. Therefore, other researchers can use the findings of this study as a reference in order to improve and dig out methods to extract information from multisource satellite images perspective. Hence, below are some of the focus areas for future study.

- The present study, did not attempt to model and estimate the carbon gain and loss as a function of each carbon pool nor each LULC, so further study is required to map the carbon stock potential of the study area
- It is recommended to further examine sentinel-1 GLCM features in terms of kernel size and shift, and analyze how their explanatory power changes.
- Study on various challenges, including accurate co-registration and dealing with spectral variation in the time series have to be addressed when designing a SAR-optical time series approach for forest carbon stock monitoring
- Further investigation on effects of timeseries multisensor image fusion model on the estimation of the forest aboveground carbon due to deforestation and degradation activities.
- A study needs to be conducted that examines the estimation of emission factors at landuse-landcover level including all carbon pools.
- The model developed here might be modified in the future by using radar longer wavelength bands such as L and P bands which have strong relationship to forest stand parameter.

REFERENCES

- Abere, F., Belete, Y., Kefalew, A., & Soromessa, T. (2017). Carbon stock of Banja forest in Banja district, Amhara region, Ethiopia: An implication for climate change mitigation. *Journal of Sustainable Forestry*, 36(6), 604–622. <https://doi.org/10.1080/10549811.2017.1332646>
- Adnan, N. A., Razi, N. M., Yusoff, Z. M., & Latif, Z. A. (2017). Spatial Temporal Estimation and Analysis of Carbon Sequestration. *Advanced Science Letters*, 23(7), 6157–6162. <https://doi.org/10.1166/asl.2017.9227>
- Adnan, N. A., Yusoff, Z. M., & Latif, Z. A. (2017). *Geospatial Temporal Estimation and Analysis of Carbon Sequestration in the Taman Botani Negara Shah Alam*. July. <https://doi.org/10.1166/asl.2017.9227>
- Amani, M., Brisco, B., Afshar, M., Mirmazloumi, S. M., Mahdavi, S., Mirzadeh, S. M. J., Huang, W., & Granger, J. (2019). A generalized supervised classification scheme to produce provincial wetland inventory maps: An application of Google Earth Engine for big geo data processing. *Big Earth Data*, 3(4), 378–394. <https://doi.org/10.1080/20964471.2019.1690404>
- Amani, M., Ghorbanian, A., Ahmadi, S. A., Kakooei, M., Moghimi, A., Mirmazloumi, S. M., Moghaddam, S. H. A., Mahdavi, S., Ghahremanloo, M., Parsian, S., Wu, Q., & Brisco, B. (2020). Google Earth Engine Cloud Computing Platform for Remote Sensing Big Data Applications: A Comprehensive Review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 5326–5350. <https://doi.org/10.1109/JSTARS.2020.3021052>
- Apley, D. W., & Zhu, J. (2019). Visualizing the Effects of Predictor Variables in Black Box Supervised Learning Models. *ArXiv:1612.08468 [Stat]*. <http://arxiv.org/abs/1612.08468>
- Argamosa, R. J. L., Blanco, A. C., Baloloy, A. B., Candido, C. G., Dumalag, J. B. L. C., Dimapilis, L. L. C., & Paringit, E. C. (2018). Modelling Above Ground Biomass Of Mangrove Forest Using Sentinel-1 Imagery. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV–3, 13–20. <https://doi.org/10.5194/isprs-annals-IV-3-13-2018>
- Barredo, J. I., Teller, A., Bastrup-Birk, A., Onaindia, M., Fernández de Manuel, B., Madariaga, I., Rodríguez-Loinaz, G., Pinho, P., Nunes, A., Ramos, A., Batista, M., Mimo, S., Cordovil, C., Branquinho, C., Grêt-Regamey, A., Bebi, P., Brunner, S. H., Weibel, B., Kopperoinen, L., ... Joint Research Centre. (2016).

- Mapping and assessment of forest ecosystems and their services: Applications and guidance for decision making in the framework of MAES*. Publications Office.
<http://bookshop.europa.eu/uri?target=EUB:NOTICE:LBNA27751:EN:HTML>
- Bekele-Tesemma, (2007). *Useful trees and shrubs for Ethiopia*. PlantUse English. (n.d.). Retrieved November 19, 2020, from <https://uses.plantnet-project.org/en/Bekele-Tesemma>, Useful trees and shrubs for Ethiopia, 2007
- Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
- Betru, T., Tolera, M., Sahle, K., & Kassa, H. (2019a). Trends and drivers of land use / land cover change in Western Ethiopia. *Applied Geography*, 104(August 2018), 83–93. <https://doi.org/10.1016/j.apgeog.2019.02.007>
- Betru, T., Tolera, M., Sahle, K., & Kassa, H. (2019b). Trends and drivers of land use/land cover change in Western Ethiopia. *Applied Geography*, 104, 83–93. <https://doi.org/10.1016/j.apgeog.2019.02.007>
- Boisvenue, C., & White, J. C. (2019). Information Needs of Next-Generation Forest Carbon Models: Opportunities for Remote Sensing Science. *Remote Sensing*, 11(4), 463. <https://doi.org/10.3390/rs11040463>
- Brack, D. (2019). *Forests and Climate Change Duncan Brack i the fourteenth session of the United Nations Forum on Forests. c*.
- Braun, A., & Hochschild, V. (2015). *Combined use of SAR and optical data for environmental assessments around refugee camps in semiarid landscapes*. 6.
- Buchhorn, M., Lesiv, M., Tsendbazar, N.-E., Herold, M., Bertels, L., & Smets, B. (2020). Copernicus Global Land Cover Layers—Collection 2. *Remote Sensing*, 12Volume 108(6), 1044. <https://doi.org/10.3390/rs12061044>
- Buchhorn, M., Smets, B., Bertels, L., De Roo, B., Lesiv, M., Tsendbazar, N.-E., Herold, M., & Fritz, S. (2020). *Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2019: Globe*. <https://doi.org/10.5281/zenodo.3939050>
- Canty, M. J., Nielsen, A. A., Conradsen, K., & Skriver, H. (2019). Statistical Analysis of Changes in Sentinel-1 Time Series on the Google Earth Engine. *Remote Sensing*, 12(1), 46. <https://doi.org/10.3390/rs12010046>
- Carrasco, L., O’Neil, A., Morton, R., & Rowland, C. (2019). Evaluating Combinations of Temporally Aggregated Sentinel-1, Sentinel-2 and Landsat 8 for Land Cover Mapping with Google Earth Engine. *Remote Sensing*, 11(3), 288. <https://doi.org/10.3390/rs11030288>
- Chave, J., Andalo, C., Brown, S., Cairns, M. A., Chambers, J. Q., Eamus, D., Fölster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J.-P., Nelson, B. W., Ogawa, H., Puig, H., Riéra, B., & Yamakura, T. (2005). Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia*, 145(1), 87–99. <https://doi.org/10.1007/s00442-005-0100-x>
- Chave, J., Réjou-Méchain, M., Burquez, A., Chidumayo, E., Colgan, M., Delitti, W., Duque, A., Eid, T., Fearnside, P., Goodman, R., Henry, M., Martinez-Yrizar, A., Mugasha, W., Muller-Landau, H., Mencuccini, M., Nelson, B., Ngomanda, A., Nogueira, E., Ortiz, E., & Vieilledent, G. (2014). Improved allometric models to

- estimate the aboveground biomass of tropical trees. *Global Change Biology*, 20, 3177–3190. <https://doi.org/10.1111/gcb.12629>
- Chen, L., Ren, C., Zhang, B., Wang, Z., & Xi, Y. (2018). Estimation of Forest Above-Ground Biomass by Geographically Weighted Regression and Machine Learning with Sentinel Imagery. *Forests*, 9(10), 582. <https://doi.org/10.3390/f9100582>
- Christiansen, M. P., Teimouri, N., Laursen, M. S., Mikkelsen, B. F., Jorgensen, R. N., & Sorensen, C. A. G. (2019). Preprocessed Sentinel-1 Data via a Web Service Focused on Agricultural Field Monitoring. *IEEE Access*, 7, 65139–65149. <https://doi.org/10.1109/ACCESS.2019.2917063>
- de Sousa, C., Fatoyinbo, L., Neigh, C., Boucka, F., Angoue, V., & Larsen, T. (2020). Cloud-computing and machine learning in support of country-level land cover and ecosystem extent mapping in Liberia and Gabon. *PLOS ONE*, 15(1), e0227438. <https://doi.org/10.1371/journal.pone.0227438>
- Dick OB, W. N. (2015). Estimating above Ground Biomass and Carbon Stock in the Lake Hawassa Watershed, Ethiopia by Integrating Remote Sensing and Allometric Equations. *Forest Research: Open Access*, 04(03). <https://doi.org/10.4172/2168-9776.1000151>
- Duncanson, L., Armston, J., Disney, M., Avitabile, V., Barbier, N., Calders, K., Carter, S., Chave, J., Herold, M., Crowther, T. W., Falkowski, M., Kellner, J. R., Labrière, N., Lucas, R., MacBean, N., McRoberts, R. E., Meyer, V., Næsset, E., Nickeson, J. E., ... Williams, M. (2019). The Importance of Consistent Global Forest Aboveground Biomass Product Validation. *Surveys in Geophysics*, 40(4), 979–999. <https://doi.org/10.1007/s10712-019-09538-8>
- Fentahun Abere, Yehualashet Belete, Alemayehu Kefalew & Teshome Soromessa (2017) Carbon stock of Banja forest in Banja district, Amhara region, Ethiopia: An implication for climate change mitigation, *Journal of Sustainable Forestry*, 36:6, 604-622, DOI: 10.1080/10549811.2017.1332646
- Filipponi, F. (2019). Sentinel-1 GRD Preprocessing Workflow. *Proceedings*, 18(1), 11. <https://doi.org/10.3390/ECRS-3-06201>
- Fortin, J. A., Cardille, J. A., & Perez, E. (2020). Multi-sensor detection of forest-cover change across 45 years in Mato Grosso, Brazil. *Remote Sensing of Environment*, 238, 111266. <https://doi.org/10.1016/j.rse.2019.111266>
- GFOI. (2013). Carbon Pools—GFOI. In *Integrating remote-sensing and ground-based observations for estimation of emissions and removals of greenhouse gases in forests: Methods and Guidance from the Global Forest Observations Initiative*. Global Forest Observations Initiative. https://www.reddcompass.org/carbon-pools?uri=s2_2_2.html&ver=v2
- Ghorbanian, A., Kakooei, M., Amani, M., Mahdavi, S., Mohammadzadeh, A., & Hasanlou, M. (2020). Improved land cover map of Iran using Sentinel imagery within Google Earth Engine and a novel automatic workflow for land cover classification using migrated training samples. *ISPRS Journal of Photogrammetry and Remote Sensing*, 167, 276–288. <https://doi.org/10.1016/j.isprsjprs.2020.07.013>
- Ghosh, S. M., & Behera, M. D. (2018). Aboveground biomass estimation using multi-sensor data synergy and machine learning algorithms in a dense tropical forest. *Applied Geography*, 96, 29–40. <https://doi.org/10.1016/j.apgeog.2018.05.011>

- Grinand, C., Rakotomalala, F., Gond, V., Vaudry, R., Bernoux, M., & Vieilledent, G. (2013). Estimating deforestation in tropical humid and dry forests in Madagascar from 2000 to 2010 using multi-date Landsat satellite images and the random forests classifier. *Remote Sensing of Environment*, *139*, 68–80. <https://doi.org/10.1016/j.rse.2013.07.008>
- Huang, X., Ziniti, B., Torbick, N., & Ducey, M. J. (2018). *Assessment of Forest above Ground Biomass Estimation Using Multi-Temporal C-band Sentinel-1 and Polarimetric L-band PALSAR-2 Data*. 15.
- Isbaex, C., & Coelho, A. M. (2021). The Potential of Sentinel-2 Satellite Images for Land-Cover/Land-Use and Forest Biomass Estimation: A Review. In *Forest Biomass—From Trees to Energy*. IntechOpen. <https://doi.org/10.5772/intechopen.93363>
- Khudinyan, M. (2019). *The Use Of Remotely Sensed Data For Forest Biomass Monitoring. A case of forest sites in North-Eastern Armenia*. Universitat Jaume I.
- Lu, X., Zheng, G., Miller, C., & Alvarado, E. (2017). Combining Multi-Source Remotely Sensed Data and a Process-Based Model for Forest Aboveground Biomass Updating. *Sensors*, *17*(9), 2062. <https://doi.org/10.3390/s17092062>
- Ma, L., Hurtt, G., Tang, H., Lamb, R., Campbell, E., Dubayah, R., Guy, M., Huang, W., Lister, A., Lu, J., O’Neil-Dunne, J., Rudee, A., Shen, Q., & Silva, C. (2021). High-resolution forest carbon modelling for climate mitigation planning over the RGGI region, USA. *Environmental Research Letters*, *16*(4), 045014. <https://doi.org/10.1088/1748-9326/abe4f4>
- Mahdianpari, M., Salehi, B., Mohammadimanesh, F., Homayouni, S., & Gill, E. (2018). 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), 43. <https://doi.org/10.3390/rs11010043>
- Martín, J. M. (n.d.). *Mapping intra- and inter-annual dynamics in wetlands*. 137.
- MEFCC, Ashebir Wondimu Zeleke, Tefera Mengistu Woldie, Florence Landsberg, & Bizuayehu Alemu Yimer. (2018). *Tree-Based Landscape Restoration Atlas of Ethiopia (version 0.0)—Introduction*. <https://eth.restorationatlas.org/>
- MEFCCC. (2017). *Ethiopia’s Forest Reference Level Submission to the UNFCCC* (Issue March).
- Mitchell, A. L., Rosenqvist, A., & Mora, B. (2017). Current remote sensing approaches to monitoring forest degradation in support of countries measurement, reporting and verification (MRV) systems for REDD+. *Carbon Balance and Management*, *12*(1), 9. <https://doi.org/10.1186/s13021-017-0078-9>
- Nations, U. (2016). *Conference of the Parties Report of the Conference of the Parties on its twenty-first session , held in Paris from 30 November to 13 December 2015 Addendum Part two: Action taken by the Conference of the Parties at its twenty-first session Contents Decis. 01194*(January), 1–36.
- Nunes, Meireles, Pinto Gomes, & Almeida Ribeiro. (2020). Forest Contribution to Climate Change Mitigation: Management Oriented to Carbon Capture and Storage. *Climate*, *8*(2), 21. <https://doi.org/10.3390/cli8020021>

- 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), 2291. <https://doi.org/10.3390/rs12142291>
- Plugge, D., Kübler, D., Neupane, P. R., Olschofsky, K., & Prill, L. (2016). Measurement, Reporting, and Verifications Systems in Forest Assessment. *Tropical Forestry Handbook*, 839–882. https://doi.org/10.1007/978-3-642-54601-3_73
- Quang, T., & Hoa, L. S. (2018). *Using Sentinel Satellite Image to Estimate Biomass of Mangrove Forest in Vinh Quang Commune, Tien Lang District, Hai Phong City*. 5, 9.
- Reiche, J. (2015). *Combining SAR and optical satellite image time series for tropical forest monitoring*. Wageningen University.
- Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T. A., Salas, W., Zutta, B. R., Buermann, W., Lewis, S. L., Hagen, S., Petrova, S., White, L., Silman, M., & Morel, A. (2011). Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences*, 108(24), 9899–9904. <https://doi.org/10.1073/pnas.1019576108>
- Santi, E., Paloscia, S., Pettinato, S., Fontanelli, G., Mura, M., Zolli, C., Maselli, F., Chiesi, M., Bottai, L., & Chirici, G. (2017). The potential of multifrequency SAR images for estimating forest biomass in Mediterranean areas. *Remote Sensing of Environment*, 12.
- Schwieder, M., Leitão, P., Teixeira, A., Pedroni, F., Sanchez, M., Bustamante, M., & Hostert, P. (2018). Landsat phenological metrics and their relation to aboveground carbon in the Brazilian Savanna. *Carbon Balance and Management*, 13. <https://doi.org/10.1186/s13021-018-0097-1>
- Su, H., Shen, W., Wang, J., Ali, A., & Li, M. (2020). *Machine learning and geostatistical approaches for estimating aboveground biomass in Chinese subtropical forests* [Preprint]. In Review. <https://doi.org/10.21203/rs.3.rs-25148/v3>
- Tavares, P., Beltrão, N., Guimarães, U., & Teodoro, A. (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), 1140. <https://doi.org/10.3390/s19051140>
- Tesfaye, M. A., Bravo-Oviedo, A., Bravo, F., & Ruiz-Peinado, R. (2016). Aboveground biomass equations for sustainable production of fuelwood in a native dry tropical afro-montane forest of Ethiopia. *Annals of Forest Science*, 73(2), 411–423. <https://doi.org/10.1007/s13595-015-0533-2>
- Thapa, R. B., Motohka, T., Watanabe, M., & Shimada, M. (2015). Time-series maps of aboveground carbon stocks in the forests of central Sumatra. *Carbon Balance and Management*, 10(1). <https://doi.org/10.1186/s13021-015-0034-5>
- The Carbon Cycle*. (2011, June 16). NASA Earth Observatory. <https://earthobservatory.nasa.gov/features/CarbonCycle>
- Udali, A. (2019). *Assessing the accuracy for area-based tree species classification using Sentinel-1 C- band SAR data*.
- Vickers, B., Trines, E., Pohnan, E., & FAO Regional Office for Asia and the Pacific. (2012). Community guidelines for accessing forestry voluntary carbon markets. Food and Agriculture Organization of the United Nations Regional Office for Asia and the Pacific. <http://books.google.com/books?id=jmGTw3HLXTMC>

Zhao, P., Lu, D., Wang, G., Wu, C., Huang, Y., & Yu, S. (2016). Examining Spectral Reflectance Saturation in Landsat Imagery and Corresponding Solutions to Improve Forest Aboveground Biomass Estimation. *Remote Sensing*, 8(6), 469. <https://doi.org/10.3390/rs8060469>

APPENDICES

ANNEX 1 Field Photos

Annex 1.1 Field Equipment



Annex 1.2 Field Crews



Annex 1.3 Field site Photos







Annex 1.4 Forest Resources Management and use related Selected photos





Annex 1.5 Tree species names cleansing



ANNEX 2 Field Based AGB and AGC Stock Calculations

Table 19: Field Survey and Plot (Tree) Level AGB Estimations

Plot 01												
Tcode	Family	Genus	Species	dbh (cm)	POM	Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB (kg)	AGB (kg/ha)	
7T1	Myrsinaceae	Maesa	lanceolata	37.6	1.3	13	1300	0.676	59706.06	59.71	298.53	
7T2	Fabaceae	Millettia	ferruginea	23.8	1.3	8	800	0.738	16585.64	16.59	82.93	
7T3	Poaceae	Arundinaria	alpina	5.6	1.3	7	700	0.63	740.37	0.74	3.70	
7T4	Poaceae	Arundinaria	alpina	6	1.3	5	500	0.63	609.98	0.61	3.05	
7T5	Poaceae	Arundinaria	alpina	9	1.3	8	800	0.63	2129.46	2.13	10.65	
7T6	Poaceae	Arundinaria	alpina	10	1.3	6	600	0.63	1975.36	1.98	9.88	
7T7	Poaceae	Arundinaria	alpina	6	1.3	6	600	0.63	728.78	0.73	3.64	
7T8	Poaceae	Arundinaria	alpina	8	1.3	6	600	0.63	1277.85	1.28	6.39	
									AGB=	83753.5	83.75	418.77

Plot 02												
Tcode	Family	Genus	Species	dbh (cm)	POM	Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB (kg)	AGB (kg/ha)	
10T1	Anacardiaceae	Rhus	ruspolii	17	1.3	7	700	0.62	6368.56	6.37	31.84	
10T2	Euphorbiaceae	Croton	macrostachyus	15	1.3	8	800	0.518	4768.12	4.77	23.84	
10T3	Euphorbiaceae	Croton	macrostachyus	18	1.3	11	1100	0.518	9287.36	9.29	46.44	
10T4	Fabaceae	Albizia	schimperiana	20	1.3	7.2	720	0.534	7770.74	7.77	38.85	
10T5	Fabaceae	Albizia	schimperiana	14	1.3	6.8	680	0.534	3663.24	3.66	18.32	
10T6	Fabaceae	Albizia	schimperiana	28	1.3	5.8	580	0.534	12135.38	12.14	60.68	
10T7	Rosaceae	Hagenia	abyssinica	24	1.3	8	800	0.5905	13561.69	13.56	67.81	
10T8	Rosaceae	Hagenia	abyssinica	9	1.3	9	900	0.5905	2242.58	2.24	11.21	
10T9	Rosaceae	Hagenia	abyssinica	30	1.3	11.5	1150	0.5905	29874.97	29.87	149.37	
10T10	Fabaceae	Albizia	schimperiana	20	1.3	9.5	950	0.534	10185.07	10.19	50.93	
									AGB=	99857.72	99.86	499.29

Plot 03												
Tcode	Family	Genus	Species	dbh (cm)	POM	Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB (kg)	AGB (kg/ha)	
8T1	Flacourtiaceae	Oncoba	spinosa	34	1.3	10	1000	0.647	36383.50	36.38	181.92	
8T2	Euphorbiaceae	Croton	macrostachyus	15.8	1.3	8	800	0.518	5277.10	5.28	26.39	
8T3	Fabaceae	Albizia	schimperiana	11.5	1.3	6	600	0.53	2192.13	2.19	10.96	
8T4	Flacourtiaceae	Oncoba	spinosa	10.5	1.3	8	800	0.647	2952.81	2.95	14.76	
8T5	Melanthaceae	Bersama	abyssinica	28	1.3	8	800	0.671	20757.00	20.76	103.78	
8T6	Euphorbiaceae	Croton	macrostachyus	17.7	1.3	9	900	0.518	7389.00	7.39	36.94	
8T7	Fabaceae	Albizia	schimperiana	20	1.3	12	1200	0.53	12699.88	12.70	63.50	
8T8	Euphorbiaceae	Croton	macrostachyus	20.3	1.3	8.5	850	0.518	9131.58	9.13	45.66	
8T9	Flacourtiaceae	Oncoba	spinosa	11.6	1.3	8	800	0.647	3586.70	3.59	17.93	
									AGB=	100369.7	100.37	501.85

Plot 04												
Tcode	Family	Genus	Species	dbh (cm)	POM	Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB (kg)	AGB (kg/ha)	
16T1	Rutaceae	Clausena	anisata	13.4	1.3	8	800	0.482	3566.10	3.57	17.83	
16T2	Moraceae	Ficus	exasperata	54.9	1.8	14.6	1460	0.377	79173.13	79.17	395.87	
16T3	Meliaceae	Trichilia	dregeana	10.3	1.3	6	600	0.482	1611.40	1.61	8.06	
16T4	Boraginaceae	Ehretia	cymosa	16	1.3	6	600	0.56	4407.22	4.41	22.04	
16T5	Meliaceae	Trichilia	dregeana	17	1.3	8	800	0.482	5674.42	5.67	28.37	
16T6	Meliaceae	Trichilia	dregeana	20	1.3	7	700	0.482	6840.62	6.84	34.20	
16T7	Meliaceae	Trichilia	dregeana	18.7	1.3	11	1100	0.482	9326.16	9.33	46.63	
									AGB=	110599.0	110.60	553.00

Plot 05											
Tcode	Family	Genus	Species	dbh (cm)	POM Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB (kg)	AGB (kg/ha)	
13T1	Poaceae	Arundinaria	alpina	5.5	1.3	8	800	0.63	814.28	0.81	4.07
13T2	Poaceae	Arundinaria	alpina	6	1.3	5	500	0.63	609.98	0.61	3.05
13T3	Poaceae	Arundinaria	alpina	6.5	1.3	5	500	0.63	713.14	0.71	3.57
13T4	Poaceae	Arundinaria	alpina	7	1.3	6.6	660	0.63	1080.63	1.08	5.40
13T5	Poaceae	Arundinaria	alpina	7.5	1.3	8.2	820	0.63	1528.18	1.53	7.64
13T6	Poaceae	Arundinaria	alpina	8	1.3	10	1000	0.63	2103.79	2.10	10.52
13T7	Boraginaceae	Ehretia	cymosa	17.4	1.3	10	1000	0.56	8546.70	8.55	42.73
13T8	Fabaceae	Albizia	grandbracteata	50	1.3	16.5	1650	0.534	104411.69	104.41	522.06
								AGB=	119808.39	119.81	599.04

Plot 06											
Tcode	Family	Genus	Species	dbh (cm)	POM Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB (kg)	AGB (kg/ha)	
19T1	Rutaceae	Vepris	dainellii	23	1.3	10.5	1050	0.7	19213.40	19.21	96.07
19T2	Asteraceae	Vernonia	amygdalina	14	1.3	9	900	0.413	3747.70	3.75	18.74
19T3	Asteraceae	Vernonia	amygdalina	25	1.3	9	900	0.413	11622.56	11.62	58.11
19T4	Euphorbiaceae	Sapium	ellipticum	43	1.3	13	1300	0.576	66363.08	66.36	331.82
19T5	Poaceae	Arundinaria	alpina	6	1.3	9	900	0.63	1082.59	1.08	5.41
19T6	Euphorbiaceae	Sapium	ellipticum	16.2	1.3	7	700	0.576	5394.81	5.39	26.97
19T7	Poaceae	Arundinaria	alpina	7	1.3	6	600	0.63	984.64	0.98	4.92
19T8	Myrtaceae	Syzygium	guineense	24	1.3	8.5	850	0.712	17271.13	17.27	86.36
								AGB=	125679.92	125.68	628.40

Plot 07											
Tcode	Family	Genus	Species	dbh (cm)	POM Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB(kg)	AGB(kg/ha)	
3T1	Fabaceae	Albizia	schimperiana	33	1.3	16	1600	0.53	44696.00	44.70	223.48
3T2	Rutaceae	Vepris	dainellii	16.6	1.3	11	1100	0.7	10638.51	10.64	53.19
3T3	Rosaceae	Hagenia	abyssinica	20	1.3	12	1200	0.5905	14112.93	14.11	70.56
3T4	Fabaceae	Milletia	ferruginea	11	1.3	6	600	0.738	2776.59	2.78	13.88
3T5	Fabaceae	Albizia	grandibracteata	13	1.3	8	800	0.534	3714.74	3.71	18.57
3T6	Rubiaceae	Rothmannia	urcelliformis	13	1.3	10	1000	0.642	5528.24	5.53	27.64
3T7	Rubiaceae	Rothmannia	urcelliformis	9	1.3	5	500	0.642	1371.02	1.37	6.86
3T8	Fabaceae	Milletia	ferruginea	32.1	1.3	12	1200	0.738	44179.09	44.18	220.90
								AGB=	127017.13	127.02	635.09

Plot 08											
Tcode	Family	Genus	Species	dbh (cm)	POM Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB(kg)	AGB (kg/ha)	
4T1	Euphorbiaceae	Sapium	ellipticum	24.7	1.3	10	1000	0.576	17407.32	17.41	87.04
4T2	Fabaceae	Milletia	ferruginea	29.9	1.3	12	1200	0.738	38461.76	38.46	192.31
4T3	Fabaceae	Milletia	ferruginea	29.9	1.3	11	1100	0.738	35330.31	35.33	176.65
4T4	Euphorbiaceae	Sapium	ellipticum	16.9	1.3	8	800	0.576	6674.80	6.67	33.37
4T5	Euphorbiaceae	Sapium	ellipticum	18.8	1.3	5	500	0.576	5194.43	5.19	25.97
4T6	Fabaceae	Milletia	ferruginea	11	1.3	6	600	0.738	2776.59	2.78	13.88
4T7	Boraginaceae	Ehretia	cymosa	14	1.3	8	800	0.56	4496.81	4.50	22.48
4T8	Fabaceae	Milletia	ferruginea	31	1.3	10	1000	0.738	34544.28	34.54	172.72
4T9	Fabaceae	Milletia	ferruginea	24.6	1.3	9	900	0.738	19846.51	19.85	99.23
4T10	Fabaceae	Milletia	ferruginea	28	1.3	9	900	0.738	25552.37	25.55	127.76
								AGB=	134416.1	190.29	672.08

Plot 09												
Tcode	Family	Genus	Species	dbh (cm)	POM	Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB(kg)	AGB (kg/ha)	
6T1	Araliaceae	Polyscias	fulva	11.4	1.3	6	600	0.44	1797.13	1.80	8.99	
6T2	Euphorbiaceae	Sapium	ellipticum	34	2.9	12	1200	0.576	38807.45	38.81	194.04	
6T3	Rubiaceae	Rothmannia	urcelliformis	13.5	1.3	7	700	0.642	4201.42	4.20	21.01	
6T4	Euphorbiaceae	Croton	macrostachyus	29	1.3	14	1400	0.518	29814.17	29.81	149.07	
6T5	Myrsinaceae	Maesa	lanceolata	20	1.3	8	800	0.676	10841.00	10.84	54.21	
6T6	Myrsinaceae	Maesa	lanceolata	18.4	1.3	8	800	0.676	9212.62	9.21	46.06	
6T7	Boraginaceae	Cordia	africana	18.9	1.3	10	1000	0.482	8676.06	8.68	43.38	
6T8	Flacourtiaceae	Oncoba	spinosa	11.9	1.3	6	600	0.647	2847.09	2.85	14.24	
6T9	Rosaceae	Hagenia	abyssinica	32	1.3	12	1200	0.5905	35323.14	35.32	176.62	
									AGB=	141520.09	141.52	707.6004

Plot 10												
Tcode	Family	Genus	Species	dbh (cm)	POM	Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB(kg)	AGB (kg/ha)	
9T1	Anacardiaceae	Rhus	ruspolii	12.5	1.3	6	600	0.62	3006.31	3.01	15.03	
9T2	Myrsinaceae	Maesa	lanceolata	15	1.3	6	600	0.676	4669.27	4.67	23.35	
9T3	Rosaceae	Hagenia	abyssinica	46	1.3	13	1300	0.5905	77559.98	77.56	387.80	
9T4	Euphorbiaceae	Croton	macrostachyus	16.2	1.3	8	800	0.518	5541.02	5.54	27.71	
9T5	Rosaceae	Hagenia	abyssinica	46	1.3	14	1400	0.5905	83377.71	83.38	416.89	
9T6	Anacardiaceae	Rhus	ruspolii	12.6	1.3	5.5	550	0.62	2804.83	2.80	14.02	
									AGB=	176959.11	176.96	884.80

Plot 11												
Tcode	Family	Genus	Species	dbh (cm)	POM	Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB(kg)	AGB (kg/ha)	
12T1	Boraginaceae	Ehretia	cymosa	14.1	1.3	10	1000	0.56	5669.21	5.67	28.35	
12T2	Boraginaceae	Ehretia	cymosa	14.3	1.3	9	900	0.56	5257.79	5.26	26.29	
12T3	Rosaceae	Hagenia	abyssinica	29	1.3	12	1200	0.5905	29147.91	29.15	145.74	
12T4	Rosaceae	Hagenia	abyssinica	32	1.3	13	1300	1.5905	100455.43	100.46	502.28	
12T5	Euphorbiaceae	Croton	macrostachyus	14.6	1.3	8	800	0.518	4523.07	4.52	22.62	
12T6	Boraginaceae	Ehretia	cymosa	12.6	1.3	9	900	0.56	4106.87	4.11	20.53	
12T7	Anacardiaceae	Rhus	ruspolii	22.2	1.3	9	900	0.62	13702.85	13.70	68.51	
12T8	Euphorbiaceae	Croton	macrostachyus	14.8	1.3	9.5	950	0.518	5493.01	5.49	27.47	
12T9	Oleaceae	Chionanthus	mildbraedii	27	1.3	10.2	1020	0.705	25719.56	25.72	128.60	
12T10	Myrsinaceae	Maesa	lanceolata	12.5	1.3	7	700	0.676	3802.13	3.80	19.01	
									AGB=	197877.82	197.88	989.39

Plot 12												
Tcode	Family	Genus	Species	dbh (cm)	POM	Height (m)	Height (cm)	Wood D. (g)	AGB (g)	AGB(kg)	AGB (kg/ha)	
1T1	Rosaceae	Hagenia	abyssinica	36	1.3	10	1000	0.5905	37207.30	37.21	186.04	
1T2	Poaceae	Arundinaria	alpina	7	1.3	5	500	0.63	824.13	0.82	4.12	
1T3	Poaceae	Arundinaria	alpina	7	1.3	5.5	550	0.63	904.48	0.90	4.52	
1T4	Euphorbiaceae	Croton	macrostachyus	54	1.3	17	1700	0.518	121269.52	121.27	606.35	
1T5	Poaceae	Arundinaria	alpina	6.5	1.3	8.5	850	0.63	1196.99	1.20	5.98	
1T6	Poaceae	Arundinaria	alpina	6	1.3	7	700	0.63	847.11	0.85	4.24	
1T7	Poaceae	Arundinaria	alpina	6	1.3	8	800	0.63	965.02	0.97	4.83	
1T8	Poaceae	Arundinaria	alpina	5	1.3	5	500	0.63	427.32	0.43	2.14	
1T9	Poaceae	Arundinaria	alpina	5	1.3	7	700	0.63	593.44	0.59	2.97	
1T10	Poaceae	Arundinaria	alpina	9	1.3	8	800	0.63	2129.46	2.13	10.65	
1T11	Poaceae	Arundinaria	alpina	8	1.3	6	600	0.63	1277.85	1.28	6.39	
1T12	Poaceae	Arundinaria	alpina	7	1.3	8	800	0.63	1303.82	1.30	6.52	
1T13	Poaceae	Arundinaria	alpina	7.5	1.3	6	600	0.63	1126.59	1.13	5.63	
1T14	Fabaceae	Albizia	grandibracteata	20	1.3	14	1400	0.534	14870.54	14.87	74.35	
1T15	Rubiaceae	Rothmannia	urcelliformis	12.1	1.3	9	900	0.642	4336.18	4.34	21.68	
1T16	Boraginaceae	Cordia	africana	27	1.6	13	1300	0.482	22485.36	22.49	112.43	
									AGB=	211765.1	211.77	1058.83

Plot 13											
Tcode	Family	Genus	Species	dbh (cm)	POM Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB (kg)	AGB (kg/ha)	
15T1	Boraginaceae	Ehretia	cymosa	30	1.3	10.3	1030	0.56	25475.13	25.48	127.38
15T2	Fabaceae	Albizia	schimperiana	27.6	1.3	9.8	980	0.53	19543.37	19.54	97.72
15T3	Fabaceae	Albizia	schimperiana	27.9	1.3	11	1100	0.53	22342.25	22.34	111.71
15T4	Boraginaceae	Ehretia	cymosa	21.6	1.3	9.5	950	0.56	12398.20	12.40	61.99
15T5	Boraginaceae	Cordia	africana	56	1.5	16	1600	0.482	114380.79	114.38	571.90
15T6	Euphorbiaceae	Sapium	ellipticum	29	1.3	8	800	0.576	19151.56	19.15	95.76
15T7	Moraceae	Trilepisium	madagascariens	19.7	1.3	8	800	0.499	7826.64	7.83	39.13
15T8	Euphorbiaceae	Sapium	ellipticum	10	1.3	5.5	550	0.576	1662.58	1.66	8.31
AGB=									222780.5	222.781	1113.903

Plot 14											
Tcode	Family	Genus	Species	dbh (cm)	POM Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB (kg)	AGB (kg/ha)	
18T1	Euphorbiaceae	Sapium	ellipticum	32	1.7	12	1200	0.576	34476.33	34.48	172.38
18T2	Euphorbiaceae	Sapium	ellipticum	43	2.1	14.5	1450	0.576	73826.62	73.83	369.13
18T3	Melanthaceae	Bersama	abyssinica	36	1.3	12.5	1250	0.671	52406.23	52.41	262.03
18T4	Fabaceae	Millettia	ferruginea	21	1.7	14	1400	0.738	22430.09	22.43	112.15
18T5	Meliaceae	Trichilia	dregeana	40	1.3	13	1300	0.482	48428.25	48.43	242.14
AGB=									231567.53	231.57	1157.84

Plot 15											
Tcode	Family	Genus	Species	dbh (cm)	POM Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB (kg)	AGB (kg/ha)	
20T1	Balanitaceae	Balanites	aegyptica	9	1.3	5	500	0.6839	1458.29	1.46	7.29
20T2	Balanitaceae	Balanites	aegyptica	30	1.3	9	900	0.6839	27142.46	27.14	135.71
20T3	Euphorbiaceae	Bridelia	grandis	14	1.3	6.5	650	0.523	3434.93	3.43	17.17
20T4	Euphorbiaceae	Bridelia	grandis	14	1.3	8	800	0.523	4206.59	4.21	21.03
20T5	Moraceae	Ficus	elastica	36	1.3	11.7	1170	0.6071	44558.33	44.56	222.79
20T6	Ulmaceae	Celtis	zenkeri	9	1.3	7	700	0.59	1753.33	1.75	8.77
20T7	Araliaceae	Schefflera	abyssinica	16	1.3	8	800	0.4052	4255.58	4.26	21.28
20T8	Araliaceae	Schefflera	abyssinica	16	1.3	8	800	0.4052	4255.58	4.26	21.28
20T9	Meliaceae	Trichilia	dregeana	28	1.3	9	900	0.482	16860.18	16.86	84.30
20T10	Anacardiaceae	Rhus	ruspolii	19	1.3	9	900	0.62	10112.47	10.11	50.56
20T11	Euphorbiaceae	Croton	macrostachyus	51	1.3	13	1300	0.518	83480.80	83.48	417.40
20T12	Araliaceae	Schefflera	abyssinica	21	1.3	8.5	850	0.4052	7676.89	7.68	38.38
20T13	Euphorbiaceae	Croton	macrostachyus	30.7	1.3	11	1100	0.518	26332.63	26.33	131.66
AGB=									235528.06	235.53	1177.64

Plot 16											
Tcode	Family	Genus	Species	dbh (cm)	POM Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB (kg)	AGB (kg/ha)	
5T1	Fabaceae	Millettia	ferruginea	17.3	1.3	9	900	0.738	9982.62	9.98	49.91
5T2	Fabaceae	Acacia	abyssinica	41	1.3	14	1400	0.826	92419.29	92.42	462.10
5T3	Flacourtiaceae	Oncoba	spinosa	32	1.3	14	1400	0.647	44888.11	44.89	224.44
5T4	Rubiaceae	Rothmannia	urcelliformis	14.5	1.3	7	700	0.642	4830.31	4.83	24.15
5T5	Fabaceae	Millettia	ferruginea	18	1.4	9	900	0.738	10786.26	10.79	53.93
5T6	Fabaceae	Millettia	ferruginea	9	1.4	7	700	0.738	2181.40	2.18	10.91
5T7	Moraceae	Trilepisium	madagascariens	21	1.3	8	800	0.499	8866.44	8.87	44.33
5T8	Euphorbiaceae	Sapium	ellipticum	26	4	16	1600	0.576	30439.43	30.44	152.20
5T9	Boraginaceae	Cordia	africana	30.1	1.3	13	1300	0.482	27799.66	27.80	139.00
5T10	Melanthaceae	Bersama	abyssinica	15	1.3	8	800	0.671	6138.22	6.14	30.69
5T11	Fabaceae	Albizia	schimperiana	19.1	1.3	9	900	0.53	8766.50	8.77	43.83
AGB=									247098.24	247.10	1235.49

Plot 17												
Tcode	Family	Genus	Species	dbh (cm)	POM	Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB(kg)	AGB(kg/ha)	
2T1	Boraginaceae	Cordia	africana	37.6	1.6	13	1300	0.482	42918.48	42.92	214.59	
2T2	Fabaceae	Albizia	grandibracteata	17.1	1.3	9	900	0.534	7116.14	7.12	35.58	
2T3	Balanitaceae	Balanites	aegyptica	22	1.3	9	900	0.6839	14815.54	14.82	74.08	
2T4	Balanitaceae	Balanites	aegyptica	16	1.3	9	900	0.6839	7957.03	7.96	39.79	
2T5	Balanitaceae	Balanites	aegyptica	28	1.3	13	1300	0.6839	33964.80	33.96	169.82	
2T6	Fabaceae	Jacaranda	mimosifolia	32	1.3	15	1500	0.4933	36847.61	36.85	184.24	
2T7	Fabaceae	Albizia	schimperiana	22.3	1.3	9.5	950	0.53	12504.25	12.50	62.52	
2T8	Fabaceae	Albizia	schimperiana	20	1.3	7.8	780	0.53	8340.71	8.34	41.70	
2T9	Rosaceae	Hagenia	abyssinica	26	1.3	8.3	830	0.5905	16435.16	16.44	82.18	
2T10	Euphorbiaceae	Croton	macrostachyus	38	1.3	16	1600	0.56	62116.25	62.12	310.58	
2T11	Asteraceae	Vernonia	amygdalina	13	1.3	7	700	0.413	2537.55	2.54	12.69	
2T12	Fabaceae	Albizia	schimperiana	29	1.3	13	1300	0.53	28360.76	28.36	141.80	
									AGB=	273914.29	273.91	1369.57

Plot 18												
Tcode	Family	Genus	Species	dbh (cm)	POM	Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB (kg)	AGB (kg/ha)	
11T1	Fabaceae	Albizia	schimperiana	12.8	1.3	7.5	750	0.534	3383.99	3.38	16.92	
11T2	Euphorbiaceae	Croton	macrostachyus	48	1.3	14	1400	0.518	79726.88	79.73	398.63	
11T3	Meliaceae	Trichilia	dregeana	11.8	1.5	5	500	0.482	1758.64	1.76	8.79	
11T4	Meliaceae	Trichilia	dregeana	15	1.5	6	600	0.482	3356.41	3.36	16.78	
11T5	Meliaceae	Trichilia	dregeana	11.8	1.5	7	700	0.482	2442.30	2.44	12.21	
11T6	Fabaceae	Albizia	schimperiana	17	1.3	7	700	0.466	4819.61	4.82	24.10	
11T7	Fabaceae	Albizia	schimperiana	46	1.3	16.5	1650	0.534	88728.46	88.73	443.64	
11T8	Anacardiaceae	Rhus	ruspolii	22.2	1.3	9	900	0.62	13702.85	13.70	68.51	
11T9	Fabaceae	Albizia	schimperiana	14	1.3	5.6	560	0.534	3030.88	3.03	15.15	
11T10	Fabaceae	Albizia	schimperiana	16	1.3	7	700	0.534	4890.50	4.89	24.45	
11T11	Fabaceae	Albizia	schimperiana	21	1.3	6.5	650	0.534	7735.19	7.74	38.68	
11T12	Euphorbiaceae	Croton	macrostachyus	24.2	1.3	11.5	1150	0.518	17284.21	17.28	86.42	
11T13	Euphorbiaceae	Croton	macrostachyus	38	1.3	12.5	1250	0.518	45240.01	45.24	226.20	
									AGB=	276099.92	276.10	1380.50

Plot 19												
Tcode	Family	Genus	Species	dbh (cm)	POM	Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB (kg)	AGB (kg/ha)	
14T1	Fabaceae	Albizia	grandibracteata	19.4	1.3	10	1000	0.534	10089.84	10.09	50.45	
14T2	Moraceae	Ficus	sur	53.7	2	17	1700	0.441	102521.53	102.52	512.61	
14T3	Euphorbiaceae	Croton	macrostachyus	29	1.3	11.5	1150	0.518	24606.11	24.61	123.03	
14T4	Moraceae	Ficus	sycomorus	16.3	1.3	10	1000	0.422	5708.35	5.71	28.54	
14T5	Moraceae	Ficus	sycomorus	70.2	2.7	17.5	1750	0.422	170441.44	170.44	852.21	
14T6	Meliaceae	Trichilia	dregeana	11.8	1.5	10	1000	0.482	3459.26	3.46	17.30	
14T7	Myrsinaceae	Maesa	lanceolata	17.7	1.3	8	800	0.676	8540.88	8.54	42.70	
									AGB=	325367.40	325.37	1626.84

Plot 20												
Tcode	Family	Genus	Species	dbh (cm)	POM	Height (m)	Height (cm)	WD(g/cm ³)	AGB (g)	AGB (kg)	AGB (kg/ha)	
17T1	Fabaceae	Albizia	grandibracteata	22	1.3	13	1300	0.534	16661.46	16.66	83.31	
17T2	Fabaceae	Albizia	grandibracteata	31	1.3	10	1000	0.534	25190.30	25.19	125.95	
17T4	Fabaceae	Albizia	grandibracteata	62	1.3	13.8	1380	0.534	133464.54	133.46	667.32	
17T5	Euphorbiaceae	Sapium	ellipticum	39	1.3	15	1500	0.576	63068.22	63.07	315.34	
17T6	Ulmaceae	Celtis	zenkeri	21	1.3	6.9	690	0.59	9037.65	9.04	45.19	
17T7	Boraginaceae	Cordia	africana	55	1.3	13	1300	0.482	90170.74	90.17	450.85	
17T8	Boraginaceae	Ehretia	cymosa	13	1.3	6	600	0.56	2938.59	2.94	14.69	
17T9	Pinaceae	Pinus	patula	7	1.3	10	1000	0.46	1192.61	1.19	5.96	
									AGB=	341724.12	341.72	1708.62

Table 20: Plot level Field AGC stock Estimations

plot#	AGB (kg/ha)	AGB (t/ha)	AGC (kg C/ha)	AGC (t C/ha)
P01	418.77	0.419	196.82	0.19682
P02	499.29	0.499	234.67	0.23467
P03	501.85	0.502	235.87	0.23587
P04	553	0.553	259.91	0.25991
P05	599.04	0.599	281.55	0.28155
P06	628.4	0.628	295.35	0.29535
P07	635.09	0.635	298.49	0.29849
P08	672.08	0.672	315.88	0.31588
P09	707.6	0.708	332.57	0.33257
P10	884.8	0.885	415.85	0.41585
P11	989.39	0.989	465.01	0.46501
P12	1058.83	1.059	497.65	0.49765
P13	1113.9	1.114	523.53	0.52353
P14	1157.84	1.158	544.18	0.54418
P15	1177.64	1.178	553.49	0.55349
P16	1235.49	1.235	580.68	0.58068
P17	1369.57	1.37	643.7	0.6437
P18	1380.5	1.38	648.83	0.64883
P19	1626.84	1.627	764.61	0.76461
P20	1708.62	1.709	803.05	0.80305

ANNEX 3 Satellite variables extracted for field plots

Table 21: Satellite variables used as input in regression analysis

Plot_code	X	Y	AGB_t/ha	AGC t/ha	NDVI201	NDWI20	EVI2019	SAVI201	glcm_Sav	glcm_ent	glcm_cor	glcm_var
P01	306599.8	615666.8	0.418768	0.196821	0.634467	-0.58128	0.396522	0.967568	8.676649	3.593232	0.298767	4.500644
P02	307005.3	614971.9	0.499289	0.234666	0.685393	-0.6063	0.418022	1.04161	9.876302	3.656427	0.609881	26.34131
P03	305982.3	615861.5	0.501849	0.235869	0.694826	-0.61183	0.428575	0.984855	35.93576	3.823574	0.548068	9.499192
P04	307887.1	611808.1	0.552995	0.259908	0.670825	-0.61364	0.430671	0.997253	19.01302	3.867884	0.372703	51.27409
P05	307921.6	613947.1	0.599042	0.28155	0.700069	-0.61665	0.48115	1.027855	19.75174	3.8953	0.407235	4.361611
P06	308270.3	614903.4	0.6284	0.295348	0.704413	-0.61938	0.484949	1.040914	24.41623	4.071869	0.45252	166.3799
P07	305103.4	615118	0.635086	0.29849	0.706064	-0.62292	0.518671	0.977059	14.03863	4.093315	0.482964	95.42049
P08	305647.1	614762.3	0.672081	0.315878	0.715167	-0.60636	0.522961	1.057543	20.20703	4.136587	0.483626	7.248757
P09	306174.7	615089.9	0.7076	0.332572	0.726479	-0.58469	0.522969	1.101702	45.10677	4.30908	0.601451	43.54327
P10	307219.6	615670.7	0.884796	0.415854	0.728201	-0.60293	0.528283	1.083552	31.92969	4.378065	0.623402	67.38863
P11	307292.9	613924	0.989389	0.465013	0.732612	-0.5984	0.535313	1.062029	29.97656	4.390553	0.652754	19.44257
P12	304696.3	615503.6	1.058826	0.497648	0.73569	-0.581	0.552077	1.059739	25.61241	4.398037	0.660918	5.972337
P13	307108.8	613109.9	1.113903	0.523534	0.740816	-0.57978	0.558535	1.09087	10.78733	4.4119	0.662947	31.81863
P14	308618.2	614239.8	1.157838	0.544184	0.745993	-0.57631	0.559857	1.091379	21.66753	4.431054	0.665695	48.93126
P15	308801.8	616074	1.17764	0.553491	0.747061	-0.53577	0.572397	1.092892	16.55208	4.587198	0.694453	8.293212
P16	306018	614199.7	1.235491	0.580681	0.748242	-0.54451	0.572771	1.098487	12.38194	4.617275	0.706753	13.04612
P17	304695.7	614255.1	1.369571	0.643699	0.74885	-0.58128	0.574674	1.099474	13.28733	4.771993	0.72302	9.259674
P18	306667.4	614275.9	1.3805	0.648835	0.753209	-0.60524	0.575468	1.101347	37.70009	4.640176	0.538301	11.29342
P19	307705.2	613748.8	1.626837	0.764613	0.757274	-0.52102	0.577895	1.069981	42.42101	4.037561	0.667916	30.72315
P20	306701.8	611067	1.708621	0.803052	0.779586	-0.5285	0.579264	1.109129	63.7704	4.694063	0.378275	5.053087

ANNEX 4 Summary of AGC and Satellite derivatives regression analysis

Table 22: AGC and NDVI regression analysis

NDVI SUMMARY OUTPUT									
<i>Regression Statistics</i>									
Multiple R	0.900271972								
R Square	0.81048962								
Adjusted R Square	0.799961269								
Standard Error	0.082425683								
Observations	20								
ANOVA									
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>				
Regression	1	0.523012517	0.523013	76.98161	6.42166E-08				
Residual	18	0.122291879	0.006794						
Total	19	0.645304397							
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>	
Intercept	-3.055745918	0.399372921	-7.65136	4.6E-07	-3.894797291	-2.216695	-3.89479729	-2.21669455	
X Variable 1	4.842993538	0.551976267	8.773916	6.42E-08	3.683334433	6.002653	3.683334433	6.002652643	
EVI SUMMARY OUTPUT									
<i>Regression Statistics</i>									
Multiple R	0.872761807								
R Square	0.76171317								
Adjusted R Square	0.748475014								
Standard Error	0.092426409								
Observations	20								
ANOVA									
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>				
Regression	1	0.491536858	0.491537	57.53922	5.18453E-07				
Residual	18	0.153767538	0.008543						
Total	19	0.645304397							
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>	
Intercept	-0.956364667	0.185841586	-5.14613	6.77E-05	-1.34680335	-0.565926	-1.34680335	-0.56592598	
X Variable 1	2.6964617	0.355477641	7.585461	5.18E-07	1.94963089	3.443293	1.94963089	3.443292511	

Table 23: AGC and SAVI and NDWI regression analysis

NDWI SUMMARY OUTPUT						
<i>Regression Statistics</i>						
Multiple R	0.759713085					
R Square	0.57716397					
Adjusted R Square	0.553673081					
Standard Error	0.123120977					
Observations	20					
ANOVA						
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>	
Regression	1	0.372446448	0.372446	24.56969	0.000101953	
Residual	18	0.272857948	0.015159			
Total	19	0.645304397				
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	3.076052909	0.531595589	5.786453	1.75E-05	1.95921202	4.192894
X Variable 1	4.491398363	0.906111641	4.956783	0.000102	2.587728446	6.395068
SAVI SUMMARY OUTPUT						
<i>Regression Statistics</i>						
Multiple R	0.758804424					
R Square	0.57578415					
Adjusted R Square	0.552216607					
Standard Error	0.1233217					
Observations	20					
ANOVA						
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>	
Regression	1	0.371556046	0.371556	24.43123	0.000105094	
Residual	18	0.27374835	0.015208			
Total	19	0.645304397				
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	-2.810778106	0.659184727	-4.26402	0.000467	-4.195673828	-1.425882
X Variable 1	3.077595317	0.622642626	4.942796	0.000105	1.769471701	4.385719

Table 24: Results of GLCM Sum-Average texture and AGC regression analysis

'SUM AVERAGE' SUMMARY OUTPUT						
Regression Statistics						
Multiple R	0.452161278					
R Square	0.204449821					
Adjusted R Square	0.160252589					
Standard Error	0.168880633					
Observations	20					
ANOVA						
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>	
Regression	1	0.131932369	0.131932	4.625851	0.045327071	
Residual	18	0.513372028	0.028521			
Total	19	0.645304397				
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	0.29727781	0.078210907	3.800976	0.001309	0.132962791	0.461593
X Variable 1	0.005855889	0.002722683	2.150779	0.045327	0.000135745	0.011576

Table 25: Results of GLCM Entropy texture and AGC regression analysis

ENTROPY SUMMARY OUTPUT						
Regression Statistics						
Multiple R	0.782385674					
R Square	0.612127344					
Adjusted R Square	0.590578863					
Standard Error	0.117920861					
Observations	20					
ANOVA						
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>	
Regression	1	0.395008466	0.395008	28.40698	4.57259E-05	
Residual	18	0.250295931	0.013905			
Total	19	0.645304397				
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	-1.297287944	0.327878459	-3.95661	0.000925	-1.986135025	-0.608441
X Variable 1	0.410794219	0.077074686	5.32982	4.57E-05	0.248866313	0.572722

Table 26: Results of GLCM correlation texture and AGC regression analysis

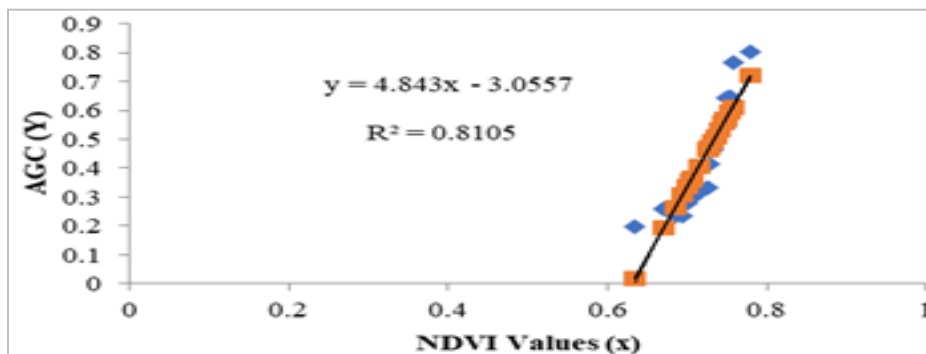
CORRELATION TEXTURE SUMMARY OUTPUT						
Regression Statistics						
Multiple R	0.457847705					
R Square	0.209624521					
Adjusted R Square	0.165714772					
Standard Error	0.16833049					
Observations	20					
ANOVA						
	df	SS	MS	F	Significance F	
Regression	1	0.135271625	0.135272	4.773986	0.042358986	
Residual	18	0.510032772	0.028335			
Total	19	0.645304397				
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	0.07372865	0.173856056	0.424079	0.676533	-0.291529369	0.438987
X Variable 1	0.660377672	0.302239913	2.184945	0.042359	0.025395176	1.29536

Table 27: Results of GLCM Variance texture and AGC regression analysis

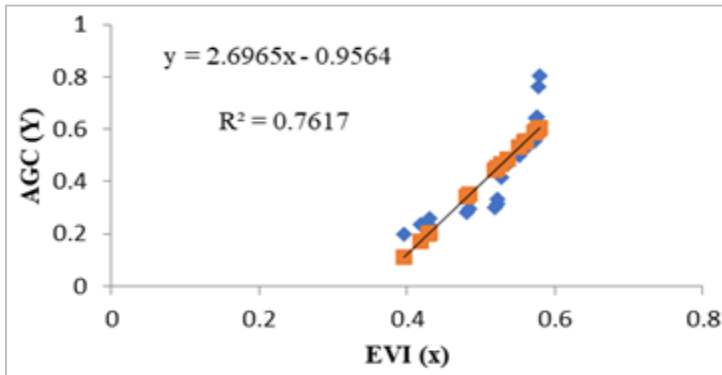
GLCM VARIANCE SUMMARY OUTPUT						
Regression Statistics						
Multiple R	0.277570303					
R Square	0.077045273					
Adjusted R Square	0.025770011					
Standard Error	0.181901491					
Observations	20					
ANOVA						
	df	SS	MS	F	Significance F	
Regression	1	0.049717653	0.049718	1.502582	0.236071143	
Residual	18	0.595586743	0.033088			
Total	19	0.645304397				
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	0.486909484	0.05335341	9.126117	3.58E-08	0.374818129	0.599001
X Variable 1	-0.00128296	0.001046632	-1.2258	0.236071	-0.003481853	0.000916

ANNEX 5 Regression Analysis Equations for Predictive Model Development

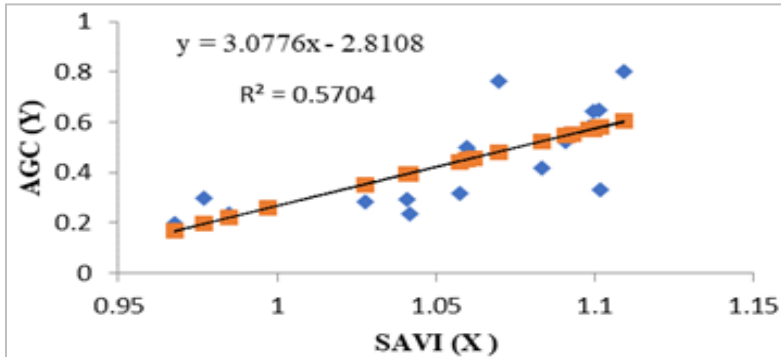
5.1 Correlation between NDVI and Aboveground Carbon Stock of Plot Stands



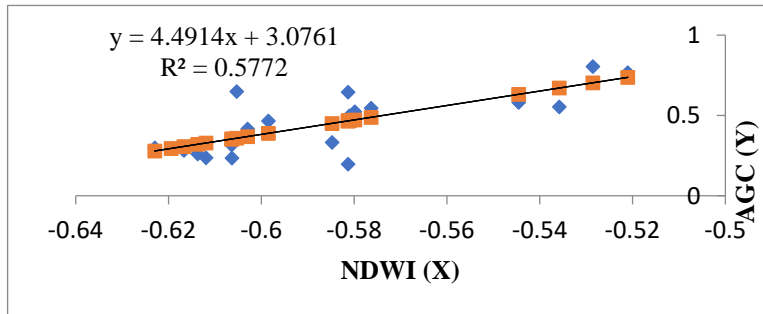
5.2 Correlation between EVI and Aboveground Carbon Stock of Plot Stands



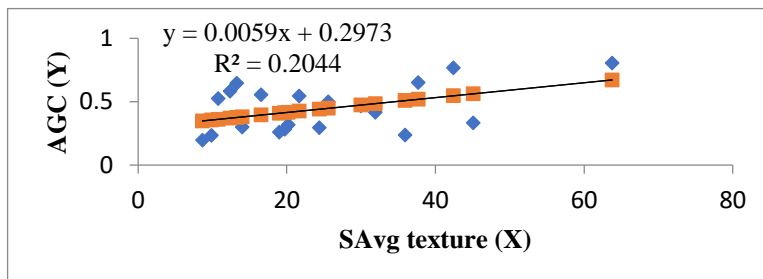
5.3 Correlation between SAVI and Aboveground Carbon Stock of Plot Stands



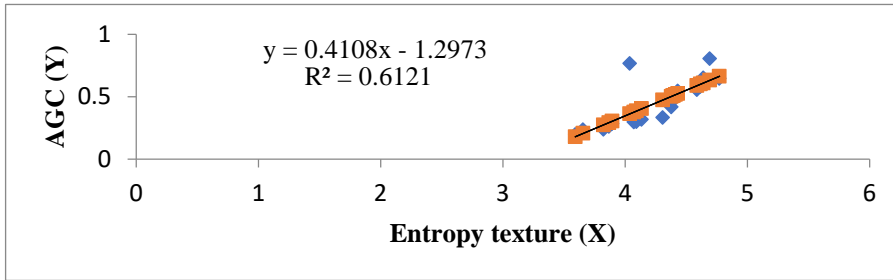
5.4 Correlation between NDWI and Aboveground Carbon Stock of Plot Stands



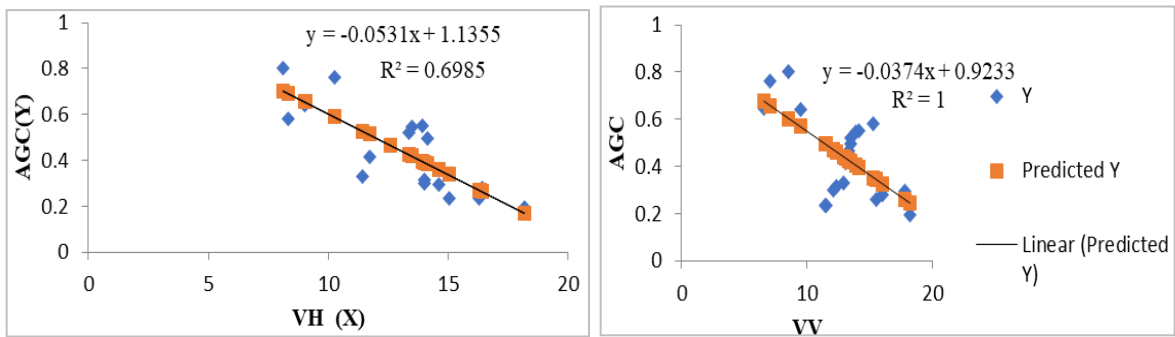
5.5 Correlation between GLCM Sum-Average and Aboveground Carbon Stock of Plot Stands



5.6 Correlation between Sentinel-1 Entropy and Aboveground Carbon Stock of Plots



5.7 Correlation between Sentinel-1 backscatters and Aboveground Carbon Stock of Plots



ANNEX 6 Sentinel-1 GRD products spatial resolution

Table 28: Modes and Spatial resolution of Sentinel-1 GRD Products

Full resolution Level-1 GRD				
Mode	Resolution range x azimuth	Pixel spacing rg x az	Number of looks	ENL
SM	9x9 m	3.5x3.5 m	2x2	3.7
High resolution Level-1 GRD				
Mode	Resolution range x azimuth	Pixel spacing rg x az	Number of looks	ENL
SM	23x23 m	10x10 m	6x6	29.7
IW	20x22 m	10x10 m	5x1	4.4
EW	50x50 m	25x25 m	3x1	2.7
Medium resolution Level-1 GRD				
Mode	Resolution rg x az	Pixel spacing rg x az	Number of looks	ENL
SM	84x84 m	40x40 m	22x22	398.4
IW	88x87 m	40x40 m	22x5	81.8
EW	93x87 m	40x40 m	6x2	10.7
WV	52x51 m	25x25 m	13x13	123.7

ANNEX 7 Sentinel-2 Spatial Resolution Spectral Bands

A. 10m Spatial Resolution Bands

B2 (490 nm), B3 (560 nm), B4 (665 nm) and B8 (842 nm)

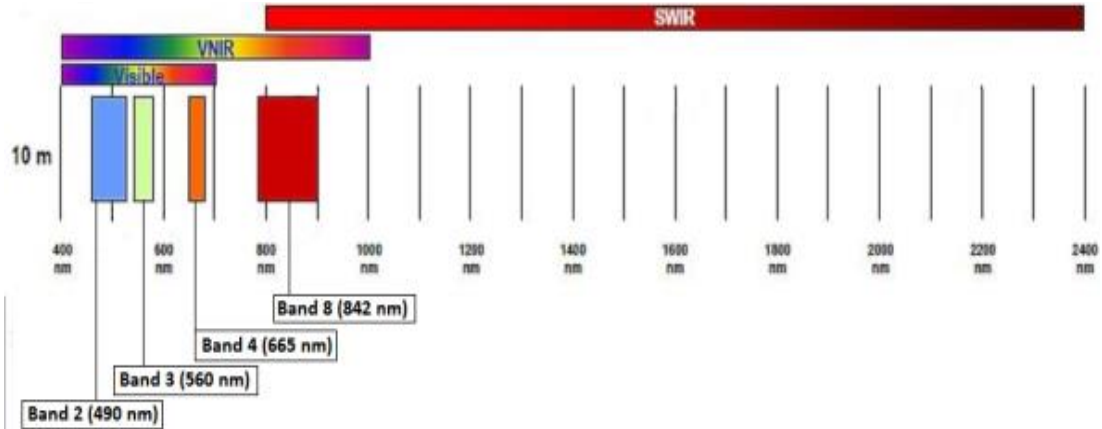


Figure 36: Sentinel-2 10m spatial resolution bands

B. 20m spatial resolution bands

B5 (705nm), B6 (740nm), B7 (783nm), B8a (865nm), B11 (1610nm) and B12 (2190nm)

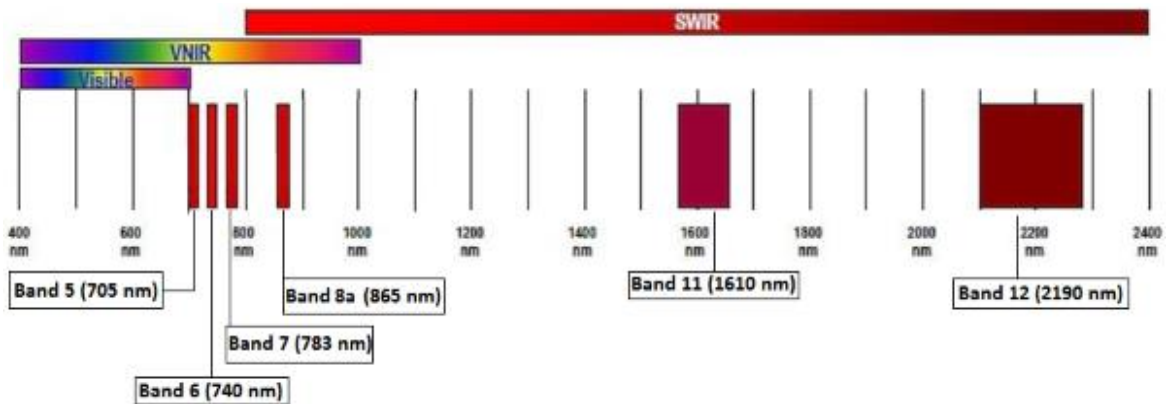


Figure 37: Sentinel-2 20m spatial resolution bands

C. 60m spatial resolution bands

B1 (443 nm), B9 (940 nm) and B10 (1375 nm)

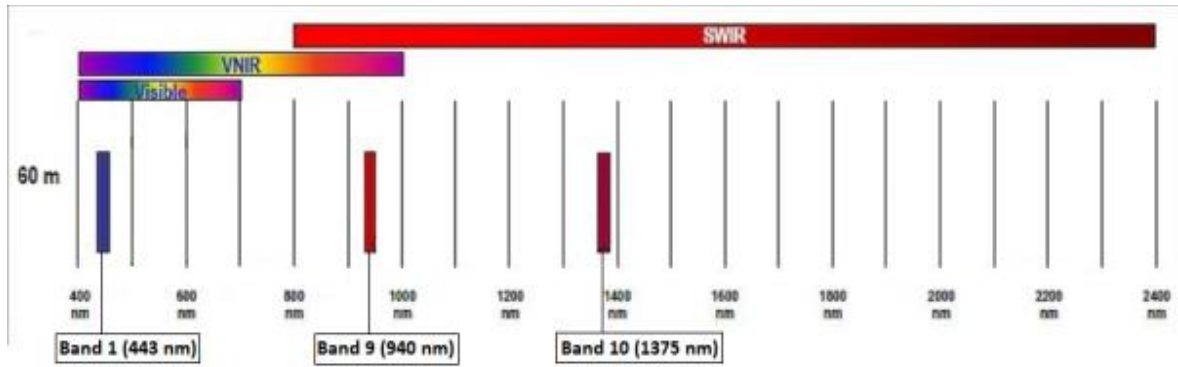


Figure 38: Sentinel-2 60m spatial resolution bands

ANNEX 8 Prototype Modules Codes

```

////////////////////////////////////
///
    JavaScript Codes for Sentinel-1, Sentinel-2 and ALOS DEM Data Processing and
    Model Input Variable Derivation //
////////////////////////////////////
///

```

Annex 8.1 Developed JavaScript Codes

```

-----
//call DefineStudyArea code
var studyArea = require('users/oytisa/FinalThesisMSc:DefineStudyArea');
var aoi = var studyArea;
// Expression to create cloudless Sentinel-2.
function maskS2clouds(image) {
    var qa = image.select('QA60');
    var mask = qa.bitwiseAnd(cloudBitMask).eq(0)
        .and(qa.bitwiseAnd(cirrusBitMask).eq(0));
    return image.updateMask(mask);
}
//----- create year list variables -----
var year_list = ['2016','2017','2018','2019', '2020'];
var month_list = ['01','02','03','04',
    '05','06','07','08',
    '09','10','11','12','year_median']
var collection20 = ee.ImageCollection('COPERNICUS/S2') ..
    .filter(ee.Filter.lt("CLOUDY_PIXEL_PERCENTAGE"))
    .filterDate('2020-01-01', '2020-12-31')
    .sort('system:time_start', true)
    );
var median20 = merged20.reduce(ee.Reducer.median()).clip(aoi);
var visParams = {bands: ['B04_median',
    'B03_median',
    'B02_median'],
    min: 300, max: 3700,
    gamma: [1.5, 1.5, 1.5]
    };
Map.addLayer(median20, visParams, 's2cloud_masked')

```

```

print(median20)
//Create geometry object encompassing area of interest
var geometry = ee.FeatureCollection("users/oytisa/shiwshiwe_bound")
Map.centerObject(geometry,11)
//*****O3+B*****
//CONSTANTS
//-----
var clipgeometry = geometry
//-----
  ///*****Sentinel-2 EVI function*****///
// Add an EVI function for the Sentinel-2 (s2) collections.
function addEVI(s2){
return s2.addBands(s2.expression(
'2.5 * ((NIR - RED) / (NIR + 6 * RED - 7.5 * BLUE + 1))',
{
'NIR': s2.select('B8'),
'RED': s2.select('B4'),
'BLUE': s2.select('B2')
}).rename('S2EVI')).float()
.clip(geometry);
});
var S2_EVI = ee.ImageCollection(s2.select("S2EVI"))
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
// calculate vegetation indices
var oytS2md = median20.clip(aoi);
var ndwi_s2 = oytS2md.expression(
  '(green-mir)/(green+mir+0.0000001)', {
    'green': oytS2md.select('B03_median'),
    'mir': oytS2md.select('B08_median'),
  });
//print(mndwi_18,'NDWI S2');
Map.addLayer(mndwi_s2,{},'MNDWI_S2');
var thresholds2 = mndwi_s2.gte(0.6);
Map.addLayer(thresholds2.clip(aoi),{'Threshold_s2'});

var evi =oytS2md.expression(
  '2.5 * ((NIR-RED) / (NIR +6 * RED -7.5* BLUE))', {
    'NIR':oytS2md.select('B08_median'),
    'RED':oytS2md.select('B04_median'),
    'BLUE':oytS2md.select('B02_median')
  });
var palette = [
  'FFFFFF', 'CE7E45', 'DF923D', 'F1B555', 'FCD163', '99B718',
  '74A901', '66A000', '529400', '3E8601', '207401', '056201',
  '004C00', '023B01', '012E01', '011D01', '011301'];
print(evi);
Map.addLayer(evi,{min: -1, max: 1, palette: palette},'S2 evi')
var s1 = ee.ImageCollection('COPERNICUS/S1_GRD');
var s2 = ee.ImageCollection('COPERNICUS/S2');
var alos = ee.Image('JAXA/ALOS/AW3D30_V1_1');
var oytAOI = ee.FeatureCollection('users/oytisa/shiwshiwe_bound');
var s1 = s1.filterBounds(oytAOI)
  .filterDate("2016-01-01", "2016-12-31")

```

```

        .filter(ee.Filter.listContains('transmitterReceiverPolarisation', 'VV'))
        .filter(ee.Filter.listContains('transmitterReceiverPolarisation', 'VH'))
        .filter(ee.Filter.eq('instrumentMode', 'TW'))
        .median()
var collection = ee.ImageCollection('COPERNICUS/S1_GRD')
    .filter(ee.Filter.eq('instrumentMode', 'TW'))
    .filter(ee.Filter.eq('resolution', 'H'))
    .filter(ee.Filter.eq('resolution_meters', 10))
    .filter(ee.Filter.eq('relativeOrbitNumber_start', 152))
    .filter(ee.Filter.date('2016-01-01', '2016-12-31'))
    .filterBounds(oytAOI)
    .map(function(image) {
        var maskedImage = image.mask().and(edge.not());
        return image.updateMask(maskedImage);
    });
    // Remove ugly edges
function maskEdge(img) {
    var mask = img.select(1).unitScale(-30, -
5).multiply(255).toByte().connectedComponents(ee.Kernel.rectangle(1,1), 100);
    return img.updateMask(mask.select(1));
}

var collectionVH = collection.map(maskEdge);

////////////////////////////////////
// Convert sigma naught to gamma naught
function toGamma0(image) {
    return
image.select(1).subtract(image.select('angle').multiply(Math.PI/180.0).cos().log10().multiply(10.0));
}

////////////////////////////////////
// Functions to convert from/to dB
function toNatural(img) {
    return ee.Image(10.0).pow(img.select(0).divide(10.0));
}

function toDB(img) {
    return ee.Image(img).log10().multiply(10.0);
}
// The Refined Lee speckle filter
function RefinedLee(img) {
    // img must be in natural units, i.e. not in dB!
    var weights3 = ee.List.repeat(ee.List.repeat(1,3),3);
    var kernel3 = ee.Kernel.fixed(3,3, weights3, 1, 1, false);
    var mean3 = img.reduceNeighborhood(ee.Reducer.mean(), kernel3);
    var variance3 = img.reduceNeighborhood(ee.Reducer.variance(), kernel3);

    // Use a sample of the 3x3 windows inside a 7x7 windows to determine gradients and directions
    var sample_weights = ee.List([[0,0,0,0,0,0,0], [0,1,0,1,0,1,0],[0,0,0,0,0,0,0], [0,1,0,1,0,1,0],
[0,0,0,0,0,0,0], [0,1,0,1,0,1,0],[0,0,0,0,0,0,0]]);

    var sample_kernel = ee.Kernel.fixed(7,7, sample_weights, 3,3, false);

```

```

var gradients = sample_mean.select(1).subtract(sample_mean.select(7)).abs();
// Create a mask for band pixels that are the maximum gradient
var gradmask = gradients.eq(max_gradient);

// duplicate gradmask bands: each gradient represents 2 directions
gradmask = gradmask.addBands(gradmask);

// Determine the 8 directions
var directions =
sample_mean.select(1).subtract(sample_mean.select(4)).gt(sample_mean.select(4).subtract(sample_m
ean.select(7))).multiply(1);
directions =
directions.addBands(sample_mean.select(6).subtract(sample_mean.select(4)).gt(sample_mean.select(
4).subtract(sample_mean.select(2))).multiply(2));
directions =
directions.addBands(sample_mean.select(3).subtract(sample_mean.select(4)).gt(sample_mean.select(
4).subtract(sample_mean.select(5))).multiply(3));
directions =
directions.addBands(sample_mean.select(0).subtract(sample_mean.select(4)).gt(sample_mean.select(
4).subtract(sample_mean.select(8))).multiply(4));
// The next 4 are the not() of the previous 4
directions = directions.addBands(directions.select(0).not().multiply(5));
directions = directions.addBands(directions.select(1).not().multiply(6));
directions = directions.addBands(directions.select(2).not().multiply(7));
directions = directions.addBands(directions.select(3).not().multiply(8));

// Mask all values that are not 1-8
directions = directions.updateMask(gradmask);
directions = directions.reduce(ee.Reducer.sum());
//var pal = ['ffffff', 'ff0000', 'ffff00', '00ff00', '00ffff', '0000ff', 'ff00ff', '000000'];
//Map.addLayer(directions.reduce(ee.Reducer.sum()), {min:1, max:8, palette: pal}, 'Directions',
false);

var sample_stats = sample_var.divide(sample_mean.multiply(sample_mean));

// Calculate localNoiseVariance
var sigmaV = sample_stats.toArray().arraySort().arraySlice(0,0,5).arrayReduce(ee.Reducer.mean(),
[0]);

// Set up the 7*7 kernels for directional statistics
var rect_weights = ee.List.repeat(ee.List.repeat(0,7),3).cat(ee.List.repeat(ee.List.repeat(1,7),4));

var diag_weights = ee.List([[1,0,0,0,0,0,0], [1,1,0,0,0,0,0], [1,1,1,0,0,0,0],
[1,1,1,1,0,0,0], [1,1,1,1,1,0,0], [1,1,1,1,1,1,0], [1,1,1,1,1,1,1]]);

var rect_kernel = ee.Kernel.fixed(7,7, rect_weights, 3, 3, false);
var diag_kernel = ee.Kernel.fixed(7,7, diag_weights, 3, 3, false);

// Create stacks for mean and variance using the original kernels. Mask with relevant direction.
var dir_mean = img.reduceNeighborhood(ee.Reducer.mean(),
rect_kernel).updateMask(directions.eq(1));
rect_kernel.updateMask(directions.eq(1));

```

```

    dir_mean = dir_mean.addBands(img.reduceNeighborhood(ee.Reducer.mean(),
diag_kernel).updateMask(directions.eq(2)));
    dir_var = dir_var.addBands(img.reduceNeighborhood(ee.Reducer.variance(),
diag_kernel).updateMask(directions.eq(2)));
  }
  // "collapse" the stack into a single band image (due to masking, each pixel has just one value in it's
directional band, and is otherwise masked)
  dir_mean = dir_mean.reduce(ee.Reducer.sum());
  dir_var = dir_var.reduce(ee.Reducer.sum());

  // A finally generate the filtered value
  var varX =
dir_var.subtract(dir_mean.multiply(dir_mean).multiply(sigmaV)).divide(sigmaV.add(1.0));

  var b = varX.divide(dir_var);
  var result = dir_mean.add(b.multiply(img.subtract(dir_mean)));
  return(result.arrayFlatten(['sum']));
}
//*****
// Combine 3 steps of speckle filtering into one simple function
function preProcessing(image) {
  var collection = toDB(RefinedLee(toNatural(image)));
  return collection;
}
/*****Shiwshiwe Time-series-Sentinel1-VH and VV
*****
// Time-series creating, stack averaging. Converting to gamma_naught
var stack1 = ee.ImageCollection(collection.VH.filterDate('2016-01-01', '2016-02-01'));
var stack1_gamma = stack1.map(toGamma0);

var stack5 = ee.ImageCollection(collection.VH.filterDate('2016-05-01', '2016-06-01'));
var stack5_gamma = stack5.map(toGamma0);
var stack5_Natural = stack5_gamma.map(toNatural);
var stack5_DB = toDB(stack5_Natural.map(RefinedLee).reduce('mean'));
//Map.addLayer(stack5_DB.clip(oytAOI), { min: -25, max: 5}, 'stack5_avg', 0);
var stack5_DB = stack5_DB.rename('stack5_VH')
var stack12_DB = stack12_DB.rename('stack12_VH')

var VH_12_gamma = stack12_Natural.reduce('mean').rename('VH12');
print(VH_12_gamma, 'VH12')
// Build a stack for all the images in the collection//
var stackedVH =stack1_DB.addBands(stack2_DB).addBands(stack3_DB).addBands(stack4_DB)
.addBands(stack5_DB).addBands(stack6_DB).addBands(stack7_DB).addBands(stack8_DB)
.addBands(stack9_DB).addBands(stack10_DB).addBands(stack11_DB).addBands(stack12_DB);
//Map.addLayer(stackedVH.clip(oytAOI), { }, 'stack_all_year', 0)
///// output GLCM Textures *****
////////////////////////////////////
var textureMeasuresVH = ['VH_asm', 'VH_contrast', 'VH_corr', 'VH_var', 'VH_idm', 'VH_savg',
'VH_ent', 'VH_sent', 'VH_dvar', 'VH_dent', 'VH_imcorr1'];

var textureMeasuresVH10 = ['VH10_savg', 'VH10_ent','VH10_corr', 'VH10_var'];
var textureMeasuresVH12 = ['VH12_savg', 'VH12_ent', 'VH12_corr', 'VH12_var'];

```

```

var glcmVH_10 = VH_10_gamma.unitScale(0,1).multiply(255).toByte().glcmTexture({ size:
4}).select(textureMeasuresVH10);
var glcmVH_12 = VH_12_gamma.unitScale(0,1).multiply(255).toByte().glcmTexture({ size:
4}).select(textureMeasuresVH12);

var VH_final
=(stack10_DB.addBands(stack12_DB).addBands(glcmVH_10).addBands(glcmVH_12)).float().clip(o
ytAOI)
print(VH_final, 'glcm')
Map.addLayer(VH_final/* .clip(oytAOI)*/, {}, 'glcmVH');
//Map.addLayer(VH_final, { 'VH_ent'}, 'VH_ent');
var alos = alos.addBands(ee.Terrain.slope(alos), ['slope'])
var composite = s2.addBands(s1).addBands(alos).addBands(VH_final).float()//.aside(print)
.reproject("EPSG:4326", null, 10)

//oytisa below
var visParams = {bands: ['B4','B3','B2'], min: 150, max: 2000}
//Map.addLayer(s2.clip(vlp_table), visParams, 'clouds masked')
Map.addLayer(composite.clip(oytAOI), visParams, 'clouds masked')
// Show the training area.
Map.addLayer(ee.Image().paint(oytAOI, 1, 2), null, 'AOI');
  /*******
  //go next for RF classification *
  /*******

//2016 classification
//Oytisa training points
var opf = ee.FeatureCollection ("users/oytisa/openForest2016")
var cf = ee.FeatureCollection ("users/oytisa/closedForest2016")
var bamboo = ee.FeatureCollection ("users/oytisa/bambooForest2016")
// separate into training and validation fractions
var training2016 = vlp_tr16.filter(ee.Filter.lt('random', 0.75));
var validation16 = vlp_tr16.filter(ee.Filter.gte('random', 0.75));
// 3. Create feature collection with spectral signatures from composite
var vlp_train = composite.sampleRegions({
  collection: training2016,
  properties: ['aoi2016'],
  scale: 10,
  tileScale: 16
});
// 4. Train a classifier
var rf2016 = ee.Classifier.randomForest(40)
  .train({
    features: vlp_train,
    classProperty: 'aoi2016',
    inputProperties: composite.bandNames()
  });
//Run the Classifier
var lulc2016 = composite.classify(rf2020);
var lulc2016vis = lulc2016.visualize(imageVisParam);
//Display the Classification
Map.addLayer(lulc2016.clip(oytAOI),
{min: 1, palette:'00b300', '006600', '33ff33', 'dfff80', 'e6e600','80ff80', 'e6ffe6', 'b3b300'}
/*['004d00','00b300','00ff00', 'fee08b', 'f46d43']*/},

```

```

'lulc2016_fusion');
//End 2016 classification

// 3. Create feature collection with spectral signatures from composite
var vlp_train = composite.sampleRegions({
  collection: training,
  properties: ['label'],
  scale: 10,
});

// 5. Apply classifier and generate classified image
var classified = composite.classify(classifier);

// 6. Assess accuracy of classification
var trainAccuracy = classifier.confusionMatrix().accuracy;
print('trainAccuracy', trainAccuracy);

// Use pre-classified imagery for accuracy assessment
var validationFc = composite.sampleRegions({
  collection: validation,
  scale: 10,
})
var confusionMatrix = validationFc.classify(classifier).errorMatrix('label', 'classification')
print(confusionMatrix)
print('testAccuracy', confusionMatrix ())
var accuracyAssessment = ee.FeatureCollection(
  [ee.Feature(
    'accuracy': confusionMatrix.accuracy()
  )]
);
Export.table.toDrive({
  collection: accuracyAssessment,
  fileNamePrefix: 'accuracyAssessment',
  fileFormat: 'CSV'
})

Map.centerObject(oytAOI, 13)
Map.addLayer(composite, min: 600, max: 3500}, 's1s2Alos_composite', true)
Map.addLayer(classified.clip(oytAOI.geometry()), imageVisParam, 'finalclassified', true)

//export image to google drive
Export.image.toDrive({
  image: composite.clip(oytAOI),
  description: 'shw2016_VIs1s2ALOS',
  folder: 'OytisaThesisGEE',
  crs: 'EPSG: 32637',
  scale: 10,
  region: oytAOI,
});

```


8.2 GEE computational time, memory usage and algorithms used during processing

Table 29: *GEE Sentinel-2 Image Computational Time, Memory and Algorithms*

Compute (Sec)	Peak Mem	Count	Description
36.618	1.7M	90	Algorithm ImageCollection.reduce with reducer Reducer.median
13.929	4.8M	1026	Algorithm Image.and computing pixels
8.265	1.3M	10686	Loading assets: COPERNICUS/S2/(...)
3.69	5.4M	35620	(plumbing)
1.252	1.4M	36	Encoding pixels to image
0.578	398k	4737	Algorithm Image.select
0.566	1.0M	18	Reprojection precalculation between EPSG:32637 and SR-ORG:6627
0.479	1.1M	57	Table query
0.359	816	4104	Reprojecting pixels from EPSG:32637 to SR-ORG:6627
0.349	1.2M	1579	Algorithm Image.updateMask
0.284	976k	273	Algorithm (user-defined function)
0.212	724k	72	Algorithm Image.paint computing pixels
0.077	648	3078	Algorithm Image.updateMask computing pixels
0.072	656k	300	no description available
0.066	2.4k	2052	Algorithm Image.bitwiseAnd computing pixels
0.058	1.5k	2052	Algorithm Image.eq computing pixels
0.05	127k	3158	Algorithm Image.bitwiseAnd
0.046	124k	3158	Algorithm Image.eq
0.034	40k	57	Table metadata
0.03	10k	288	Algorithm Image.clip computing pixels
0.024	65k	1579	Algorithm Image.and
0.023	544	432	Algorithm Image.visualize computing pixels
0.017	12k	226	Algorithm Image.visualize
0.015	464	144	Algorithm Image.constant computing pixels
0.01	5.2k	390	Algorithm Image.constant
0.009	14k	57	GeoTables transform
0.006	345k	111	Algorithm ImageCollection.reduce
0.004	9.4k	114	Algorithm Image.clip
0.004	7.1k	57	Algorithm Image.paint
0.003	800	22	Listing collection
0.002	5.5k	57	Algorithm Image.mask
0.002	5.4k	57	Algorithm Image.byte
0.002	69k	12	Algorithm Collection.geometry
0.001	200	36	Algorithm Image.byte computing pixels

0.001	240	36	Algorithm Image.mask computing pixels
0.001	6.8k	3	Algorithm Geometry.centroid
-	1.7M	126	Algorithm ImageCollection.reduce computing pixels
-	168k	244	Loading assets: COPERNICUS/S2
-	24k	333	Algorithm Collection.map
-	15k	120	Algorithm Collection.loadTable
-	14k	240	Loading assets: users/oytisa/shiwshiwe_bound
-	13k	111	Algorithm ImageCollection.load
-	7.6k	160	Loading assets: users/oytisa
-	7.4k	76	Loading assets: COPERNICUS
-	6.6k	108	Algorithm TypedImageCollection.Constructor
-	5.6k	111	Algorithm Filter.dateRangeContains
-	5.5k	108	Algorithm Filter.gt
-	5.5k	3	Algorithm Filter.lt
-	5.4k	111	Algorithm ErrorMargin
-	5.3k	3	Algorithm Feature
-	2.3k	37	Algorithm DateRange
-	1.2k	37	Algorithm Reducer.median
-	256	111	Expression evaluation
-	64	4104	Algorithm Image.load computing pixels
67.138			