UNIVERSITÀ DEGLI STUDI DI PADOVA

Department of Land, Environment Agriculture and Forestry

Second Cycle Degree (MSc)

in Forest Science

Integrating Sentinel-1 SAR Polarimetry for Mangrove Forest Biomass Estimation: A Case Study in West Kalimantan Province, Indonesia

Supervisor

Prof. Francesco Pirotti

Co-supervisor

Dr. Erico Kutchartt

Submitted by

Giusti Ghivarry

Student no.

2081370

ACADEMIC YEAR

2023/2024

TABLE OF CONTENTS

ABSTRACT

Mangrove forests play a crucial role in ecosystems by offering essential environmental services that benefit nearby communities. They act as cultural resources, offering ecotourism, education, and preserving cultural heritage. Moreover, mangroves safeguard the coastlines from waves and surges by absorbing their force and preventing erosion. Furthermore, they hold economic significance due to their ability to store blue carbon. Restoration initiatives can generate revenue through carbon markets, while conservation prevents potential carbon loss. Given the potential ecosystem services that mangrove forests can provide, the availability of reliable biomass information is crucial. Not only to support protection and conservation efforts but also to maintain the potential economic value of mangrove forest ecosystems. The remote sensing approach, especially with active systems such as Synthetic Aperture Radar (SAR), is an attractive option due to its ability to capture the Earth's surface regardless of the atmospheric condition. Nevertheless, the typical backscatter method loses its reliability as biomass levels increase beyond a certain threshold, resulting in incomplete data due to excessive signal strength that overwhelms the receiver.

This study emphasized the potential utilization of scattering mechanism parameters from polarimetric decomposition on Sentinel-1 to estimate mangrove biomass in West Kalimantan Province, Indonesia, in 2020. This study integrated the decomposition parameters of entropy, anisotropy, and alpha angle from H/A/α decomposition with backscatter parameters as dependent variables. Along with biomass information derived from the Global Ecosystem Dynamics Investigation (GEDI) as independent variables. Utilizing machine learning approach provided by H2O autoML, correlations were established between SAR parameters and GEDI-derived biomass to develop a predictive model. Then, various configurations of SAR parameters (both individually and in groups) were assessed using statistical metrics, including R^2, RMSE, MSE, and MAE. Additionally, comparisons were conducted with biomass prediction products from other projects to assess the model`s predictive accuracy. The information obtained will be useful for large-scale mangrove biomass mapping with an acceptable level of accuracy, which will be useful in mangrove forest protection and conservation activities.

LIST OF ACRONYM AND ABBREVIATION

A	Anisotropy
AGB	Aboveground Biomass
AGBD	Aboveground Biomass Density
AGV	Aboveground vegetation
ANN	Artificial Neural Network
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
C2	2 x 2 covariance matrix
CCI	Climate Change Initiative
ChL	Leaf Chlorophyll Content
DBH	Diameter at Breast Height
DEM	Digital Elevation Model
Dual-Pol	Dual Polarization
EOADP	Earth Observation Application Development Program
ESA	European Space Agency
GEDI	Global Ecosystem Dynamics Investigation
GEE	Google Earth Engine
GMW	Global Mangrove Watch
GPS	Global Positioning System
GRD	Ground Range Detected
H	Entropy
HP	Horizontal Polarization
HH	Horizontally Transmitted – Horizontally Received
HV	Horizontally Transmitted – Vertically Received
IDAN	Intensity-Driven Adaptive Neighborhood
ILS	Improved Lee Sigma
IMU	Inertial Measurement Unit
IPF	Instrument Processing Facility
ISS	International Space Station
IW	Interferometric Wide Swath
LAI	Leaf Area Index
LiDAR	Light Detection and Ranging

LULC	Land Use and Land Cover
MAPE	Mean Absolute Percentage Error
MMAF	Ministry of Marine Affairs and Fisheries
MoEF	Ministry of Environment and Forestry
NASA	National Aeronautics and Space Administration
ORNL DAAC	Oak Ridge National Laboratory Distributed Active Archive Center
Quad-Pol	Quad Polarization
R^2	Coefficient of Determination
RF	Random Forest
RF-AGA	Random Forest Adaptive Genetic Algorithm
RMSE	Root Mean Squared Error
RS	Remote Sensing
SAR	Synthetic Aperture Radar
SLC	Single Look Complex
SRTM	Shuttle Radar Topography Mission
SVM	Support Vector Machine
T3	3 x 3 coherency matrix
TLS	Terrestrial Laser Scanning
TOPS InSAR	Terrain Observation with Progressive Scans Interferometric SAR
TOPSAR	Terrain Observation with Progressive Scans SAR
VP	Vertical Polarization
VH	Vertically Transmitted – Horizontally Received
VV	Vertically Transmitted – Vertically Received
α	Alpha Angle

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

1.1.Background

Mangrove forests are one of the most important ecosystems because they can provide environmental services that support the lives of local communities. Culturally, the mangrove forest provides benefits to the surrounding community from its tangible and intangible components that are interconnected with each other [1]. These include ecotourism, education, cultural identity, and heritage. And despite the potential negative influences or "ecosystem disservices" of mangroves, local communities are more likely to gravitate towards the benefits offered as their well-being solely depends on mangroves. On the other hand, mangrove ecosystems also have a protective function in coastal areas against waves and surges, both directly and indirectly [2]. The components of mangroves, such as their root systems directly intercept and decelerate the force of waves and surges while providing soil stabilization and erosion control in coastal areas. And while mangroves offer the potential for tsunami protection, an accurate assessment also requires considering factors beyond their influence.

Promoting mangrove restoration and conservation can not only support local well-being culturally and physically but also economically by optimizing the ecosystem's potential. Based on a report from the World Bank [3], mangrove forests in Indonesia could generate revenue through carbon markets due to their effectiveness in storing and capturing blue carbon. Numerical measurements showed that mangrove restoration can reduce CO_2 emissions by up to 38.5 ± 12.7 tCO_2e per year per hectare. In addition, mangrove conservation also contributed by avoiding the potential loss of 25% of mangrove carbon stocks, leading to an emission reduction of approximately 271 tCO_2e per year per hectare. Blue carbon conservation is attractive because it is cost-effective and offers co-benefits such as fish habitat, water filtration, and protection from storms [4]. However, Indonesia's lax regulations allow the destruction of mangrove forests, hindering emission reduction efforts. Nonetheless, Indonesia still has the potential to participate in international initiatives that trade carbon credits from these ecological assets [3].

Aside from illustrating the economic potential of mangrove forests, aboveground biomass (AGB) estimation and mapping practices play a crucial role in supporting conservation and protection efforts. Following the development of technology, remote sensing approaches have become an optional tool given the harsh conditions of mangrove ecosystems that pose

obstacles to obtaining the necessary information. However, according to [5], mangrove biomass estimation studies using remote sensing approaches still lagged compared to terrestrial forests. One reason was the limitation of suitable remote sensing (RS) data to identify the required biophysical parameters, such as the leaf chlorophyll content (ChL) and the leaf area index (LAI). Indeed, the biophysical structure of mangroves can be estimated with remote sensing data. However, accurate results can only be obtained with the help of fine-scale data, whose availability is limited.

Recent studies have tried various RS data sources and methodologies to estimate mangrove biomass, and active systems such as SAR are an attractive option. According to [6], the sensitive sensor response of active systems to surface roughness and dialectic properties helped in assessing mangrove biophysical parameters. Revealing the correlation between backscatter coefficients and structural features enables precise characterization and estimation of diverse mangrove species. Despite that, the relationship between backscatter and aboveground vegetation (AGV) becomes difficult to identify after reaching the saturation point, as supported by the result in [7]. This caused the reflected radar signal to be too strong, exceeding the dynamic range of the SAR receiver, thereby clipping the signal and causing information loss. Alternative approaches to addressing this challenge have been explored, one of which is applying the polarimetric decomposition method. Research conducted by [8], [9], and [10] demonstrated the potential of the polarimetric decomposition method to estimate AGB in forest areas with no obvious saturation tendency. Outlining the potential of this approach for forestry applications, particularly in estimating AGB within forested areas. In this study, the potential of scattering mechanism parameters from the polarimetric decomposition method on Sentinel-1 data to estimate biomass density in mangrove forests is discussed. Focusing on the assessment of the model's capability to estimate mangrove forest biomass as well as the identification of the saturation tendency of the parameters with respect to the variation of the biomass values, the results are presented.

1.2.Problem Statements

Mangroves not only play an important role in the sustainability of the ecosystem but also in the welfare of the surrounding communities, and their conservation and protection will be beneficial to maintaining the sustainability of this ecosystem. Such efforts can be made efficiently if accurate information is available. Remote sensing approaches, especially with active systems, are an attractive option for mangrove biomass mapping. Active systems

transmit their signal to the Earth's surface and then record the reflection from a receiver. Coupled with longer wavelengths, these systems can capture the Earth's surface regardless of weather conditions and time of day.

Traditionally, active remote sensing systems utilize the amplitude of backscatter captured by the receiver to estimate the level of biomass on the Earth's surface. However, this approach has the drawback that when the backscatter is too strong and exceeds the capacity of the receiver, the recorded signal becomes clipped and thus "saturated". This leads to incomplete information being obtained.

Another alternative is to use the scattering mechanism parameters from the polarimetric decomposition method instead of signal backscatter to estimate the biomass. Studies focusing on this approach have shown the potential to estimate biomass without a clear saturation tendency, such as the study conducted by [8], [9], and [10]. However, existing studies used quad-polarization SAR data, which is not freely available to the public. Furthermore, studies that utilized dual-polarization SAR data, which are more commonly available to the public, still mostly focus on agricultural studies, such as research by [11], [12], and [13]. The study conducted by [14] focused on estimating biomass but not on the mangrove ecosystem. While other studies, such as those carried out by [15] and [16], were carried out in mangrove forests, the objective was not to estimate biomass.

1.3.Research Questions and Objectives

This study attempted to build a biomass density estimation model for mangrove forests in the coastal area of West Kalimantan Province, Indonesia, during the year 2020. To do so, scattering mechanism parameters from the polarimetric decomposition method of Sentinel-1 imagery have been integrated into the model development. The use of these parameters, which are based on phase information instead of amplitude, is an attempt to identify this alternative approach's capability for mangrove forest biomass estimation. Based on these statements, the research questions that need to be addressed are described below.

1) To what extent can scattering mechanism parameters from the polarimetric decomposition method be integrated with the GEDI-derived biomass estimation model over mangrove forests?
2) How well do the scattering mechanism parameters quantify the variance of biomass information derived from GEDI in the produced model compared to the backscatter parameters?

3) How does the predictive capability of the developed model incorporating scattering mechanisms compare to existing models from similar projects?

To answer the research questions above, the following objectives have been developed:

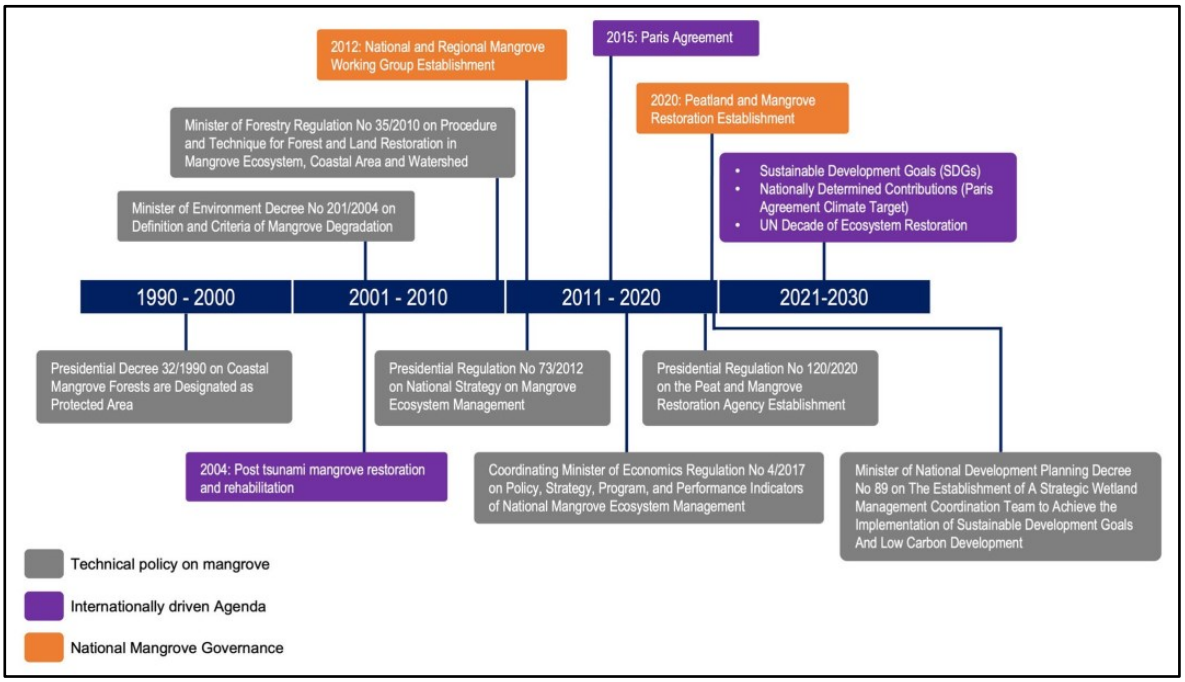
1) Develop a mangrove forest biomass density estimation model based on backscatter and scattering mechanism parameters from Sentinel-1 in correlation with biomass information derived from GEDI, both individually and simultaneously.
2) Identify the effectiveness of scattering mechanism parameters and backscatter parameters in capturing the variability of biomass information derived from GEDI data within the biomass estimation model, both visually and from statistical metrics.
3) Compare the predictive capabilities of the developed model incorporating scattering mechanisms with existing biomass estimation models employed in similar studies.

LITERATURE REVIEW

2.1.National Strategies in Indonesia Mangrove Conservation

According to [17], Mangrove forests are unique ecosystems found on tropical and subtropical coastlines. Despite their simpler structure compared to rainforests, mangrove trees have adapted to survive in saltwater and even reproduce with the help of tides. These adaptations allow them to become vital ecosystems that can provide physical, social, and economic services to local communities. Therefore, mangrove forest protection and conservation efforts are important to maintain its sustainability. Driven by a growing awareness of the threats to mangrove ecosystems, international agreements have been established to promote their conservation and sustainable use [18]. These agreements play an important role in protecting vast expanses of mangrove forests worldwide. In addition, individual countries have undertaken various conservation efforts, which effectively complement these international frameworks. According to [19], Indonesia formally adopted mangrove conservation in 1990, a commitment that evolved after the 2004 tsunami when mangroves were replanted to restore coastal defences, though with limited success. Subsequent laws in 2007 aimed to better manage coastal areas and safeguard mangroves, culminating in a 2012 national strategy engaging diverse stakeholders to rehabilitate 3.49 million hectares by 2045. However, in 2020, a shift occurred with the establishment of the Peatland and Mangrove Restoration Agency (BRGM), aiming for a more centralized approach and a bold goal of rehabilitating 600,000 hectares by 2024.

Figure 1: Graph by Sasmito et al. (2023) describing Indonesia's mangrove protection efforts and policies from 1990 to 2030, as well as related global environmental goals [19].



Indonesia conserves mangrove ecosystems through a national strategy and regulations overseen by the Ministry of Environment and Forestry (MoEF) and the Ministry of Marine Affairs and Fisheries (MMAF) [20]. Where its conservation efforts not only target the protection of biodiversity, carbon stocks, and coastlines but also emphasize community involvement through ecotourism in mangrove management. However, a study by [19] conducted in 2021 revealed that Indonesia has 193,367 ha of restorable mangrove land, but 75% was outside protected forests, complicating restoration. This area represents just 30% of the government's target (600,000 ha), categorized by opportunity levels (high, medium, and low) based on land use, ownership, and future carbon storage potential. According to [21], mangrove forests also face environmental challenges due to complex interactions with a variety of wildlife, not all of which are beneficial. Pests such as crabs, insects, snails, and barnacles can damage and kill mangrove plants, especially at the seedling stage, potentially reducing biodiversity. Furthermore, anthropogenic marine debris (AMD) in a form of solid waste such as plastic and fabric is also a major threat to the health of mangroves in Indonesia. Nonetheless, there is an opportunity for Indonesia to become a leader in mangrove

conservation by demonstrating its contribution to the international environmental agenda [19].

2.2.Technological Advancement for Mangrove Forest Mapping

Remote sensing has become an optional approach in mangrove forest mapping efforts, given its effectiveness in obtaining the necessary information. With passive sensors, mangroves can be distinguished by analyzing the textural and spectral characteristics of their canopy and leaves [22]. Conversely, in active sensors, especially in SAR systems, the internal and external properties of the plant are measured based on variations in its backscattering coefficient. This makes interpretation more complex but offers additional insights into plant structure and moisture content, highlighting the importance of accurate and detailed data for effective mangrove management.

Effective conservation strategies rely on accurate and up-to-date information on mangrove distribution and health. Aligned with this view, research by [23] employed advanced remote sensing to map global mangroves at 30-meter resolution. Motivated by the limited knowledge of global mangrove distribution at the time due to inaccurate and low-resolution satellite data, inconsistent sources, and fragmented local studies. This map provides a more accurate picture of mangrove distribution, offering a valuable tool to protect and conserve these critically important ecosystems. Following the development of technology, global mangrove distribution datasets have begun to be improved with the help of more sophisticated methods. In 2018, a global basemap of mangrove extent from Earth Observation data for 2010 was released under the Global Mangrove Watch initiative [24]. The resulting map was highly accurate and had an improved level of detail, down to approximately 1 ha. And due to the autonomous and reproducible nature of the methodology utilized, it allows for the development of a mangrove land monitoring system. Moreover, utilizing the multi-temporal recording capability of remote sensing data, the study by [25] developed a multi-temporal global mangrove distribution dataset. With the temporal coverage of the dataset from 1996 to 2020, divided into several epochs, the dynamics of the mangrove ecosystem can be identified. A similar attempt was made by [26], resulting in a 10-meter resolution map of global mangrove distribution in 2020, namely High-resolution Global Mangrove Forests. The resulting map was an improvement over the existing dataset, utilizing object-based analysis and random forest classification on Sentinel-2 imagery. This

dataset allowed for a more comprehensive analysis of the conservation and threat status of mangrove forests.

According to [5], remote sensing approaches for mangrove mapping have been applied since 1956. Over time, technological advances have enabled the extraction of more detailed information, including thematic data such as carbon stock and biomass estimates. One example is the study conducted by [27], which integrated unmanned aircraft platforms (UAVs) with satellite imagery to estimate the aboveground biomass (AGB) of mangrove forests. This study innovated by combining satellite data with low-cost drone surveys for accurate, large-scale monitoring of young mangrove forests. Another innovation is to utilize cloud computing, as has been done in the study by [28]. Who proposed a tool called Google Earth Engine Mangrove Mapping Methodology (GEEMMM) to help track and protect mangrove forests. This tool has been tested in Myanmar's mangrove forests, providing accurate results, and can be used to track changes in mangrove cover over time.

2.3.Polarimetric Decomposition Utilization of Quad-Pol Data

The initial focus was to evaluate the polarimetric decomposition method for quad-pol data. A study by [8] evaluated the polarimetric parameters of Quad-Pol C-bands Gaofen-3 and RADARSAT-2 as well as L-bands ALOS-2/PALSAR-2 for AGB estimation over two forest sites. The model revealed an improvement in the accuracy of AGB estimation compared to traditional backscatter coefficients. Overall, the best results were the leaf component at site one, producing R^2, RMSE, and rRMSE values of 0.64, 1.27 t/ha, and 28.84% without obvious saturation tendency. As for total AGB, the model from site one resulted in higher accuracy with an rRMSE value of 34.38%, compared with the model from site two with an rRMSE value of 43.85. A similar study also conducted by [10], but a random forest adaptive genetic algorithm (RF-AGA) model with feature selection was used instead of the regression model of the above-mentioned study. The model, which was applied in coniferous forests, produced AGB estimates without obvious saturation, producing R^2 and RMSE values of 0.93 and 10.42 t/ha. A study by [29] also utilized machine learning concepts, especially random forest (RF) and artificial neural network (ANN), in model generation. However, this study used sample plots of diameter at breast height (DBH) and tree height obtained from point clouds generated using a Terrestrial Laser Scanner (TLS). The model was evaluated in the Barkot Forest area of India and gave the best results in RF, with an R^2 of 0.94 and an RMSE of 59.72 t/ha. Moreover, parameters from Yamaguchi decomposition resulted in the best

correlation with R^2 values of 0.05, 0.55, and 0.2 for surface, double-bounce, and volume scattering mechanisms.

A study by [9] evaluated the performance of decomposition parameters on ALOS-2 in the forest of North China. Produced an R^2 value similar to the study [29], with the parameters from the Yamaguchi decomposition showing a relatively stronger correlation compared to the other. The model produced from decomposition parameters also has a higher saturation point compared to the model from the backscatter coefficient. Moreover, the final model produced from the RF algorithm with combined input also showed R^2 and RMSE values of 0.90 and 14.24 t/ha. A study by [30] also evaluated the parameters of Yamaguchi decomposition, especially on C-band RADARSAT-2 and X-band TerraSAR-X over boreal forests. The resulting model had R^2 and RMSE values of 0.52 and 58.1 t/ha for C-bands and 0.47 and 60.6 t/ha for X-bands. In addition, the C-Bands model did not show a saturation tendency, whereas there was a clear saturation tendency in the X-Bands model. In addition, the relationship between in-situ biomass data and decomposition parameters derived from Pauli and H/A/α decomposition in ALOS/PALSAR over tropical deciduous forests was also investigated by [31]. The resulting biomass prediction generated from the exponential regression for the parameters above showed a moderate correlation of 0.56 with an RMSE of 29.13 t/ha.

2.4.Polarimetric Decomposition Utilization of Dual-Pol Data

The previous explanation focused on applying the polarimetric decomposition method to quad-pol data. Given the limited access, an evaluation of this method against publicly available data has also been carried out. A study by [14] utilized the entropy and α parameters of the H/A/α decomposition (entropy, anisotropy, alpha angle) together with other parameters in the AGB estimation in the Brazilian Caatinga biome. This resulted in the optimal model during intermediate phenological periods (vegetation transitions from green to dry) with R^2 and RMSE values of 0.72 and 8.33 t/ha. Followed by the green period (the vegetation is green and active) with an R^2 of 0.72 and an RMSE of 8.40 t/ha. The same methodology was also applied in the report by [32] but focused on the impact of various polarimetric filters on the AGB estimation results. The best estimation results were achieved by combining entropy from IDAN filtering with α from ILS, resulting in an R^2 value of 0.32.

Although studies on forest biomass estimation using polarimetric decomposition of dual-pol SAR data are still limited, the methodology itself has been evaluated in other fields. A study

by [11] utilized Sentinel-1 data covering Uttarakhand, India, for classification with scenarios such as agricultural land, forest cover, urban structure, and snow cover. The results showed that the reflected polarized signal differs from the broadcast signal due to orientation effects, which may be applicable to land use and land cover (LULC) and biophysical parameter estimation. Utilizing the same data, a study by [12] employed the entropy and α parameters, then used K-means unsupervised classification for optimal target decomposition. The classification revealed typical dispersal mechanisms at various growth stages of plants such as ginger, tobacco, rice, cabbage, and pumpkin. Also, temporal VH polarization backscattering analysis consistently aligns with the H-α plan, offering valuable insights into dynamic crop growth scattering mechanisms. The H/A/α decomposition of the multi-temporal Sentinel-1A dataset was also evaluated in a study by [33] to monitor important biophysical parameters of rice plants in the Anand and Kheda Region, India. Identified the correlation, especially from Primordia and Booting Stages, with a maximum R^2 of 0.6 for entropy and vegetation water content. In addition, the increase in correlation through multivariate regression can also be seen from the combination of the backscatter coefficient and decomposition parameters. A similar approach was evaluated by [34] to monitor the development of wheat and barley crops in North-East Germany. Wheat plant height showed the highest R^2 values for entropy and anisotropy of 0.64 and 0.61, respectively. In contrast, the multivariate regression model consistently produced better R^2 values for plant height, wet and dry biomass, and vegetation water content. Meanwhile, a study conducted by [13] showed a lower level of correlation in the biophysical parameters of cotton and peanuts. Entropy was most correlated with cotton plants height and peanut leaf area index (LAI), with R^2 values of 0.43 and 0.49. Meanwhile, α showed the same value for both plants in the H- α plane, thus posing a challenge in distinguishing the two.

Moving on to the application of the H/A/α decomposition method, several studies have been carried out. Entropy and α have been used in a study by [35] as input to a support vector machine (SVM) to differentiate secondary forests in the Amazon Forest. The classification resulted in a Kappa index of 0.70 and an overall accuracy of 79.58%, outperforming the texture features and COSMO-SkyMed (CSM) backscatter coefficients used as a comparison. However, in dense mangrove forests, the entropy parameters serve as a strong indicator but may not capture certain scattering phenomena, such as double reflections. This argument was based on a study by [15], which explained the α effectiveness of emphasizing dense mangrove zones with volume scattering and medium and low entropy values. For other

applications, such as the classification of wetlands in China's Poyang lake, it has been carried out by [36]. A decision tree (DT) classifier was used with the backscatter coefficient and decomposition parameters as input, either individually or combined. The best result was generated from combined parameters of 77.29%, but the forest class showed lower producer and user accuracy (57.80% and 45.07%). A similar occurrence also happened in the classification of mangrove tree species in South Bali, Indonesia, as reported by [16], who used RF as a classifier. The highest accuracy at the genus level was demonstrated by combined parameters from H/A/α decomposition and the dual-polarization radar vegetation index (DpRVI) of 65%. Meanwhile, at the species level, the highest overall accuracy was obtained from the combination of backscatter coefficient and H/A/α decomposition parameters of 57%.

2.5.Applicability of GEDI Data for Forest Biomass Estimation

Biomass research in forest areas is also enriched by the potential of the Global Ecosystem Dynamics Investigation (GEDI) as additional data. Utilization of this product can help the process of mapping and estimating biomass in the mangrove ecosystem. Another reason for the delay of this study compared to that in terrestrial forests. Specifically, the lack of reliable ground truth data is due to the harsh condition of the mangrove ecosystem, as explained in [5]. According to [37], GEDI is a NASA project that involves the use of laser devices on the International Space Station. The aim is to measure the three-dimensional structure of forests and other ecosystems. The information generated by this instrument was then reprocessed into a variety of products, one of which was biomass.

GEDI has been used for various studies, one of which was by [38], which built a forest inventory dataset over France in 2020 named FORMS. The deep learning model was applied to GEDI data together with Sentinel-1 and Sentinel-2 imagery to generate high-resolution maps of forest canopy height (FORMS-H) and aboveground biomass (FORMS-B). Validation showed better performance compared to independent data sets, with errors less than 3 meters for height and below 60 Mg per hectare for biomass. Besides for forest inventory purposes, GEDI data has also been used for more specific thematic topics, such as in the research conducted by [39]. Where the GEDI data was combined with Landsat imageries to calculate biomass loss in charcoal-related degraded forests from 2007 to 2019 in Southern Mozambique. The RF model produced an R^2, RMSE, and rRMSE values of 0.62, 7.05 Mg/ha, and 42% in a 10-fold cross-validation dataset. In addition, total AGB loss over

10 years was calculated at 2.12 ± 0.06 Megatons, 6.35 ± 2.56% less than field-based estimation. This framework highlighted that GEDI and Landsat data fusion through predictive modeling can quantify the AGBD dynamic of low biomass forests.

RESEARCH MATERIALS AND METHODOLOGY

3.1.Study Area

Indonesia is an archipelago that is home to mangrove ecosystems of varying extents [40]. These mangrove forests thrive along the coastlines and estuaries of major islands, including Java, Sumatra, Kalimantan, Sulawesi, and Papua, forming diverse ecosystems. Based on calculations from Global Mangrove Watch, mangrove habitat in Indonesia by 2020 was ~ 2.95 million ha [25]. Meanwhile, calculations based on satellite image interpretation by the MoEF indicate that the area of mangrove forests in 2021 was ~3.36 million ha [41]. Of which ~27% was classified as protected forest, ~22% was classified as conservation forest, and the rest was classified as production forest or other land use areas. Additionally, potential mangrove habitat was also identified at ~760,000 ha, resulting in a total area of identified mangrove habitat in 2021 of ~4.1 million ha.

This extensive mangrove coverage plays an important role in supporting Indonesia's high biodiversity, which includes a variety of species uniquely adapted to this environment. According to a study conducted by [20] in 2024, Indonesia has high biodiversity with 240 identified species, most of which are found on Java Island. This is inversely proportional to the number found on Maluku Island, but this could be due to the majority of surveys being conducted on Java Island. Additionally, 48 of the 240 species were identified as true mangrove species, including trees, shrubs, herbs, ferns, and even palm species. These included species such as *Rhizophora spp.*, *Avicennia spp.*, and *Sonneratia spp*. [20].

The biomass estimation model was carried out over the mangrove forest area in the west of West Kalimantan province, specifically in the regencies of Kubu Raya and North Kaong. According to the classification scheme by [19], this site belongs to the medium and low classes of potential restoration areas. Moreover, based on calculations from Global Mangrove Watch data [25], the mangrove forest used in this study has an area of ~ 90,000 ha. The mangrove forest covers the coastline of Kubu Raya and North Kayong Regency. This includes the mangrove forest on Padangtikar Island in Kubu Raya Regency and the

mangrove forest on Maya Island in North Kaong Regency. The study site in question is shown on the map below.

Figure 2: Location of study sites (highlighted with red circles) within the geographic distribution of areas with potential restoration mapped at 1° grid scale map courtesy of a study by [19].

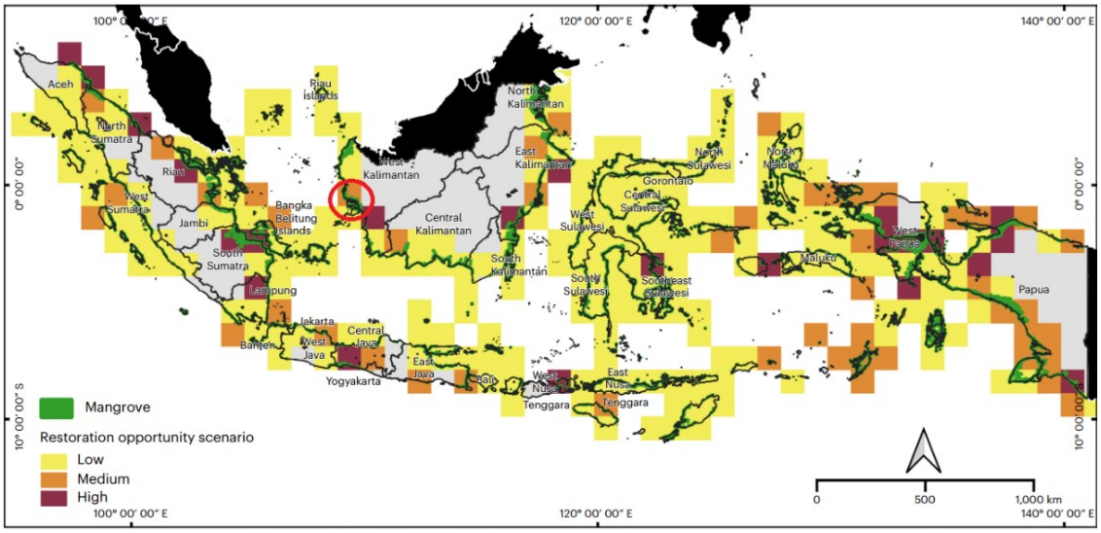
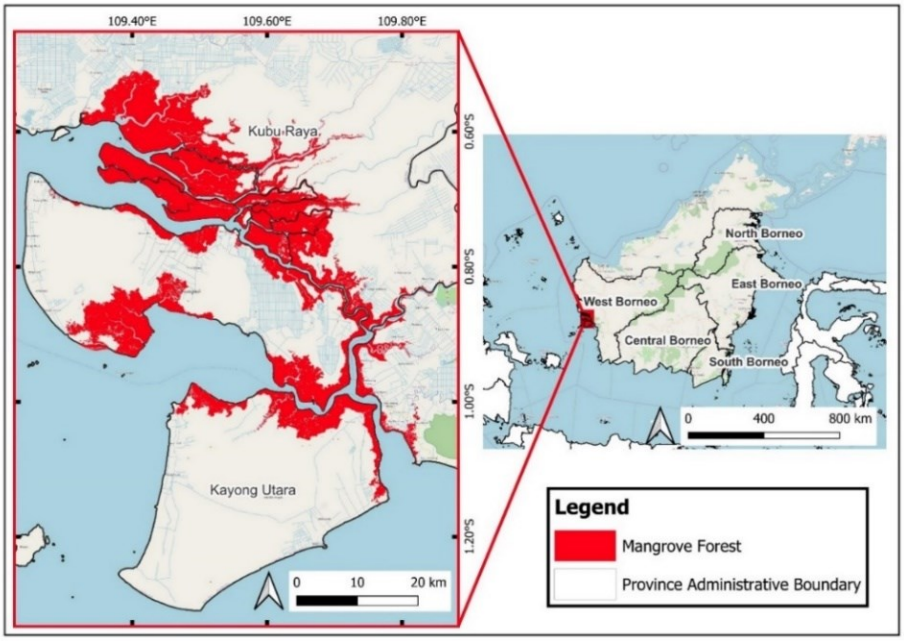


Figure 3: Map of the study site, located in Kubu Raya District and Kubu Raya Regency, West Kalimantan Province, marked with a red box. The study site covered not only the coastal area of the main island but also parts of the coast of two smaller islands (Padangtikar Island in Kubu Raya District and Maya Island in North Kaong District).



3.2.Tools

To support the research process, several software and platforms have been used. The purpose of using several software/platforms was to maximize the processing and analysis of the data. SNAP is a software that focuses on satellite image processing, developed, and published by the European Space Agency (ESA). This software was used to process Sentinel-1 L1 SLC data to extract decomposed polarimetric parameters. It was assisted by the Radarsat Polarimetric Toolkit Module plugin developed by the Canadian Space Agency Earth Observation Application Development Program (EOADP). In addition, the Google Earth Engine (GEE) platform, which is an online-based platform for processing and analyzing geospatial data developed by Google, has also been used [42]. This platform was used for pre-processing geospatial data such as Sentinel-1 L1 GRD and GEDI L4A. Finally, Finally, the package "h2o" developed by H2O.ai which is available in the RStudio software has also been used [43]. This program was used for model building and assessing its predictive capabilities as well as the correlation between AGB and radar parameters.

3.3.Materials

In this study, mangrove forest biomass in Kubu Raya and Kayong Utara Regency, West Kalimantan, was estimated using remote sensing data. The remote sensing data used were Sentinel-1 Level-1 synthetic aperture radar (SAR) images with Single Look Complex (SLC) and Ground Range Detected (GRD) formats. To enrich the data sources for estimation purposes, several images acquired during 2020 have been used. Sentinel-1 SAR imagery has advantages in terms of its ability to penetrate clouds and vegetation, so it can be used to obtain information on mangrove forest biomass.

Samples for model training and validation were obtained from the Global Ecosystem Dynamic Investigation (GEDI) Level 4A (L4A) Raster Aboveground Biomass Density (AGBD) Version 2.1. This laser-based GEDI data is able to show the three-dimensional structure of vegetation, which can be further processed to obtain biomass information. In addition, a mangrove forest boundary map from Global Mangrove Watch (GMW) was used, which aims to determine the boundaries of mangrove and non-mangrove forests. This map has a high accuracy, so it can be used to ensure that the training and validation samples are in mangrove forest areas. In addition, other estimation models, such as the Climate Change Initiative (CCI) Biomass Map by ESA, were also used as a comparison with the estimation

model produced. A brief explanation of the data used and the purpose of using the data is shown in Table 1 as follows:

Table 1: Recapitulation of the utilized geospatial data

Source	Level	Pixel Size (m)	Description
Sentinel-1 L1 SAR	SLC	15	Polarimetric Decomposition parameter extraction
Sentinel-1 L1 SAR	GRD	10	Backscatter Parameter extraction
GEDI	L4A AGB Density	25	Lidar base AGBD information
GMW Mangrove Extent	Yearly Mosaic	10	Mangrove forest extent
CCI Biomass Map	Yearly Mosaic	100	Global-scale AGB map for comparison

The explanation of the utilized data described previously is only an overview, and for a complete description of the utilized data, it is explained as follows:

3.3.1. Sentinel-1 Level-1 Single Look Complex (SLC) and Ground Range Detected (GRD) Format

Sentinel-1 Level-1 SAR was the remote sensing image used in this study to build the mangrove forest biomass estimation model. According to the Sentinel-1 user guide provided by ESA [44], Sentinel-1 is an imaging radar mission that provides continuous imagery around the clock, both day and night. With its sun-synchronized orbital system and polar proximity, recording can be done in repeated cycles of 12 days in one pass (ascending or descending). In addition, the Sentinel-1 has four main acquisition modes, one of which is the Interferometric Wide swath (IW) mode. IW is the main way of acquiring land and fulfills most of the service requirements. This method makes it possible to acquire data with a large swath width of 250 kilometers with a moderate spatial resolution of 5 x 20 meters in a single view. Additionally, IW uses the Terrain Observation with Progressive Scans SAR (TOPSAR) technique to capture three sub-plots. Where the radar beam is electronically controlled in range and azimuth direction during each burst transmission, resulting in homogeneous image quality across the swath and better radiometric resolution [44].

Also based on the same user manual [44], the Sentinel-1 uses a single C-band synthetic aperture radar (SAR) instrument to image the Earth's surface. With a center frequency of 5.405 GHz and a wavelength of about 5.56 cm, the C-band offers a good balance between spatial resolution and penetration depth compared to other bands such as the L-band or X-band. In addition, the Sentinel-1 can send and receive radar signals in two different orientations or polarizations (HH+HV, VV+VH). This is possible because its design has one switchable transmission chain and two separate receiver chains to receive signals in Horizontal (H) and Vertical (V) polarizations. For its distribution, it is carried out at three levels, one of which is level-1. Level-1 data is the basic product generated from raw instrument data (Level-0) by the Instrument Processing Facility (IPF). Which can be further processed into SLC or GRD formats, offering different levels of processing and extraction of thematic information [44].

The development of the mangrove ecosystem biomass estimation model was carried out using two file formats of Sentinel-1 Level-1 SAR images, namely SLC and GRD formats. These two data formats were obtained from different platforms, and the processing process was also different. Sentinel-1 SLC was obtained from ESA's Copernicus Browser website, while Sentinel-1 GRD is available in the Google Earth Engine (GEE) dataset catalog. According to [42], this platform can provide a geospatial data repository of up to petabytes of information, all pre-processed to reduce the complexity associated with data management. In addition, the catalog is continuously updated with new data from ongoing missions and user contributions, ensuring the latest information is always available.

According to the user manual for Sentinel-1 L1 SLC [45], the SLC format file is a product based on the tilt and azimuth ranges during data acquisition. Which is processed to produce a single view in each dimension and full signal bandwidth. Each pixel of this product contains amplitude and phase information, represented by complex values in the I and Q bands with a value range of 16 bits. Different characteristics are seen in other Sentinel-1 SAR L1 format files that have been used, namely the GRD format. According to the user guide for Sentinel-1 L1 GRD [46], this format was obtained by applying detection, multi-looking, and projection to the ground surface area using an earth model. Resulting in an image with square-shaped pixels and relatively even resolution across the image due to multi-looking processing that represents only the signal strength. Moreover, there are several distribution modes provided for both data, but for this study, the IW distribution format has been used. One IW Product consists of several small images called sub-swaths, constructed

from smaller processing units called bursts [46]. The technical specifications of both formats are shown in Table 2 as follows:

Table 2: Technical properties of the Sentinel-1 SAR Level-1 Interferometric Wide Swath product for SLC and GRD formats

Image Properties	SLC Format			GRD Format		
	IW1	IW2	IW3	IW1	IW2	IW3
Spatial Resolution (rg x az) (m)	2.7x22.5	3.1x22.7	3.5x22.6	20.4x22.5	20.3x22.6	20.5x22.6
Pixel spacing (rg x az) (m)	2.3x14.1	2.3x14.1	2.3x14.1	10x10		
Incidence angle (º)	32.9	38.3	43.1	32.9	38.3	43.1
Pixel value	Complex			Magnitude detected		
Bits per pixel	16					
Polarization options	Single (HH or VV) or Dual (HH+HV or VV+VH)					
Ground range coverage (Km)	251.8					

Source: Sentinel-1 SAR User Guide (https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-1-sar)

3.3.2. Global Ecosystem Dynamics Investigation (GEDI)

According to [37], NASA launched GEDI in 2018, a mission aiming to use the LiDAR Instrument on the International Space Station (ISS) to measure the forest ecosystem's three-dimensional structure. The GEDI instrument employs lasers that are fired 242 times per second, with one laser split into two beams for a total of four beams on the ground. The instrument can achieve high-density data collection by rapidly shifting the light beams to create eight data paths spaced 600 meters apart, covering a total swath of 4.2 kilometers. The LiDAR beam records the reflected laser energy as a function of height, enabling the extraction of vegetation structure information. And signal processing identifies the ground surface as well as analyzes the distribution of laser energy within the canopy, determining canopy height and density even in dense forests. GEDI offers detailed three-dimensional earth data with varying levels of processing (levels 0 to 4) in both individual footprint and summarized grid formats, details of which are shown in Table 3 [37].

For this study, the GEDI04_A product, which provided biomass density information in footprint form, was used. According to [47], the GEDI04_A provides an estimate of

aboveground biomass density (AGBD) in Mg/ha for each laser footprint sampled by the GEDI instrument. The data covers the recording period from 2019 to 2023 in the International Space Station (ISS) observable latitude (51.6° N and S) strip. And according to [37], to ensure accurate location data, GEDI employs a combination of GPS, the Inertial Measurement Unit (IMU), and three-star trackers to achieve a geolocation accuracy of 10 meters. The theoretical foundation of the algorithm for obtaining AGBD information from GEDI is explained in [48]. Here, the GEDI04_A uses the GEDI02_A and 13 models to predict AGBD for different crops and regions observed by the ISS. These models were built using data from more than 8,500 simulated GEDI pulses and field measurements of AGBD in 21 countries. This process enabled the GEDI04_A to provide estimates of AGBD on a global scale for each of the areas covered by the GEDI.

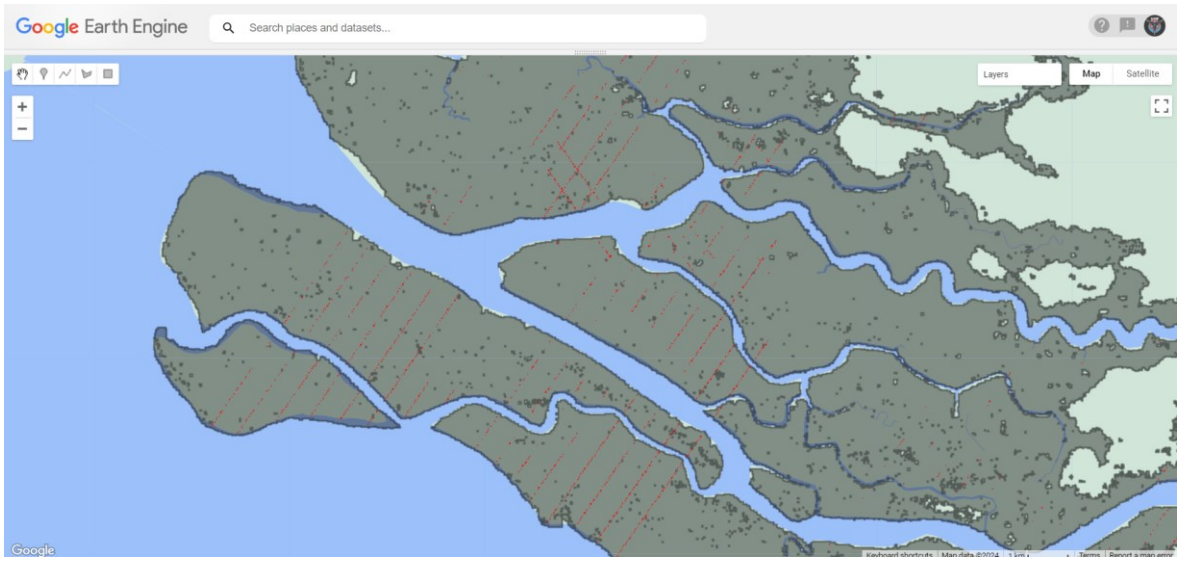
Table 3: The description of the available GEDI data products with highlighted rows indicates the products that have been used in this study.

Product/ Data file	**Description**	**First Data Delivery After Initial On-orbit Checkout**	**Data Latency**
GEDI01_B	Level 1B Geolocated Waveforms and Fitted Parameters	First 2 months of L1B released at 6 months	4 months in monthly intervals
GEDI02_A	Level 2A Footprint Elevation and Height Metrics	First 2 months of L2A released at 6 months	4 months in monthly intervals
GEDI02_B	Level 2B Footprint Cover and Vertical Profile Metrics	First 2 months of L2B released at 6 months	4 months in monthly intervals
GEDI03_A	Level 3A Gridded Land Surface Metrics	Populated with first 2 months of L2 data at 6 months	4 months in monthly intervals
GEDI04_A	Level 4A Footprint Aboveground Biomass Density	First 12 months of data at 17 months	6 months after global sampling required to meet L1 requirement
GEDI04_B	Level 4B Gridded Aboveground Biomass Density	First 12 months of data at 17 months	6 months after global sampling required to meet L1 requirement

Source: R. Dubayah et al. 2020 [37]

While the GEDI can be accessed through NASA's Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC) data archive. This study used the GEDI Level-4A Raster Aboveground Biomass Density (AGBD) Version 2.1 obtained through 2020, which is available in the GEE data catalog. According to [49], this dataset provides AGBD predictions and standard error estimates for each geolocated raster footprint with a spatial resolution of 25 meters. The AGBD information is also expressed in units of Mg/ha and organized into monthly composites. The GEDI Level-4A Raster Aboveground Biomass Density (AGBD) Version 2.1 data display is shown in Figure 4 as follows:

Figure 4: View of the GEDI L4A Raster AGBD V2.1 data over GMW`s mangrove extend (represented in grey polygon) used as study site, shown with zoom level 13. The straight line of red dots represents the footprint of the GEDI L4A Raster AGBD V2.1.



Source: Google Earth Engine

The GEDI Level-4A data has been assessed for accuracy and influencing factors by [50]. The assessment was conducted using data from the National Ecological Observatory Network (NEON), which showed a tendency to underestimate the AGBD with moderate errors in most areas. Furthermore, the simulated waveform strategy deviation (SWSD) used in GEDI L4A had the greatest influence on accuracy compared to other factors, such as forest type and vegetation complexity. The performance of GEDI Level-4A has also been assessed using measured national forest inventory (NFI) data and aerial laser scanning surveys (ALS) over a large area in Spain [51]. While the model performed adequately, using forest type

(FT) as a spatial domain corrected model biases. Adding forest cover as a predictor improved the model, with improvements ranging from under 1% in sparse old-growth oak forests to over 20% in Mediterranean pine forests. This evaluation showed that GEDI L4A has a reasonable, albeit imperfect, biomass prediction capability, although the spatial advantage of GEDI is worth considering for its application.

3.3.3. Global Mangrove Extent by Global Mangrove Watch

To define the boundaries of mangrove forests, the global mangrove extent proposed by Bunting et. al. was used [24]. This data was obtained by using the Extreme Random Tree classifier on ALOS PALSAR and Landsat satellite sensor data to classify mangrove forests over previously created habitat masks. Through this method, the global mangrove area in 2010 was obtained with a high accuracy rate of 94%. This data was the first iteration, which was then updated to improve the quality and completeness of the mangrove extent, as reported in [52]. The overall accuracy of the map increased to 95.1%, and the mapped mangrove area increased from 137,600 km^2 to 140,260 km^2. Providing a more reliable resource for global mangrove conservation and sustainable utilization.

The consistent and automated application of the method made it possible to build a mangrove monitoring system, as done by [25]. Using a map-to-image approach to track changes over 24 years, from 1996 to 2020, an overall mangrove forest cover map was generated. The resulting map has an accuracy rate of 87.4%, with a lower change detection accuracy rate of between 58% and 60%. Nonetheless, this data is the most comprehensive record of global mangrove change to date and can be used for various conservation efforts. However, the authors of [24] suggest that the minimum recommended mapping unit for these data is 1 ha. And among others, higher errors often occur in disturbed areas and due to data gaps due to cloud cover or missing data.

In this study, Global Mangrove Extend V3.0 2020 data is used, which can be downloaded from the Global Mangrove Watch (GMW) website. This data is the latest version resulting from improvisations made by Bunting et al. [25]. This data was used as a mask so that only areas with a high probability of being mangrove forests were processed. The high accuracy of this data was expected to provide an accurate picture of the boundaries of mangrove and non-mangrove land cover. That way, the biomass estimation model process used can be more effective and precise.

3.3.4. Climate Change Initiative (CCI) Biomass Map

The European Space Agency's (ESA) Climate Change Biomass Initiative (CCI) is a project to provide a multi-temporal global map of aboveground biomass. As described in [53], the data consist of global per-pixel estimates of AGB (Mg/ha) with a spatial resolution of 100 meters or smaller, together with its standard deviation. The data is based on the BIOMASAR algorithm, which was applied to C-band and L-band SAR imagery to obtain AGB information representing the aboveground living dry biomass. In addition, the latest version released was version 4, where instead of three data periods (2010, 2017, and 2018), two data periods were added, namely 2019 and 2020.

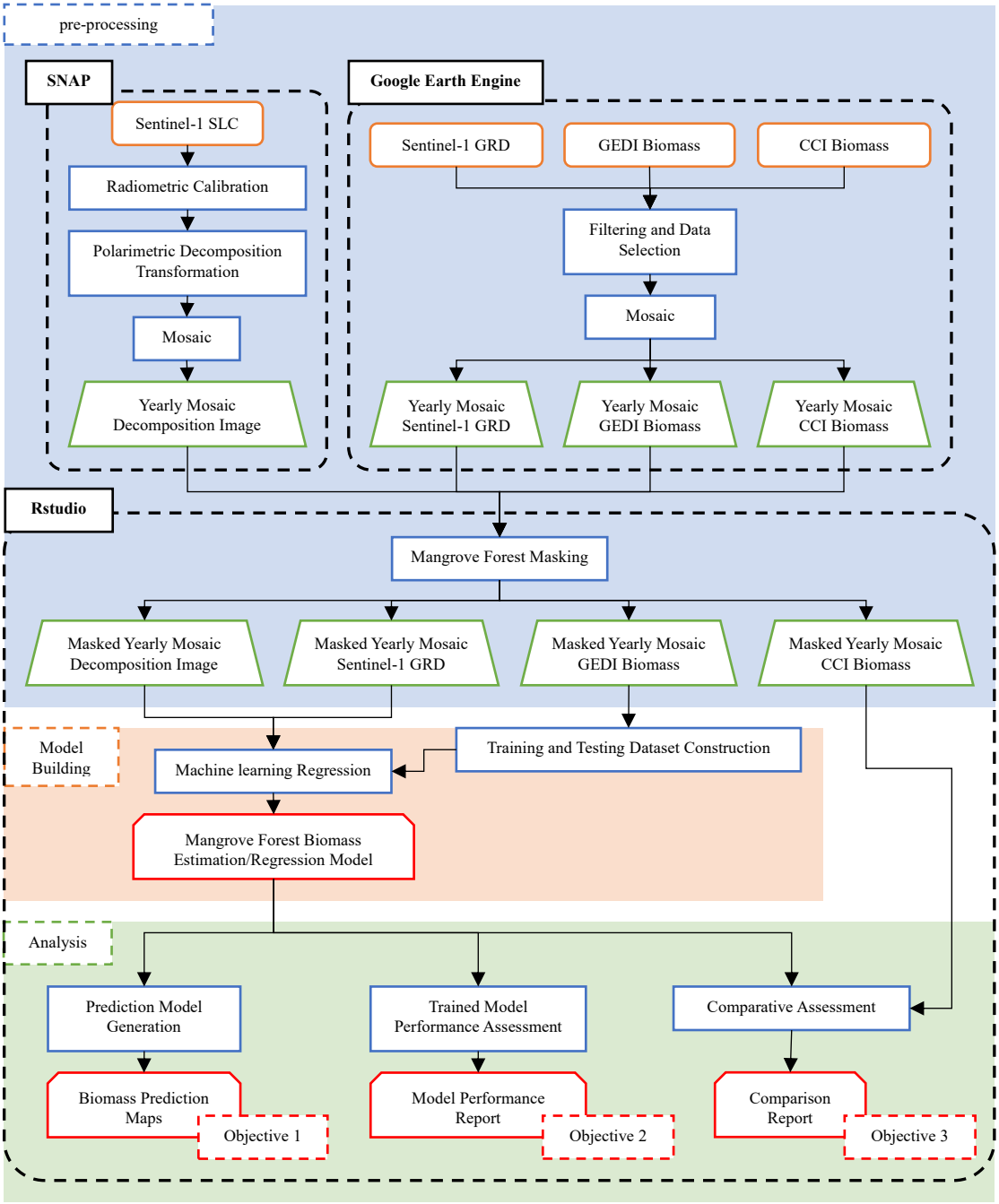
In this study, the AGB map of 2020 was used for comparison with the developed biomass estimation model. Based on the dataset user guide [53], additional forest masks are required to obtain statistical parameters specific to forest land, as the CCI Biomass map covers all land covers. Additionally, it is worth noting that while the model predicts that backscattering will increase along with biomass, the opposite was observed for mangrove forests. This led to an underestimation of mangrove biomass on the map compared to the specific mangrove product. This possible difference in values was used for analysis to enrich the estimation model.

3.4.Methods

Generally, the workflow carried out in this study is divided into three interconnected stages, namely data pre-processing, model building, and analysis. The pre-processing stage was carried out with the aim of tidying up the data so that it could represent accurate and precise information. This stage also includes tidying up satellite images to obtain the SAR parameters, as well as processing and selecting GEDI data to obtain training and validation samples. The next step was to build a model for estimating biomass in mangrove ecosystems. In this process, the prediction model is developed using the supervised machine learning algorithm H2O AutoML, developed by H2O.ai [43]. Where SAR parameters and GEDI-derived biomass information were employed as dependent and independent variables, respectively. Also, to identify the decomposed polarimetric parameters capability in the estimation process, several scenarios were created based on the number of variables used in the modeling. The last step was analysis, in which the estimation model was analyzed for performance and accuracy as well as compared with biomass estimation products from other initiatives. In addition, the characteristics of changes in the values of SAR parameters with

respect to variations in biomass density are identified. The explained steps are illustrated in Figure 5, as follows:

Figure 5: Mangrove Forest Biomass Model Generation Workflow



The explanation above is a brief description of the steps taken to analyze the potential of polarimetric decomposition parameters for mangrove biomass estimation. And for a detailed discussion, each stage that has been carried out is explained as follows.

3.4.1. Sentinel-1 Level-1 GRD data Pre-processing

At this stage, pre-processing does not need to be done. This is because the Sentinel-1 SAR GRD that has been stored in the GEE dataset catalog has been pre-processed with the Sentinel-1 Toolbox built into this platform [54]. The pre-processing steps consisted of three interconnected processes. Specifically, thermal noise removal, radiometric calibration, and terrain correction. Lastly, the log scaling was employed to transform the result into a decibel scale.

Although this data has been natively pre-processed by GEE, further steps to reduce the dimensionality of the data still need to be taken. The aim was to obtain a collection of Sentinel-1 Level-1 SAR GRD images with spatial and temporal dimensions consistent with other datasets. To achieve this goal, a filter function has been used, so that only image collections recorded during 2020 and within the specified extents are selected. In addition, the filter function was also used to constrain the internal parameters of this dataset, the details of which are described in Table 4 as follows:

Table 4: Image properties for the filter and select functions of the Sentinel-1 SAR GRD image collection in the GEE platform

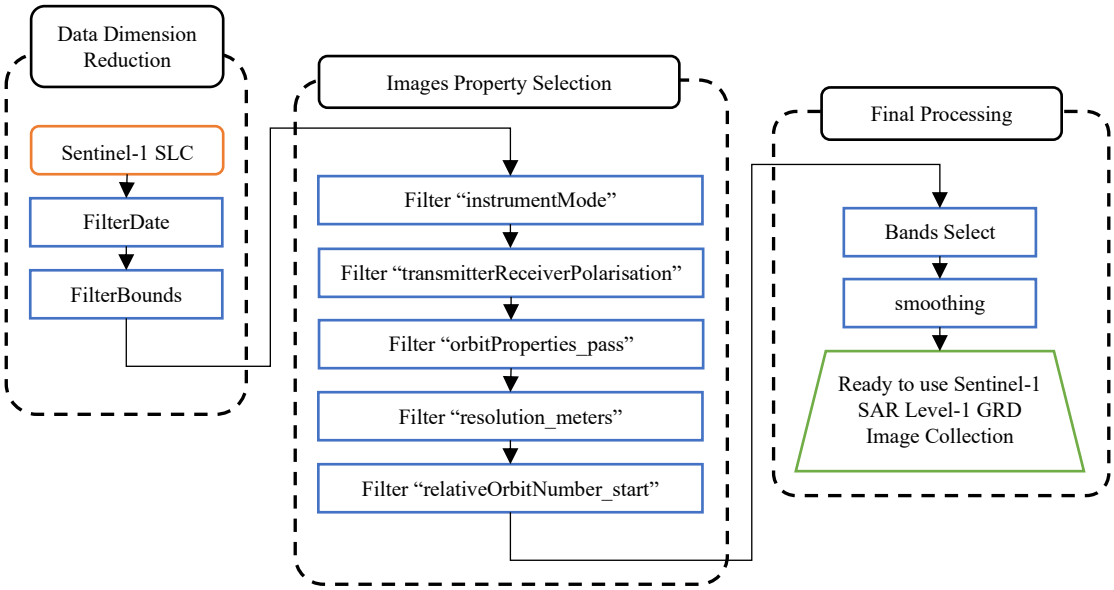
Image Property Name	Value Selected	Description
instrumentMode	IW	Indicates the recording mode.
transmitterReceiverPolarisation	VV, VH	Indicates transmit/receive polarization for data
orbitProperties_pass	DESCENDING	Indicates the direction of the orbit
resolution_meters	10	Indicates the spatial resolution of the selected data
relativeOrbitNumber_start	149	characterizes the oldest line in the image data indicated by its relative orbit number.

Source: Google Earth Engine Sentinel-1 GRD Level-1 Dataset Catalog (https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S1_GRD#image-properties)

Next, the select function was used so that only the VV and VH bands were interested. Then smoothing was also performed to reduce noise in the form of radar speckle. The results obtained from these steps were a set of Sentinel-1 SAR Level-1 SAR GRD images with dimensions that match the other datasets. From these stages, a collection of 31 images was

recorded from the 5th January 2020 to 30th December 2020. The Sentinel-1 Level-1 GRD data preprocessing steps are shown in Figure 6.

Figure 6: Sentinel-1 GRD Level-1 data pre-processing



3.4.2. Sentinel-1 Level-1 SLC Pre-Processing Procedure

The Sentinel-1 Level-1 SLC products that have been used were images obtained during 2020, with a total of 31 images downloaded from the official ESA website. In the process, a configuration has been chosen in the form of Sentinel-1A IW mode with VV and VH polarization channels on relative orbit number 149 on a descending orbital trajectory. The goal was to obtain enough images in one observation period that were spatially consistent with each other while being able to represent a whole year worth of phenomena.

The SLC format has the ability to store complete radar information, including phase and amplitude information [44]. By analyzing the phase information in the scattering matrix, polarimetric decomposition can unlock details beyond the strength of the reflected signal in SAR data. According to [55], the scattering matrix, denoted by S, is important for its ability to fully characterize the scattering behavior of a target. Added to this, this vector condenses the target's complex response to various combinations of horizontally and vertically polarized electromagnetic waves. By analyzing the intricate details within the scattering matrix, the polarimetric decomposition method can separate the reflected signal into its

fundamental parts, as described by [56]. Each of them represents the scattering mechanism of the object based on its scattering matrix, such as surface, dihedral, and volume scattering. And by understanding the scattering mechanism of its corresponding object, the characteristics of the object can be identified.

A. H/A/α Decomposition

H/A/α decomposition, also known as Cloude-Pottier Decomposition, is a polarimetric decomposition method proposed by Cloude and Pottier [57]. This eigenvalue/eigenvector-based decomposition method assumes that each cell has a dominant scattering mechanism. And this parameter is obtained through a smoothing algorithm based on second-order statistics. This method utilizes a three-level Bernoulli statistical model to estimate the average target scattering matrix parameters. This is achieved through the analysis of the 3 x 3 coherency matrix (T3). This analysis provides a basic invariant description of the scattering by decomposing it into scattering process types and their relative magnitudes, which are represented as eigenvectors and eigenvalues, respectively. Through this method, the estimation of the average target scattering matrix parameters does not require the assumption of a particular statistical distribution for the data, thus avoiding the limitations caused by the predefined physical constraints of the model [57].

Modifications of this method have been made by [58] and [59], where the matrix (T3) is simplified to be analyzed on dual polarimetric SAR data. However, for this study, a 2 x 2 covariance matrix (C2) was chosen instead. Based on [55], the covariance matrix has a role in the accurate representation of the backscatter signal, especially after speckle reduction through spatial averaging. This matrix condenses the averaged properties of multiple resolution cells into a single entity, formed by calculating the outer product of the averaged target vectors. However, due to the transmission and receiver limitations of dual polarimetric SAR data, it reduces the captured information to a covariance matrix containing only those elements. The coherence matrix modified by [58] (equation 1) as well as the covariance matrix modified by [55] (equation 2), which were further modified for application to Sentinel-1 SLC data, are shown below.

$$[T2] = \frac{1}{2}\begin{bmatrix} |S_{HH} + S_{VV}|^2 & (S_{HH} + S_{VV})(S_{HH} - S_{VV})^* \\ (S_{HH} - S_{VV})(S_{HH} + S_{VV})^* & |S_{HH} - S_{VV}|^2 \end{bmatrix} \quad (1)$$

$$[C2] = \begin{bmatrix} \langle S_{VV} S_{VV}^{*} \rangle & \langle S_{VV} S_{VH}^{*} \rangle \\ \langle S_{VH} S_{VV}^{*} \rangle & \langle S_{VH} S_{VH}^{*} \rangle \end{bmatrix} \quad (2)$$

Where S_{rt} denotes the complex scattering amplitude at the receive and transmit polarizations, and |.| denotes the determinant of a matrix. Whereas ⟨.⟩ represents the average of the ensemble, and (*) represents the conjugate complex. The eigenvalues and eigenvectors of this matrix are then used to decompose the scattering parameters into entropy (H), anisotropy (A), and alpha angle (α).

According to [60], entropy (H) is a measure of randomness or disorder in the scattering system. This value is calculated based on the eigenvalue (λ) of the coherence matrix (T) using the Von Neumann formula (equation 3) shown as follows:

$$H = -\sum_{i=1}^{n} p_i \log_n p_i \quad (3)$$

Where;

$$p_i = \frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j} \quad (4)$$

Where p_i denotes the pseudo-probability derived from the eigenvalue of λ_i, while n denotes the logarithmic basis determined from the number of available polarimeters (2 for dual polarimeters and 3 for full polarimeters). From this formula, entropy will be obtained with a range from 0 to 1. A low entropy indicates a system with a dominant scattering mechanism, while the opposite indicates a more random and complex scattering process. As entropy increases, the ability to distinguish between different scatterers using polarimetric data decreases. And in the extreme case (H = 1), the scattered information becomes a completely random perturbation.

The second parameter is anisotropy (A), which is used in conjunction with entropy to describe the randomness of the scattering. According to [60], anisotropy is calculated from the ordered eigenvalues ($\lambda_1 > \lambda_2 > \lambda_3 > 0$), which are obtained from decomposing the scattering matrix (equation 5), described as follows:

$$A = \frac{\lambda_2 - \lambda_3}{\lambda_2 + \lambda_3} \tag{5}$$

While entropy describes overall randomness, anisotropy focuses on the relative strengths of the second (λ_2) and third (λ_3) eigenvalues. According to [56], anisotropy has a value range of 0 to 1. A value of 0 indicates deterministic scattering or perfect polarization scattering, while a high value indicates the presence of a dominant secondary scattering process. The usefulness of anisotropy increases with higher entropy values because its effect on noise is reduced, as mentioned in [60]. This allows anisotropy to help distinguish different types of scattering that may have similar overall randomness (entropy) but with different eigenvalue distributions.

According to [61], while entropy and anisotropy are derived from eigenvalues, alpha angle (α) and beta angle (β) are obtained from eigenvectors. Where the α parameter is an indicator of the type of scattering mechanism, the β parameter describes the physical orientation of the object with respect to the line of sight. For this study, focus is only given to α, as its value can be easily related to the physical aspects behind the scattering process [60]. And for the calculation, it is based on the average of the eigenvector scattering angle (υ), shown in equation 6, which is the result of a modification done by [59], shown as follows:

$$\alpha = \sum_{i=1}^{n} p_i \alpha_i \tag{6}$$

where;

$$\alpha_i = \cos^{-1}(|v_{1i}|) \quad (7)$$

This parameter directly connects the scattering mechanism of the medium, where it is possible to observe the physical properties of the observed medium. According to [60], the α parameter has a range of 0º to 90º, where the characterized scattering mechanism changes as its value increases. At 0º, light is scattered purely on the surface, resembling geometric optics. As the angle increases, the scattering mechanism transitions through various models, including dipole scattering and double-bounce scattering, until it reaches pure metal surface scattering at 90º. However, due to necessary reductions made due to limitations in dual polarimetric SAR data, the parameter range α is reduced to 0º to 45º, as mentioned by [59].

B. Polarimetric Decomposition Parameter Extraction

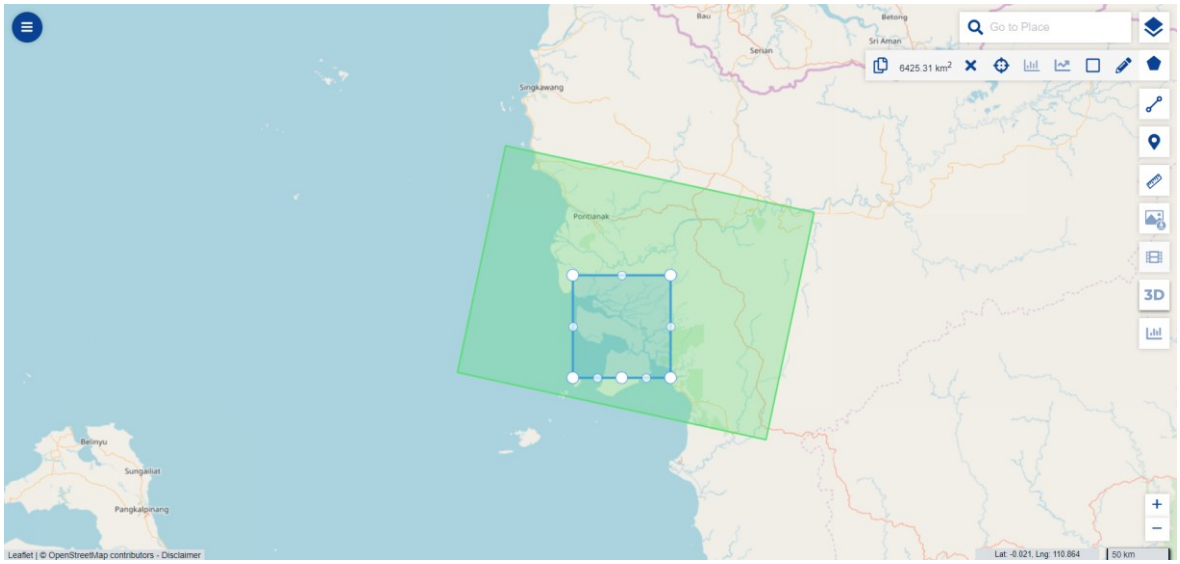
In practice, the pre-processing and extraction procedure of decomposed polarimetric parameters for the Sentinel-1 SLC was performed on the basis of the procedure by Mandal et. al. [62]. Broadly speaking, the steps are divided into two major modules. The first module aims to calibrate the radiometric aspects of the image, while the second module focuses on the extraction of the decomposed polarimetric parameters. The whole process is performed in SNAP software, with the help of the Radarsat Polarimetric Toolkit Module plugin developed by the Canadian Space Agency EOADP.

In the first module, refinement of the satellite position and velocity is required, as the orbit information available from the SAR product metadata is generally inaccurate. This can be done by downloading and applying the correct orbit file to the SAR product. Once a SAR product with accurate orbit information has been obtained, the next step requires splitting each sub swath with selected bursts into separate products. This step can be performed with the TOPSAR Split operator and was the first processing step in the TOPS InSAR processing chain. And since the study area of this study was in only one slice (Figure 4), the slice merging operation with the TOPSAR Slice Assembly operator was not necessary.

The TOPSAR Split operator results in two products, named IW1 and IW2, respectively, which were further used. Both products need to go through a radiometric calibration process so that the pixel values can be directly related to the radar backscatter at the scene. And to be used in the decomposition process, the output needs to be stored as "complex data". This is because raw SAR images contain radiometric distortions that hinder accurate analysis.

And applying radiometric correction can ensure the pixel values reflect the true nature of the surface. Next, the bursts in each corrected product need to be combined according to their zero doppler time. This was followed by merging the two products (IW1 and IW2) into one seamless image. These two processes were performed with the help of the Sentinel-1 TOPSAR Deburst and Merge operators. The next step was to crop the image according to the size of the study area that had been determined. This is because the original spatial dimensions of the corrected product were quite large, while the analyzed area was only limited to the predetermined study area (Figure 7). With the help of the subset operator, the output will be a Sentinel-1 Level-1 SLC product with reduced spatial dimensions that only covers the study area.

Figure 7: ESA's Copernicus Browser (https://browser.dataspace.copernicus.eu/) website view of the study area (blue rectangle) within the original coverage of the Sentinel-1 Level-1 SAR SLC product (green rectangle).



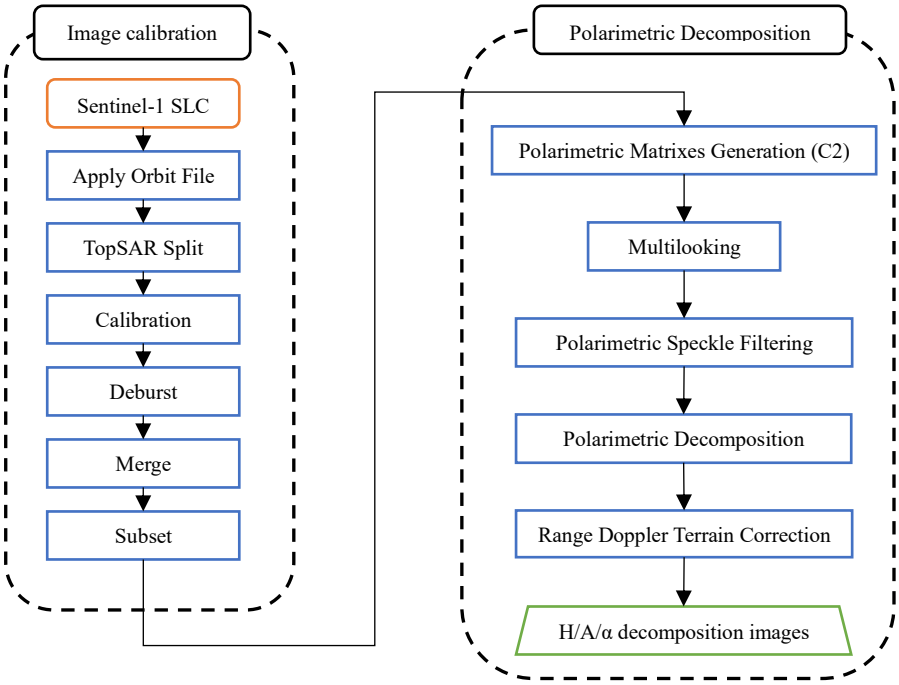
After the radiometrically corrected image has been obtained, the next step can be carried out. The second module aimed to extract the 2 x 2 covariance matrix (C2), which was then applied to the H/A/α decomposition method to obtain the parameters of the scattering mechanism. The C2 matrix can be obtained with the help of the Polarimetric Matrix Generation operator. Which was then continued with the multilook operation. SAR products are often grainy due to speckle noise. To improve their clarity, the multilook operation incoherently combines several images as if they correspond to different views of the same

scene. This operation resulted in a clearer image with more user-friendly pixels, where the shape changed to a square with an average size of 14 x 14 meters.

Furthermore, the Polarimetric Speckle Filtering operation was applied to the multilook image. SAR images are naturally covered by coarse noise called speckle, which makes image interpretation difficult. And by applying speckle filtering, the noise can be reduced, thereby improving image quality, and making features easier to identify. For this study, the Refined Lee Filter method with a window size of 7 x 7 has been used. According to [60], the core concept of this method is that neighbouring pixels that have similar characteristics to the center pixel are found and then used for noise reduction. The filtering process is performed by detecting the direction of an edge and then using the pixels aligned along that edge to filter out the noise while maintaining the sharpness of the edge itself.

The output of the polarimetric speckle filtering operation was then used as the input for the polarimetric decomposition operation. For this study, the H/A/α decomposition method proposed by Cloude and Pottier [57], which was later modified by [58] and [59] to be applicable to dual polarimetric SAR data, has been used. In this Polarimetric decomposition operator, the H/A/α decomposition for dual polarimetric SAR data is named "H-Alpha Dual Pol Decomposition". And in practice, processing with a window size of 5 x 5 has been selected. From this operation, an image with bands containing scattering mechanism information named Entropy (H), Anisotropy (A), and Alpha Angle (α) was obtained. The next step was to correct the distortions in the image due to variations in ground height and satellite sensor angle. To do so, the Range Doppler Terrain Correction (RDTC) operator has been used. RDTC corrects these distortions by referring to a digital elevation model (DEM) and satellite orbit data, essentially making the image geometrically accurate when compared to the real world. In practice, the reference DEM used was SRTM 3sec, as well as resampling the DEM and image by using Bilinear Interpolation. From this configuration, an image with bands containing geometrically corrected scattering mechanism information with a spatial resolution of 15 meters was obtained. Which can then be uploaded to the GEE platform for further processing. The Sentinel-1 L1 SLC data processing steps to obtain the polarimetric decomposition image are shown simply in Figure 8, as follows:

Figure 8: Sentinel-1 SLC Level-1 data pre-processing to obtain the polarimetric decomposition image



3.4.3. GEDI Data Filtering and Selection for Samples Collection

During the process of building the model for estimating biomass in mangrove forests, the availability of GEDI products was essential. For this study, GEDI L4A Rasterized AGBD Density V2.1 stored in GEE's dataset catalog has been used. The data used was based on observations from January to December 2020. In addition, the biomass GEDI was distributed as a geolocated laser raster footprint with a diameter of 25 meters.

Because this product is stored in the GEE dataset catalog, the selection and filtering were done in the JavaScript programming language. The process was adapted from the processing steps performed by [63], where the variable (AGBD), which reflects the estimated GEDI-derived AGBD information (Mg/ha), was used. The aim was to ensure that only biomass estimation information was extracted from this product. In addition, to ensure that the biomass information extracted was valid and met data quality standards, a masking process based on the variables (l4_quality_flag) and (degrade_flag) was used. It was determined that products with pixel values equal to zero of both variables would be retained, while others were removed. This process also excludes unreliable GEDI measurements. This was identified by the relative standard error in the GEDI-derived AGBD being greater than 50%,

expressed as (((agbd_se) / (agbd) × 100) > 50) [63]. Where (agbd_se) reflects the standard error value of the GEDI-derived AGBD information.

The filtered product was then clipped according to the mangrove forest boundary from GMW, resulting in a product that covered only mangrove forest land cover. Through this step, a filtered product with a total of ~ 6,000 rasterized footprints was obtained, which were then split for model training and validation purposes.

Figure 9: Zoomed-in view of the Google Earth Engine website (https://code.earthengine.google.com/) showing footprints from the GEDI L4A raster AGBD V2.1 recorded during 2020 used as samples. (A) before filtering (B) after filtering.

3.4.4. Mangrove Biomass Estimation Model Generation and Evaluation

H2O AutoML is a machine learning-based algorithm developed by H2O.ai that has been released since June 2017, and is part of the H2O framework introduced in 2013 [43]. It offers a robust solution for a variety of machine learning tasks, including regression, binary classification, and multi-class classification on tabular datasets. H2O AutoML adopted a comprehensive approach with stacked ensemble models and generates an easy-to-understand model leaderboard. Where users can select the best model based on their preferences, such as performance, training time, or prediction speed, which can be interpreted with built-in and third-party tools [43]. This phase was done in RStudio with the package "h2o" available in the software.

The mangrove forest biomass prediction model was built using a dataset consisting of a grid of overlapping input data in the form of Sentinel-1 GRD images and images from the polarimetric decomposition process. The amplitude information was obtained from Sentinel-1 GRD, while the phase information was obtained from the polarimetric image of the H/A/α method decomposition. Biomass data from GEDI was integrated as an independent variable in the modelling process. Two grid resolution sizes were selected to build this dataset, namely 25 meters which corresponds to the resolution of GEDI data, and 100 meters, considering the horizontal accuracy level of GEDI data [64]. Resampling and aggregating pixels of the input data using median and mean were also considered to group the pixels into a coarser-resolution grid dataset. A summary of the setup that has been used to build the dataset for model training can be found in Table 5, as follows:

Table 5: configuration of the dataset used to train the prediction model

No	Configuration Description
1	Dataset resampled and aggregated pixel values with the median method and a 25-meter grid size.
2	Dataset resampled and aggregated pixel values with the median method and a 100-meter grid size.
3	Dataset resampled and aggregated pixel values with the mean method and a 25-meter grid size.
4	Dataset resampled and aggregated pixel values with the mean method and a 100-meter grid size.

Additionally, model training for biomass prediction was carried out following five scenarios based on the varying use of SAR parameters, namely backscatter and polarimetric decomposition. The breakdown of the scenarios and the parameters included are shown in Table 6. Scenarios one and two utilized backscatter parameters, both topographically uncorrected and corrected, respectively. The topographically uncorrected backscatter parameters were coded " TPUncorrected", and "TPCorrected" for the corrected parameters, placed next to the polarization type. Scenario three utilized parameters from the polarimetric decomposition process. Scenarios four and five combined both types of parameters, with the difference that scenario four employed backscatter parameters with uncorrected topography, while scenario five employed corrected ones. The purpose of this varied scenario was to examine the performance of the prediction model when the different parameters were applied.

Table 6: Scenarios to generate mangrove forest biomass prediction model, including the parameters used for each scenario.

Scenarios	Parameters
1	VV TPUncorrected, VH TPUncorrected
2	VV TPCorrected, VH TPCorrected
3	Entropy, Anisotropy, Alpha Angle
4	VV TPUncorrected, VH TPUncorrected, Entropy, Anisotropy, Alpha Angle
5	VV TPCorrected, VH TPCorrected, Entropy, Anisotropy, Alpha Angle

The generated model needs to have its predictive ability considered so that the information it provides can be trusted. For this purpose, several accuracy metrics have been used, namely the coefficient of determination (R^2), root mean squared error (RMSE), mean absolute error (MAE), and mean squared error (MSE), which are shown as follows:

$$R^2 = 1 - \frac{\sum_{i-1}^{n}(\gamma_i - \hat{\gamma}_i)^2}{\sum_{i-1}^{n}(\gamma_i - \bar{\gamma})^2} \tag{8}$$

$$RMSE = \sqrt{\frac{\sum_{i-1}^{n}(\gamma_i - \hat{\gamma}_i)^2}{n}} \tag{9}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\gamma_i - \hat{\gamma}_i)^2 \tag{10}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|(\gamma_i - \hat{\gamma}_i)| \tag{11}$$

Where n indicates the number of samples, γ_i represents the AGBD derived from GEDI, $\hat{\gamma}_i$ represents the AGBD value predicted by the model, and $\bar{\gamma}$ represents the mean of the AGBD value derived from GEDI.

RESULTS

4.1. SAR Parameters against GEDI-derived Biomass Information

4.1.1. GEDI L4A AGBD Outlier Removal

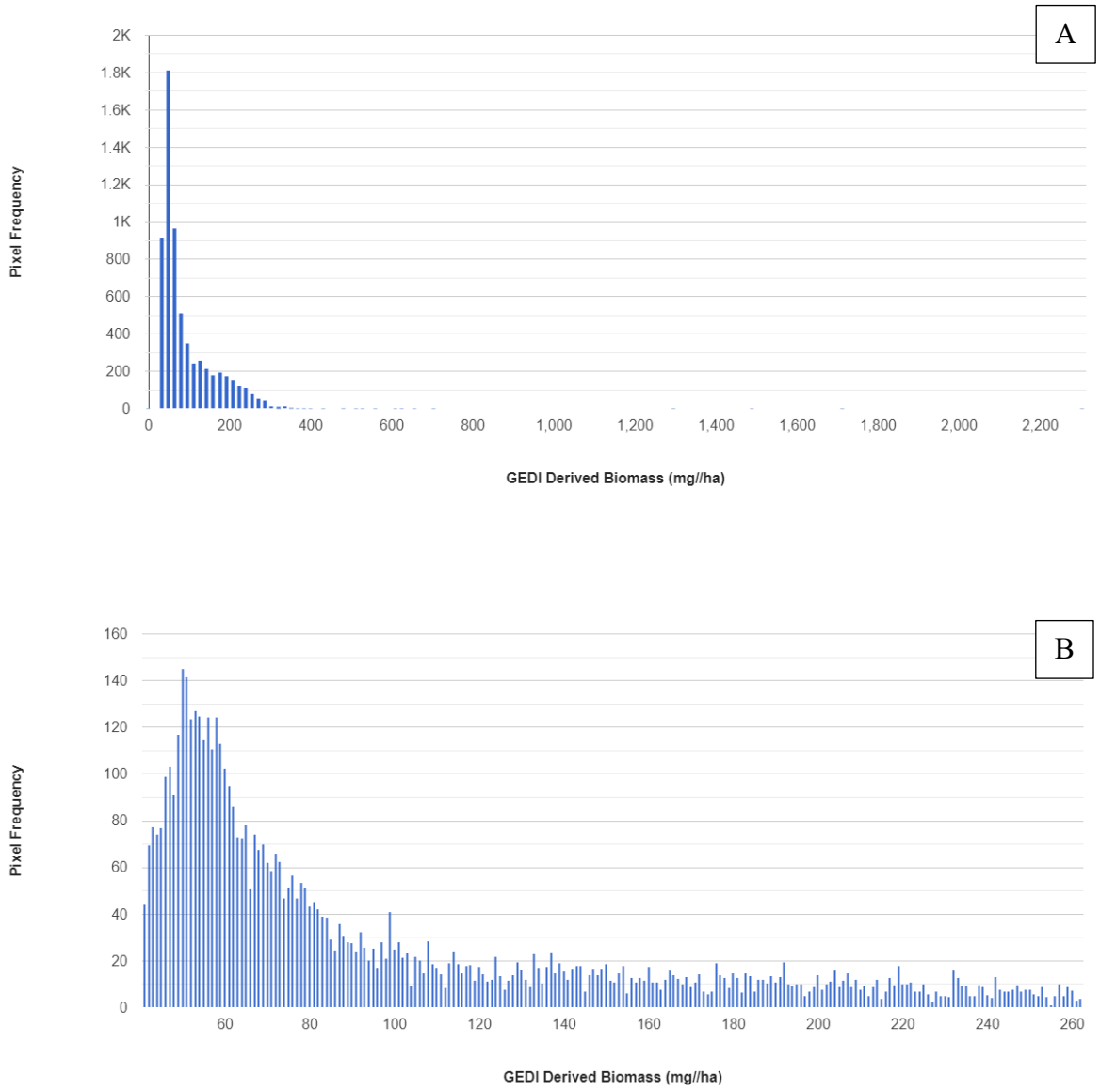
Biomass prediction was conducted by utilizing biomass information from GEDI as a predictor; hence, a tidying up process was necessary to provide accurate information. In the process, outliers were removed by excluding the top and bottom 2% of the data population from the pre-processed data. The changes in the values of the GEDI data after the subtraction process are shown in Table 7 below.

Table 7: Statistical summaries of biomass information derived from GEDI data before and after outlier removal, units in Mg/ha

Type of GEDI Data	Min	Max	Range	Mean	Med1ian	Standard Deviation
Before outlier removal	11.805	2304.295	2292.491	101.575	71.433	79.412
After outlier removal	41.471	262.953	221.482	97.513	72.449	56.587

GEDI data where the outliers have not been removed has a large range of values compared to those that have been removed. The GEDI data that has not been removed has a range of values up to 2292.491 Mg/ha, with minimum and maximum values of 11.805 Mg/ha and 2304.295 Mg/ha, respectively. While the GEDI whose outliers have been removed has a value range of 221,482 Mg/ha, and with minimum and maximum values of 41,471 Mg/ha and 262,953 Mg/ha, respectively. The mean value of the GEDI data before and after the outliers were removed also decreased, from 101,575 Mg/ha to 97,513 Mg/ha. Meanwhile, the median increased from 71,433 Mg/ha to 72,449 Mg/ha. However, the standard deviation decreased after the outliers were removed, from 79,412 Mg/ha to 56,587 Mg/ha. In addition, the distribution of the GEDI data pre and post outlier removal has also been represented in graphical form, as shown in Figure 10, as follows:

Figure 10: Histogram of GEDI L4A AGBD data where (A) before outlier removal and (B) after outlier removal.



4.1.2. SAR Parameter and GEDI-Derived Biomass Data Plotting

The model training used datasets constructed from overlapping preprocessed GEDI data with Sentinel-1 GRD (both uncorrected and topographically corrected) and decomposed polarimetric images. Following the grouping method and grid size configuration as described in Table 5, four datasets with different distribution patterns were obtained. The data distribution of the datasets using the four predefined configurations is shown in Table 8 as follows:

Table 8: Summary data distribution of dataset with varied grouping method and grid size. Different configuration of grouping method and resolution is indicated in bold number with blue highlight

Configuration 1 (Dataset with grouping method in median and 25-meter grid size)					
Parameters	Min	Max	Mean	Median	Standard Deviation
VV TPUncorrected	-14.176	-4.266	-7.701	-7.455	1.066
VH TPUncorrected	-21.311	-10.747	-14.267	-13.997	1.038
VV TPCorrected	-13.890	-3.139	-6.611	-6.377	1.070
VH TPCorrected	-20.624	-9.617	-13.176	-12.920	1.042
Entropy	0.377	0.880	0.716	0.715	0.028
Anisotropy	0.402	0.854	0.605	0.607	0.028
Alpha Angle	9.109	29.128	19.840	19.758	1.312
GEDI-derived AGBD	41.465	262.953	106.617	92.045	53.236
Configuration 2 (Dataset with grouping method in median and 100-meter grid size)					
Parameters	Min	Max	Mean	Median	Standard Deviation
VV TPUncorrected	-19.038	-3.097	-7.728	-7.445	1.122
VH TPUncorrected	-25.408	-9.830	-14.293	-13.991	1.098
VV TPCorrected	-17.862	-2.790	-6.641	-6.358	1.115
VH TPCorrected	-24.224	-9.645	-13.206	-12.911	1.090
Entropy	0.499	0.881	0.716	0.714	0.027
Anisotropy	0.397	0.778	0.605	0.607	0.027
Alpha Angle	14.186	29.164	19.848	19.748	1.283
GEDI-derived AGBD	41.471	262.747	104.739	91.315	50.590
Configuration 3 (Dataset with grouping method in mean and 25-meter grid size)					
Parameters	Min	Max	Mean	Median	Standard Deviation
VV TPUncorrected	-14.327	-4.239	-7.707	-7.459	1.063
VH TPUncorrected	-20.978	-10.926	-14.275	-14.000	1.036
VV TPCorrected	-14.061	-3.119	-6.617	-6.377	1.067
VH TPCorrected	-20.711	-9.806	-13.184	-12.922	1.040
Entropy	0.410	0.875	-0.716	0.714	0.027
Anisotropy	0.404	0.832	0.602	0.605	0.027
Alpha Angle	10.036	29.378	19.968	19.870	1.278
GEDI-derived AGBD	41.465	262.953	106.617	92.045	53.237
Configuration 4 (Dataset with grouping method in mean and 100-meter grid size)					
Parameters	Min	Max	Mean	Median	Standard Deviation
VV TPUncorrected	-18.564	-3.205	-7.747	-7.456	1.127
VH TPUncorrected	-25.068	-10.277	-14.314	-14.005	1.107
VV TPCorrected	-17.276	-2.984	-6.659	-6.374	1.120
VH TPCorrected	-23.780	-9.837	-13.226	-12.919	1.099
Entropy	0.507	0.873	0.716	0.715	0.027
Anisotropy	0.409	0.763	0.602	0.605	0.027
Alpha Angle	14.249	28.949	19.986	19.880	1.268
GEDI-derived AGBD	41.471	262.747	104.754	91.472	50.504

Pairs of configurations one and three used the same resolution of 25 meters but with different grouping methods, the median and the mean, respectively. A reduction in the range of values was observed for the backscatter parameters, both with uncorrected and corrected topographies, apart from the corrected VH parameter. Meanwhile, the decomposed polarimetric parameters as well as the biomass parameters showed an increase in the range of values. However, a different pattern was shown for pairs of configurations two and four, which have a greater resolution of 100 meters for both. There was an increase in the range of values for all parameters from the input image after the data frame configuration changed from median to mean.

On the other hand, there was a reduction in the range of values for both topography-corrected and uncorrected backscatter parameters in pairs of configurations one and two. This pattern was observed when the resolution was altered without changing the grouping method. The same pattern was also observed in pairs of configurations three and four when the mean was applied instead of the median. Additionally, the range of values of the decomposed polarimetric parameters and biomass increased after the grouping method was switched from median to mean. This pattern was observed in both pairs of configurations one and two, as well as in configurations three and four.

However, a different pattern was exhibited in the mean, median, and standard deviation for all configurations. There was no significant change with different clustering methods or spatial resolutions.

4.2. Mangrove Forest Biomass Estimation Model

4.2.1. Mangrove Forest Model Training

Model training for the prediction of mangrove forest biomass has been conducted based on a machine learning approach using the H2O platform in RStudio software. In this step, both backscatter and decomposed polarimetric SAR parameters were utilized as dependent variables, while biomass information from GEDI was used as independent variables.

The model training process utilized data frames from configurations one to four described earlier. And for each configuration, the prediction model was trained following predetermined scenarios. Where these scenarios consist of variations in the combination of SAR parameters of both backscatter and decomposed polarimeters. The goal was to determine the effect of these parameters, especially the decomposed polarimetric parameters,

on the prediction results. A summary of the statistical metrics of the trained model that has been validated is shown in Table 9 as follows:

Table 9: Summary statistical metrics of trained models. Bold numbers with blue highlight indicate different grouping method and pixel resolution of data frame utilized to train the models.

Configuration 1 (Dataset with grouping method in median and 25-meter grid size)					
Scenario	Parameters	MAE	MSE	RMSE	R^2
1	VV TPUncorrected, VH TPUncorrected	38.763	2435.544	49.351	0.15
2	VV TPCorrected, VH TPCorrected	38.986	2429.215	49.287	0.15
3	Entropy, Anisotropy, Alpha Angle	43.390	2751.406	52.454	0.04
4	VV TPUncorrected, VH TPUncorrected, Entropy, Anisotropy, Alpha Angle	38.567	2398.013	48.970	0.16
5	VV TPCorrected, VH TPCorrected, Entropy, Anisotropy, Alpha Angle	38.387	2415.251	49.145	0.16
Configuration 2 (Dataset with grouping method in median and 100-meter grid size)					
Scenario	Parameters	MAE	MSE	RMSE	R^2
1	VV TPUncorrected, VH TPUncorrected	35.845	2091.000	45.727	0.20
2	VV TPCorrected, VH TPCorrected	36.359	2126.869	46.118	0.19
3	Entropy, Anisotropy, Alpha Angle	40.490	2459.277	49.591	0.06
4	VV TPUncorrected, VH TPUncorrected, Entropy, Anisotropy, Alpha Angle	35.770	2070.991	45.508	0.21
5	VV TPCorrected, VH TPCorrected, Entropy, Anisotropy, Alpha Angle	35.563	2089.123	45.707	0.20
Configuration 3 (Dataset with grouping method in mean and 25-meter grid size)					
Scenario	Parameters	MAE	MSE	RMSE	R^2
1	VV TPUncorrected, VH TPUncorrected	38.922	2399.738	48.987	0.16
2	VV TPCorrected, VH TPCorrected	38.853	2396.906	48.958	0.16
3	Entropy, Anisotropy, Alpha Angle	43.427	2749.317	52.434	0.04
4	VV TPUncorrected, VH TPUncorrected, Entropy, Anisotropy, Alpha Angle	38.490	2385.075	48.837	0.17
5	VV TPCorrected, VH TPCorrected, Entropy, Anisotropy, Alpha Angle	38.141	2405.571	49.047	0.16
Configuration 4 (Dataset with grouping method in mean and 100-meter grid size)					
Scenario	Parameters	MAE	MSE	RMSE	R^2
1	VV TPUncorrected, VH TPUncorrected	35.611	2062.802	45.418	0.21
2	VV TPCorrected, VH TPCorrected	35.613	2092.386	45.743	0.20
3	Entropy, Anisotropy, Alpha Angle	40.215	2433.964	49.335	0.07
4	VV_TPUncorrected, VH_TPUncorrected, Entropy, Anisotropy, Alpha Angle	35.075	2048.701	45.263	0.22
5	VV TPCorrected, VH TPCorrected, Entropy, Anisotropy, Alpha Angle	35.474	2060.278	45.390	0.21

Configuration one employed the median grouping method with a 25-meter grid size to train the prediction model. With this configuration, Scenario four, which used all parameters with the backscatter parameter being the one with uncorrected topography, performed the best. The R^2, RMSE, MSE, and MAE values were 0.16, 48.970, 2398.013, and 38.567 respectively. while scenario five, which utilized backscatter parameters with corrected topography, showed a better MAE of 38.387. Followed by scenario two, then one with R^2, RMSE, MSE, and MAE values of 0.15, 49.287, 2429.215, and 38.986, respectively, for scenario two, and 0.15, 49.351, 2435.544, and 38.763, respectively, for scenario one. Finally, scenario three, which only utilized decomposed polarimetric parameters, exhibited the lowest prediction performance with R^2, RMSE, MSE, and MAE values of 0.04, 52.454, 2751.406, and 43.390, respectively.

Configuration two also employed the median grouping method, but with a larger grid size of 100-meters to train the model. Demonstrating the same pattern as configuration one, scenario four performed best with R^2, RMSE, and MSE values of 0.21, 45.508, and 2070.991, respectively. Followed by scenario five, which had lower performance when looking at the R^2, RMSE, and MSE values, but the MAE value was greater than in scenario four, which was 35,563. In addition, scenario one, which used a combination of backscatter parameters with uncorrected topography, performed better than scenario two, which used backscatter parameters with corrected topography. Where the R^2, RMSE, MSE, and MAE values for scenario one were 0.15, 49.412, 2441.585, and 38.915, respectively. Compared to scenario two, which had R^2, RMSE, MSE, and MAE values of 0.19, 46.118, 2126.869, and 36.359, respectively. Finally, scenario three had the lowest performance compared to the other scenarios, with R^2, RMSE, MSE, and MAE values of 0.06, 49.591, 2459.277, and 40.490, respectively.

As for configuration three, in which the grouping method used was the mean and with a grid size of 25 meters has been used to train the model. In this configuration, scenario four also had the best performance with R^2, RMSE, and MSE values of 0.17, 2385.075, and 48.837, respectively. While scenario five had a better MAE value compared to scenario four with a value of 38,141. However, looking at other statistical metrics, scenario two followed by scenario one performed better. Where the R^2, RMSE, MSE, and MAE values for scenario two were 0.16, 48.958, 2396.906, and 38.853, respectively, and 0.16, 48.987, 2399.738, and

38.922, respectively for scenario one. Whereas, scenario three has the lowest performance where the R^2, RMSE, MSE, and MAE values were 0.04, 52.434, 2749.317 and 43.427, respectively.

The last was configuration four for prediction model training, where the grouping method was the mean method but with a grid size of 100 meters. In this configuration, scenario four showed the best performance compared to the others with R^2, RMSE, MSE, and MAE values of 0.22, 45.263, 2048.70, and 35.075, respectively. Followed by scenario five with R^2, RMSE, MSE, and MAE values of 0.21, 45.390, 2060.278, and 35.474, respectively. Scenarios one and two had lower performance where R^2, RMSE, MSE, and MAE were 0.21, 45.418, 2062.802, and 35.611, respectively for scenario one and 0.20, 45.743, 2092.386, and 35.613, respectively for scenario two. Finally, scenario three has the lowest performance where the R^2, RMSE, MSE, and MAE values were 0.07, 49.335, 2433.964, and 40.215, respectively.

4.2.2. Prediction Map Generation

Once the trained model has been obtained, prediction of biomass values in mangrove forests can be performed on the input data. From the previous step, four configurations of trained models have been obtained, which differ from the grouping method and grid resolution of the dataset used for model training. And in each configuration, the model has been trained by following scenarios based on variations in the number of parameters used in the prediction process.

The parameters in question were backscatter parameters, both topography-corrected and uncorrected, as well as decomposed polarimetric parameters. And each scenario was formed from variations in parameter usage, both in groups or individually determined in such a way that the influence of decomposed polarimetric parameters on prediction results could be identified. With such an arrangement, several mangrove forest prediction models in raster format were obtained based on a variety of trained model configurations and different parameter usage scenarios. A summary of the pixel metric statistics describing the distribution of biomass values for each prediction model is provided in Table 10. For readability, the statistical metrics are grouped according to the configuration of the dataset used to train the model, which is then detailed for each scenario.

Table 10: Pixel statistic summary of predictive model generated from trained model with different grouping method and resolution setting which indicated in bold number and highlighted in blue color with units in Mg/ha.

Configuration 1 (Dataset with grouping method in median and 25-meter grid size)					
Scenario	Parameters	min	max	mean	Standard Deviation
1	VV TPUncorrected, VH TPUncorrected	-120.53	153.97	274.5	94.7
2	VV TPCorrected, VH TPCorrected	-173.63	154.73	328.35	97.52
3	Entropy, Anisotropy, Alpha Angle	57.04	156.08	99.04	109.46
4	VV TPUncorrected, VH TPUncorrected, Entropy, Anisotropy, Alpha Angle	6.26	170.22	163.97	97.14
5	VV TPCorrected, VH TPCorrected, Entropy, Anisotropy, Alpha Angle	41.8	175.26	133.47	94.95
Configuration 2 (Dataset with grouping method in median and 100-meter grid size)					
Scenario	Parameters	min	max	mean	Standard Deviation
1	VV TPUncorrected, VH TPUncorrected	-34.56	163.54	198.1	97.83
2	VV TPCorrected, VH TPCorrected	10.05	163.98	153.92	95.4
3	Entropy, Anisotropy, Alpha Angle	58.68	151.4	92.72	105.19
4	VV TPUncorrected, VH TPUncorrected, Entropy, Anisotropy, Alpha Angle	50.39	182.71	132.32	99.13
5	VV TPCorrected, VH TPCorrected, Entropy, Anisotropy, Alpha Angle	41.78	162.72	120.94	95.38
Configuration 3 (Dataset with grouping method in mean and 25-meter grid size)					
Scenario	Parameters	min	max	mean	Standard Deviation
1	VV TPUncorrected, VH TPUncorrected	37.65	168.63	130.98	98.46
2	VV TPCorrected, VH TPCorrected	-14.55	178.27	192.82	99.3
3	Entropy, Anisotropy, Alpha Angle	56.76	171.93	115.17	109.91
4	VV TPUncorrected, VH TPUncorrected, Entropy, Anisotropy, Alpha Angle	-45.23	175.62	220.86	97.25
5	VV TPCorrected, VH TPCorrected, Entropy, Anisotropy, Alpha Angle	-43.07	168.54	211.62	93.53
Configuration 4 (Dataset with grouping method in mean and 100-meter grid size)					
Scenario	Parameters	min	max	mean	Standard Deviation
1	VV Uncorrected, VH Uncorrected	22.85	181.09	158.25	96.68
2	VV TPCorrected, VH TPCorrected	21.32	163.48	142.16	94.96
3	Entropy, Anisotropy, Alpha Angle	46.97	174.78	127.8	105.92
4	VV TPUncorrected, VH TPUncorrected, Entropy, Anisotropy, Alpha Angle	19.44	174.8	155.36	94.88
5	VV TPCorrected, VH TPCorrected, Entropy, Anisotropy, Alpha Angle	48.29	176.44	128.15	97.1

The dataset with configuration one has been used to train the model to build mangrove forest biomass prediction models. The resulting prediction models had mean values from 94.7

Mg/ha to 109.5 Mg/ha. The prediction model with scenario one that utilized the backscatter parameter with corrected topography had the longest range of predicted biomass values at 328.3 Mg/ha. This was followed by scenario two which utilized the backscatter parameter with uncorrected topography at 274.5 Mg/ha. Scenarios four and five that combine backscatter parameters with decomposed polarimetric parameters instead had a shorter range of predicted biomass values, where scenario four, which included backscatter parameters with uncorrected topography, had a range of predicted biomass values of 164.0 Mg/ha. While scenario five, which instead used the backscatter parameter with corrected topography, had a range of predicted biomass values of 133.5 Mg/ha. Finally, scenario three, which only utilized decomposed polarimetric parameters, had a predicted biomass value range of 99,0 Mg/ha.

The resulting mangrove biomass prediction model with configuration two has an average value with a range of 95.4 Mg/ha to 105.2 Mg/ha. The scenario that only utilized the backscatter parameter produced a model with the highest range of biomass values compared to the other scenarios. Scenario one had a range of 198.1 Mg/ha, while scenario two had a range of 153.9 Mg/ha. On the other hand, the scenarios that combined the backscattering parameter and the decomposed polarimetric parameter had a shorter range of biomass values, with scenario four at 132.3 Mg/ha and scenario five at 120.9 Mg/ha. Finally, scenario three, which only utilized the decomposed polarimetric parameter, had the shortest range of biomass values, at 92.7 Mg/ha.

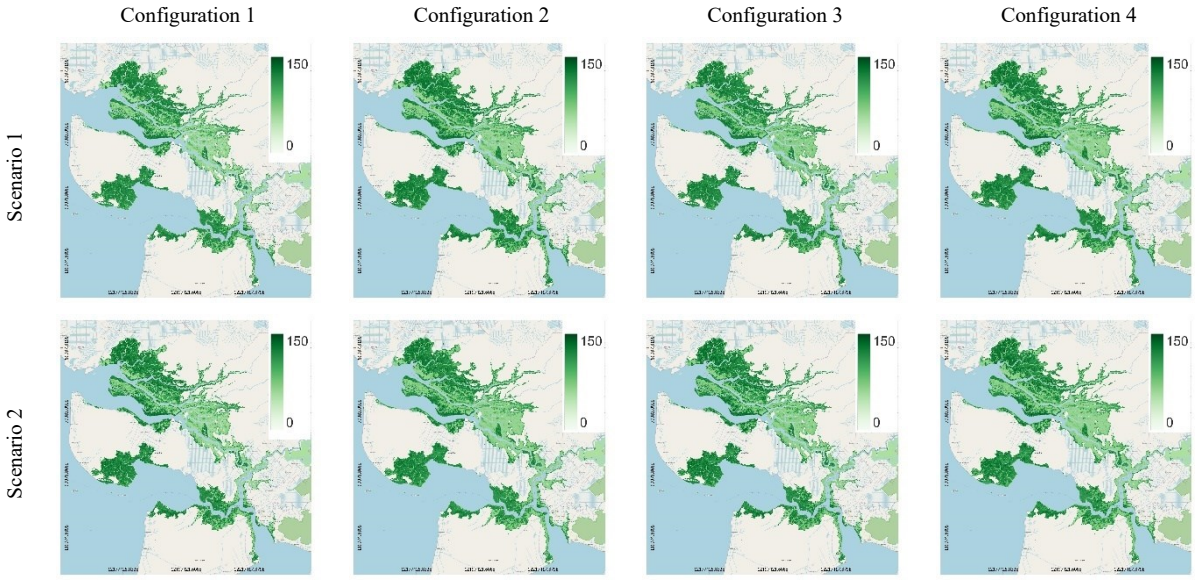
With configuration three, the resulting prediction model has a different pattern, with mean values from 93.5 Mg/ha to 109.9 Mg/ha. Additionally, the range of predicted biomass values was wider than the previous configuration. Scenario four followed by scenario five had the highest range of values with 220.9 Mg/ha and 211.6 Mg/ha, respectively. Scenario two, which utilized the topography-corrected backscatter parameter, had a shorter range of values, 192.8 Mg/ha. This was followed by scenario one that used the uncorrected topographic backscatter parameter with a value range of 131.0 Mg/ha. While scenario three produced a prediction model with the shortest range of values at 115.2 Mg/ha.

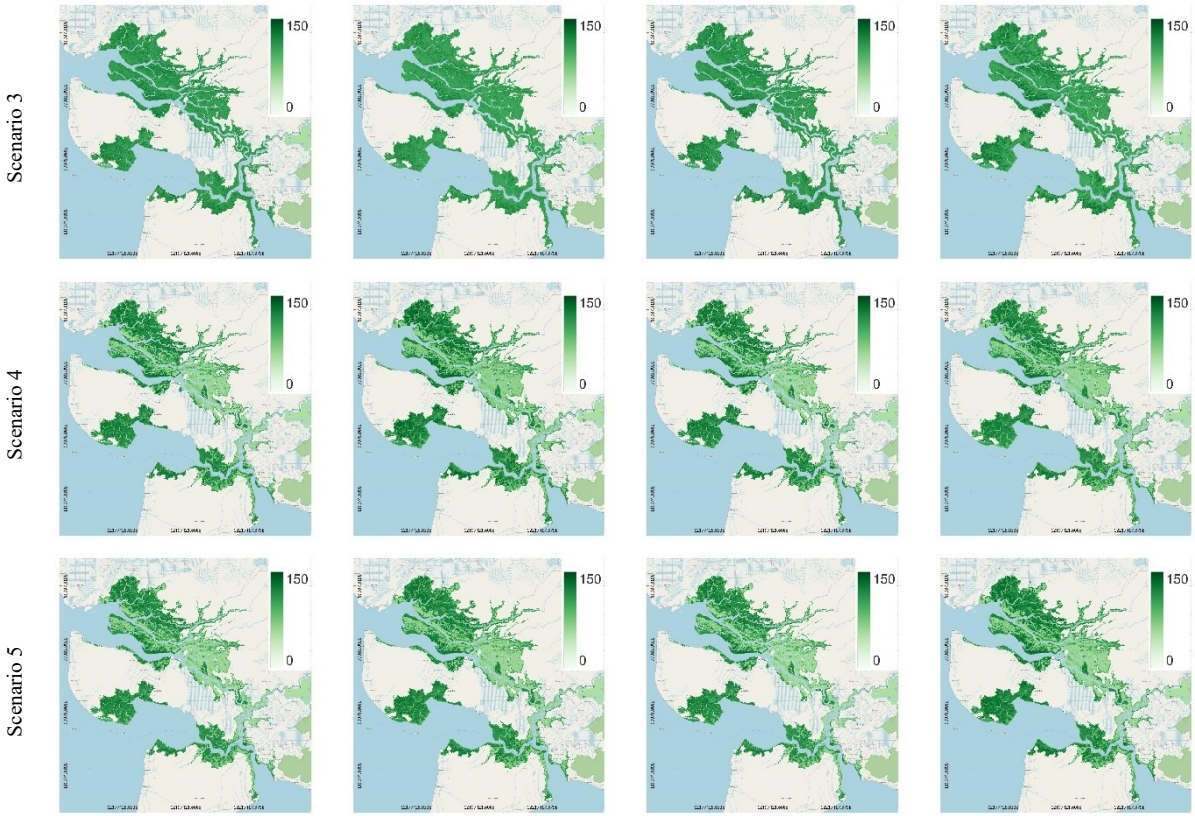
The dataset with configuration four resulted in a prediction model with a shorter range of biomass values with an average value of 95.0 Mg/ha to 105.9 Mg/ha. The prediction model of scenario one which utilized backscatter parameters with uncorrected topography had the longest range of values at 158.2 Mg/ha. Followed by the model from scenario four that used

the backscatter parameter with uncorrected topography combined with decomposed polarimetric parameters with a value of 155.4 Mg/ha. Scenario two, which only included the topography-corrected backscatter parameter, had a shorter range of values, 142.2 Mg/ha. Scenario five, which combined backscatter parameters with topography corrected by decomposed polarimetry, decreased the range of biomass prediction values to 128.1 Mg/ha. Finally, scenario three which utilized only the decomposed polarimetric parameter had a predicted biomass value range of 127.8 Mg/ha.

Additionally, the mangrove forest biomass prediction maps constructed from the trained model of the dataset with varying configurations are shown in Figure 11. Where each mangrove forest biomass prediction map has been created using parameters following the scenarios described in Table 5. The resulting maps had a uniform value range, from 0 to 150 Mg/ha with two different spatial resolutions following the datasets used to train the model. The maps using datasets configuration one and three have a resolution of 25 meters, while the maps using datasets configuration two and four have a resolution of 100 meters.

Figure 11: Mangrove Forest Prediction Maps for each scenario of parameter usage and dataset configuration for model training, units in Mg/ha.

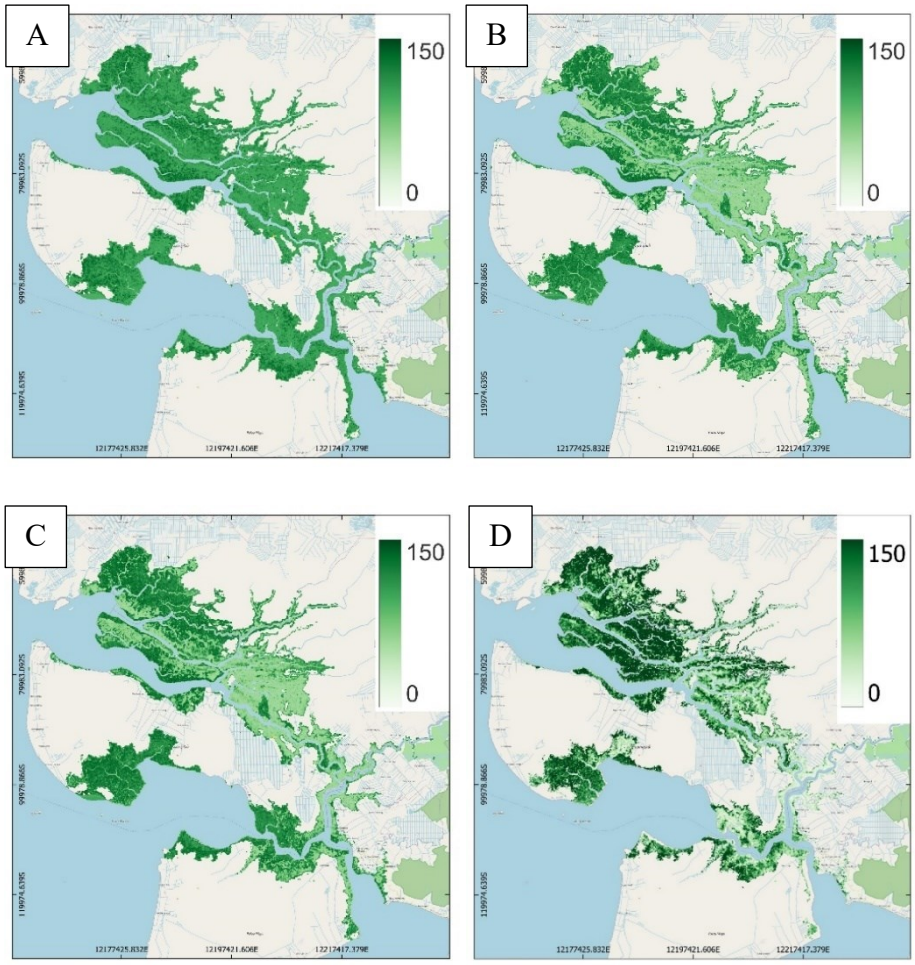




4.2.3. Comparison Against Model from Other Initiative

The resulting biomass prediction model was then compared with ESA's CCI biomass prediction model. This comparison allowed the predictive performance of the developed model relative to widely known biomass predictions models to be assessed. Prior to the analysis, the lower and upper 2% of the data population were removed to reduce the possibility of outliers being included in the analysis. Through this process, the biomass model of ESA's CCI was obtained with a range of values from 11.0 to 258.0 Mg/ha with mean and median of 117.0 Mg/ha and 109.0 Mg/ha, respectively. However, for comparison purposes, this range of values has been adjusted as shown in Figure 12, as follows:

Figure 12: Side-to-side Comparison of biomass prediction values where (A) Developed Prediction model in scenario four with dataset with configuration four (B) Scenario Four with configuration four, (C) Scenario Five with configuration four, and (D) Biomass prediction model from CCI by ESA. Units in Mg/ha.



The difference in the predicted biomass from the developed model compared to the CCI prediction model from ESA has also been carried out. This is done by subtracting the predicted values of the developed models with ESA's CCI prediction model. In order to obtain a more representative difference, the prediction models from datasets one and three needed to be resampled following the pixel resolution of the CCI prediction model. As for the models that use dataset configurations two and four, they are left with the actual situation.

As shown in Table 11, summary statistics of the difference in predicted biomass at each value range between the developed prediction model and ESA's CCI prediction model are presented. Negative values indicate that the biomass predicted by the developed model was lower than that predicted by ESA's CCI at the same pixel position. Whereas positive values indicate that the predicted biomass value in the developed model is higher when compared

to the prediction model from ESA's CCI. The statistical metrics such as minimum, maximum, mean and standard deviation ranges that can explain the distribution are shown below:

Table 11: Summary of Difference of Predicted Biomass between Prediction Model and CCI ESA Biomass Map in Mg/ha. Differentiated based on the scenario of parameter usage and the configuration of the dataset used to train the model.

Configuration 1 (Dataset with grouping method in median and 25-meter grid size)					
Scenario	Parameters	min	max	Range	Mean
1	VV TPUncorrected, VH TPUncorrected	-200.68	128.14	328.82	-23.55
2	VV TPCorrected, VH TPCorrected	-200.05	129.47	329.52	-20.76
3	Entropy, Anisotropy, Alpha Angle	-168.74	140.61	309.35	-9.11
4	VV TPUncorrected, VH TPUncorrected, Entropy, Anisotropy, Alpha Angle	-196.86	144.80	341.66	-21.19
5	VV TPCorrected, VH TPCorrected, Entropy, Anisotropy, Alpha Angle	-196.56	128.62	325.17	-23.54
Configuration 2 (Dataset with grouping method in median and 100-meter grid size)					
Scenario	Parameters	min	max	Range	mean
1	VV TPUncorrected, VH TPUncorrected	-195.66	141.02	336.68	-17.88
2	VV TPCorrected, VH TPCorrected	-191.30	129.23	320.53	-20.34
3	Entropy, Anisotropy, Alpha Angle	-161.13	132.26	293.39	-11.50
4	VV TPUncorrected, VH TPUncorrected, Entropy, Anisotropy, Alpha Angle	-195.11	147.06	342.17	-16.55
5	VV TPCorrected, VH TPCorrected, Entropy, Anisotropy, Alpha Angle	-195.80	131.79	327.59	-20.46
Configuration 3 (Dataset with grouping method in mean and 25-meter grid size)					
Scenario	Parameters	min	max	Range	mean
1	VV TPUncorrected, VH TPUncorrected	-192.25	128.06	320.31	-20.04
2	VV TPCorrected, VH TPCorrected	-192.41	129.97	322.39	-19.18
3	Entropy, Anisotropy, Alpha Angle	-156.77	153.92	310.69	-8.78
4	VV TPUncorrected, VH TPUncorrected, Entropy, Anisotropy, Alpha Angle	-197.45	131.18	328.62	-21.17
5	VV TPCorrected, VH TPCorrected, Entropy, Anisotropy, Alpha Angle	-199.35	130.14	329.49	-24.87
Configuration 4 (Dataset with grouping method in mean and 100-meter grid size)					
Scenario	Parameters	min	max	Range	mean
1	VV TPUncorrected, VH TPUncorrected	-195.29	162.19	357.48	-18.72
2	VV TPCorrected, VH TPCorrected	-196.54	133.61	330.15	-20.62
3	Entropy, Anisotropy, Alpha Angle	-164.50	139.94	304.45	-9.92
4	VV TPUncorrected, VH TPUncorrected, Entropy, Anisotropy, Alpha Angle	-198.70	139.07	337.76	-21.11
5	VV TPCorrected, VH TPCorrected, Entropy, Anisotropy, Alpha Angle	-199.54	136.79	336.34	-18.32

The difference in predicted biomass values between the model developed with the dataset configuration one and the prediction model from CCI has a range of mean values from -23.6 Mg/ha to -9.1 Mg/ha. Where the farthest minimum range was observed in the model with scenario one with a value of -200.7 Mg/ha. While the closest minimum range was observed in the model with scenario three with a value of -168.7 Mg/ha. The model with scenario one showed a difference in value with the closest maximum value range with a value of 128.1 Mg/ha. While the farthest maximum value range was recorded in the model with scenario four with a value of 144.8 Mg/ha.

Whereas in configuration two, the developed biomass prediction model had a narrower value difference compared to the prediction model from CCI. Though not significant, the mean values range from -20.5 Mg/ha to -11.5 Mg/ha. The model with scenario five had the value difference with the farthest minimum range with a value of -195.8 Mg/ha. While the opposite condition was seen in scenario three with a value of -161.1 Mg/ha. The closest maximum value difference was seen in scenario two with a value of 129.2 Mg/ha. While the opposite condition was seen in scenario four with a value of 147.1 Mg/ha.

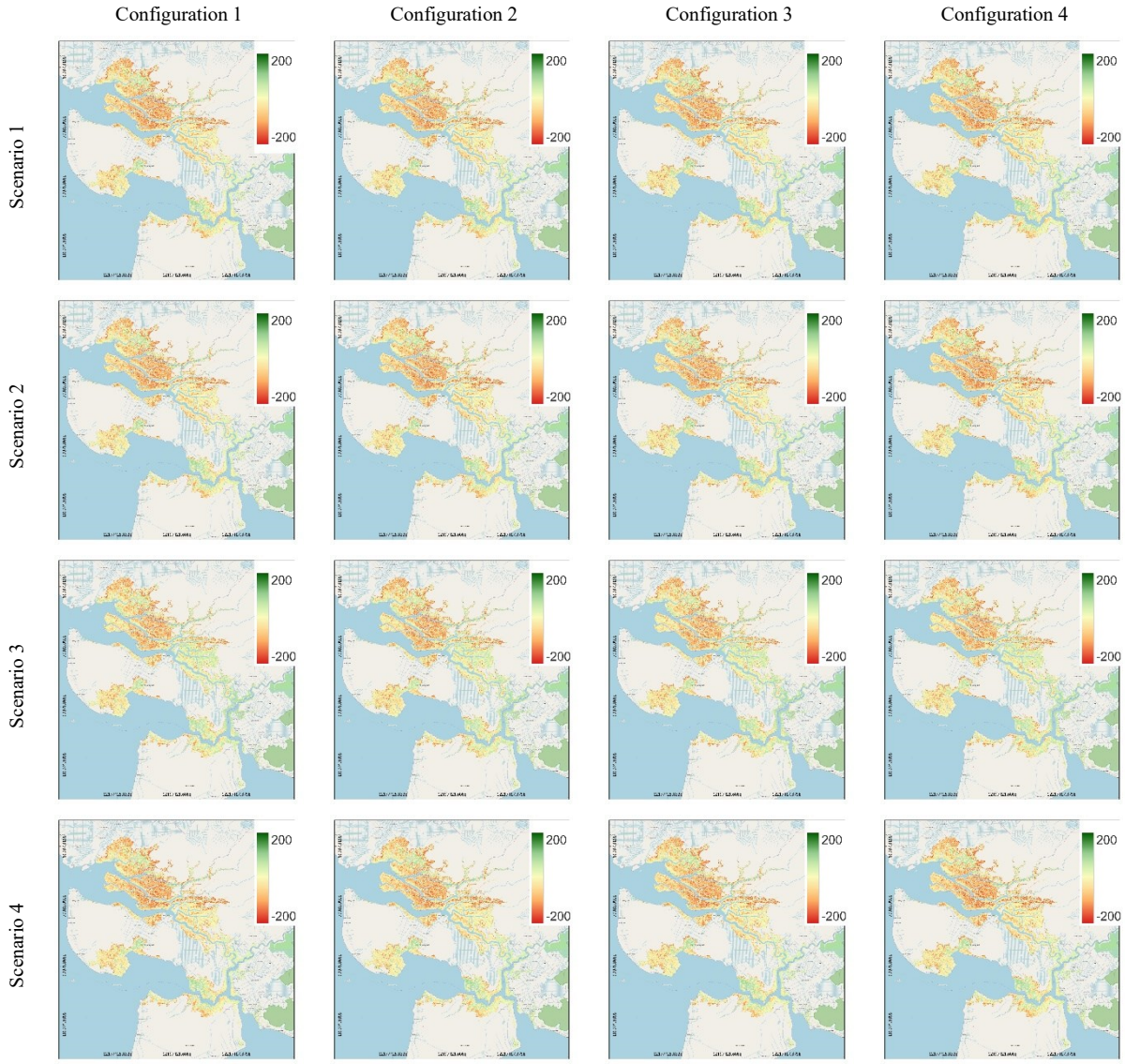
For the models based on the dataset with configuration three, the pattern is quite different. The mean values range from -24.9 Mg/ha to -8.8 Mg/ha. The difference in value at the farthest minimum range was observed in the model with scenario five with a value of -199.3 Mg/ha, whereas scenario -161.1 Mg/ha showed the opposite situation. For the difference with the closest maximum range, it was seen in the model with scenario one with a value of 128.1 Mg/ha. And the furthest maximum range was seen in the model with scenario three with a value of 153.9 Mg/ha.

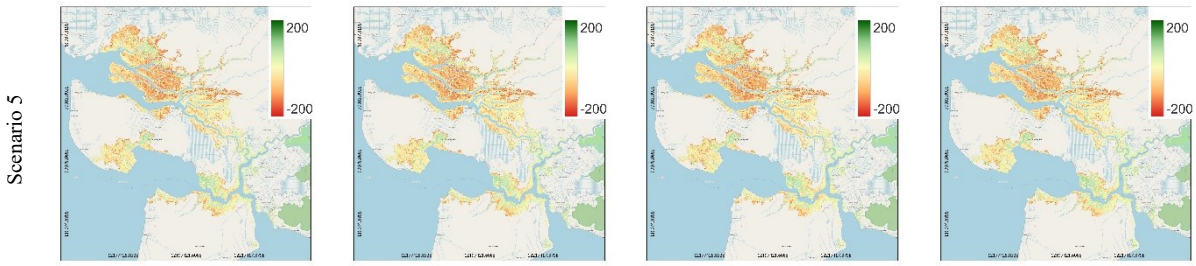
The dataset with configuration four produced biomass prediction models with a different patterned value difference, where the mean value ranges from -21.1 Mg/ha to -9.9 Mg/ha. The model with scenario five has the furthest minimum predicted biomass value difference with a value of -199.5 Mg/ha, whereas scenario three shows the opposite condition with a value of -164.5 Mg/ha. Additionally, the difference in predicted biomass values at the farthest maximum range was seen in the model with scenario one with a value of 162.2 Mg/ha. Meanwhile, in scenario two, there was a difference in the predicted biomass value at the closest maximum range with a value of 133.6 Mg/ha.

Aside from the description of the distribution of differences in predicted biomass values between the two models, the differences in predicted biomass values are also shown in

Figure 13. The maps shown are differentiated based on the parameter usage scenarios and dataset configurations used for model training. The resulting maps had a pixel resolution of 100 meters, which corresponds to the native resolution of the CCI biomass prediction model. In additional, the difference value range shown on the map was set at -200 to 200 with units of Mg/ha. Negative values represented a lower value difference between the predicted model and the model from CCI, while positive values represented the opposite condition.

Figure 13: Predicted Biomass Difference between Prediction Model against ESA`s CCI Biomass Map





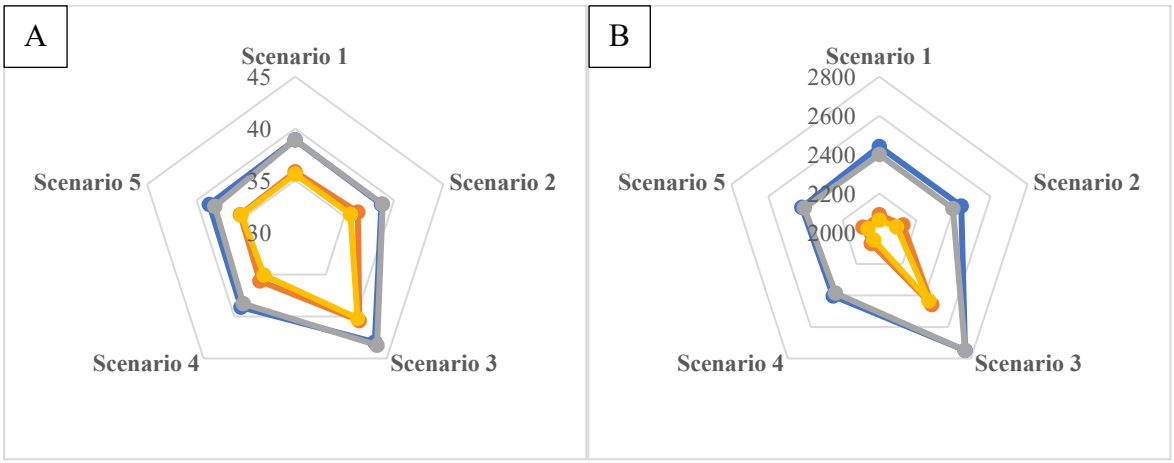
DISCUSSION

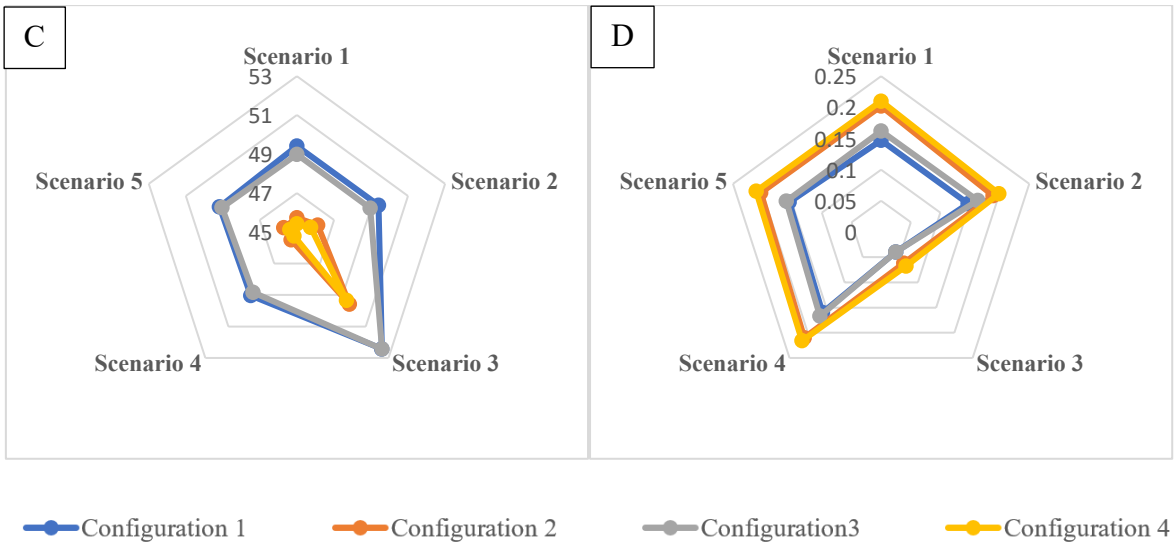
5.1. Mangrove Forest Biomass Estimation Model Performance

5.1.1. Statistic Metric-Based Assessment

This study investigated the capability of decomposed polarimetric parameters when integrated to predict mangrove forest biomass. For this purpose, parameter usage scenarios (Table 6) were introduced into the training process of the mangrove forest biomass prediction model. In addition, the dataset used to train the model was also set up following the configuration described in Table 5 to see the effect of the configuration of the dataset on the modelling results. From such an arrangement, a trained model was obtained for each parameter usage scenario and based on the variation of dataset configurations whose summary statistical metrics are represented in Figure 14 as follows:

Figure 14: Graph Representation of the Statistical Metrics of the Trained Model for Each Scenario Grouped by Dataset Configuration for Model Training, where (A) MAE, (B) MSE, (C) RMSE, and (D) R^2.





According to Figure 11, employing a dataset with a finer level of grid detail resulted in a lower performing prediction model. The decrease in prediction performance based on the statistical metric was consistent for each model related to the applied scenario. Which suggests that when there are no data fields that can represent actual physical conditions, the inclusion of datasets that provide generalized information (from coarser grid sizes) is more advisable. And while the grouping method does not seem to have a significant effect on the model's prediction performance, models utilizing datasets with a mean grouping method provide better prediction performance.

Incorporating all parameters in the model training resulted in better albeit insignificant results. This pattern was consistent regardless of the configuration of the dataset used for model training. Indicating that the incorporation of phase information from decomposed polarimetric parameters can improve the prediction performance of the model. Which was consistent with the results obtained by studies from [8] and [33] although the materials, methods, and objects of study were different from this study. And while there was an overall improvement in prediction performance, the model that utilized backscatter parameters with uncorrected topography resulted in better prediction performance. Showing that in some cases (especially for mangrove biomass prediction), topography correction can reduce the predictive performance of the model. And although their implementation can improve the performance of the prediction model, the decomposed polarimetric parameters themselves have a low correlation with biomass. This is in line with the conclusion drawn in the study by [32], where in general the decomposed polarimetric parameters have a low correlation

with the AGB. Evident in the low statistical metric values of models that only utilized this parameter, regardless of the dataset configuration. SAR signals with specific frequencies and wavelengths can detect a particular component of the forest [65]. Using SAR data with one type of wavelength (*i.e.*, C-band from Sentinel-1) reduces the scope of forest components that can be sensed. Potentially reducing the accuracy, and increasing the uncertainty of values in the model.

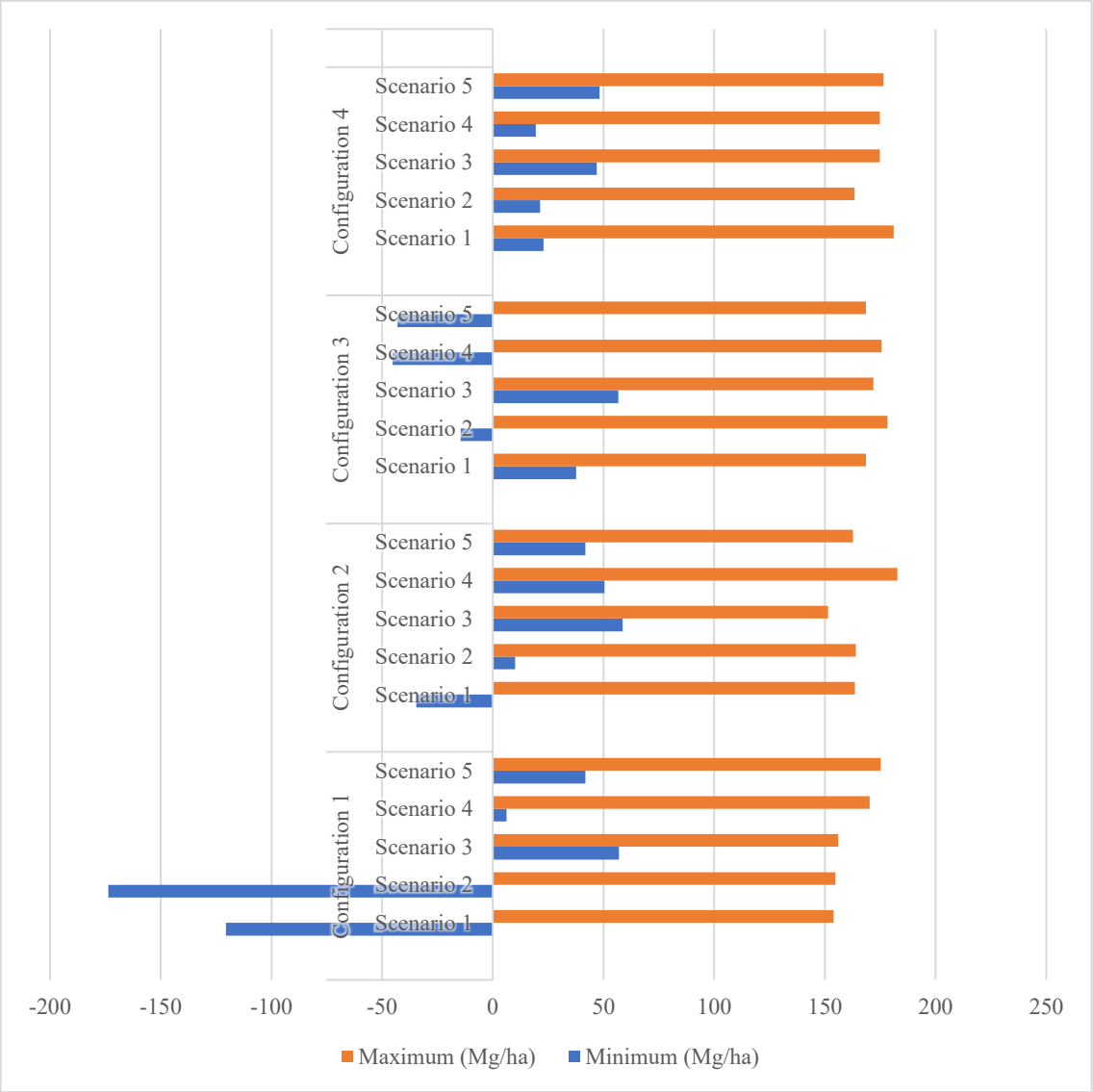
While the values of the statistical metrics obtained from the predicted models were low, this could be explained. Apart from the different study objects, these studies that produced high statistical metric values such as those by [8], [9], and [30] utilized quad-polarization data. Which has a wider range of phase information that can be decomposed than the dual-polarization Sentinel-1 data employed in this study. In quad- polarization data, phase information can be decomposed on the basis of its coherence matrix (T3) given the availability of polarization [60]. From this matrix, a variety of polarimetric decomposition approaches can be performed (*e.g.*, Yamaguchi decomposition), which is not possible with dual-polarization data that have limited phase information.

Additionally, utilizing GEDI-derived biomass information played a role in why the statistical metric values of the developed model were rather low. Considering that studies by [31] and [33] that used the same type of data, despite their different study objects, produced models with higher statistical metric values. Although the usage of biomass information from GEDI would greatly assist the modelling process, given the availability and scope of the data. However, it should be noted that the biomass information is derived from the signal captured by the lidar instrument [37], which still lacks the accuracy compared to field data. Moreover, the object of interest itself, the mangrove, has biophysical characteristics that make analysing it with remote sensing data more complicated [22].

5.1.2. Shift In Predicted Biomass Value Range After Polarimetric Decomposition Parameter Inclusion into The Model

The next step after the trained model has been obtained was to predict the biomass value of the study area. The prediction model was developed based on the dataset with the configuration described in Table 5, which was built following the parameter usage scenario described in Table 6. The obtained prediction model has varying biomass range values, of which the distribution is depicted in the graph in Figure 15 as follows:

Figure 15: Graphical representation of the predicted biomass value range (Mg/ha) of each scenario grouped by dataset configuration for model training.



Regardless of the scenario of parameter usage in the prediction model development, only the models from the dataset of configuration four did not give negative values to the prediction results. For models that used the dataset with configuration one, scenarios that used only the backscatter parameter (whether topography-corrected or not) produced models with negative values. While the dataset of configuration two predicted negative values only in scenario one, which utilized backscatter parameters with uncorrected topography. Meanwhile, with dataset of configuration three, only scenarios one and three did not predict negative values.

While this pattern was only seen in models that employed datasets with the median grouping method, the inclusion of decomposed polarimetric parameters in the prediction models increased the maximum range of predicted biomass values. The largest maximum range was seen in scenario four with dataset configuration two, while the smallest minimum range was seen in scenario four with dataset configuration one. Both prediction models used a combination of backscatter with uncorrected topography with decomposed polarimetric parameters.

The opposite pattern was shown exclusively in models that used the dataset with configuration three. Combining backscatter parameters (both topography-corrected and uncorrected) with decomposed polarimetric parameters caused the models to predict negative values. This condition was seen not only in scenario four, which utilized backscatter parameters with uncorrected topography, but also in scenario five, which utilized backscatter parameters with corrected topography. And regardless of the dataset configuration that was used, scenario three had stable minimum and maximum range values. Where the longest maximum range and the shortest minimum range were seen in the model using dataset configuration four.

Attention is turned to the changes in the minimum and maximum predicted values of the prediction model after the decomposed polarimetric parameters were added. Where the distribution was represented in the form of a heatmap with three colour levels as shown in Figure 16. While the green colour represents the highest maximum range value that can be recorded by the model, the opposite condition applies to the minimum value.

Scenarios one and two only utilized backscatter parameters with the difference that scenario two used topography-corrected parameters, while scenario one did not. Whereas scenarios four and five combined decomposed polarimetric parameters with backscatter parameters in the modelling process, both uncorrected and corrected topography, respectively. Therefore, changes in the range of values can be seen between scenario one and scenario four, and scenario two and scenario five.

The addition of decomposed polarimetric parameters did not consistently increase the maximum range or decrease the minimum range of biomass predicted by the model. The increase in the maximum value range was only observed in scenario four when datasets with configurations two and three were utilized. Additionally, when the dataset with configuration four was employed, the increase in the maximum range was only seen in scenario five. Only

the models that used the dataset with configuration one consistently experienced an increase in the maximum range after adding the decomposed polarimetric parameters to the model. Moreover, the models that utilized the dataset with configuration one also obtained the maximum value range with the largest change compared to the other models.

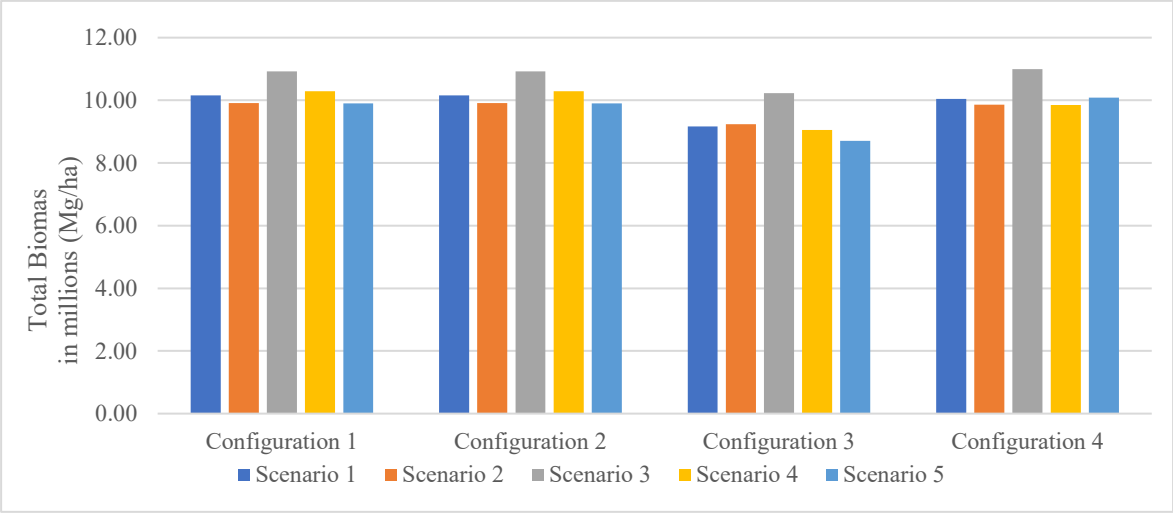
Figure 16: Heatmap representation of the maximum and minimum range of predicted models (in Mg/ha).

	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
Maximum Range (Mg/ha)					
Configuration 1	154.0	154.7	156.1	170.2	175.3
Configuration 2	163.5	164.0	151.4	182.7	162.7
Configuration 3	168.6	178.2	171.9	175.6	168.5
Configuration 4	181.1	163.5	174.8	174.8	176.4
Minimum Range (Mg/ha)					
Configuration 1	-120.5	-173.6	57.0	6.3	41.8
Configuration 2	-34.6	10.0	58.7	50.4	41.8
Configuration 3	37.6	-14.6	56.8	-45.2	-43.1
Configuration 4	22.9	21.3	47.0	19.4	48.3

The addition of decomposed polarimetric parameters also did not consistently reduce the minimum range of predicted biomass values. Out of all the combinations of scenarios and dataset configurations, only scenario four that used the dataset with configuration four experienced a decrease in the minimum value range. Scenarios four and five that used the datasets with configurations one and three also technically decreased the minimum value, but the predicted biomass was at a negative value. The exception was seen in scenario four that used the dataset with configuration one. In which, the minimum range obtained was the lowest when compared to the other models despite initially being recorded as a negative value.

Additionally, the total predicted biomass of each model was also calculated. And to facilitate the comparison process, the pixel resolution of the models' utilizing datasets with configurations one and three was also adjusted to that of the models from datasets with configurations two and four. A graphical representation of the predicted total biomass for each model is shown in Figure 17, as follows:

Figure 17: Graphical representation with supplementary tables of Total Biomass Predictions from the developed model (units in Mg/ha) for each scenario, grouped by data set configuration.



	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
Configuration 1	10,159,333.00	9,906,947.00	10,924,518.00	10,294,613.00	9,905,216.00
Configuration 2	10,159,333.00	9,906,947.00	10,924,518.00	10,294,613.00	9,905,216.00
Configuration 3	9,164,928.00	9,242,199.00	10,230,293.00	9,049,953.00	8,704,135.00
Configuration 4	10,039,851.00	9,861,496.00	10,999,729.00	9,853,135.00	10,084,304.00

Looking at the changes in total predicted biomass before and after the addition of decomposed polarimetric parameters, only three models showed an increase in the total amount of predicted biomass. The models in question were the model with scenario four that used the dataset with configuration one, the model with scenario four that used the dataset with configuration two and the model with scenario five that used the dataset with configuration. The model that used the dataset with configuration three consistently showed a decrease after the decomposed polarimetric parameters were added.

While the addition of decomposed polarimetric parameters did not show a consistent change in total predicted biomass, a clearer pattern was seen in the pixel clustering method for dataset formation. Particularly the models from datasets with configurations one and two that use the median clustering method to obtain new pixel values. This consistently resulted in an increase in the total biomass value, as seen in scenario four, and a reduction in the total biomass value in scenario five. Whereas the dataset that used the mean grouping method, especially the one with 100-meter grid size, showed a comparable pattern of predicted changes in total biomass. And the dataset with configuration three using the mean clustering

method with a 100-meter grid size consistently experienced a reduction in predicted biomass values after decomposed polarimetric parameters were added.

5.2. Comparative Assessment Against Other Product

The predictive performance of the developed model was also assessed by comparing it with prediction models from other initiatives, particularly with ESA's CCI prediction model. Through comparison of the maximum and minimum values of the predicted biomass, the ability of the model to predict biomass relative to the CCI model can be determined.

Looking back to Figure 16, regardless of the scenario and configuration of the dataset used, it was observed that the model with scenario four utilizing the dataset with configuration two was able to predict mangrove biomass up to 182.7 Mg/ha. While the lowest predicted biomass value, provided that negative values were not considered, was shown by the model with scenario four utilizing the dataset with configuration one, which amounted to 6.3 Mg/ha. It was also noted that the range of predicted biomass values from the prediction model by CCI was from 11.0 Mg/ha to 258.0 Mg/ha. This indicated that the model could predict biomass values at the lower end of the minimum range, but not vice versa. However, it is important to note that these highest maximum and lowest minimum biomass values are from models with different scenarios and dataset configurations.

The lower predicted biomass relative to that produced by the CCI prediction model was also reflected in the difference in predicted total biomass values. The predicted total biomass values of the developed model range from 8,704,135 Mg/ha to 10,999,729 Mg/ha. Whereas the CCI predicted biomass prediction model with 100-meter pixel resolution predicted a total mangrove biomass of 10,177,282 Mg/ha. Except for the model with scenario four that used a dataset with configuration two, the resulting models had lower total biomass relative to the predicted model from CCI. The low biomass values predicted by the developed model were in line with those found in a study by [50], where biomass was generally underestimated by GEDI L4A. The biomass values obtained were also relatively lower when compared to biomass prediction models from other projects. Additionally, the model with scenario three showed consistently high predicted values of total biomass. However, these results cannot be taken into consideration as the models have low predictive performance according to statistical metrics.

The range of the difference in biomass values provides an insight into the relative predicted biomass values between the developed prediction model and the CCI prediction model.

Figure 18 represents the range with different colours. Where the green colour represents the shortest range and the red colour represents the opposite. While Figure 19 illustrates the minimum and maximum limits of this value difference. Negative minimum values indicate that the predicted biomass in the developed model is lower than in the model from CCI, and positive values indicate the opposite condition.

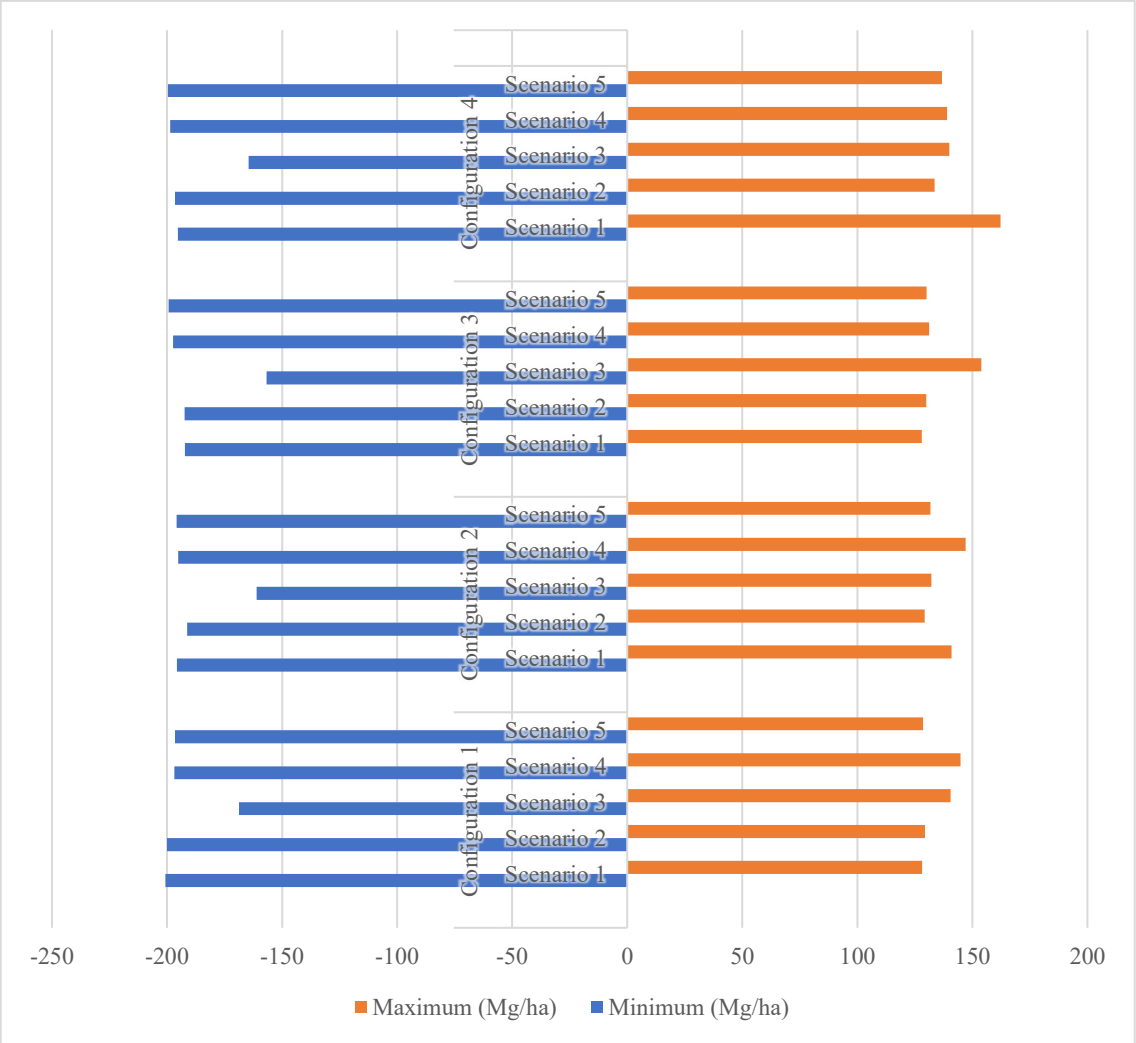
Figure 18: Heatmap representation of the range from difference (in Mg/ha) between developed prediction models against biomass model from CCI by ESA.

	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
Configuration 1	328.8	329.5	309.3	341.7	325.2
Configuration 2	336.7	320.5	293.4	342.2	327.6
Configuration 3	320.3	322.4	310.7	328.6	329.5
Configuration 4	357.5	330.1	304.4	337.8	336.3

The range of differences in predicted biomass values between the two models did not exhibit a consistent pattern. The reduction in the difference in predicted biomass values indicated that the biomass predicted by the developed model was close to the biomass predicted by ESA's CCI biomass model. The values shown can give an idea of the extent of the difference in predicted biomass values between the two models, although the direction of the difference is not considered. Of these, only the model with scenario five that used the dataset with configuration one and the model with scenario four that used the dataset with configuration four showed a reduction in the range of difference values. Which suggests that, except for these two models, the addition of the decomposed polarimetric parameter increases the range of differences in the predicted biomass of the two models.

In terms of the difference in the predicted biomass values of the two models, adding decomposed polarimetric parameters appeared to show an inconsistent pattern, either increasing the maximum range or decreasing the minimum range. Only the models from the four-configuration dataset did not show any increase in the predicted biomass values. And models from datasets with configuration three show an increase, although not significant. Whereas the models from the dataset with configuration one, especially in scenario four, showed the largest increase in predicted biomass values.

Figure 19: Column graphical representation of the maximum and minimum difference of predicted biomass value between the developed model and ESA's CCI biomass model.



Adding the decomposed polarimetric parameter to the mangrove biomass prediction also did not consistently reduce or increase the minimum range of the difference in predicted biomass between the two models. Out of the four configurations, only the models from the dataset with configuration one demonstrated a reduction in the minimum range after the decomposed polarimetric parameter was added. Indicating that with the addition of the decomposed polarimetric parameter to the prediction process, the minimum value of the predicted biomass became closer relative to the CCI biomass model.

CONCLUSIONS

This study aimed to identify the impact of decomposed polarimetric parameters from Sentinel-1 SAR images to estimate mangrove forest biomass. Biomass information from

GEDI was utilized as an independent variable, while SAR parameters were used as dependent parameters. In the process, the model building followed a parameter usage scenario, in which the backscatter parameters and decomposed polarimetric parameters were used both individually and in groups. Moreover, to account for the horizontal accuracy of the GEDI data, a newly generated dataset obtained from resampling and aggregation of the input data was calculated and then plotted on a coarser resolution grid. This dataset was then used as the basis for the mangrove biomass modelling.

Through the arrangements described, mangrove forest biomass prediction models have been obtained. Consistent with findings from studies by [8] and [33], It was found that integrating the decomposed polarimetric parameters with the backscatter parameters can improve the predictive performance of the resulting models. The model that utilized backscatter parameters with uncorrected topography gave the best results for biomass prediction. Additionally, utilizing datasets with a mean clustering method with a coarser grid size was seen to produce better performing models. However, utilizing the decomposed polarimetric parameters individually gave the model with the lowest predictive performance compared to the others. This indicates a poor relationship between the decomposed polarimetric parameters and the biomass information. Though this parameter can improve the predictive performance when integrated with other parameters.

Furthermore, comparison with ESA's CCI biomass prediction model showed that the developed model generally predicted lower biomass values. This was consistent with findings from a study by [50], which found a tendency for GEDI to underestimate AGBD. Reflected both in the difference between the maximum and minimum ranges of predicted biomass, as well as in the total predicted biomass. This difference was smaller for the coarser resolution model (100 meters) than for the finer resolution model (25 meters).

For future studies, it would be advisable to assess in more depth the correlation of decomposed polarimetric parameters with biomass with varying density levels. Especially in mangrove forest ecosystems, in which the variability is caused by intrinsic factors of the mangrove plants themselves. Furthermore, integrating field data that can show the actual information more accurately would be highly recommended to obtain a better performing model.

REFERENCES

[1] K. Bimrah, R. Dasgupta, and I. Saizen, "Cultural Ecosystem Services of Mangroves: A Review of Models and Methods," in *Assessing, Mapping and Modelling of Mangrove Ecosystem Services in the Asia-Pacific Region*, Springer, Singapore, 2022, pp. 239–250. doi: 10.1007/978-981-19-2738-6_13.

[2] D. E. Marois and W. J. Mitsch, "Coastal protection from tsunamis and cyclones provided by mangrove wetlands – a review," *Int. J. Biodivers. Sci. Ecosyst. Serv. Manag.*, vol. 11, no. 1, pp. 71–83, Jan. 2015, doi: 10.1080/21513732.2014.997292.

[3] B. T. Van Zanten *et al.*, *The Economics of Large-scale Mangrove Conservation and Restoration in Indonesia: Internal Document*. World Bank, 2021. doi: 10.1596/37605.

[4] S. A. Dalimunthe, I. A. P. Putri, and A. P. S. Prasojo, "Depicting Mangrove's Potential as Blue Carbon Champion in Indonesia," in *Assessing, Mapping and Modelling of Mangrove Ecosystem Services in the Asia-Pacific Region*, R. Dasgupta, S. Hashimoto, and O. Saito, Eds., Singapore: Springer Nature, 2022, pp. 167–181. doi: 10.1007/978-981-19-2738-6_9.

[5] L. Wang, M. Jia, D. Yin, and J. Tian, "A review of remote sensing for mangrove forests: 1956–2018," *Remote Sens. Environ.*, vol. 231, p. 111223, Sep. 2019, doi: 10.1016/j.rse.2019.111223.

[6] T. D. Pham, N. Yokoya, D. T. Bui, K. Yoshino, and D. A. Friess, "Remote Sensing Approaches for Monitoring Mangrove Species, Structure, and Biomass: Opportunities and Challenges," *Remote Sens.*, vol. 11, no. 3, Art. no. 3, Jan. 2019, doi: 10.3390/rs11030230.

[7] N. Joshi *et al.*, "Understanding 'saturation' of radar signals over forests," *Sci. Rep.*, vol. 7, no. 1, Art. no. 1, Jun. 2017, doi: 10.1038/s41598-017-03469-3.

[8] P. Zeng, W. Zhang, Y. Li, J. Shi, and Z. Wang, "Forest Total and Component Above-Ground Biomass (AGB) Estimation through C- and L-band Polarimetric SAR Data," *Forests*, vol. 13, no. 3, Art. no. 3, Mar. 2022, doi: 10.3390/f13030442.

[9] Z. Liu, O. O. Michel, G. Wu, Y. Mao, Y. Hu, and W. Fan, “The Potential of Fully Polarized ALOS-2 Data for Estimating Forest Above-Ground Biomass,” *Remote Sens.*, vol. 14, no. 3, Art. no. 3, Jan. 2022, doi: 10.3390/rs14030669.

[10] Y. Hu, Y. Nie, Z. Liu, G. Wu, and W. Fan, “Improving the Potential of Coniferous Forest Aboveground Biomass Estimation by Integrating C- and L-Band SAR Data with Feature Selection and Non-Parametric Model,” *Remote Sens.*, vol. 15, no. 17, Art. no. 17, Jan. 2023, doi: 10.3390/rs15174194.

[11] A. Bhatt, S. Pathak, M. Kr. Aghwariya, P. Kr. Pal, A. Kumar, and T. Goel, “Model Based Decomposition of Alakananda River Basin using SAR Data,” in *2023 First International Conference on Microwave, Antenna and Communication (MAC)*, Mar. 2023, pp. 1–6. doi: 10.1109/MAC58191.2023.10177102.

[12] S. Salma, N. Keerthana, and B. M. Dodamani, “Target decomposition using dual-polarization sentinel-1 SAR data: Study on crop growth analysis,” *Remote Sens. Appl. Soc. Environ.*, vol. 28, p. 100854, Nov. 2022, doi: 10.1016/j.rsase.2022.100854.

[13] R. Dave *et al.*, “Sensitivity Analysis of Crop Biophysical Parameters Using Multi-Temporal Dual–Polarization Sentinel -1 C-Band SAR Data,” in *2021 IEEE International India Geoscience and Remote Sensing Symposium (InGARSS)*, Dec. 2021, pp. 214–217. doi: 10.1109/InGARSS51564.2021.9791994.

[14] J. B. de Jesus, T. M. Kuplich, Í. D. de C. Barreto, F. L. Hillebrand, and C. N. da Rosa, “Estimation of aboveground biomass of arboreal species in the semi-arid region of Brazil using SAR (synthetic aperture radar) images,” *J. Arid Land*, vol. 15, no. 6, pp. 695–709, Jun. 2023, doi: 10.1007/s40333-023-0017-4.

[15] S. Fatnassi, M. Yahia, T. Ali, and M. Mortula, “Polarimetric SAR Characterization of Mangrove Forest Environment in the United Arab Emirates (UAE),” *Int. J. Adv. Comput. Sci. Appl. IJACSA*, vol. 14, no. 3, Art. no. 3, 42/30 2023, doi: 10.14569/IJACSA.2023.0140380.

[16] M. F. Ghazali and K. WIkantika, “Pre-assessment of the Potential of Dual Polarization of Sentinel 1 Data for Mapping the Mangrove Tree Species Distribution in South Bali, Indonesia,” in *2021 7th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR)*, Nov. 2021, pp. 1–6. doi: 10.1109/APSAR52370.2021.9688441.

[17] D. M. Alongi, “Present state and future of the world’s mangrove forests,” *Environ. Conserv.*, vol. 29, no. 3, pp. 331–349, Sep. 2002, doi: 10.1017/S0376892902000231.

[18] S. S. Romañach *et al.*, “Conservation and restoration of mangroves: Global status, perspectives, and prognosis,” *Ocean Coast. Manag.*, vol. 154, pp. 72–82, Mar. 2018, doi: 10.1016/j.ocecoaman.2018.01.009.

[19] S. D. Sasmito, M. Basyuni, A. Kridalaksana, M. F. Saragi-Sasmito, C. E. Lovelock, and D. Murdiyarso, “Challenges and opportunities for achieving Sustainable Development Goals through restoration of Indonesia’s mangroves,” *Nat. Ecol. Evol.*, vol. 7, no. 1, pp. 62–70, Jan. 2023, doi: 10.1038/s41559-022-01926-5.

[20] Rahman *et al.*, “A review on the biodiversity and conservation of mangrove ecosystems in Indonesia,” *Biodivers. Conserv.*, vol. 33, no. 3, pp. 875–903, Mar. 2024, doi: 10.1007/s10531-023-02767-9.

[21] V. B. Arifanti *et al.*, “Challenges and Strategies for Sustainable Mangrove Management in Indonesia: A Review,” *Forests*, vol. 13, no. 5, Art. no. 5, May 2022, doi: 10.3390/f13050695.

[22] C. Kuenzer, A. Bluemel, S. Gebhardt, T. V. Quoc, and S. Dech, “Remote Sensing of Mangrove Ecosystems: A Review,” *Remote Sens.*, vol. 3, no. 5, Art. no. 5, May 2011, doi: 10.3390/rs3050878.

[23] C. Giri *et al.*, “Status and distribution of mangrove forests of the world using earth observation satellite data,” *Glob. Ecol. Biogeogr.*, vol. 20, no. 1, pp. 154–159, 2011, doi: 10.1111/j.1466-8238.2010.00584.x.

[24] P. Bunting *et al.*, “The Global Mangrove Watch—A New 2010 Global Baseline of Mangrove Extent,” *Remote Sens.*, vol. 10, no. 10, Art. no. 10, Oct. 2018, doi: 10.3390/rs10101669.

[25] P. Bunting *et al.*, “Global Mangrove Extent Change 1996–2020: Global Mangrove Watch Version 3.0,” *Remote Sens.*, vol. 14, no. 15, Art. no. 15, Jan. 2022, doi: 10.3390/rs14153657.

[26] M. Jia *et al.*, “Mapping global distribution of mangrove forests at 10-m resolution,” *Sci. Bull.*, vol. 68, no. 12, pp. 1306–1316, Jun. 2023, doi: 10.1016/j.scib.2023.05.004.

[27] J. A. Navarro, N. Algeet, A. Fernández-Landa, J. Esteban, P. Rodríguez-Noriega, and M. L. Guillén-Climent, “Integration of UAV, Sentinel-1, and Sentinel-2 Data for Mangrove Plantation Aboveground Biomass Monitoring in Senegal,” *Remote Sens.*, vol. 11, no. 1, Art. no. 1, Jan. 2019, doi: 10.3390/rs11010077.

[28] J. M. M. Yancho, T. G. Jones, S. R. Gandhi, C. Ferster, A. Lin, and L. Glass, “The Google Earth Engine Mangrove Mapping Methodology (GEEMMM),” *Remote Sens.*, vol. 12, no. 22, Art. no. 22, Jan. 2020, doi: 10.3390/rs12223758.

[29] A. Singh *et al.*, “Aboveground Forest Biomass Estimation by the Integration of TLS and ALOS PALSAR Data Using Machine Learning,” *Remote Sens.*, vol. 15, no. 4, Art. no. 4, Jan. 2023, doi: 10.3390/rs15041143.

[30] H. J. Persson, R. Mukhopadhyay, I. Huuva, and J. E. S. Fransson, “Comparison of Boreal Biomass Estimations Using C- and X-Band Polsar,” in *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, Jul. 2022, pp. 5555–5558. doi: 10.1109/IGARSS46834.2022.9884506.

[31] S. Sinha, “H/A/α Polarimetric Decomposition Of Dual Polarized Alos Palsar For Efficient Land Feature Detection And Biomass Estimation Over Tropical Deciduous Forest,” *Geogr. Environ. Sustain.*, vol. 15, no. 3, Art. no. 3, Oct. 2022, doi: 10.24057/2071-9388-2021-095.

[32] J. B. de Jesus, T. M. Kuplich, Í. D. de C. Barreto, and D. C. Gama, “Dual polarimetric decomposition in Sentinel-1 images to estimate aboveground biomass of arboreal caatinga,” *Remote Sens. Appl. Soc. Environ.*, vol. 29, p. 100897, Jan. 2023, doi: 10.1016/j.rsase.2022.100897.

[33] Rucha B. Dave, Koushik Saha, Amit Kushwaha, Manisha Vithalpura, Nidhin P., and Abishek Murugesan, “Analysing the potential of polarimetric decomposition parameters of Sentinel–1 dual-pol SAR data for estimation of rice crop biophysical parameters,” *J. Agrometeorol.*, vol. 25, no. 1, Feb. 2023, doi: 10.54386/jam.v25i1.2039.

[34] K. Harfenmeister, S. Itzerott, C. Weltzien, and D. Spengler, “Agricultural Monitoring Using Polarimetric Decomposition Parameters of Sentinel-1 Data,” *Remote Sens.*, vol. 13, no. 4, Art. no. 4, Jan. 2021, doi: 10.3390/rs13040575.

[35] B. H. Kiyohara and E. E. Sano, “Evaluation of polarimetric data and texture attributes in SAR images to discriminate secondary forest in an area of amazon rainforest,” *Ciênc. Florest.*, vol. 33, p. e71235, Aug. 2023, doi: 10.5902/1980509871235.

[36] G. Zhou, Z. Wang, H. Miao, C. Jiang, and G. Jing, “Wetland Classification in Poyang Lake Using Dual-polarization Synthetic Aperture Radar Data with Feature Combination,” *Sens. Mater.*, vol. 33, no. 12, p. 4607, Dec. 2021, doi: 10.18494/SAM.2021.3590.

[37] R. Dubayah *et al.*, “The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth’s forests and topography,” *Sci. Remote Sens.*, vol. 1, p. 100002, Jun. 2020, doi: 10.1016/j.srs.2020.100002.

[38] M. Schwartz *et al.*, “FORMS: Forest Multiple Source height, wood volume, and biomass maps in France at 10 to 30 m resolution based on Sentinel-1, Sentinel-2, and Global Ecosystem Dynamics Investigation (GEDI) data with a deep learning approach,” *Earth Syst. Sci. Data*, vol. 15, no. 11, pp. 4927–4945, Nov. 2023, doi: 10.5194/essd-15-4927-2023.

[39] M. Liang, L. Duncanson, J. A. Silva, and F. Sedano, “Quantifying aboveground biomass dynamics from charcoal degradation in Mozambique using GEDI Lidar and Landsat,” *Remote Sens. Environ.*, vol. 284, p. 113367, Jan. 2023, doi: 10.1016/j.rse.2022.113367.

[40] C. Kusmana, “Distribution and Current Status of Mangrove Forests in Indonesia,” in *Mangrove Ecosystems of Asia*, Springer, New York, NY, 2014, pp. 37–60. doi: 10.1007/978-1-4614-8582-7_3.

[41] H. F. Nurofq *et al.*, *Status Hutan & Kehutanan Indonesia 2022: Menuju FOLU Net Sink 2030*. Ministry of Environment and Forestry, 2022. [Online]. Available: https://statistik.menlhk.go.id/sisklhkX/assets/publikasi/file/00_FULL_SoIFo%202022%20B5_IDN_31.08.23_opt.pdf

[42] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore, “Google Earth Engine: Planetary-scale geospatial analysis for everyone,” *Remote Sens. Environ.*, vol. 202, pp. 18–27, Dec. 2017, doi: 10.1016/j.rse.2017.06.031.

[43] E. LeDell and S. Poirier, “H2O AutoML: Scalable Automatic Machine Learning,” 2020.

[44] “User Guides - Sentinel-1 SAR - Sentinel Online,” Sentinel Online. Accessed: Mar. 04, 2024. [Online]. Available: https://copernicus.eu/user-guides/sentinel-1-sar

[45] ESA, “Single Look Complex - Sentinel-1 SAR Technical Guide - Sentinel Online,” Sentinel Online. Accessed: Mar. 17, 2024. [Online]. Available: https://copernicus.eu/technical-guides/sentinel-1-sar/products-algorithms/level-1-algorithms/single-look-complex

[46] ESA, “Ground Range Detected - Sentinel-1 SAR Technical Guide - Sentinel Online,” Sentinel Online. Accessed: Mar. 17, 2024. [Online]. Available: https://copernicus.eu/technical-guides/sentinel-1-sar/products-algorithms/level-1-algorithms/ground-range-detected

[47] R. O. Dubayah *et al.*, “GEDI L4A Footprint Level Aboveground Biomass Density, Version 2.1,” *ORNL DAAC*, Mar. 2022, doi: 10.3334/ORNLDAAC/2056.

[48] J. R. Kellner, J. Armston, and L. Duncanson, “Algorithm Theoretical Basis Document for GEDI Footprint Aboveground Biomass Density,” *Earth Space Sci.*, vol. 10, no. 4, p. e2022EA002516, 2023, doi: 10.1029/2022EA002516.

[49] Google Earth Engine Developer, “GEDI L4A Raster Aboveground Biomass Density, Version 2.1 | Earth Engine Data Catalog,” Google for Developers. Accessed: Mar. 05, 2024. [Online]. Available: https://developers.google.com/earth-engine/datasets/catalog/LARSE_GEDI_GEDI04_A_002_MONTHLY

[50] D. Jia, C. Wang, C. R. Hakkenberg, I. Numata, A. J. Elmore, and M. A. Cochrane, “Accuracy evaluation and effect factor analysis of GEDI aboveground biomass product for temperate forests in the conterminous United States,” *GIScience Remote Sens.*, vol. 61, no. 1, p. 2292374, Dec. 2024, doi: 10.1080/15481603.2023.2292374.

[51] A. Pascual *et al.*, “Assessing the performance of NASA’s GEDI L4A footprint aboveground biomass density models using National Forest Inventory and airborne laser scanning data in Mediterranean forest ecosystems,” *For. Ecol. Manag.*, vol. 538, p. 120975, Jun. 2023, doi: 10.1016/j.foreco.2023.120975.

[52] P. Bunting, A. Rosenqvist, L. Hilarides, R. Lucas, and N. Thomas, “Global Mangrove Watch: Updated 2010 Mangrove Forest Extent (v2.5),” *Remote Sens.*, vol. 14, p. 1034, Feb. 2022, doi: 10.3390/rs14041034.

[53] M. Santoro and O. Cartus, “ESA Biomass Climate Change Initiative (Biomass_cci): Global datasets of forest above-ground biomass for the years 2010, 2017, 2018, 2019 and 2020, v4.” [object Object], p. 5183 Files, 302039459020 B, 2023. doi: 10.5285/AF60720C1E404A9E9D2C145D2B2EAD4E.

[54] Google Earth Engine Developer, “Sentinel-1 SAR GRD: C-band Synthetic Aperture Radar Ground Range Detected, log scaling | Earth Engine Data Catalog | Google for Developers.” Accessed: Mar. 13, 2024. [Online]. Available:

https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S1_GRD

[55] K. Conradsen, A. A. Nielsen, and H. Skriver, "Determining the Points of Change in Time Series of Polarimetric SAR Data," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 3007–3024, May 2016, doi: 10.1109/TGRS.2015.2510160.

[56] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou, "A tutorial on synthetic aperture radar," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 1, pp. 6–43, Mar. 2013, doi: 10.1109/MGRS.2013.2248301.

[57] S. R. Cloude and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 1, pp. 68–78, Jan. 1997, doi: 10.1109/36.551935.

[58] S. Cloude, "The Dual Polarization Entropy/Alpha Decomposition: A PALSAR Case Study," vol. 644, p. 2, Mar. 2007.

[59] K. Ji and Y. Wu, "Scattering Mechanism Extraction by a Modified Cloude-Pottier Decomposition for Dual Polarization SAR," *Remote Sens.*, vol. 7, no. 6, Art. no. 6, Jun. 2015, doi: 10.3390/rs70607447.

[60] J.-S. L. Pottier Eric, *Polarimetric Radar Imaging: From Basics to Applications*. Boca Raton: CRC Press, 2017. doi: 10.1201/9781420054989.

[61] M. Ouarzeddine, B. Souissi, and A. Belhadj-Aissa, "Target Detection and Characterization using H/alpha Decomposition and Polarimetric Signatures," in *2006 2nd International Conference on Information & Communication Technologies*, Apr. 2006, pp. 395–400. doi: 10.1109/ICTTA.2006.1684402.

[62] D. Mandal *et al.*, "Sentinel-1 SLC Preprocessing Workflow for Polarimetric Applications: A Generic Practice for Generating Dual-pol Covariance Matrix Elements in SNAP S-1 Toolbox." Preprints, Nov. 30, 2019. doi: 10.20944/preprints201911.0393.v1.

[63] E. A. Muñoz, A. Zozaya, and E. Lindquist, “Fusion of Sentinel 1 and Alos Palsar Data to Separate Palm Oil Plantations from Forest Cover Mapping using Pauli Decomposition Approach,” *IEEE Lat. Am. Trans.*, vol. 20, no. 6, pp. 921–930, Jun. 2022, doi: 10.1109/TLA.2022.9757374.

[64] Y. Shendryk, “Fusing GEDI with earth observation data for large area aboveground biomass mapping,” *Int. J. Appl. Earth Obs. Geoinformation*, vol. 115, p. 103108, Dec. 2022, doi: 10.1016/j.jag.2022.103108.

[65] M. Santoro and O. Cartus, “Research Pathways of Forest Above-Ground Biomass Estimation Based on SAR Backscatter and Interferometric SAR Observations,” *Remote Sens.*, vol. 10, no. 4, Art. no. 4, Apr. 2018, doi: 10.3390/rs10040608.