

Leaflet No. 14/ 2022

The Republic of the Union of Myanmar
Ministry of Natural Resources and Environmental Conservation
Forest Department



**Estimating Aboveground Biomass of Two Different Forest Types in Myanmar
from Sentinel-2 Data with Machine Learning and Geostatistical Algorithms**

Phyo Wai
Staff Officer
Forest Department

February, 2023

Contents

	Pages
Abstract	i
စာတမ်းအကျဉ်းချုပ်	ii
1. Introduction	1
2. Objectives	3
3. Materials and Methods	3
3.1 Study area	3
3.2 Data Collection and Processing Methods	4
3.3 Aboveground Biomass Detection Methods	9
3.4 Accuracy Assessment	11
4. Results	12
4.1 Variable Importance and Selections	12
4.2 Validation Metrics for RF and SGB Models	12
4.3 Semi variogram Analysis Results of RF-Derived Residuals	13
4.4 Forest AGB Mapping Results based on RFOK and RFCK Models	16
5. Discussion	17
5.1 Sensitivity of Sentinel-2 Derivatives to AGB	17
5.2 Sensitivity of Topographic Variables to AGB	19
5.3 Comparison between Models	19
5.4 Effects of Forest Management on AGB in the Study Sites	20
5.5 Attainment for SDG and REDD+	21
6. Conclusion	21
7. Acknowledgements	22
8. References	23

List of Tables

	Pages
Table 1	5
Table 2	6
Table 3	7
Table 4	9
Table 5	13
Table 6	14
Table 7	15

List of Figures

	Pages
Figure 1	4
Figure 2	8
Figure 3	12
Figure 4	14
Figure 5	15
Figure 6	16
Figure 7	17

Abstract

Estimating Aboveground Biomass of Two Different Forest Types in Myanmar from Sentinel-2 Data with Machine Learning and Geostatistical Algorithms

Phyo Wai, Staff Officer

Forest Department

Abstract

The accurate estimation of spatially explicit forest aboveground biomass (AGB) provides an essential basis for sustainable forest management and carbon sequestration accounting, especially in Myanmar, where there is a lack of data for forest conservation due to operational limitations. This study mapped the forest AGB using Sentinel-2 (S-2) images and Shuttle Radar Topographic Mission (SRTM) based on random forest (RF), stochastic gradient boosting (SGB) and Kriging algorithms in two forest reserves (Namhton and Yinmar) in Myanmar, and compared their performance against AGB measured by the traditional methods. Specifically, a suite of forest sample plots were deployed in the two forest reserves, and forest attributes were measured to calculate the plot-level AGB based on allometric equations. The spectral bands, vegetation indices (VIs) and textures derived from processed S-2 data and topographic parameters from SRTM were utilized to statistically link with field-based AGB by implementing random forest (RF) and stochastic gradient boosting (SGB) algorithms. Followed by an evaluation of the algorithmic performances, RF-based Kriging (RFK) models were employed to determine the spatial distribution of AGB as an improvement of accuracy against

RF models. The study's results showed that textural measures produced from wavelet analysis (WA) and vegetation indices (VIs) from Sentinel-2 were the strongest predictors for evergreen forest reserve (Namhton) AGB prediction and spectral bands and vegetation indices (VIs) showed the highest sensitivity to the deciduous forest reserve (Yinmar) AGB prediction. The fitted models were RF-based ordinary Kriging (RFOK) for Namhton forest reserve and RF-based co-Kriging (RFCK) for Yinmar forest reserve because their respective R^2 , whilst the RMSE values were validated as 0.47 and 24.91 AGB t/ha and 0.52 and 34.72 AGB t/ha, respectively. The proposed random forest Kriging framework provides robust AGB maps, which are essential to estimate the carbon sequestration potential in the context of REDD+. From this particular study, we suggest that the protection/disturbance status of forests affects AGB values directly in the study area; thus, community-participated or engaged forest utilization and conservation initiatives are recommended to promote sustainable forest management.

Keywords: SRTM; random forest; stochastic gradient boosting; random forest Kriging; Wavelet analysis

Remote sensing နည်းပညာအသုံးပြု၍ မြန်မာနိုင်ငံရှိ သစ်တောကြိုးဝိုင်း(၂)ခု၏ မြေပေါ်ဇီဝ ဒြပ်ထုမြေပုံ(Aboveground Biomass Map)များ ရေးဆွဲခြင်း

ဖြိုးဝေ၊ ဦးစီးအရာရှိ

သစ်တောဦးစီးဌာန

စာတမ်းအကျဉ်း

သစ်တောများ၏ နေရာဒေသအလိုက်တည်ရှိနေသော မြေပေါ်ဇီဝဒြပ်ထု (Aboveground Biomass)တန်ဖိုးများ ခန့်မှန်းတွက်ချက်ခြင်းသည် ရေရှည်တည်တံ့သော သစ်တောစီမံအုပ်ချုပ်ခြင်းနှင့် ကာဗွန်သိုလျှောင်နိုင်မှု ပမာဏကို တွက်ချက်ခြင်းများအတွက် အရေးပါသော အခန်းကဏ္ဍမှပါဝင်လျက် ရှိပါသည်။ ယခုဆောင်ရွက်သော သုတေသနသည် သစ်တောကြိုးဝိုင်း (၂)ခုဖြစ်သော ပူတာအိုမြို့နယ်ရှိ နန်ထုံကြိုးဝိုင်း(အမြဲစိမ်းသစ်တော)နှင့် မိုးမိတ်မြို့နယ်ရှိ ယင်းမာကြိုးဝိုင်း(ရွက်ပြတ်ရောနှောတော)များတွင် နမူနာစာရင်း ကောက်ယူထားသော သစ်ပင်များ၏ ရင်စို့လုံးပတ်နှင့် အပင်အမြင့်များမှ တွက်ချက်ရရှိလာသော ဇီဝဒြပ်ထုတန်ဖိုးများနှင့် Sentinel-2 ဂြိုဟ်တုဓာတ်ပုံများမှ ထုတ်ယူရရှိသော variables များကို machine learning နည်းပညာများအသုံးပြု၍ ကြိုးဝိုင်း(၁)ခုချင်းစီ၏ မြေပေါ် ဇီဝဒြပ်ထုမြေပုံအား ရေးဆွဲခြင်းဖြစ်ပါသည်။ ထို့ပြင် machine learning model များ၏ ခန့်မှန်းတွက်ချက် ရရှိမှုအပေါ် Kriging ဟုခေါ်သော နေရာအလိုက် ဆက်နွယ်မှု (spatial auto-correlation)ကို အသုံးပြု၍ တိကျမှန်ကန်မှုကို ဆန်းစစ်ခဲ့ပါသည်။ သုတေသနတွေ့ရှိချက်များမှာ Sentinel-2 ဂြိုဟ်တုဓာတ်ပုံမှ ထုတ်ယူရရှိသော vegetation indices များနှင့် textures အချက်အလက်များသည် သစ်တောကာဗွန်ခန့်မှန်း တွက်ချက်ရာတွင် အလွန်အသုံးဝင်ပြီး machine learning model များသည် ယခင်အသုံးပြုခဲ့ ကြသော တွက်ချက်ခန့်မှန်း နည်းပညာများ(conventional models)ထက် တိကျမှန်ကန်မှုသာ လွန်ကောင်းမွန်ကြောင်း တွေ့ရှိရပါသည်။ ခန့်မှန်းတွက်ချက်မှုအရ နန်ထုံကြိုးဝိုင်း (အမြဲစိမ်းတော)၏ ဇီဝဒြပ်ထု တန်ဖိုးမှာ တစ်ဟက်တာလျှင် ၉၄.၃ တန်မှ ၁၃၉.၈၃ တန်ထိရှိပြီး ယင်းမာကြိုးဝိုင်း (ရွက်ပြတ် ရောနှောတော)၏ ဇီဝဒြပ်ထုတန်ဖိုးမှာ တစ်ဟက်တာလျှင် ၃၂.၈၈ တန်မှ ၁၈၅.၆၅တန်ထိ ခန့်မှန်းတွက်ချက်ရရှိခဲ့သဖြင့် ရွက်ပြတ်ရောနှောတောသည် ကာဗွန်သိုလျှောင်

နိုင်မှုပမာဏ ပိုမိုမြင့်မားကြောင်း တွေ့ရှိရပါသည်။ ယခုသုတေသနတွင် အသုံးပြုခဲ့သော machine learning model (random forest) နှင့် kriging နည်းပညာအား တွဲဖက်အသုံးပြုခြင်း သည် အလားအလာ ကောင်းသောနည်းပညာ တစ်ခုဖြစ်ပြီး REDD+ စီမံကိန်း၏ အဓိက စံချိန် စံညွှန်းဖြစ်သော သစ်တောကာဗွန်သိုလျှောင် နိုင်မှုကိုခန့်မှန်း တွက်ချက်မှုများတွင် အသုံးပြု နိုင်မည်ဖြစ်ပါသဖြင့် ရေရှည်တည်တံ့သော သစ်တောစီမံအုပ်ချုပ်မှုမှတစ်ဆင့် စဉ်ဆက်မပြတ် ဖွံ့ဖြိုးတိုးတက်မှု ရည်မှန်းချက် (sustainable development goal-15) ကို ဖြည့်ဆည်းနိုင်မည် ဖြစ်ပါသည်။

Estimating Aboveground Biomass of Two Different Forest Types in Myanmar from Sentinel-2 Data with Machine Learning and Geostatistical Algorithms

1. Introduction

Global climate change, which has become a major environmental problem for all nations, mainly results from anthropogenic fossil fuel combustion and land-use changes (Bonan 2008). As the biggest carbon pool of terrestrial ecosystems, forests play a major role in the global carbon sequestration process, which contributes to climate change mitigation (Wolosin and Harris 2018). Overall, forests store approximately 45% of terrestrial carbon globally (Bonan 2008), most of which is stored in trees in the major form of aboveground biomass (AGB) that accounts for 44% of the total biomass through the process of photosynthesis (Rodríguez-Veiga et al. 2017). Understanding the spatial and temporal dynamics of AGB is the most critical step in quantifying carbon stocks and fluxes from forests (Gibbs et al. 2007). Hence, it is necessary to develop robust methods to estimate AGB for calculating carbon stocks, which is an essential indicator of the ‘reducing emission from deforestation and forest degradation’ (REDD)+ initiative as well as sustainable forest management (Basuki et al. 2009) (Dang et al. 2019).

AGB estimation with direct field surveys or that uses of species-specific allometric equations based on destructive sampling is accurate and widely used. However, these methods are difficult, expensive, time-consuming and not viable for wide regions due to limited samples and incomplete spatial coverage (Lu et al. 2016). Remote sensing images have become efficient data sources for AGB predictions in diverse-scale landscapes by providing different spatial, spectral and temporal resolutions of images. Recent studies have used various remote sensing data types to map the biomasses of different forests (Y. Li et al. 2020) (Shen et al. 2018) (Mon and Myint 2015). Currently, the prevailing high-resolution light detection and ranging (LiDAR) systems can capture the three-dimensional information of vertical forest structures and are well suited for forest biomass estimation since they reduce the spectral saturation problem (Xue 2015) (Pham et al. 2019) against those optical remote sensing images. However, LiDAR data have an operational limitation for large area estimation due to their increased imaging cost, data processing cost and spatial limitation (Vaglio Laurin et al. 2014). Additionally, LiDAR systems do not provide infrared signals, which undermine their capability to analyze the vegetation status. On the other hand, low-resolution satellite images (e.g., AVHRR, MODIS and SPOT Vegetation) have an advantage for AGB prediction in large areas since only one scene covers a wide area of interest. However, their accuracy is the lowest compared to moderate or high spatial data due to their plot-pixel matching differences (Yuan et al. 2016). Therefore, the medium resolution satellite images (e.g., Sentinel and Landsat) were increasingly applied for forest AGB estimation at different spatial scales for their free accessibility and high suitability to landscape scale analysis (H. Su et al. 2020) (López-Serrano et al. 2019).

The European Space Agency (ESA) launched the Sentinel-2A (S-2) multispectral satellite in 2015 following the SPOT and Landsat missions to monitor terrestrial surface (Addabbo et al. 2016). The S-2 has a wide swath at 290 km with 13 multispectral bands including four bands at 10 m, six bands at 20 m, and three bands at 60 m, respectively. Therefore, it provides data on land surface reflectance for many different wavelengths, such as Landsat 8. Some bands of S-2 have better resolution over Landsat 8 (10 m overrides 30 m of Landsat 8). Additionally, the level-1 TOA radiance or reflectance product from the S-2 satellite (S-2 L1C) can be improved to a level-2 product as the BOA reflectance (S-2 L2A) using free atmospheric correction tools in the Sentinel Application Platform (SNAP) software (Gascon, Ramoino, and Deanos 2017).

Particularly, for the presence of longer wavelength red-edge bands in S-2 data, it is extremely useful in vegetation monitoring. Although many studies have explored the application of S-2 optical satellite imagery in biomass mapping, there is still a present saturation problem since it lacks the capability to vertically penetrate dense forests (H. Li et al. 2022). For example, an optimal combination of reflectance, VIs and textural variables of S-2 has been employed in existing AGB mapping in an attempt to reduce this saturation effect (Safari and Sohrabi 2020).

S-2 derived VIs can minimize atmospheric saturation to some extent and better distinguish vegetation characteristics (e.g., moisture content) and forest AGB, especially in the relatively simply structured forests (Adamu et al. 2021). However, the sensitivity of VIs to AGB varies among forest types and structures (Taddese et al. 2020). In a canopy complex forest where spectral features cannot identify the vegetation structure (e.g., canopy depth), texture features may improve the prediction accuracy of AGB thanks to their sensitivity to the horizontal arrangement patterns of canopies and their shadows. For example, Li et al. explored the performances of S-2 textures in AGB estimation for the mature broadleaved forest with complex canopy layers (L. Li et al. 2020). Pandit et al. found that if proper processing techniques were used, texture features could mitigate the saturation problem of S-2 spectral data to some extent in the mature forest AGB prediction (Pandit, Tsuyuki, and Dube 2019). Different texture processing techniques including the principal component analysis (PCA), gray level co-occurrence matrix (GLCM) and wavelet analysis (WA) were applied in existing scientific works. PCA has been proven to be a method of reducing the dimensionality and increasing the interpretability of satellite datasets and can hold over 80% of all representative original information (Jolliffe, Cadima, and Cadima 2016); thus, these generated principal components may have a stronger relationship with biomass than individual original spectral bands, which are somehow independent of different biophysical conditions. Recent studies have claimed that textural measures derived from the GLCM and WA of satellite images could improve forest cover differentiation and forest AGB estimation (Ghosh and Behera 2018) (Ouma, Tetuko, and Tateishi 2008). Additionally, topographic features such as elevation, slope, and aspect are significantly related with the growth and distribution pattern of forests, and thus are of great significance in AGB estimations. Recent research findings have proven that variables derived from SRTM data at 30 m resolution are helpful for estimating forest biomass and have great influence over the spatial distribution of AGB (Chen et al. 2019). Elevation is especially well correlated with forest AGB since it provides information about the forest distribution and site attributes (Y. Wang, Zhang, and Guo 2021). As proven by numerous studies for different forest types with varying terrains, the topography may depend on the reflectivity of the specific forest site and affect the AGB estimation (Simard et al. 2006).

Empirical regression techniques were widely used in the early studies of remote sensing-based AGB estimation, considering the normality of the modeling datasets (Ahmad, Gilani, and Ahmad 2021). Therefore, simple linear or multiple linear regression models for practical AGB estimation are limited because of the complex non-linear relationships between forest AGB and remote sensing variables. Non-parametric models, also called machine-learning algorithms, do not require a strictly linear assumption between the response and covariates due to the independence of their data distribution. Machine learning models such as artificial neural network (ANN), support vector machine (SVM), random forest (RF), and stochastic gradient boosting (SGB) are popular non-parametric methods for identifying complex relationships between the predictors and forest AGB (Ye et al. 2021). Among these, RF and SGB are efficient machine learning methods proposed by Breiman (Breiman 2001) and Friedman (Friedman 2002).

and have been successfully used in forest AGB estimations. Although these models estimating AGB well, the major drawback of these models lies in their ignorance of the spatial autocorrelation of sample plots. Kriging interpolation provides the linear unbiased prediction of variables based on the variogram model and is best applied to minimize the spatial variation error between samples in AGB estimation.

Although these existing studies have had varying degrees of success in estimating forest AGB in different forest ecosystems with varying structural complexities, they regard AGB as an independent spatial biophysical variable when creating AGB prediction models, whereas as one of the more important variables in biogeochemical cycles, forest AGB not only has its own randomness in distribution, but also has structurized characteristics in space. Thus, the modeling techniques of the classic statistics applied in the vast majority of existing works that do not consider the auto-correlation information of forest AGB cannot adequately capture spatial variations in AGB, which necessitates an improvement upon existing modeling means by introducing geostatistical analytical methods, such as Kriging interpolation.

In recent decades, radical demographic, economic and social changes in Myanmar have placed considerable pressure on its forest resources (C. Wang and Myint 2016). According to the Food and Agricultural Organization (FAO), in 2015, the forest area in Myanmar was 42.92% of the total country, which had decreased from 45.04% in 2010 (Conservation, Ministry of Natural Resources and Environmental Conservation, and Forest Department 2020). To cover this loss, the forestry sector in Myanmar has been implementing the Myanmar reforestation and rehabilitation plan (MRRP) under the REDD+ scheme starting from 2017 to 2026 in the areas of forest degradation. As a REDD+ scheme, Myanmar predicts its forest reference emission level (FREL) based on preliminary information from the reference year 2005 to 2017, which needs to be periodically updated by integrating carbon improvement from reforestation programs based on new knowledge, methods and trends in the future (Forest Department and UNREDD Program 2018). Hence, the reliable method and data sources for mapping AGB by forest types are essential for Myanmar's future FREL calculation of REDD+ since local AGB maps are also the basis for the extension of estimates to larger areas using remote sensing approaches. To date, however, no systematic research has been conducted to predict the spatial distribution of AGB by forest types, especially in inaccessible areas of northern and eastern Myanmar because of operational limitations and the consequent lack of data, technology and appropriately qualified individuals. In this context, the optimal integration of remote sensing data and modeling algorithms may fill this gap.

2. Objectives

The overall goal of this study was to evaluate the performance improvement of overlaying geostatistical interpolation onto machine learning modeling based on S-2 and SRTM in mapping the AGB of two forest reserves in Myanmar. Additionally, the robust AGB maps generated from this work were also expected to support the strategic development of carbon sequestration-aimed forestry management efforts in Myanmar.

3. Materials and Methods

3.1. Study Area

Two forest reserves in Myanmar, namely Namhton (NH) and Yinmar (YM), were selected as case studies (Figure 1). They are located in the northern and central-eastern parts of

the country and have been formally protected by the Forest Department and by the 1992 Forest Law since 1995 and 2003, respectively.

The NH forest reserve area is approximately 19,418 hectares and the dense evergreen forest type dominates this region, geographically spanning from 97°13'00"E, 27°23'30"N to 97°24'30"E, 27°13'00"N. It is situated in the Putao Township, Myitkyina District, of northern Kachin State. The terrain in the region is mountainous, with an altitude ranging from 416 m to 1951 m. The annual average temperature of the area is 22.13 °C, with its average annual precipitation ranging from 1218 mm to 2800 mm; snowfall used to be heavy in the northern part, at higher altitudes of the mountainous regions. Evergreen tree species such as *Quercus glauca*, *Macaranga denticulata*, *Michelia champaca*, *Shorea assamica*, and *Ficus cuspidata*, are typical tree species of NH. A village called “Namhton Ku” is located in the center of this reserved forest as an encroachment, where forest-related field data collection was not performed.

The YM is a deciduous forest reserve-dominated region with an area of 20,163 hectares which is located between latitude 23°7'30"N–23°17'00"N and longitude 96°18'30"E–96°33'30"E, in Moemaik Township, Kyaukme District of Shan State. The mean annual temperature of this region is approximately 27 °C with an average annual rainfall ranging from 1000 mm to 1500 mm. The terrain is relatively flat, ranging from 128 m to 261 m according to the analysis of the SRTM DEM data. The vegetation in the YM region is predominantly Dipterocarpaceae species such as *Dipterocarpus alatus*, *Hopea odorata*, *Dipterocarpus tuberculatus*, and *Shorea obtuse*. The vegetation in this area is mainly deciduous, losing its leaves during the dry season. Additionally, bush fires, which frequently occur during the dry season, further reduce the available foliage. Biomass or carbon stocks during the dry season months can therefore dramatically differ from those during the rainy season for the same area.

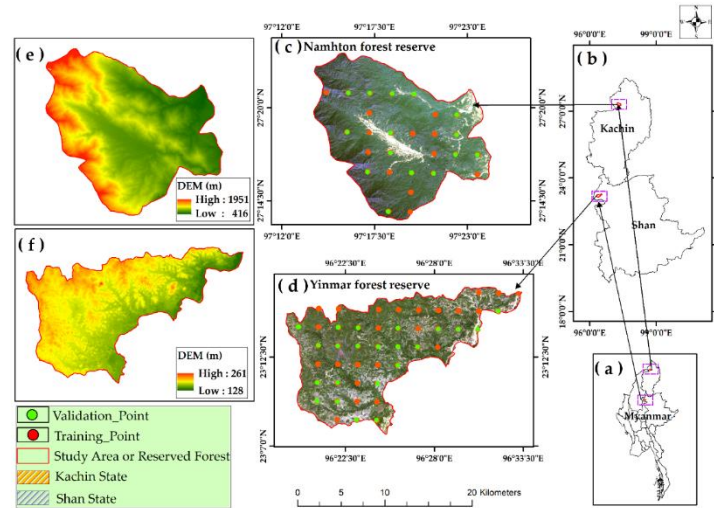


Figure 1. Location of the study area: (a) country boundary; (b) location of the two forest reserves; (c) the Sentinel-2 true color image (collected in 26 January 2017) attached with field sample plots in NH; (d) the Sentinel-2 true color image (collected in 5 February 2017) attached with field sample plots in YM; (e) the DEM of NH; and (f) the DEM of YM.

3.2. Data Collection and Processing Methods

3.2.1. Sentinel-2 Images Pre-Processing and Indices Extraction

The study site is located in two eco-regions (northern and central eastern) of Myanmar. Frequent rains and cloud contamination exist in the study sites, which highly restrict the availability of images collected in the peak season of vegetation growth (i.e., June–September). Therefore, two S-2 L1C MSI satellite images, with tile numbers of T47RLL and T46QHL acquired on 26 January 2017 and 5 February 2017, respectively, were downloaded from the European Space Agency. Available online: <https://www.scihub.copernicus.eu> (accessed Mar 21, 2022). These images are composed of 100 km² tiles with UTM/WGS84 projection. The descriptive information of the images is summarized in Table 1. The atmospheric correction of the two S-2 L1C scenes was performed with the Sen2Cor plugin in SNAP software to reduce the atmospheric, adjacency, and slope effects. In the process, TOA reflectance images were converted into surface reflectance images with aerosol-free and noise reduction. Then, all 20 m spectral bands were resampled to 10 m using the nearest neighbor strategy. Bands 1, 9, and 10 were not suitable for AGB estimation and excluded from the analysis (Maung and Sasaki 2020).

Table 1. Descriptive information of the images used in the analysis

Image/ Product	Tile Number and Acquisition Date	Cloud %	Bands Used for Modeling	Spatial Resolution (m)	Central Wavelength (nm)
S-2 L1C Product	T47RLL on 26 January 2017	3.63	B2 (blue)	10	490
			B3 (green)	10	560
			B4 (red)	10	665
	T46QHL on 5 February 2017	0.18	B5 (red edge)	20	705
			B6 (red edge)	20	740
			B7 (red edge)	20	783
			B8 (NIR)	10	842
			B8A (red edge)	20	865
			B11 (SWIR1)	20	1610
			B12 (SWIR2)	20	2190

In addition to the spectral bands, VIs were calculated based on the original reflectance bands in the raster calculator tool. The plot-level vegetation index mean values were extracted using the zonal statistics tool of ArcGIS, and using the plot size (0.08 ha) to match the AGB calculations. The detailed formulas for VIs calculation are described in Table 2.

3.2.2. SRTM Data Pre-Processing and Variables Extraction

SRTM topographic data were downloaded from the USGS EROS Data Center. Available online: <https://www.earthexplorer.usgs.gov/> (accessed Mar 23, 2022). These elevation data offer worldwide coverage of void-filled data at a resolution of one arc-second (approximately 30 m) and a high-resolution global dataset. These topographic data were first reprojected into UTM/WGS84 since the projection systems of these data are GCS/WGS84. To match with S-2 spectral bands, they were also resampled to 10 m spatial resolution using the nearest neighbor method in the ArcGIS package. Then, from this resampled dataset, two forest reserves boundaries were clipped and the elevation, slope and aspect were similarly extracted using the zonal statistics tool in ArcGIS (Table 2).

Table 2. List of Sentinel-2-derived variables and topographic factors used in AGB modeling

Satellite Data		Bands and Indices	Formula
Sentinel 1-2 Level-2A 10 m- resolution	Multispectral bands	Band 2	BLUE
		Band 3	GREEN
		Band 4	RED
		Band 5	RE1
		Band 6	RE2
		Band 7	RE3
		Band 8	NIR
		Band 8A	RE4
		Band 11	SWIR1
		Band 12	SWIR2
	Vegetation indices (Broad bands)	NDVI	$\text{NIR} - \text{RED} / \text{NIR} + \text{RED}$
		SAVI	$1.5 \times (\text{NIR} - \text{RED}) / (\text{NIR} + \text{RED} + \text{L})$
		EVI	$2.5 \times (\text{NIR} - \text{RED} / \text{NIR} + 2.4\text{RED} + 1)$
		GNDVI	$(\text{NIR} - \text{GREEN}) / (\text{NIR} + \text{GREEN})$
		WDVI	$(\text{NIR} - 0.5 \times \text{RED})$
		SR	$(\text{NIR} / \text{RED})$
		NDWI	$\text{NIR} - \text{SWIR2} / \text{NIR} + \text{SWIR2}$
		NDI45	$(\text{RE1} - \text{RED}) / (\text{RE1} + \text{RED})$
		MTCI	$(\text{RE2} - \text{RE1}) / (\text{RE1} - \text{RED})$
	Vegetation indices (Narrow red- edge bands)	RENDVI	$\text{NIR} - \text{RE1} / \text{NIR} + \text{RE1}$
		REEVI	$2.5 \times (\text{NIR} - \text{RE1} / \text{NIR} + 2.4\text{RE1} + 1)$
Resampled SRTM DEM(10 m)	Elevation	Ele	-
	Slope	Slope	-
	Aspect	Asp	-

3.2.3. Texture Features Extraction

Principal component analysis (PCA) can be used to remove correlated or redundant information in the satellite images and simultaneously reduce their dimensionality [34]. The first three principal components were produced as the potential image variable for modeling AGB. The first principal component (PC1) was used for texture extraction as it contained over 80% of the original spectral information. When extracting textural features, the gray level co-occurrence matrix (GLCM) method and wavelet decomposition method were applied. Among these, the GLCM textures including the mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, and correlation were extracted with different window sizes (3×3 , 5×5 , 7×7) from the PC1 image in the ENVI5.3 package (Table 3). In this study, the GLCM-based textures derived from a 7×7 window size were selected as predictive variables for AGB estimation after the correlations of different window sizes' textures with the measured AGB were tested.

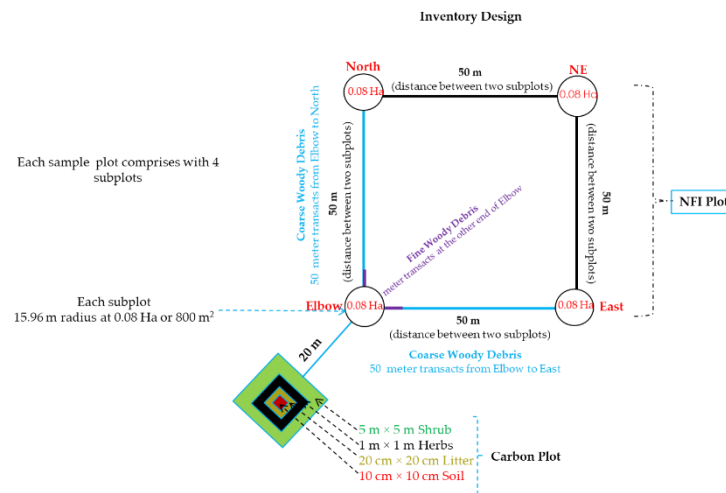
Table 3. Gray level co-occurrence matrix-based textural measures extracted in the current work

Data	GLCM Texture	Formula	Reference
PC1 image from 10 m resolution bands of S-2 L2A	mean	$\sum_{i,j=0}^{N-1} iP_{i,j}$	(Haralick, Shanmugam, and Dinstein 1973)
	variance	$\sum_{i,j=0}^{N-1} iP_{i,j} (1 - \mu_i)$	
	homogeneity	$\sum_{i,j=0}^{N-1} iP_{i,j} / (1 + (i - j)^2)$	
	contrast	$\sum_{i,j=0}^{N-1} iP_{i,j} (i - j)$	
	dissimilarity	$\sum_{i,j=0}^{N-1} iP_{i,j} i - j $	
	entropy	$\sum_{i,j=0}^{N-1} iP_{i,j} \ln P_{i,j}$	
	second moment	$\sum_{i,j=0}^{N-1} iP_{i,j}^2$	
	correlation	$\sum_{i,j=0}^{N-1} (i(\sum_{i,j=0}^{N-1} ijP_{i,j}^2 - \mu_i \mu_i) / \sigma_i^2 \sigma_i^2)$	

Additionally, the wavelet analysis was also considered to be an effective means of extracting textures. Wavelet transformation is a multi-resolution analysis tool for image signal processing, which has two distinct abilities: subtle variation in spectral features in the original data can be detected at different scales and the useful information can be represented by fewer wavelet features by compressing data. The wavelet analysis produces four basic components including the approximation image, horizontal detail, vertical detail and diagonal detail images of which the latter three are usually regarded as helpful textural measures. In this study, the Coiflet discrete wavelet function was chosen after repeated tests with different mother wavelets (the Haar wavelet, Daubechies (dbN) and Symlets (symN)) in the Matlab package because it had the highest correlation with AGB. Thus, based on the first principal component (PC1), a three-level decomposition strategy was implemented through programming in the Matlab environment to generate 9 detailed images as independent textural variables for AGB modeling. Finally, two types of textures derived from GLCM-based and wavelet analysis were included in the AGB modeling in this study.

3.2.4 Field Forest Inventory Data

Myanmar's Forest Department conducted the national forest inventory (NFI) data collection in February and March of 2017 with the financial support of the Finnish government and according to the guidelines of the FAO technical team. These secondary forest inventoried



Source: Planning and Statistics Division, Forest Department of Myanmar.

Since the species-specific allometric equations are not available for the study region, we had to use the unpublished national-level coarse allometric equations for evergreen forest and broad-leaved forest to calculate the plot-level AGB. The AGB formulas were as follows:

For YM Deciduous, $AGB = \rho_2 \times \exp(-0.667 + 1.78 \ln(DBH) + 0.207(\ln(DBH))^2 - 0.0281(\ln(DBH))^3)$

Using the national level biomass expansion factors Equations (1) and (2), AGB was computed for each tree and summed up by each plot to obtain plot-level AGB. Finally, the plot-level field AGB values were converted from kg/plot into ton/ha. Therefore, the AGB unit in this study was t/ha. Table 4 shows the descriptive statistics of the field-observed AGB.

Table 4. Descriptive statistics of the field-measured AGB (t/ha) for the two forest reserves

Forest Type	Number of Sample Plots	AGB (t/ha) Value Range	Median	Mean	Std. Deviation	Number of Sample Plot Used in Modeling	
						Training	Validation
Evergreen	88	0.57–151.64	38.40	49.00	39.206	71	17
Deciduous	170	2.74–215.24	100.02	98.00	51.73	140	30

3.3. Aboveground Biomass Detection Methods

3.3.1. Prediction Model Establishment

In this study, Random Forest (RF) and stochastic gradient boosting (SGB) models were first performed for AGB prediction. Then, based on the better one (RF or SGB), the predicted residuals (the difference between the observed AGB and the model-predicted AGB) were further analyzed and compared using ordinary Kriging (OK) and the co-Kriging (CK) to separate the structured components hidden in the residuals, followed by adding the better structured components onto the better model predictions to obtain the final AGB predictions. The detailed modeling approaches are summarized as follows:

3.3.2. Random Forest and Stochastic Gradient Boosting Models

The RF model is a bagging algorithm which enhances accuracy and reduces overfitting and bias (Liu et al. 2018). SGB is a boosting ensemble method with low sensitivity to outliers, with the ability to deal with unbalanced training datasets. Both models are non-parametric modeling approaches which have a performance superior to those of other machine learning techniques such as the K-nearest neighbor (KNN), support vector machine (SVM), and the multivariate adaptive regression splines (MARS) (Wu et al. 2016) (Zhang et al. 2020), which is increasingly being applied to satellite-based biomass mapping.

In view of these advantages, in the current study, the RF and SGB models were performed as the first attempt to predict the AGB of the two forest reserves (NH and YM). The RF model was implemented in the “randomForest” package within R Studio. This package supports the chart that illustrates the GI-index and OOB error rate to determine the most important modeling variables. From this comprehensive chart, preference variables can be selected for a prediction model to reduce the complexity and load of computation. In the RF regression analysis, the variables’ importance ranking was determined by out-of-bag (OOB) error and node-purity percent (IncNodePurity). The first variable importance analysis was calculated by randomly permuting each predictor variable and computing the associated reduction in predictive performance using the out-of-bag (OOB) error. The second most important variable was estimated by determining the decrease in node impurities attributable to each predictor variable. Larger IncNodePurity and %IncMSE indicate higher model accuracy in terms of ranking variable importance. Parameters such as the number of trees (ntree), the number of variables used to split the tree at each node (mtry), and node size are adjustable for the RF model. For RF-prediction models in this study, after multiple tests, the ntree, mtry and nodesize were specified to be 600, 4, and 5 for the evergreen forest (NH) and 1400, 2, and 5 for the deciduous forest (YM), respectively.

The SGB algorithm is based on the combination of the regression tree and boosted algorithms to predict the response variable. This algorithm also reduces the chance of overfitting

by introducing an element of stochasticity due to its flexibility and high predictive performance. Unlike RF, a tree is constructed from a different random sub-sample of the dataset in the SGB model, producing an incremental improvement in the model. In this study, all steps of the SGB analysis were implemented in the “gbm” package in the R-studio. The “vip” function in the “gbm” package was also used for the selection of important variables for the SGB model. The SGB model adjustment includes the distribution, interaction depth, bagging fraction, shrinkage rate, and training fraction. The out-of-bag method was used for determining the optimal number of boosting iterations. The maximum iteration tree was stopped at 420 in the evergreen forest AGB (NH) SGB-prediction model and 600 at the deciduous forest AGB (YM) SGB-prediction model, as more iteration tree numbers could no longer improve models' accuracy. The interaction depth, also known as the maximum number of possible interactions, was set to 3 and 4 nodes for the NH and YM SGB-models, respectively. The bagging fraction controlling the fraction of the training data, the shrinkage rate controlling the learning speed of the algorithm, and the training fraction randomly selected for calculating each tree were set to 10, 0.03, and 10 for both SGB models (NH and YM).

Once the RF and SGB models were created for the two forest reserves, their predictive performances in AGB estimation were compared to determine the better model; then, based on this model, all the residuals resulting from the model were further analyzed by implementing OK and CK autocorrelation algorithms to separate the structurized component or trend item hidden in the residuals.

3.3.3. Random Forest-Based Kriging Model

In this study, RF performed better than SGB in both the NH and YM AGB prediction models in terms of accuracy evaluations; thus, to improve the accuracy of RF models or finding the spatial correlation of AGB samples, RF-based OK and CK analyses (RFOK and RFCK) were also performed as subsequent steps. Since the RF model does not consider spatial autocorrelation among the AGB sample plots, AGB is actually a typical item with relatively high spatial autocorrelation; thus, a combination of RF and Kriging (RFK) was potentially an effective and more reliable means of determining the spatial distribution of AGB in this study. Specifically, a regression-Kriging technique was used to extract the structured components of the residuals obtained from the RF regression. As the procedure of Kriging interpolation, the modeling semivariogram is important to determine the accuracy and reliability of the estimates. Kriging includes ordinary Kriging (OK) and co-Kriging (CK) in which OK is a suitable interpolation method for the uneven distribution of terrain and climatic variation events, while CK is the best method for improving the accuracy of target prediction [22]. OK is a linear estimation method suitable for inherently stationary random fields which satisfies the isotropic hypothesis and fully considers spatial parametric non-stationarity as well as the effects of environmental variables derived from the benefits of RF. It is a widely used geostatistical technique that generates an optimal unbiased estimated surface employing a semivariogram based on regionalized variables. The interpolation formula of OK is as follows:

$$Z_{OK}^*(x_0) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (3)$$

where $Z_{OK}^*(x_0)$ is the residual value of the AGB to be estimated at location x_0 , n is the number of sample points used for interpolation, $Z(x_i)$ is the AGB residual of site i , and λ_i is the weighting coefficient at point i .

CK is an improvement over the OK method and deals with multivariate problems [22]. Since the study areas include two reserved forests with different terrains, AGB is definitely affected by elevation. Thus, the elevation factor was used as a co-variable for interpolation. This was expressed as follows:

$$Z_{2,CK}^*(x_0) = \sum_{i=1}^{N_1} \lambda_{1i} Z_1(x_{1i}) + \sum_{j=1}^{N_2} \lambda_{2j} Z_2(x_{2j}) \quad (4)$$

where $[Z_{2,CK}]^*(x_0)$ is the residual values of AGB to be estimated, $Z_1(x_{1i})$ is the AGB residual of the site i , λ_{1i} is the weighting coefficient of site i , $Z_2(x_{2j})$ is the elevation of site j , and λ_{2j} is the weighting coefficient of site j .

Variograms are an effective tool for analyzing the spatial variation and structure of target predictors for the reliable estimation of AGB. The performance of the semivariogram was assessed by the coefficient of determination (R^2) and the root mean square error (RMSE). The larger the R^2 , the smaller the RMSE and nugget effect, and the better the fitting performance was.

In RF-Kriging modeling, the estimated residual value of each sample point was calculated by subtracting the RF-derived predicted AGB value from the field-observed AGB value. It can be calculated as follows:

$$Z(x_i) = C(x_i) - C_{RF}^{\wedge}(x_i) \quad (5)$$

Final AGB prediction by RFOK or RFCK method was acquired by the following equation:

$$C_{RFOK/RFCK}^{\wedge}(x_i) = C_{RF}^{\wedge}(x_i) + Z_K^{\wedge}(x_i) \quad (6)$$

where $[C^{\wedge}]_{RFOK/RFCK}(x_i)$ is the predicted AGB at site i using RFOK or RFCK, where $[Z^{\wedge}]_K(x_i)$ is the AGB residual value of site i , $C(x_i)$ is the observed AGB of site i , $[C^{\wedge}]_{RF}(x_i)$ is the RF-based predicted AGB at site i .

Finally, the maximum likelihood classifier in ENVI Classic 5.3 was applied to classify the forest and non-forest areas of the study area. The resulting classified forest area was used as a mask to obtain the forest AGB maps of the study area.

3.4. Accuracy Assessment

The training and validation sets were determined as 80 and 20% of the sampled data using stratified sampling for all statistical analyses. Model performances were assessed based on the coefficient of determination- R^2 , the mean absolute error (MAE), the root mean square error (RMSE), the RMSE%, bias, and bias%, based on the validation set (20% of samples). In addition, the relative improvement (RI) index for assessing RFOK and RFCK over RF models was also performed (7):

$$RI = \frac{RMSE_{RF} - RMSE_{RFOK/RFCK}}{RMSE_{RF}} \quad (7)$$

Where, $RMSE_{RF}$ is the root mean square error from RF predicted model, $(\sqrt{RMSE_{RF}^2 - (RMSE_{RFOK}^2 + RMSE_{RFCK}^2)})$ is the root mean square error from RFOK and RFCK models, respectively.

4. Results

4.1. Variable Importance and Selections

In the RF variable importance analysis, the %IncMSE by OOB error and %IncNodePurity by Gini index are indicators of the importance ranking. The topmost important variables picked up by the evergreen NH and deciduous YM RF models are shown in Figure 3a, 3b. For the generalizing the model and reducing the computation load, only the top 10 variables were selected as the final RF model inputs. S-2 derived the reflectance, Vis and textures from the wavelet decomposition, and the topographic variables were included in the list. In the RF model, topographic variables and textures from WA were major contributors to the evergreen NH forest AGB estimation, while VIs showed better sensitivity to deciduous AGB.

In the variables analysis for the SGB model, the lowest RMSE could be found for an interaction depth of 3, a shrinkage ratio of 0.03, and 420 iterations in the evergreen forest reserve. For the deciduous forest reserve, an interaction depth of 4, shrinkage ratio of 0.03, and 600 trees were the best parameters for model fitting. Figure 3c, 3d show the most important variables of SGB models for evergreen NH and deciduous YM AGB estimation. In SGB models, topographic variables and textures from WA were identified as major contributors to the NH evergreen AGB estimation, while VIs had better sensitivity for the deciduous AGB prediction. According to the variable importance analyses in the RF and SGB models, the variables did not differ a great deal in terms of their sensitivity to AGB in same forest types.

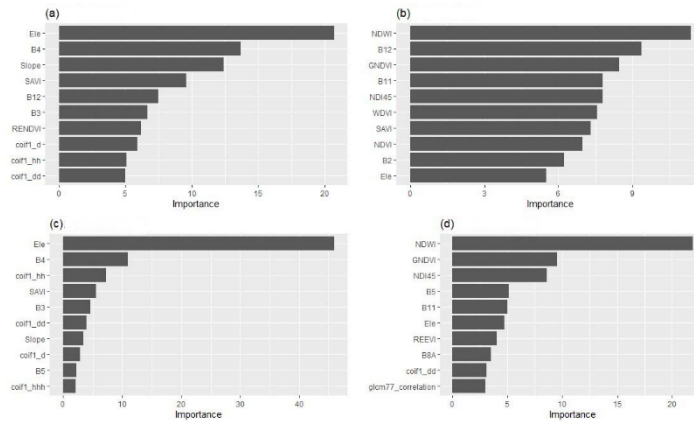


Figure 3. The importance ranking of the top 10 variables identified by (a) evergreen RF model; (b) deciduous RF model; (c) evergreen SGB model; and (d) deciduous SGB model

4.2. Validation Metrics for RF and SGB Models

Among the sample plots, 80% were used for training the models (RF and SGB). For the RF evergreen model training, the optimum model accuracy was obtained from the following parameters, $n_{tree} = 600$, $m_{try} = 4$ with $nodesize = 5$, considering the important predictors. The adjusted parameters values of $n_{tree} = 1400$, $m_{try} = 2$ and $nodesize = 5$ were selected for the RF deciduous model. The RMSEs values of 11.17, 17.36 t/ha and corresponding RMSE% 23.06 and 17.45 were obtained while the R^2 values were 0.95 and 0.97 for evergreen and deciduous training models. It was also observed that the bias and bias% of the evergreen training models

were -0.59 and -1.20% , while the deciduous model obtained -0.011 and -0.11% , respectively. The smallest RMSE value was observed in the evergreen forest since the AGB values of the training plots in this type were smaller than those of the deciduous forest (Table 4).

Depending on the shrinkage ratio, the interaction depth, and the number of regression trees, the two SGB model performances were evaluated. The SGB evergreen model resulted in an R^2 value of 0.98 and an RMSE of 4.62 t/ha with the corresponding RMSE%, bias and bias% of 9.45, -1.24 and -2.54 , respectively. In the trained model of SGB for the deciduous forest, the R^2 , RMSE, RMSE%, bias and bias% were 0.97, 8.83, 8.88, -3.02 and -3.04 .

In addition to model training, 20% of the sample plots were used for model validation. Table 5 show the validation metrics for RF and SGB models. The R^2 values of the NH evergreen RF model and NH SGB model were 0.47 and 0.35, respectively, while RMSEs values of those models were 25.45 and 32.02 t/ha, respectively. Meanwhile, the R^2 and RMSE values of the YM deciduous RF model and YM SGB model were estimated to be 0.38, 0.35 and 40.23, 41.85 t/ha. By considering the validation models' metrics, the RF models had higher R^2 , lower RMSE, and bias values than SGB models in the biomass estimation at each growth stage, indicating that the RF models can provide more accurate biomass estimations than SGB; thus, the residuals from the RF model were further analyzed to attempt to extract the structured components from the residuals to possibly improve the ultimate prediction accuracy of AGB.

Table 5. Validation metrics of RF, SGB, RFOK, and RFCK models based on 20% of the sampled data

Forest Type	Model	R^2	RMSE(t/ha)	RMSE%	MAE(t/ha)	Bias	Bias%	RI
NH Evergreen	RF	0.47	25.45	51.44	22.45	0.15	0.29	-
NH Evergreen	SGB	0.35	32.02	64.72	27.03	2.48	5.03	-
NH Evergreen	RFOK	0.47	24.91	50.34	22.19	3.35	6.76	0.021
NH Evergreen	RFCK	0.46	25.75	52.04	57.67	-49.27	-99.57	-0.011
YM Deciduous	RF	0.38	40.23	44.09	33.07	6.18	6.77	-
YM Deciduous	SGB	0.35	41.85	45.88	33.00	6.91	7.58	-
YM Deciduous	RFOK	0.52	34.84	38.19	27.51	0.30	0.33	0.134
YM Deciduous	RFCK	0.52	34.72	38.06	27.47	0.06	0.06	0.137

4.3. Semivariogram Analysis Results of RF-Derived Residuals

Residuals of the RF-predicted AGB were derived by subtracting the RF-predicted AGB from the field-measured AGB. Table 6 shows the statistics of the residuals. The mean residual values for the evergreen and deciduous AGB were -0.91 and -1.60 t/ha, respectively. In addition, the residual value range of the deciduous forest (82.43 t/ha) was much higher than that of the evergreen forest (34.52 t/ha). The standard deviation of the evergreen residuals (15.38 t/ha) was smaller than that of the deciduous (23.12 t/ha). The residual values were shown in different colors and sizes based on their distribution in Figure 4. As shown in the histogram distribution in Figure 4, the residual values of the evergreen forest were not close to normal distribution while deciduous residuals were approximately normally distributed. Nevertheless, the procedure of the semivariogram analyses for both forest reserves was performed for testing the accuracy of the models.

Table 6. Residual's statistics derived from the RF model prediction

Forest Type	Residual Mean (t/ha)	Std. Deviation (t/ha)	Value Range (t/ha)	Skewness	Kurtosis
NH Evergreen	-0.91	15.38	-45.90–34.52	-0.31	3.57
YM Deciduous	-1.60	23.12	-75.97–82.43	-0.03	4.50

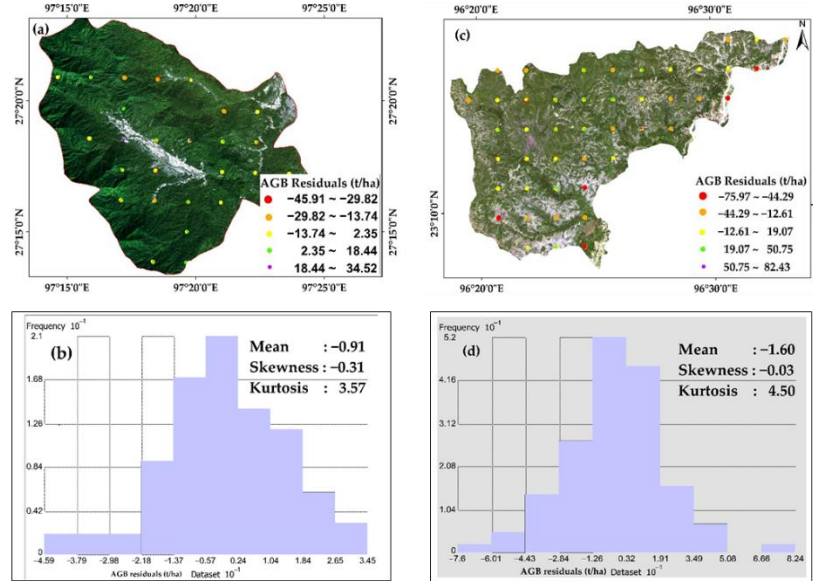


Figure 4. Spatial distributions of the residuals and corresponding frequency histograms; (a,c) are RF-based AGB-residuals for NH evergreen and YM deciduous, respectively; (b,d) are frequency histograms of AGB-residuals for NH evergreen and YM deciduous, respectively

After the normality of the residuals was verified in SPSS, these residual values could be used for semivariogram analysis based on the ordinary Kriging (OK) and the co-Kriging (CK) in the semivariance analysis. The model with the largest R^2 and smallest RMSE was determined as the optimal analytical function of the semivariogram. As a result, Gaussian function was picked up (Table 7). Furthermore, elevation was used as a co-variable in the CK model for the better estimation of AGB. Table 7 and Figure 5 show the modeled semivariogram and semivariance models using OK and CK analyses. Evergreen forest models were poorly fitted; the R^2 values were much lower than those of deciduous models (0.19-OK and 0.15-CK versus 0.58-OK and 0.62-CK). The nugget value in the OK model of the evergreen AGB residuals was slightly smaller than that of the corresponding CK model, indicating stronger spatial homogeneity. In the CK model of deciduous residuals, the nugget value was also smaller than that of the OK model. Moreover, the ratio of nugget and sill (nugget/sill) determined the variation of spatial autocorrelation between the AGB sample plots. From the OK models, the smaller nugget/sill ratio (0.95) exhibited stronger spatial homogeneity than the CK model with nugget/sill (0.99) for evergreen residuals. In the deciduous residuals, CK gave a smaller nugget/sill (0.74), indicating a stronger spatial correlation by considering elevation as a co-variable. This meant that deciduous forest AGB varied closely in space with the terrain variations. The improvement was not obvious in the evergreen OK and CK models compared to the evergreen RF model since the residuals

were not normally distributed because of the existence of a spatial gap between the evergreen sample plots that restrict the principle of spatial correlation (Figure 4a). Thus, the OK model could not serve to fit the spatial patterns of the residuals in the evergreen forest AGB, while the CK model performed better in terms of fitting the spatial patterns of the residuals in the deciduous forest AGB in the study area when considering their R^2 and RMSE values. Nevertheless, these residuals were used to interpolate their structured components in the evergreen and deciduous forests, respectively.

Table 7. Parameter estimations for the semivariogram analysis based on Gaussian function

Model Parameter	Theoretical Model	Nugget	Sill	Nugget/Sill	Range (m)	R^2	RMSE (t/ha)
NH Evergreen OK	Gaussian	138.53	145.09	0.95	99.57	0.19	12.42
NH Evergreen CK	Gaussian	249.63	249.75	0.99	9611	0.15	18.69
YM Deciduous OK	Gaussian	245.93	324.58	0.76	10123	0.58	24.38
YM Deciduous CK	Gaussian	239.97	326.26	0.74	9890	0.62	22.65

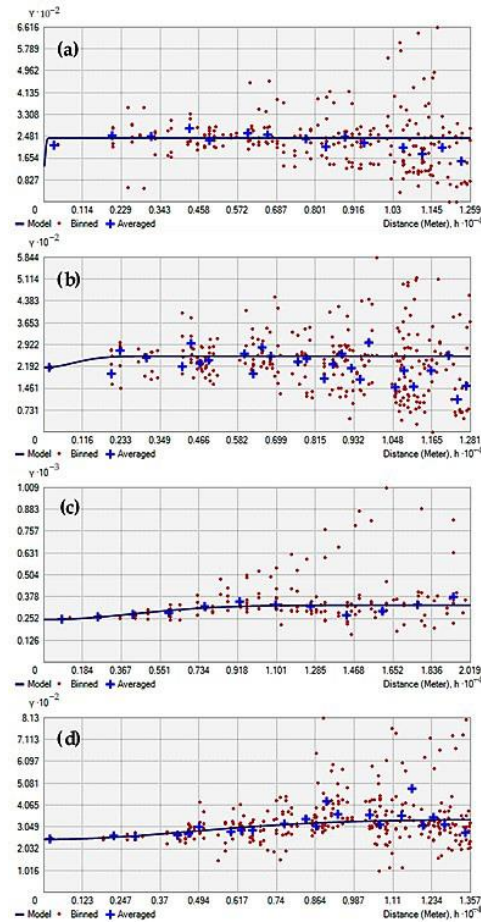


Figure 5. Empirical semivariograms and covariance models for the two forest reserves' RF-derived residuals; (a,c) are the semivariogram models for NH evergreen and YM deciduous using OK analysis; (b,d) are the semivariogram models for NH evergreen and YM deciduous using CK analysis with a covariable of elevation. The vertical axis is the $\frac{1}{2}$ variance (γ) and covariance (C) of the two positions as the distance increases

4.4. Forest AGB Mapping Results Based on RFOK and RFCK Models

The predicted AGB of the two forest reserves were obtained from the RFOK for the evergreen and RFCK for the deciduous forests, and the validation performances with 20% sample data based on equation (6) were correspondingly derived. Table 5 shows the validation accuracy improvements of RFOK and RFCK in relation to the initial RF model.

Although the R^2 value (0.47) of the RFOK model for evergreen did not increase compared to the original RF model, its RMSE, RMSE% and MAE all decreased with different magnitudes, and its RI value was 0.02, indicating a slight improvement in AGB prediction. However, the RFOK had the worst predictive performance in the evergreen forest. For the deciduous forest, the RFCK model outperformed other two RFOK and RF models, particularly in relation to the RF, as the RFCK's R square value increased from 0.38 to 0.52 and its RMSE decreased from 40.23 t/ha to 34.72 t/ha, with an RI value of 13.7%. However, compared to RFOK, RFCK only took on a very tiny improvement in the prediction accuracy.

In addition to the model evaluation, the generalization ability of the model was also considered. The AGB value range of the predicted map could reflect the model's robustness to some extent. The range of AGB values predicted for the evergreen using the RF model was 88.75–129 t/ha. The AGB prediction value range from RFOK was 94.3–139.83 t/ha and the RFCK had a value range of 94.06–139.62 t/ha, respectively. The AGB prediction value range in the deciduous was 40–176 t/ha, 30.41–187.84 t/ha, and 32.88–185.65 t/ha for RF, RFOK, and RFCK. These variations in the value range clearly indicated an improved generalization ability of RFOK and RFCK, with higher robustness. The largest evergreen AGB values were found in the northern and western boundaries of NH while the AGB in the central and eastern parts were sparsely distributed. In this area, the low AGB values were found close to a village and flat area. Deciduous AGB was covered with large values in some parts of northern YM, and they were evenly distributed in the reserved area, except in the southeastern boundary which is close to a village. The AGB maps derived from all performed models are shown in Figures 6 and 7.

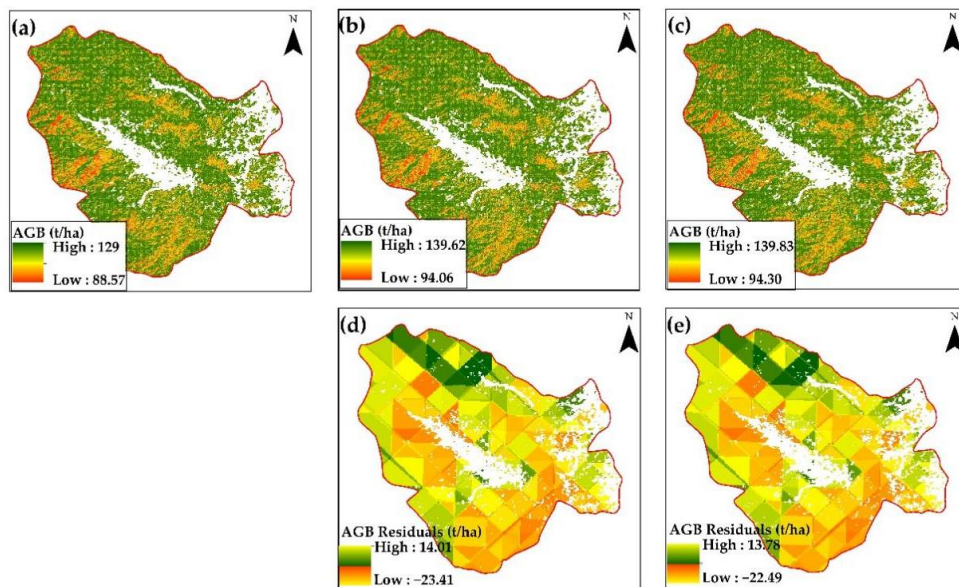


Figure 6. The estimated AGB of evergreen forests in the NM from (a) RF, (b) RFOK, and (c) RFCK models, and the calculated residuals for (d) RFOK and (e) RFCK models

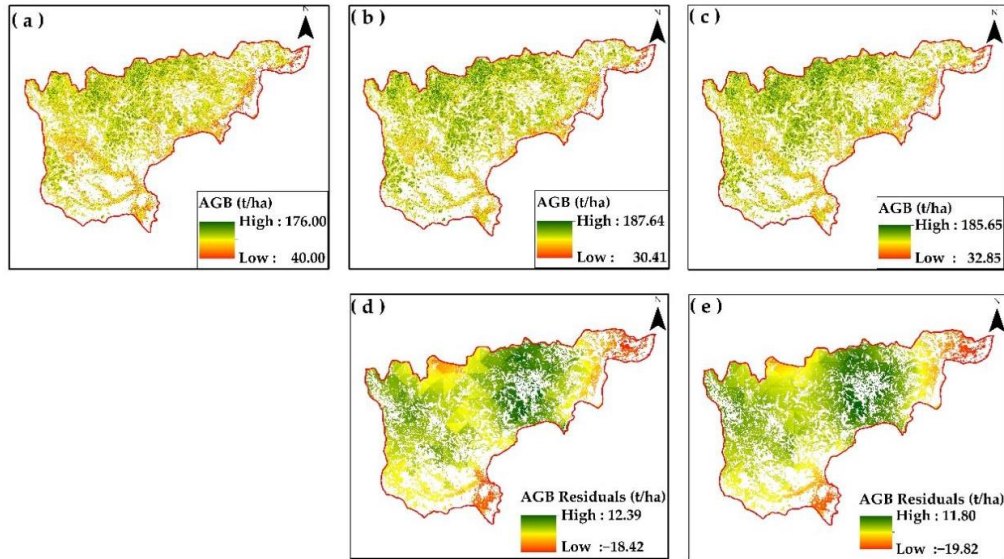


Figure 7. The estimated AGB of deciduous forests in the YM from (a) RF, (b) RFOK, and (c) RFCK models, and the calculated residuals for (d) RFOK and (e) RFCK models.

5. Discussions

To the best of our knowledge, this is the first study evaluating the capability of S-2 satellite data for AGB mapping in evergreen and deciduous forest reserves in Myanmar. AGB mapping in these forest reserves is very fundamental for carbon strategies, especially for Myanmar REDD+ mechanisms which need to report the carbon improvement of forests from reforestation programs for future FREL submission. Moreover, the prediction methods proposed in this study are noticeably relevant for future forest management practices in Myanmar, which previously exhibited a lack of robust AGB estimation methods for forest types. We proved that THE combined use of S-2, its derivatives, and topographic parameters, in tandem with proper modeling techniques, could improve AGB estimation.

5.1. Sensitivity of Sentinel-2 Derivatives to AGB

The correct selection of variables contributing to a model is critical for AGB estimation. In most AGB mapping studies with S-2 derivatives, four novel wavebands in the red-edge region and near-infrared region (NIR) (B5, B6, B7 and B8A) showed good performance (Pandit, Tsuyuki, and Dube 2018). These reflectance regions offer unprecedented spectral signatures which are highly sensitive to the biophysical and biochemical responses of vegetation that are critical for measuring vegetation characteristics such as biomass. However, in this study, the classical and short-wave infrared bands (B2, B3, B4, B11, and B12) outperformed these aforementioned spectral bands. The excellent performance of these bands is that carbon and nitrogen-containing metabolites reach their reflectance peak in wavebands between 440 nm and 570 nm due to the nature of the forests in the study area (B2 and B3) and concurred with the previous study (Ferwerda, Skidmore, and Mutanga 2005). Forest canopy in the evergreen forest can uptake maximum chlorophyll absorption due to non-deciduous phenomena. B3 and B4 in the S-2 have strong sensitivity to chlorophyll in evergreen vegetation, while B2 can effectively distinguish vegetation and soil background in the deciduous forest where the reflectivity of soil is

apparent because of the simple canopy structure. Even though vegetation can reflect the maximum energy at NIR despite the fact that it is unable to provide any information on the soil under the vegetation, SWIR bands in S-2 can distinguish the vegetation and soil to some extent. A recent finding by Chen et al. verified that SWIR spectral bands (B11) could efficiently detect the moisture content of vegetation (Chen et al. 2019). Moreover, Dang et al. proved that the broad bands (B11, B12) of S-2 were the best response variables of AGB prediction with an R^2 of 0.81 and an RMSE of 36.67 Mg t/ha (Dang et al. 2019). Hence, the high sensitivity of SWIR bands to biomass observed in this study seems plausible, which is in agreement with the existing studies.

The VIs used in this study (NDVI, GNDVI, SAVI, SR, RENDVI, NDI45, and NDWI) were primary contributors to the AGB estimation of forests since they have the ability to maximize the sensitivity of vegetation characteristics and minimize soil background reflectance and atmospheric effects. Balidoy et al. proved that the SR and NDVI of S-2 data were the most effective biomass predictors, providing the highest accuracy ($R^2 = 0.89$; RMSE = 5.69 Mg/ha)(Balidoy Baloloy et al. 2018). Ghosh et al. found that the effectiveness of NDVI, GNDVI, SAVI indices of S-2 for dense tropical AGB mapping had an R^2 value of 0.6 and an RMSE value of 79.45 t/ha for the teak forest(Ghosh and Behera 2018). The findings of Pandit et al. were entirely consistent with those of this study. The set of 24 variables including NDVI, GNDVI, RENDVI, SAVI, and SR produced overall plausible and strongly explained variable values, with $R^2 = 0.81$ and RMSE = 1.07 kg/m(Pandit, Tsuyuki, and Dube 2018). Although VIs produced from traditional broad bands can reduce the saturation problem in simple canopy forest, they are less sensitive to complex forest stands with high biomass values(Mngadi, Odindi, and Mutanga 2021). In this regard, the red-edge bands derived indices are highly sensitive to such kind of dense vegetation structures and relatively less prone to spectral saturation. For example, the standard NDVI from B4 and B8 is less effective than RENDVI from red-edge bands (B5, B8) in AGB estimation (Forkuor et al. 2020) and hence red-edge-derived indices can be effectively applied in dense vegetation cover (e.g., RENDVI is sensitive to the NH evergreen forest AGB in the current work). SWIR bands are related with nitrogen, lignin, and cellulose, capable of retrieving canopy structural attributes and biomass. Canopy water content index from SWIR bands (e.g., NDWI in this work) is highly sensitive to deciduous forest AGB but not to evergreen AGB because canopy structure in deciduous forest is relatively simple, which is in agreement with the previous findings of Ewald (Fassnacht et al. 2021). They pointed out that in the very dense canopy plantation, SWIR indices could not effectively improve AGB estimation compared to other indices. This study claims that red-edge indices are suitable for complex canopy AGB retrievals while SWIR indices are useful for simple canopy forest AGB estimation.

The textural variables are strongest candidates of evergreen forest AGB, especially the textures (coif1-d, coif1-dd, and coif1-hh) extracted from Coiflect wavelet analysis of the PC1 image. They could obviously improve the AGB estimation as the horizontal structures of the evergreen forest can be effectively characterized by them. Previous research had shown that texture measures have the potential to improve AGB estimation, especially for complex vegetation structures where canopies' reflectance values tend to be saturated but the horizontal structures represented by textural indices still have differences. If proper processing techniques are used, textural measures could improve the prediction accuracy of AGB models. According to Eckert et al., textures were much better to capture the various forest canopy structures of the forest strata than the spectral reflectance or band ratios, due to their sensitivity to the spatial aspects of the canopy shadow(Eckert 2012). Cutler et al. argued that the textures extracted from

GLCM method and the WA of satellite images yielded better results in the AGB estimation and forest type classification (Cutler et al. 2012). Our results concur with aforementioned studies. However, in this study, GLCM-based textures were not highly correlated with AGB. The reason might be that the window size (7×7) could not reduce the border effects of pixels to attain original spectral values and thus, texture window size determination should be considered in accordance with the types of satellite data in future studies. We conclude that the wavelet decomposition analysis of satellite images might improve evergreen forest AGB estimation because it produces more suitable textures to effectively depict evergreen forest horizontal structures to reduce the saturation problem of spectral signals.

5.2. Sensitivity of Topographic Variables to AGB

Topographic features (elevation and slope) were also strong predictors of AGB in this study. The vital role of topography is also an important factor influencing AGB because water and sunlight storage change in function of topography (Y. Su et al. 2016). Slope inclination was also a strong factor affecting AGB in the evergreen forest, since the terrain in this area is mountainous with different sunlight and water detention levels by its varying geomorphological characteristics while deciduous forest is not affected by slope by growing its trees over sandy soil flat surface. This finding is consistent with the proof of Hamere et al., which claimed that AGB carbon, BGB carbon, and the total carbon density trend showed a decrease as the slope increased due to the little vegetation cover in very steep slope areas (Yohannes and Soromessa 2015). In addition, in the accessible flat area of nearby settlements and stream banks, the anthropogenic effects might decrease AGB values. Therefore, topographic factors ultimately affect the AGB observed in this study.

5.3. Comparison between Models

This study evaluated the two modeling techniques (RF and SGB) for AGB mapping in two forest reserves and indicated a saturation problem of S-2 d satellite data, thus causing the presence of bias in the AGB prediction models. The NH evergreen RF model estimated the smaller AGB value of 129 t/ha than the observed AGB value of 151.64 t/ha. Similarly, the estimation of the YM deciduous forest AGB value (176 t/ha) was smaller than the field-observed AGB value (215.24 t/ha). The scatterplots in the validation metrics of this study indicated the limitation of the classical wavelength bands in the S-2 MSI sensor when dealing with saturation in high biomass stands. From the important variables ranking, the candidates of the classical wavelength region and topography such as B3, B4, Ele, Slope, and one vegetation index SAVI in the evergreen forest, and B11, B12, NDWI, and GNDVI in the deciduous forest were high ranked while no red-edge reflectance was correlated with AGB. This ranking affects the performance of models since the improvement in red-edge bands features was relatively larger than that in classical bands and topographic variables. This assumption was proven by previous studies of (Forkuor et al. 2020) and (Nuthammachot et al. 2020). Additionally, Chen et al. suggested that the broadleaved forests with AGB values above 160 t/ha could be underestimated because of the saturation problem in S-2 satellite data (Chen et al. 2018). The observed AGB value in the YM deciduous forest was 215.24 t/ha and thus should agree with the finding of Chen. An almost similar ranking in variables was observed in the two SGB models but the two models could not improve the estimation when considering their performances in prediction. Thus, these SGB models occurred and similar saturation problem was found in the RF models.

In order to optimize the estimation, the Kriging interpolation analysis of the residuals from RF models (RFK) was further employed since the RF showed better performance than SGB in this study. Our study claimed that the ordinary Kriging of RF's residuals (RFOK) performed better than other tested models in NH evergreen, while co-Kriging of RF's residuals (RFCK) with covariance elevation (Ele) provides a better result than other models in YM deciduous AGB prediction. The limited contribution of the accuracy of the RFOK model for the NH evergreen forest reserve was due to the poor spatial autocorrelation between AGB samples which occurred due to the spatial gaps between the sample plots. In the YM deciduous forest, RFCK achieved good prediction results, however, the spatial correlations of the current AGB were also weaker than previous studies of (H. Su et al. 2020) and (Chen et al. 2019) on forest AGB mapping based on the integration of multi-sensor and Advanced Land Observing Satellite (ALOS) data. It was denoted that AGB residuals from the integration of the MSI and SRTM data model in this study obtained a lower spatial correlation than that built by the integration of MSI and ALOS indices. This also proves that AGB estimation from the combined MSI and SRTM data was only suitable for the simple structure forest stand (e.g., deciduous forest in this study).

Overall, we conclude that the accuracies of AGB prediction could be improved to a certain extent by Kriging methods by reducing the spatial heterogeneity between AGB samples. In the future, the accuracy of evergreen forest AGB estimation in this study might be improved with LiDAR data since it can penetrate the forest canopy to a certain depth, so that its variables are suitable for extracting vertical vegetation structures with their sensitivity to the biomass of vegetation and the roughness of land cover surfaces. The determinant of the spatial setting and a sufficient number of sample plots should be considered in future AGB studies to maintain the Kriging accuracy. Additionally, the study area is located in the regions where frequent rains and clouds exists, highly restricting the availability of images collected in the vegetation growth peak season (e.g., June–September). Therefore, we had to use cloud-free images acquired in January or February, which falls outside the ideal time window for characterizing vegetation attributes. This limitation may affect the accuracy of forest AGB prediction model.

In addition, an allometric equation for in situ AGB calculation may be another factor undermining accuracy. This study used a national-level coarse allometric equation which was based on an existing inventory dataset and pantropical equation (Chave et al., 2005, Chen et al., 2013, and IPCC, 2003) for AGB calculation since there have been no species-specific equations developed for this study area. In the near future, developing species-specific allometric equations through limited destructive sampling should prioritize the carbon accounting and climate change response studies in Myanmar because these create more accurate in situ plot-level AGB measurements, laying a solid foundation for the remote sensing-based regional estimation of AGB. In general, AGB estimated from the RF model could yield acceptable results of validated $R^2 = 0.47$, $RMSE = 25.45$ t/ha for evergreen and $R^2 = 0.38$, $RMSE = 40.23$ t/ha for deciduous from S-2 derivatives, topographic variables and ancillary information.

5.4. Effects of Forest Management on AGB in the Study Sites

Population growth has led to a high demand for forest products, unsustainable forest management practices, and high deforestation rates, thus causing forest cover loss. The extent of the forest cover loss depends on the forest protection status with different rules applying to public and reserved areas (Tun, Stefano, and Volkova 2016). Forest protection typically reduces the conversion of natural land cover types to alternative uses and often results in positive outcomes (including reduced deforestation rates and the maintenance of forest cover) compared

to unprotected sites. In Myanmar, intact forests are gradually decreasing to only 38% of the country's forests due to the rapid political and economic changes. The study site comprises two protected forests under Myanmar forest law. However, the expansion of the human population and the need for more agricultural lands tend to encroach into these areas, especially in the evergreen forest presently under study. Encroachment in protected forests for agricultural lands, food, and fuels is directly correlated with loss in AGB values. On the other hand, deciduous forests' AGB values might be following a decreasing trend because the demands for commercial timber species are increasing and AGB sources are gradually decreasing. In this context, the spatial agreement of AGB was observed in the estimated AGB maps derived from the RFOK evergreen and RFCK deciduous models. For example, the small AGB values were estimated in the forest area closest to the villages and cultivated lands while large AGB values were distributed in the high-elevation forest of the NH evergreen forest reserve. A similar finding was observed in the YM deciduous forest. To sustain the forest AGB in these areas, community-based forest management is suggested to reduce these pressures as this would meet the needs of forestry products for forest dwellers.

5.5. Attainment for SDG and REDD+

The UN SDGs set out the commitment of the international community to rid the world of poverty and hunger and achieve sustainable development in its three dimensions—the economic, social and environmental facets. In addition to using standardized national official data sources as the basis for monitoring and reporting on these goals and targets, geospatial information and global/regional datasets have been identified as viable replacement and complementary data sources for achieving these SDGs. As the indicator of SDG 15, the accurate estimation and monitoring of aboveground biomass stocks need to be achieved. In this regard, the results of this study are applicable and useful for the attainment of SDG 15, especially for Myanmar where freely available optical data are preferred to map biomass stocks, and will greatly assist in the deriving, monitoring, and reporting of carbon stock changes in a timely and accurate manner.

Myanmar has been implementing the REDD+ project to achieve SDG goals and targets through sustainable forest management practices since 2013. The REDD+ objective is to find an accurate method of biomass estimation that is also cost-effective. Based on the results obtained in this study, S-2-derived derivatives (spectra, VIs, textures) and the topographic features of SRTM (elevation, slope) have potential in the forest biomass estimation of two forest reserves. In addition, the methods used in this study are viable and compatible software has been developed (e.g., SNAP), in such a way that REDD+ can apply it at a larger scale, including the national and regional levels. The outcomes of this study can surely assist the evaluation of carbon stock changes via reforestation programs that will be included in the upcoming FREL calculation under the REDD+ agenda of Myanmar.

6. Conclusions

This study investigates the performance of S-2 MSI derivatives and SRTM DEM topographic data with field ancillary information based on two machine learning models (RF and SGB) in mapping the forest AGB of two forest reserves (namely the NH evergreen and YM deciduous forests) in Myanmar. In addition, the RF-based Kriging (RFK) was employed for improving the prediction accuracy to find a spatial correction between the AGB samples. Based on these findings, it is concluded that:

- (1) S-2-derived reflectance, VIs, and textures are effective in predicting the AGB of the two forests if the proper processing techniques are applied;
- (2) The RFOK model in the evergreen forest and RFCK model in the deciduous forest provided a more realistic spatial distribution of AGB by considering the spatial correlation than the RF and SGB models with $R^2 = 0.47$, RMSE = 24.91 t/ha and $R^2 = 0.52$, RMSE = 34.72 t/ha due to their spatial correlation between AGB sample plots;
- (3) The extraction of textures from wavelet analysis (WA) is suggested to improve estimation for the forests with a complex structure and saturation problems;
- (4) In future studies, the accuracy may be improved by combining both the active and passive remotely sensed data to characterize complex forest structures to better estimate the forest AGB and understand their spatial distributions.

7. Acknowledgments

Special thanks need to go to the European Space Agency for the provision of the remotely sensed images used in the work and the Forest Department of Myanmar for providing the forest inventory data.

8. References

- Adamu, Bashir et al. 2021. "Evaluating the Accuracy of Spectral Indices from Sentinel-2 Data for Estimating Forest Biomass in Urban Areas of the Tropical Savanna." *Remote Sensing Applications: Society and Environment* 22(February): 100484. <https://doi.org/10.1016/j.rsase.2021.100484>.
- Addabbo, Pia et al. 2016. "Contribution of Sentinel-2 Data for Applications in Vegetation Monitoring." *Acta IMEKO* 5(2): 44–54.
- Ahmad, Adeel, Hammad Gilani, and Sajid Rashid Ahmad. 2021. "Forest Aboveground Biomass Estimation and Mapping through High-Resolution Optical Satellite Imagery—a Literature Review." *Forests* 12(7): 1–20.
- Balidoy Baloloy, Alvin et al. 2018. "ESTIMATION of MANGROVE FOREST ABOVEGROUND BIOMASS USING MULTISPECTRAL BANDS, VEGETATION INDICES and BIOPHYSICAL VARIABLES DERIVED from OPTICAL SATELLITE IMAGERIES: RAPIDEYE, PLANETSCOPE and SENTINEL-2." *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 4(3): 29–36.
- Basuki, T. M., P. E. van Laake, A. K. Skidmore, and Y. A. Hussin. 2009. "Allometric Equations for Estimating the Above-Ground Biomass in Tropical Lowland Dipterocarp Forests." *Forest Ecology and Management* 257(8): 1684–94.
- Bonan, Gordon B. 2008. "Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests." *Science* 320(5882): 1444–49. www.sciencemag.org.
- Breiman, Leo. 2001. "Random Forests." *Machine Learning* 45: 5–32.
- Chen, Lin et al. 2018. "Estimation of Forest Above-Ground Biomass by Geographically Weighted Regression and Machine Learning with Sentinel Imagery." *Forests* 9(1): 582.. 2019.
- Chen, Lin, Wang, Yeqiao "Optimal Combination of Predictors and Algorithms for Forest Above Ground Biomass Mapping from Sentinel and SRTM Data." *Remote Sensing* 11(4): 414.
- Conservation, Ministry of Natural Resources and Environmental, Ministry of Natural Resources and Environmental Conservation, and Forest Department. 2020. 53 Forestry In Myanmar. ed. U Ohm (Ministry of Natural Resources and Environmental Conservation) Winn. Naypyitaw, Myanmar: Ministry of Natural Resources and Environmental Conservation. https://www.forestdepartment.gov.mm/sites/default/files/Documents/Forestry_in_Myanmar_2020_0.pdf.
- Cutler, M. E.J., D. S. Boyd, G. M. Foody, and A. Vetrivel. 2012. "Estimating Tropical Forest Biomass with a Combination of SAR Image Texture and Landsat TM Data: An Assessment of Predictions between Regions." *ISPRS Journal of Photogrammetry and Remote Sensing* 70(1): 66–77. <http://dx.doi.org/10.1016/j.isprsjprs.2012.03.011>.

- Dang, An Thi Ngoc et al. 2019. "Forest Aboveground Biomass Estimation Using Machine Learning Regression Algorithm in Yok Don National Park, Vietnam." *Ecological Informatics* 50(1574–9541): 24–32. <https://doi.org/10.1016/j.ecoinf.2018.12.010>.
- Department, Forest, and UN-REDD Program. 2018. Forest Reference Level (FRL) of Myanmar. Naypyitaw, Myanmar.
- Eckert, Sandra. 2012. "Improved Forest Biomass and Carbon Estimations Using Texture Measures from WorldView-2 Satellite Data." *Remote Sensing* 4(4): 810–29.
- Fassnacht, Fabian Ewald et al. 2021. "Using Sentinel-2 and Canopy Height Models to Derive a Landscape-Level Biomass Map Covering Multiple Vegetation Types." *International Journal of Applied Earth Observation and Geoinformation* 94: 102236. <https://www.science direct.com/science/ article/ pii/S0303243420308795> (August 17, 2021).
- Ferwerda, Jelle G, Andrew K Skidmore, and Onesimo Mutanga. 2005. "Nitrogen Detection with Hyperspectral Normalized Ratio Indices across Multiple Plant Species." *International Journal of Remote Sensing* 26(18): 4083–95. <https://doi.org/10.1080/01431160500181044>.
- Forkuor, Gerald et al. 2020. "Above-Ground Biomass Mapping in West African Dryland Forest Using Sentinel-1 and 2 Datasets - A Case Study." *Remote Sensing of Environment* 236(October 2019): 111496. <https://doi.org/10.1016/j.rse.2019.111496>.
- Friedman, Jerome H. 2002. "Stochastic Gradient Boosting." *Computational Statistics and Data Analysis* 38(4): 367–78. <https://www.sciencedirect.com/science/article/pii/S0167947301000652>.
- Gascon, Ferran, Fabrizio Ramoino, and Yves-louis Deanos. 2017. "Sentinel-2 Data Exploitation with ESA's Sentinel-2 Toolbox." *EGU General Assembly 19*: 19548. <https://meeting organizer. copernicus.org/EGU2017/EGU2017-19548.pdf>.
- Ghosh, Sujit Madhab, and Mukunda Dev Behera. 2018. "Aboveground Biomass Estimation Using Multi-Sensor Data Synergy and Machine Learning Algorithms in a Dense Tropical Forest." *Applied Geography* 96(1): 29–40.
- Gibbs, Holly K., Sandra Brown, John O. Niles, and Jonathan A. Foley. 2007. "Monitoring and Estimating Tropical Forest Carbon Stocks: Making REDD a Reality." *Environmental Research Letters* 2: 045023.
- Haralick, R M, K Shanmugam, and I Dinstein. 1973. "Textural Features for Image Classification." *IEEE Transactions on Systems, Man, and Cybernetics* SMC-3(6): 610–21.
- Jolliffe, Ian T, Jorge Cadima, and Jorge Cadima. 2016. "Principal Component Analysis: A Review and Recent Developments Subject Areas: Author for Correspondence:" *THE ROYAL SOCIETY A* 374(20150202): 1–16.
- Li, Hantao, Tomomichi Kato, Masato Hayashi, and Lan Wu. 2022. "Estimation of Forest Aboveground Biomass of Two Major Conifers in Ibaraki Prefecture, Japan, from PALSAR-2 and Sentinel-2 Data." *Remote Sensing* 14(3): 468.

- Li, Long et al. 2020. "Estimating Urban Vegetation Biomass from Sentinel-2A Image Data." *Forests* 11(2): 1–24.
- Li, Yingchang, Mingyang Li, Chao Li, and Zhenzhen Liu. 2020. "Forest Aboveground Biomass Estimation Using Landsat 8 and Sentinel-1A Data with Machine Learning Algorithms." *Scientific Reports* 10(1): 99–52.
- Liu, Yanan, Weishu Gong, Xiangyun Hu, and Jianya Gong. 2018. "Forest Type Identification with Random Forest Using Sentinel-1A, Sentinel-2A, Multi-Temporal Landsat-8 and DEM Data." *Remote Sensing* 10(6): 1–25.
- López-Serrano, Pablito M. et al. 2019. "Modeling of Aboveground Biomass with Landsat 8 OLI and Machine Learning in Temperate Forests." *Forests* 11(1): 11.
- Lu, Dengsheng et al. 2016. "A Survey of Remote Sensing-Based Aboveground Biomass Estimation Methods in Forest Ecosystems." *International Journal of Digital Earth* 9(1): 63–105. <http://dx.doi.org/10.1080/17538947.2014.990526>.
- Maung, Win Sithu, and Jun Sasaki. 2020. "Assessing the Natural Recovery of Mangroves after Human Disturbance Using Neural Network Classification and Sentinel-2 Imagery in Wunbaik Mangrove Forest, Myanmar." *Remote Sensing* 13(1): 52. <https://www.mdpi.com/2072-4292/13/1/52>.
- Mngadi, Mthembeni, John Odindi, and Onesimo Mutanga. 2021. "The Utility of Sentinel-2 Spectral Data in Quantifying above-Ground Carbon Stock in an Urban Reforested Landscape." *Remote Sensing* 13(21): 4281.
- Mon, M S, and A A Myint. 2015. "Estimating above Ground Biomass of Tropical Mixed Deciduous Forests Using Landsat ETM+ Imagery for Two Reserved Forests in Bago Yoma Region, Myanmar." In *Multi-Scale Forest Biomass Assessment and Monitoring in the Hindu Kush Himalayan Region: A Geospatial Perspective*, ed. H Murthy, M S R, Wesselman, S, Gilani. , 165–77.
- Nuthammachot, Narissara, Askar Askar, Dimitris Stratoulis, and Pramaditya Wicaksono. 2020. "Combined Use of Sentinel-1 and Sentinel-2 Data for Improving above-Ground Biomass Estimation." *Geocarto International* 0(0): 1–11. <https://doi.org/10.1080/10106049.2020.1726507>.
- Ouma, Yashon O., J. Tetuko, and R. Tateishi. 2008. "Analysis of Co-occurrence and Discrete Wavelet Transform Textures for Differentiation of Forest and Non-forest Vegetation in Very-high-resolution Optical-sensor Imagery." *International Journal of Remote Sensing* 29(12): 3417–56.
- Pandit, Santa, Satoshi Tsuyuki, and Timothy Dube. 2018. "Estimating Above-Ground Biomass in Sub-Tropical Buffer Zone Community Forests, Nepal, Using Sentinel 2 Data." *Remote Sensing* 10(601): 1–18. 2019.
- Pandit, Santa, Tsuyuki, Satoshi, Dube, Timothy "Exploring the Inclusion of Sentinel-2 MSI Texture Metrics in above-Ground Biomass Estimation in the Community Forest of Nepal." *Geocarto International* 35(16): 1832–49.

- Pham, Tien D et al. 2019. "Remote Sensing Approaches for Monitoring Mangrove Species, Structure, and Biomass: Opportunities and Challenges." *Remote Sensing* 11(3): 230.
- Rodríguez-Veiga, Pedro et al. 2017. "Quantifying Forest Biomass Carbon Stocks From Space." *Springer* 3: 1–18.
- Safari, Amir, and Hormoz Sohrabi. 2020. "Integration of Synthetic Aperture Radar and Multispectral Data for Aboveground Biomass Retrieval in Zagros Oak Forests, Iran: An Attempt on Sentinel Imagery." *International Journal of Remote Sensing* 41(20): 8069–95. <https://doi.org/10.1080/01431161.2020.1771789>.
- Shen, Wenjuan et al. 2018. "Annual Forest Aboveground Biomass Changes Mapped Using ICESat/GLAS Measurements, Historical Inventory Data, and Time-Series Optical and Radar Imagery for Guangdong Province, China." *Agricultural and Forest Meteorology* 259(November 2017): 23–38. <https://doi.org/10.1016/j.agrformet.2018.04.005>.
- Simard, Marc et al. 2006. "Mapping Height and Biomass of Mangrove Forests in Everglades National Park with SRTM Elevation Data." *Photogrammetric Engineering and Remote Sensing* 72(3): 299–311.
- Su, Huiyi et al. 2020. "Machine Learning and Geostatistical Approaches for Estimating Aboveground Biomass in Chinese Subtropical Forests." *Forest Ecosystems* 7(1): 1–20. <https://doi.org/10.1186/s40663-020-00276-7>.
- Su, Yanjun et al. 2016. "Spatial Distribution of Forest Aboveground Biomass in China: Estimation through Combination of Spaceborne Lidar, Optical Imagery, and Forest Inventory Data." *Remote Sensing of Environment* 173(1): 187–99. <http://dx.doi.org/10.1016/j.rse.2015.12.002>.
- Taddese, Habitamu et al. 2020. "Use of Remotely Sensed Data to Enhance Estimation of Aboveground Biomass for the Dry Afromontane Forest in South-Central Ethiopia." *Remote Sensing* 12(20): 1–23.
- Tun, Kyaw, Julian Stefano, and Liubov Volkova. 2016. "Forest Management Influences Aboveground Carbon and Tree Species Diversity in Myanmar's Mixed Deciduous Forests." *Forests* 7: 217.
- Vaglio Laurin, Gaia et al. 2014. "Above Ground Biomass Estimation in an African Tropical Forest with Lidar and Hyperspectral Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 89: 49–58. <http://dx.doi.org/10.1016/j.isprsjprs.2014.01.001>.
- Wang, Chuyuan, and Soe W. Myint. 2016. "Environmental Concerns of Deforestation in Myanmar 2001–2010." *Remote Sensing* 8(9): 728.
- Wang, Yueting, Xiaoli Zhang, and Zhengqi Guo. 2021. "Estimation of Tree Height and Aboveground Biomass of Coniferous Forests in North China Using Stereo ZY-3, Multispectral Sentinel-2, and DEM Data." *Ecological Indicators* 126: 107645. <https://doi.org/10.1016/j.ecolind.2021.107645>.

- Wolosin, Michael, and Nancy Harris. 2018. Ending tropical deforestation Tropical Forests and Climate Change: The Latest Science. Washington, DC. wri.org/ending-tropical-deforestation.
- Wu, Chaofan et al. 2016. “Comparison of Machine-Learning Methods for above-Ground Biomass Estimation Based on Landsat Imagery.” *Journal of Applied Remote Sensing* 10(3): 035010. <https://doi.org/10.1117/1.JRS.10.035010>.
- Xue, Bowei. 2015. “Lidar and Machine Learning Estimation of Hardwood Forest Biomass in Mountainous and Bottomland Environments.” University of Arkansas, Fayetteville. <http://scholarworks.uark.edu/etd/1274>.
- Ye, Quanping et al. 2021. “Aboveground Biomass Estimation of Black Locust Planted Forests with Aspect Variable Using Machine Learning Regression Algorithms.” *Ecological Indicators* 129: 107948. <https://doi.org/10.1016/j.ecolind.2021.107948>.
- Yohannes, Hamere, and Teshome Soromessa. 2015. “Carbon Stock Analysis along Slope and Slope Aspect Gradient in Gedo Forest: Implications for Climate Change Mitigation.” *Journal of Earth Science & Climatic Change* 06(09): 6–11.
- Yuan, Xiuliang et al. 2016. “Estimation of Above-Ground Biomass Using MODIS Satellite Imagery of Multiple Land-Cover Types in China.” *Remote Sensing Letters* 7(12): 1141–49. <https://doi.org/10.1080/2150704X.2016.1219458>.
- Zhang, Yuzhen et al. 2020. “An Evaluation of Eight Machine Learning Regression Algorithms for Forest Aboveground Biomass Estimation from Multiple Satellite Data Products.” *Remote Sensing* 12(24): 4015.