ISSN: 2755-6190

Research Article

Open Access Journal of Artificial Intelligence and Technology

Biomass Carbon Map over Forests in Malaysia Produced from Landsat-Based Aboveground Carbon Density Indicator (ACDI) and a Collection of 12 Years Inventory Data

Hamdan Omar[1*], Muhamad Afizzul Misman[1], Valeria Linggok[2], Suhaini Haron[3], Ahmad Ashrin Mohd Bohari[4], and Mohammad nor Firdaus Sariee[4]

[1]*Forest Research Institute Malaysia (FRIM), 52109 Kepong, Selangor, Malaysia*

[2]*Sabah Forestry Department, 90000 Sandakan, Sabah, Malaysia*

[3]*Malaysia Forest Fund, IOI City Tower Two, 62502, Putrajaya*

[4]*Forest Department Sarawak, Baitul Makmur II, Medan Raya, Petra Jaya, 93050 Kuching, Sarawak, Malaysia*

[*]**Corresponding author**

Hamdan Omar, Forest Research Institute Malaysia (FRIM), 52109 Kepong, Selangor, Malaysia.

Received: July 11, 2025; **Accepted:** July 16, 2025; **Published:** July 23, 2025

ABSTRACT

The accurate estimation of biomass carbon in forests is of paramount importance for effective forest management and mitigating climate change. This study presents a novel approach to produce a high-resolution map of biomass carbon over forests in Malaysia using the Aboveground Carbon Density Indicator (ACDI) and a comprehensive collection of 12 years of inventory data, i.e., from 2011 to 2023. The ACDI was derived based on several vegetation indices (VIs) that were produced from the original Landsat images to indicate the level of aboveground biomass carbon (AGC) stock in the forested areas. The ACDI was then integrated with ground-based measurements and became a single indicator for estimating AGC. This calculation was conducted on Google Earth Engine (GEE) platform to match the date of field observation with the satellite imagery datasets. Latest dates of Landsat imagery were used to produce a forest cover map throughout the country, which comprised three major forest types, which are dry inland forest, mangrove forest, and peat swamp forest. Results indicated significant spatial variations in AGC across Malaysia's forests. Based on the estimates, a 30-metre resolution, wall-to-wall map of AGC across the entire forested region of Malaysia has been created. The total AGC in all types of forests in Malaysia was estimated at 3.0 billion Mg C with an attainable accuracy of about 80%. The presented methodology showcases the value of combining advanced remote sensing techniques with long-term inventory data, which contributes to the understanding of carbon dynamics in Malaysian forests and promotes effective strategies for mitigating climate change through better-informed forest conservation and management practices.

Keywords: Aboveground Carbon Stock, Tropical Forest, Landsat, Malaysia Ecosystem, Spectral Indices

Introduction

Malaysia is considered one of the highest forest carbon countries in the world due to its significant forested areas and the carbon-rich nature of its forests. Several factors contribute to Malaysia's status as a country with substantial forest carbon, that include: home to vast tropical rainforests, high plant diversity, has extensive peat swamp forests and extensive mangrove ecosystems along its coastlines. While Malaysia's forests are rich in carbon, they have faced deforestation, habitat loss, and land-use changes [1]. Malaysia has implemented various conservation measures and forest management practices to protect its forests and their carbon stocks [2]. This includes establishing protected areas and national parks.

Malaysia has made commitments under international agreements like the United Nations Framework Convention on Climate Change (UNFCCC) to reduce emissions from deforestation and forest degradation (REDD+) [3]. Malaysia has also been involved in carbon offset projects, where the country can earn carbon credits by reducing deforestation and forest degradation, as well as implementing rehabilitation, reforestation, and afforestation initiatives.

Citation: Hamdan Omar, Muhamad Afizzul Misman, Valeria Linggok, Suhaini Haron, Ahmad Ashrin Mohd Bohari, et al. Biomass Carbon Map over Forests in Malaysia Produced from Landsat-Based Aboveground Carbon Density Indicator (ACDI) and a Collection of 12 Years Inventory Data. Open Access J Artif Intel Tech. 2025. 1(2): 1-17. DOI: doi.org/10.61440/OAJPSD.2025.v1.06

In addition to being an essential part of forest ecosystems, forest biomass is also important for mitigating climate change, storing carbon, and preserving biodiversity [4]. Accurate assessment and forecast of forest biomass are essential for understanding the effects of climate change, managing forests, and accounting for carbon emissions. Remote sensing technologies, like satellite data from Landsat have revolutionised the prediction of forest biomass by providing crucial insights into the characteristics of forests and changes in land cover [5].

Landsat satellites, launched by the National Aeronautics and Space Administration (NASA) and the U.S. Geological Survey, have been providing high-resolution and multispectral imagery of the Earth's surface since 1972 [6]. Landsat data have been widely employed for various environmental and land use applications due to their long-term data archive, consistent data quality, and global coverage [7]. Landsat satellites capture data in different spectral bands, allowing researchers to analyse land cover, vegetation, and biomass across diverse landscapes. The entire historical Landsat archive has been opening for public access since 2008 [8]. As such, the Landsat archive has become one of the most valuable and cost-effective remotely sensed data sources supporting worldwide land/forest research and monitoring activities.

Among the advantages of using Landsat for biomass estimations are (i) large coverage from specific landscapes, regional to global scales, (ii) temporal and spatial scales; provide the advantage of temporal consistency, allowing for long-term biomass change monitoring in specific time-series, and (iii) sensitivity to environmental changes, which Landsat data can capture changes in forest biomass due to factors like disturbances (e.g., forest fires and logging) and climate-related stressors [9]. This sensitivity enables better understanding of the impacts of these changes on forest ecosystems.

Although Landsat data is a valuable resource for monitoring and estimating forest biomass, it has some challenges when it comes to biomass estimation in tropical regions, especially Malaysia [10]. One of the main challenges is that optical sensors are sensitive to the amount of light reflected by the vegetation, which is influenced by the structure and density of the forest canopy. However, the relationship between the amount of light reflected and the biomass is not straightforward, as it can be affected by factors such as the species composition, age, and health of the trees [11]. Moreover, clouds and atmospheric conditions can interfere with the accuracy of optical data acquisition, which can lead to incomplete or inconsistent data [12]. Another important limiting factor to direct biomass carbon modelling lies in the lack of repeated and coincident field reference data at different times [13].

Several attempts have been placed to overcome these limitations and the approaches taken can be categorised into two, which are (i) diversifying uses of spectral and vegetation indices and (ii) applying machine learning and statistical models [14,15]. These indices are used as predictor variables to estimate forest biomass, indirectly. Machine learning techniques, including Random Forest, Support Vector Machines (SVM), artificial neural network (ANN), and regression models, have been combined with Landsat data to predict forest biomass [16-18]. These models use spectral information, vegetation indices, and other environmental variables to establish relationships between the data and biomass estimates [19]. These techniques have demonstrated their efficiency in predicting forest biomass at various scales, from local to regional. Another popular solution is to combine Landsat-based data with datasets from other sensors, both optical and synthetic aperture radar (SAR) and even integrate with light detection and ranging (LiDAR)-based data [19-25]. Eventually, each approach offers different levels of difficulties and challenges [14].

This study aimed at producing reliable AGC estimates at national scale, pixel-based, wall-to-wall at acceptable spatial resolution produced from a single satellite with consistent observations that is able to represent the forest types and physical conditions of the forests over time. To achieve this, the following objectives are drawn; (i) to derive the Aboveground Carbon Density Indicator (ACDI) using the Google Earth Engine (GEE) platform, (ii) to produce AGC estimation models from the ACDI, and (iii) to map the AGC at 30-m pixel resolution for the entire forests across Malaysia.

Materials and Methods

The Study Area

Malaysia is a country in Southeast Asia, located just north of the Equator. It is composed of two non-contiguous regions: Peninsular Malaysia and East Malaysia (Figure 1). The country has a total area of about 330,803 km². Malaysia currently has about 18 million ha of forests [26]. Major forest types in Malaysia are lowland dipterocarp forest, hill dipterocarp forest, upper hill dipterocarp forest, oak-laurel forest, montane ericaceous forest, peat swamp forest and mangrove forest. In addition, there are also smaller areas of freshwater swamp forest, Melaleuca swamp forest, heath forest, transitional forest, forest on limestone and forest on quartz ridges. These are lies on lands that have various topographic features that elevate from the coastal floodplains up to 2187 m a.s.l of Tahan Mount in Peninsular Malysia and the majestic Kinabalu Mount in Sabah, with 4095 m a.s.l. Considering the composition of these forests in Malaysia, the types can be generalised into three types, which are dry inland, peat swamp and mangrove forests.

Timber production is also one of the commodities in Malaysia where State Governments are depending greatly on the forest resources for generating and sustaining the economy [27]. Malaysia is practising sustainable forest management (SFM) to balance timber production with conservation efforts. This approach aims to maintain forest carbon stocks while allowing for responsible logging. Harvesting only for merchantable timbers at certain controlled cutting limits. There are also forest plantations, established with certain timber tree species, developed to support timber supplies and meet the industrial demands.

Methodology

The framework of methodology was developed based on six major pillars, which are (i) collection of field datasets at sample plots, (ii) derivation of ACDI, (iii) development of AGC estimation models, (iv) production of seamless mosaic images, (v) forest delineation and forest types classification, and (vi) map production. Previously, there was no study that produce a

spatial information on forest biomass carbon at a country level. This study was designed in a way to produce a framework that integrate inventory data with Landsat images covering various types and conditions of forests. The first step was to match the field data collection date with the derived ACDI from the Landsat images. Google Earth Engine was used to execute this calculation and identify trends in data to create the ACDI. the derived ACDI was then correlated with the inventoried AGC. Clean images are crucial to produce a wall-to-wall map of AGC; therefore, a seamless mosaic of the Landsat images was produced. Steps ii and iv above were performed on the GEE platform, while the remaining processes were conducted separately by using image processing and GIS software, i.e., ERDAS Imagine®, Exelis ENVI Software, and Esri's ArcGIS Desktop. Overall framework of the methodology is illustrated in Figure 2 and the detailed processes involved are described as follow.

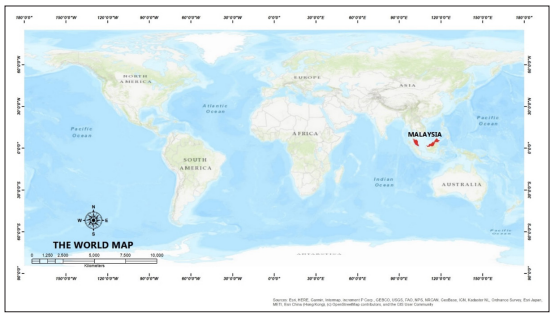


Figure 1: Map showing location of Malaysia on the World Map.

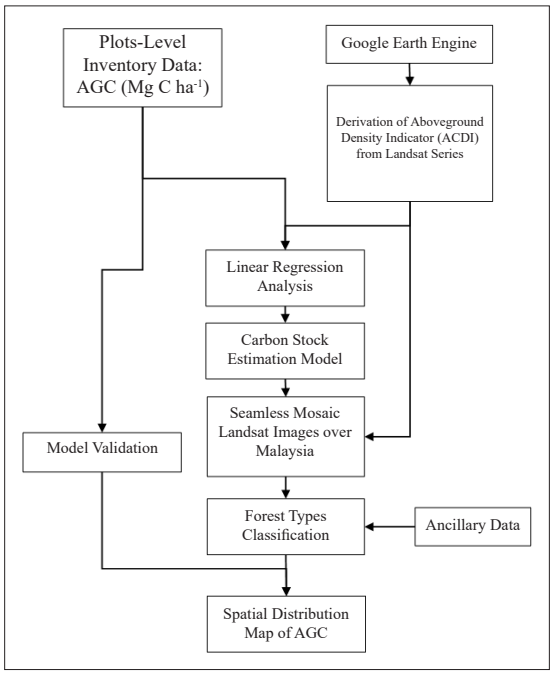


Figure 2: Flowchart of the methodology adopted in the study.

Collection of Field Inventory Data

Sampling work has been started since 2011 at several locations focused on lowland and hill dipterocarp forests in Peninsular Malaysia [28]. The work was carried out occasionally depending on available research projects that have been undergoing since then until year 2023, covering all forest types in Malaysia (Table 1 and Table 11 in Appendix B). The applied forest inventory design was stratified random, where sampling plots were distributed according to the forest types and covering all stands conditions of the forests (i.e., virgin forest, totally protected areas, logged forests, secondary forest, and degraded areas). This was considered to ensure all variations of biomass carbon are captured in the samples. Locations of the sample plots are depicted in Figure 3.

Table 1: Summary of the total number of sample plots.

Forest type	No. of sample plots		Total
	Data used for modelling	Data used for validation	
Dry inland forest	2,970	350	3,320
Peat swamp forest	1,125	75	1,200
Mangrove forest	1,750	50	1,800
Total	5,845	475	6,320

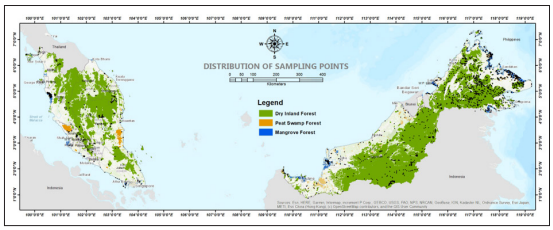


Figure 3: Distribution of sampling points on the classified forest types map.

Modified sampling design was applied in this study, following the standard operating procedure (SOP) that has been developed by Winrock International which complies to the Intergovernmental Panel on Climate Change (IPCC) protocols [29,30]. The design was then customised to suit forest conditions and management practices in Malaysia [31,32]. The sampling designs are divided into three, which are corresponding to dry inland forest, peat swamp forest and mangrove forest. The design of the sampling plots was done in clusters, where random samples were chosen after the population is split up into groups according to the types of forests and strata. Cluster sampling is a probability sampling method used when the population is large and geographically dispersed.

Estimation of biomass carbon was based on the published allometric equations found in the literature, suitable to the corresponding types of forests in Malaysia. Aboveground biomass (AGB) of the measured trees in the sample plots was first estimated before the values were converted to AGC. The estimation of AGB that was calculated at tree-level was converted to the plot-level, where the measurement is reported in Megagram (Mg) per hectare, Mg ha-1. This estimation was then converted into a biomass carbon unit of AGC by multiplying the AGB with 0.47, which is the constant carbon fraction and reported in Mg C ha-1 [30].

Dry Inland Forest

A cluster comprises four sampling plots and the distance between plots is 100 m as shown in Figure 16 in Appendix A. The plot was designed in a circular shape with smaller nests inside. The biggest nest measures 20 m in radius, followed by the smaller

nests measuring 12 m and 4 m (Figure 19 in Appendix A). The sizes of trees are measured according to the nest sizes, which is summarised in Table 10 in Appendix A. Depending on the nest size, it indicates that not all trees are measured in a single plot. In addition to these nests, there is another small nest measuring 2 m in radius, which is used to count the saplings. The clustering of multiple plots at one sampling unit allows field crews to sample a larger area per sampling point. The sampling system is designed in such a way to make the data collection processes easier, faster, reliable and representative for a forest stratum. The distance of the trees is controlled by using a Distance Measurement Equipment (DME) that utilises sonar waves to communicate with a transponder that is installed at centre of the plot. Therefore, in reality the nests with particular radius do not exist on the ground.

The estimation of AGB on dry inland forest was calculated based on an allometric equation that was developed for the dry inland forest [33]. The allometric equation is expressed as follow;

$$AGB = exp[-1.803-0.976E+0.976\ ln\ (\rho)+2.673\ ln\ (D)\ -0.0299[ln\ (D)]^2\ ] \quad (1)$$

where AGB denotes the estimated biomass of a tree (kg tree^{-1}), D is diameter at breast height (dbh) of each tree (cm), ρ is wood specific gravity or wood density (typical average value for all Southeast Asia's tree species is 0.57 g cm^{-3}) and E is bioclimatic variable, which is available at http://chave.upstlse.fr/pantropicalallometry.htm [34].

Peat Swamp Forest

Peat swamp forests are terrestrial wetland ecosystems with low nutrient levels and highly acidic soil (pH less than 4.0) [35]. Ecologically, peat swamp forests have organic soil horizons, or peat that can receive water and nutrients exclusively from flooding and groundwater or from rainfall. In the tropics, peat formation is influenced by high rainfall rates, minimal drainage, and high temperatures with little seasonal change. According to peat swamp forests are typically submerged during the rainy season, which encourages anaerobic conditions that influence the rates and pathways of decomposition and accumulation [36]. Peat soils are described as having at least 50 centimetres of thickness and a content of organic matter greater than 65% in tropical ecosystems. The layout of sampling plots for peat swamp forest is depicted in Figure 16 and Figure 19 in Appendix A [37]. The sizes of trees are measured according to the nest sizes, which is summarised in Table 10 in Appendix A. The allometric equation for the estimation of AGB in peat swamp forest is expressed as

$$AGB = 0.136D^{2.51} \quad (2)$$

Mangrove Forest

Mangroves are defined as an association of halophytic trees, shrubs and other plants growing in brackish to saline tidal waters of tropical and subtropical coastlines [38]. Mangroves are generally restricted to the tidal zone. As such, mangroves in fringe areas will be inundated by practically all high tides, while those at the higher topographic boundaries may be flooded only during the highest of tides (spring tides) or during storm surges. Mangroves are typically found along tropical and subtropical coastlines between about 25° N and 25° S.

Mangrove forest has its own habitat, which is unique in terms of ecology, standing structure and species composition. Therefore, the sampling method for mangrove forest is designed specifically for the mangroves. The sampling can be organised in a cluster, comprising 6 plots (Figure 18 in Appendix A). The layout of sampling plots for the mangrove forest is depicted in Figure 19 in Appendix A [39]. The sizes of trees were measured according to the nest sizes, which is summarised in Table 10 in Appendix A. The allometric equation adopted for the calculation of AGB in mangrove forest is expressed as

$$AGB = 0.251\rho D^{2.46} \quad (3)$$

where ρ is wood specific gravity or wood density (average value for all mangroves tree species is 0.752 g cm^{-3}).

Production of Seamless Mosaics, Cloud-Free Images Over Malaysia

The production of cloud-free images over Malaysia was done using GEE. In this study, a Top-of-Atmosphere's (ToA) cloud-free mosaic image for Malaysia in the year 2023 was generated using Landsat 8 and Landsat 9 satellite imagery obtained from the "LANDSAT/LC08/C02/T1_TOA" and "LANDSAT/LC09/C02/ T1_TOA" collections. The use of data that were acquired between 2022 and 2023 is to maximise the chance of getting clear images with less cloud covers. A cloud masking approach was applied to the selected images using the "QA_PIXEL" band. This band was used to mask pixels containing dilated clouds, cirrus clouds, and cloud shadows. This cloud masking process was crucial for excluding cloudy or obscured pixels, resulting in a cleaner and more accurate composite image. The composite image was generated using the median value for each pixel across the selected cloud-masked images. The median composite method was chosen because it is simple for calculation and its robustness against outliers and its ability to reduce the influence of noise and artifacts in the final image. Finally, the cloud-free mosaic image over Malaysia of year 2023 was created by mosaicking the individual median composite images.

Forest Cover and Types Classifications

Forest is defined as "a portion of land larger than 0.5 ha and has trees with a height of more than five metres and has a tree canopy cover of more than 10 percent or with trees that can meet these criteria". This definition is based on the UN Food and Agriculture Organization's (FAO), which is adopted by the Malaysian government Laws of Malaysia - National Forestry Act 1984 (Amended, 2006). However, there are different types of forests in Malaysia, such as inland mixed dipterocarp forest, peat swamp forest, and mangrove forest, which have different characteristics and functions. Therefore, the definition of a forest may vary depending on the context and the purpose of the classification. Inland mixed dipterocarp forest, which is divided into several layers according to the land elevations, i.e., lowland dipterocarp forest (< 300 m), hill dipterocarp forest (300 - 750 m), upper-hill dipterocarp forest (750 - 1200 m), oak-laurel forest (1200 – 1500 m), montane ericaceous forest (>1500 m), are dominant in Malaysia [40].

All dryland forests are included in this category. It includes all primary and secondary forests that meet the defined threshold. It would, thus, also include the dwarf Montane and Sub-Montane forests growing on the thin soils of mountain summits and ridges of the interior of the peninsula. The dry inland forest in Malaysia is mostly dominated by trees from the Dipterocarpaceae family, hence the term 'dipterocarp' forests. The dipterocarp forest occurs on dry land just above sea level to an altitude of about 900 m. The dipterocarp specifically refers to the fact that most of the largest trees in this forest belong to one plant family known as Dipterocarpaceae. It was so called because their fruits have seeds with two wings (di = two; ptero = wing; carp = seed) [41].

The peat swamp forest resides on peatlands behind the mangrove forest, where less salty soils are present. On the other hand, mangrove forest normally exists near to the coastal and estuarine areas where the forest is influenced by tidal waves. Tidal forest where the genera Rhizophora, Bruguiera and Avicennia are most common. Smaller sections of casuarina/beach forest, freshwater swamp forest, melaleuca swamp forest, heath forest, limestone forest, and quartz ridge forest are also present. In Sabah, there is another type of vegetation zone, known as sub-alpine vegetation, which occurs only at the elevation of > 3500 m a.s.l., at the peak of Kinabalu Mount [31].

Image classification was executed on the mosaic images to delineate these forest types. To ensure that the classification is done correctly, spatial information such as Permanent Reserve Forest (PRF) boundaries, management regimes, and locations of various ecosystems are necessary. The training areas were manually created based on this ancillary information aided by the sampling plots information. In this case, the image classification was performed to delineate forests from other land features. Maximum likelihood image classification algorithm was utilised to execute the classification.

The most difficult aspect of image classification was dealing with large amounts of data and producing classification results with minimum uncertainty [42]. Pixel format classification results have been converted to shapefile vector format (.shp) for further analysis and post-classification recognition processes. Further editing and refining were conducted manually over the shapefile to ensure that the classification results are clean and only cover the forested areas.

Development of ACDI

The ACDI is a metric developed on the premise that there exists a direct correlation between the density of a forest's canopy, or the amount of foliage and branches in its upper layers, and the quantity of carbon stored in the forest's biomass. This relationship is rooted in the principle that a denser canopy typically implies a more extensive and carbon-rich vegetation structure. The ACDI is used to estimate the amount of carbon stored in a forest, which is important for evaluating forest carbon sink capacities. As such, the ACDI will serve as a valuable tool for estimating the amount of carbon sequestered in a forest ecosystem by analysing its AGC. The development of ACDI is based on the Forest Canopy Density (FCD) model that was established by and modified by [43-45].

The FCD model used SI, advanced vegetation index (AVI), Bare-soil Index (BI) and Thermal Index (TI). An inspection was conducted on this model and found that ambiguities exist at the grassland and the shrublands, especially burn scars areas where the FCD is found to have higher values than that of forested areas [46]. This effect needs to be eliminated and the only solution to this is by suppressing the values to a level that is representative to the actual physical condition on the ground. Therefore, the other indices were incorporated into the equation which each index is able represent the real conditions of forests and the ACDI is thus developed, which can be expressed as

$$ACDI = \left[\frac{NDVI \times NBR \times SI}{SAVI \times 10 \times MNDWI \times SWIR \times EVI}\right] \times 2 \quad (4)$$

where each image variable is summarised in Table 2. All indices were calculation based on Top of Atmosphere (ToA) reflectance values.

The vegetation indices used in the ACDI were chosen with care to highlight the forest areas, distinguish them from other features, and show variations under different circumstances. The Normalised Difference Vegetation Index (NDVI) is a widely-used metric for quantifying the health and density of vegetation using sensor data. The Shadow Index (SI) is used to derive information about various landscape phenomena, including vegetation health and land classifications. However, the specific purpose or application of the Shadow Index is for detecting and correcting for shadows in optical satellite imagery. On the other hand, the Normalised Burn Ratio (NBR) is a radiometric measure of burn severity that was originally developed using Landsat Thematic Mapper data. The NBR is a widely used index for monitoring environmental changes, particularly those related to fire intensity and burn severity.

The SAVI is a vegetation index that is designed to minimise the influence of soil brightness on the vegetation signal. It is particularly useful in areas where vegetative cover is low. In contrast, the IO can be used to estimate the presence of iron oxide in various landscapes, such as wetlands. The ratio presented in IO is also used as a geological index used for identifying rock features that have experienced oxidation of iron-bearing sulphides. However, in this case the IO was included in the equation to differentiate forest cover especially in wetlands areas [47]. On the other hand, the MNDWI is a spectral index used for several purposes, such as enhancement of open water features that is particularly useful in built-up areas as it can reduce or even remove built-up land. It is also used to analyse water bodies such as rivers, lakes, and dams. In this case, the MNDWI was included to diminish built-up area features that are often correlated with open water in other indices. Finally, EVI was included in the equation as one of the multiplicative indicators in the denominator. This "optimised" vegetation index aims to improve vegetation monitoring by decoupling the canopy background signal and minimising atmospheric impacts, hence increasing the vegetation signal's sensitivity in high biomass regions. It thus enhanced the vegetation health and density of vegetation.

The ACDI equation was then applied to the Landsat-8 Operational Land Imaging (OLI) for the year 2023. This process is similar

to the production of a seamless mosaic of Landsat images over Malaysia as described earlier. However, an additional step was applied to include the ACDI formula to the image. This process was also conducted on the GEE platform.

Table 2: Image variables that were used to develop ACDI.

Image variable	Full name	Formula	Reference
NDVI	Normalised Difference Vegetation Index	[(NIR – R)/ (NIR + R)]	[48]
NBR	Normalised Burn Ratio	[(NIR – SWIR)/ (NIR + SWIR)]	[49]
SI	Shadow Index	[(1 – B) (1 – G) (1 – R)]1/3	[50]
SAVI	Soil-Adjusted Vegetation Index	[(NIR – R)/ (NIR+R+L)] *[1+L]	[51]
IO	Iron Oxide Index	R/B	[52]
MNDWI	Modified Normalised Difference Water Index	[(G – SWIR)/ (G + SWIR)]	[53]
EVI	Enhanced Vegetation Index	GF× [(NIR – R)/ (NIR + C1 × R – C2 × B + L)	[54]

B = blue wavelength channel, G = green wavelength channel, R = red wavelength channel, NIR = near infrared wavelength channel, SWIR = short wave infrared wavelength channel, GF = Gain Factor, L = the canopy background adjustment that addresses non-linear, differential NIR and red radiant transfer through a canopy. The coefficients adopted are: L=1, C1 = 6, C2 = 7.5, and GF = 2.5.

Development of Agc Estimation Models

The linear relationship between AGC and the ACDI is a fundamental connection in the assessment of carbon content in terrestrial ecosystems. AGC represents the total carbon stored in the aboveground biomass of trees. ACDI, on the other hand, is a metric used to express this carbon content relative to a unit of area, typically per hectare or square metre. The extraction process was conducted on the GEE platform where a specific program code was created to extract the ACDI values from Landsat data that match the date (or year) of the field inventory data. This is to ensure that the value of AGC is true at the specific time, because the forest can change over time.

The linear relationship between AGC and ACDI is straightforward: as the aboveground carbon content increases in a given area, the ACDI value for that area also increases proportionally. Simple linear regression is a statistical method used to estimate the relationship between two quantitative variables. It is preferred over other regression models to measure the strength of the relationship between AGC and ACDI. Simple linear regression is also preferred when only one independent variable, (i.e., ACDI) is available. In addition to the simple linear regression, logarithmic regression (also known as log-linear regression) regression was also applied to model the relationship between AGC and ACDI using logarithmic transformations. In this case ACDI is the predictor for AGC, where the linear relationship between these two variables can be expressed as

$$y = mx \quad (5)$$

$$y = m.log(x) \quad (6)$$

where y denotes AGC, x is the ACDI, and m is the equation's coefficient. Both x and y variables intercept at 0, which means that the line passes through the origin (0, 0) of the plane, where ACDI is 0 when AGC is 0 or no vegetation (cleared land and water bodies).

Models' Validation

Some of the sample plots data were used separately for validation (Table 1). The validation plots are those measurements that have been conducted recently in the year 2023 to match the AGC map that was produced for the year 2023. To check the accuracy of the estimates, root mean square error (RMSE) was calculated. In this case, the accuracy is a measure of the error between a derived/predicted AGC from the ACDI and the actual AGC measured on the ground. RMSE is a useful metric and commonly used for comparing the fit of different regression models. It is calculated as the square root of the variance of the residuals, which are the differences between the observed data values and the predicted values from the model. The calculation can be expressed as follows:

$$RMSE = \sqrt{\sum \frac{(AGC_p - AGC_r)^2}{n}} \quad (7)$$

where RMSE is the root mean square error of the estimated AGC (± Mg C ha^{-1}), AGC_p and AGC_r are the predicted and reference AGC, respectively, and n is the sample size (i.e., number of validation plots).

In additional to the RMSE, the accuracies of the estimates were also measured in terms symmetric mean absolute percentage error (SMAPE). SMAPE is a commonly used metric for measuring the percentage accuracy between forecasted and actual values. It is particularly used to assess the performance of a forecasting model, and it has a preference for symmetrical errors. The adjusted SMAPE values typically range from 0% to 100% [55]. A lower SMAPE indicates a better forecast accuracy, while a higher SMAPE indicates a less accurate forecast. SMAPE is calculated as follows:

$$SMAPE = \frac{100}{n} \sum_{n=1}^{n} \frac{|AGC_p - AGC_r|}{|AGC_R| + |AGC_p|} \quad (8)$$

Thematic Map Production

The empirical equations that have been derived from the regression analysis were applied to the ACDI images by using ERDAS Imagine® Model Maker tools. Each equation was applied to produce estimated AGC according to the forest types. Since the model produced was made according to the type of forest, the processes were repeated for three times, each for dry inland forest, peat swamp forest, and mangrove forest. Each resulting AGC image was then cropped to match the forest type. Then the three images were recomposed to produce a single image containing the AGC value according to the type of forest. The mosaiced product was a single-layer image with pixel values representing AGC, producing a wall-to-wall map

of AGC throughout Malaysia at 30-m resolution. By using this map, AGC at any location can be determined and statistics of AGC within any polygon can be extracted.

Maps Comparison

Forest biomass carbon map that has been produced from this study was also compared with the products that have been produced previously by Baccini and Hensen [56,57]. These maps were produced from different data sources, mainly from Landsat imageries in years 2011-2012. To match the dates between these maps and the products of this study, another layer of AGC map was produced for the year 2012, by using the prediction models derived from this study. These datasets were stacked together and pixel-to-pixel comparison was made based on AGC values on these maps.

Results and Discussion

Summary of the Sample Plots Data

The field inventory work that has been conducted covered a wide range of forest types and conditions, from severely degraded areas to the highly pristine, virgin forest. Non-tree spots within the sampling areas, with AGC value of 0 were also included as samples. Statistics of the AGC that were measured at sample plots are summarised in boxplots as shown in Figure 4.

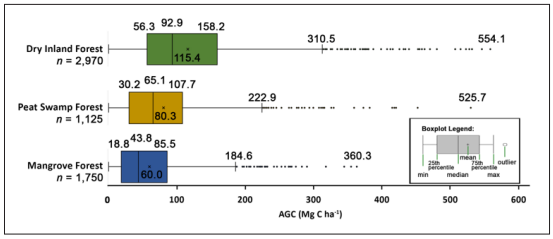


Figure 4: Boxplots summarising the sample plots data.

Seamless Mosaics, Cloud-Free Images Over Malaysia

A seamless mosaic, cloud-free images of Landsat images over Malaysia has been produced as shown in Figure 5. A 2-year threshold (i.e., 2022-2023) was found the most optimal for the production of cloudless images containing clear pixels of more than 99%. The median composite method that was chosen to produce this product was found as the most appropriate technique to produce quality data. Each pixel represents the "typical" appearance of the area over the time period. This technique was found useful for reducing noise and removing clouds or other atmospheric disturbances that are present in individual images [58].

The Classified Forest Cover and Types

From the classification, it was found that the total area of forests in Malaysia in 2023 was about 18 million ha with dry inland forest being dominant at 93.3% (Table 3). The classification accuracy was assessed by using the most recent sampling plot data and compared with the classified pixels in a confusion matrix. The resulted accuracy was reported as user's and producer's accuracy at 92.3 % and 89.6%, respectively. The forest types classification results were produced in shapefile (.shp) format so that further analysis and statistical extractions can be carried out over the AGC map. Figure 6 shows the map of forest types that have been classified from the image.

Table 3: Extents of forests in Malaysia produced from image classification (2023).

Forest type	Extent (ha)	Percentage (%)
Dry inland forest	16,859,417	93.3
Mangrove forest	547,564	3.0
Peat swamp forest	655,422	3.6
Total	**18,062,403**	**100.0**

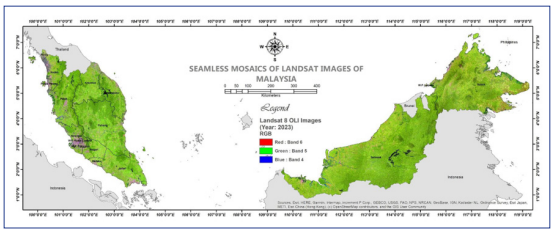


Figure 5: Seamless mosaic, cloud-free images of Landsat over Malaysia of year 2023.

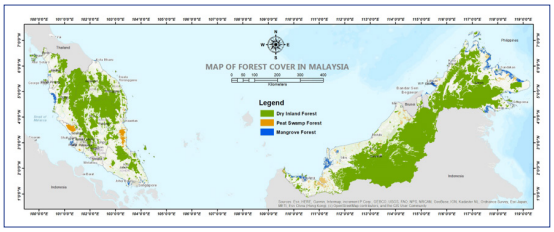


Figure 6: Map showing forest cover and types over Malaysia, produced form the image classification.

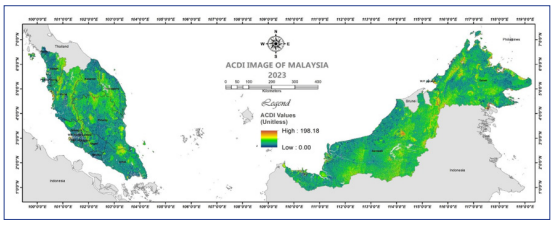


Figure 7: Map showing spatial distribution of ACDI over Malaysia, derived from the Landsat mosaic images.

Summary of the ACDI

The ACDI that have been derived from the Landsat images for the year 2023 ranged from about 0 to 200. However, the values are mainly concentrated at values ranging from 0 to 50 (Figure 8) and minor pixels containing values exceeding 100. The statistic of ACDI is summarised in Table 4 and the spatial distribution of ACDI is presented on the map in Figure 7. The histogram represents all terrestrial features in Malaysia, which includes all categories of land use/cover. While water bodies, bare lands and built-up areas have relatively low ACDI values, vegetation covers tend to have higher values. In this case, all vegetation including forests and agricultural lands are mixed and some of them are sharing the same values of ACDI. Therefore, the forest cover and types classification are crucial and took the primary part in the processes.

Table 4: Basic statistics of ACDI values over Malaysia for the year 2023.

Min	Max	Mean	Median	Mode	Std. Dev.
0.00	198.18	25.34	22.46	19.36	14.77

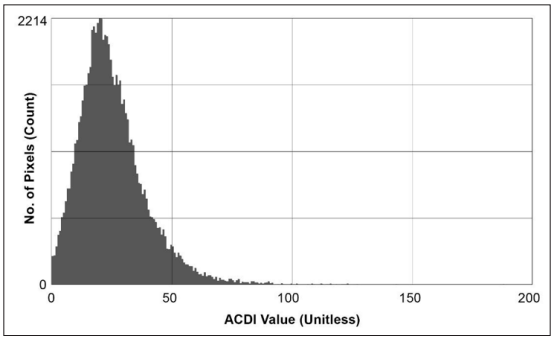


Figure 8: Histogram of ACDI distribution over Malaysia.

AGC Estimation Models

Scatterplots of AGC against ACDI have been produced with linear correlations created for all forest types. Referring to the scatterplot it is obvious that the ACDI demonstrated different responses towards the AGC. The slope of the linear regression line, which is steeper than that of peat swamp and mangrove forest, indicates that the dry inland forests exhibit a larger proportional relationship between ACDI and AGC. Mangrove forests, on the other hand, have the least gradient. This is because wetlands (i.e., peat swamps and mangrove forests) and dry inland forest have different soil properties [59]. The ACDI formula has normalised the impact on soil properties. In contrast, peat swamps and mangrove forests appear darker due to the reflectance in the infrared region that has been absorbed by the moisture as it interacts with the wetlands [60]. Mineral soil beneath dry inland forest tends to become brighter because the amount of reflectance mainly comes from the canopy of the trees [61]. The derived AGC estimation models are summarised in Tables 5 and the scatterplots of correlations between AGC and ACDI are depicted in Figure 9.

Logarithmic transformations regression performed almost similar to the simple linear regression. However, the prediction model was better for inland forest. In contrast, the models showed lower performance in mangrove forest as compared to the simple linear regression. This reflects the complexity different forest types and hence the behaviour is unpredicted in a nonlinear form. Furthermore, the developed ACDI has taken this variability into consideration by normalizing the unevenness that occur within each forest type by using various vegetation indices. The derived AGC estimation models from logarithmic transformations are summarised in Tables 6.

It is desirable and expected that the AGC has a perfect linear relationship with ACDI. However, after the analysis was carried out, it was found that the relationship is still divergent and this happens due to several factors. The first factor is the coordinates of the location of the sampling plot which is not very accurate and the position of the plot which does not fall exactly on the actual pixel locations. Another factor is the use of allometric equations that do not relate forest canopy information in AGC calculations, whereas the information extracted from satellite data is based on the forest canopies. In addition, the spatial resolution of Landsat data at 30-m accuracy includes many mixed features in a pixel when compared to the plot sizes used in the census, especially for peat swamp and mangrove forests where the plot sizes are smaller than the pixel resolution. Nonetheless, the correlations exhibit significance, with r2 values surpassing 0.5, and are deemed acceptable, given the huge amount of field data available to reveal the true relationship between AGC and the ACDI.

Table 5: Summary of AGC estimation models derived from the regression analysis.

Forest Type	Empirical Equation*	Number of samples (n)	Correlation Coefficient (r^2)	Adjusted r^2
Overall forest types	AGC = 2.1187*ACDI	5,845	0.4897	0.4898
Dry inland forest	AGC = 3.3763*ACDI	2,970	0.6275	0.6276
Peat swamp forest	AGC = 2.3133*ACDI	1,125	0.5787	0.5791
Mangrove Forest	AGC = 1.0815*ACDI	1,750	0.6230	0.6232

*All correlations are significant at probability value, $p < 0.05$.

Table 6: Summary of AGC estimation models derived from the logarithmic transformations function.

Forest Type	Empirical Equation*	Number of samples (n)	Correlation Coefficient (r^2)	Adjusted r^2
Overall forest types	AGC = 26.1995*log (ACDI)	5,845	0.629	0.631
Dry inland forest	AGC = 31.6466*log (ACDI)	2,970	0.698	0.702
Peat swamp forest	AGC = 21.2150*log (ACDI)	1,125	0.588	0.592
Mangrove Forest	AGC = 14.4267*log (ACDI)	1,750	0.501	0.519

Statistics Extracted from the AGC Map

The AGC map that has been produced from the estimation models contained pixel values ranging from about 0 to 450. The histogram shows that there are two distinct regions of distributions, creating two different peaks, which reflect the estimated AGC for forests and other vegetative covers (Figure 10). Although the estimation is not valid for vegetation other than forests, all pixels will contain AGC values once the model is applied to the ACDI image. The statistic of the AGC is summarised in Table 7 and the spatial distribution of AGC is portrayed on map in Figure 13. It was estimated that the total AGC in the entire forests in Malaysia was at 3.0 billion Mg C, which was a sum of 2.87 billion Mg C, 71.9 million Mg C, and 56.86 million Mg C from dry inland, peat swamp and mangrove forests, respectively. Given the entire forests in Malaysia is divided into three types, the averages AGC estimated for dry inland, peat swamp and mangrove forests are 171.45 ± 67.00 Mg C ha^{-1}, 109.51 ± 60.78 Mg C ha^{-1}, and 91.50 ± 76.18 Mg C ha^{-1}, respectively. Total AGC was also calculated based on the forested areas found throughout the country by using the

AGC map. The carbon stock profile for each state in Malaysia has been determined. A summary of the AGC profile is given in Table 9 and shown in Figure 14. This information is very useful in determining the carbon stock capacity at the national, states, and project or site-specific levels.

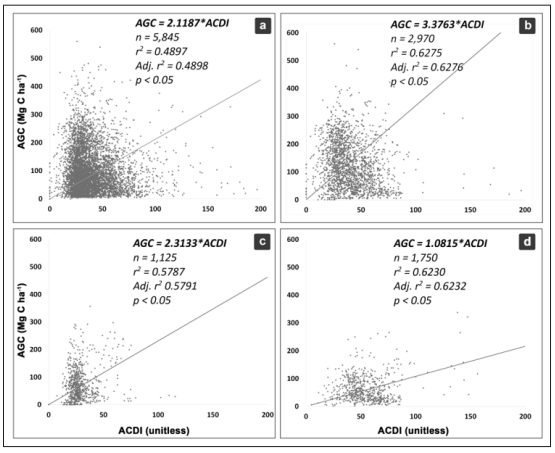


Figure 9: Scatterplots of correlations between AGC and ACDI for (a) overall forest types, (b) dry inland forest, (c) peat swamp forest, and (d) mangrove forest.

Table 7: Basic statistics of AGC values (Mg C ha-1) throughout Malaysia for the year 2023.

Min	Max	Mean	Median	Mode	Std. Dev.
0.00	448.79	126.72	151.35	59.83	61.98

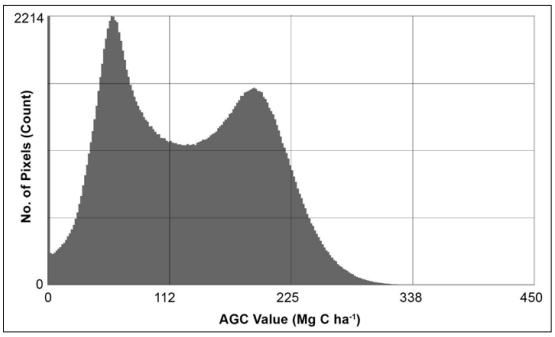


Figure 10: Histogram of AGC distribution over Malaysia.

The performance of the AGC map produced from this study was measured by extracting the profiles of AGC at different forest types and conditions. Twelve areas (Figure 11) have been selected to demonstrate the variations of AGC distribution, which are summarised in Table 9. The spatial distribution of AGC over these areas are depicted in Figure 15, which represent dry inland forest, mangrove forest and peat swamp forest, respectively. These areas are among areas that are known for their functions.

A comprehensive review on the aboveground carbon stock at various forests in Malaysia was reported by [62]. The values vary according to the forest types and conditions and most of the reported AGC values are agreeable with the values estimated in this study. Similar situation occurs in mangrove forest, where the range of AGC is agreeable to that of reported by [60,63].

Previous report estimated that the total aboveground biomass carbon in 2015 was at 2.248 billion Mg C, with an average of 154.78 Mg C ha^{-1}, within a forested area of 18.278 million ha [64]. This was somehow lower than that estimated by the current study. However, it was estimated that the total AGC in all forest types in Malaysia was at 3.15 billion Mg Cover the year 2020 [10]. This is agreeable with that found in the current study.

It was reported that the total AGC in the lowland and hill dipterocarps forests in Peninsular Malaysia was at 775,884,956 Mg Cover the year 2015 [65]. AGC in about 5.25 million ha of the dry inland forest, excluding montane forest, in Peninsular Malaysia was estimated at 855,970,674 Mg C and 833,141,077 Mg Cover the year 2016 [66,67]. Current study found that the total AGC in the dry inland forest in Peninsular Malaysia was at 921,731,750 Mg C. The estimates were slightly higher because it includes montane forests, which has elevation > 1200 m a.s.l.

A study found the Totally Protected Areas (TPA) forest has among the highest carbon densities in Sabah, averaging 165 Mg C ha^{-1}, Maliau Basin with 220 ± 69 Mg C ha^{-1}, and Danum Valley with 207 ± 71 Mg C ha^{-1}. Other forest reserves that are in intact condition yielded even higher carbon densities, with Imbak Canyon producing the highest mean stock of 229 ± 81 Mg C ha^{-1} [68]. These estimates are very close to that found in this study with an average of 215.72 Mg C ha^{-1} in Maliau Basin forest landscape (Table 10-B).

It was estimated that AGC in Endau Rompin National Park in Johor wat at an average of 281 Mg C ha^{-1} [69]. Assuming that forest condition in this area is similar to the Greater National Park, this study estimated the AGC in this kind of forest was at 203.61 Mg C ha^{-1} (Table 10-A), which is lower than that was estimated by them. However, it was justified that the allometric equation that was used in the study is different from that of used by this study. This can contribute to the final AGC estimates. In contrast, it was indicated that the carbon stock in production forest at the production area ranged between 24.6 and 265.8 Mg C ha^{-1} with the mean at 166.8 Mg C ha^{-1} [70]. This is comparable with that found in this study with an average at 192.55 Mg C ha^{-1} (Table 10-D).

Another assessment looked into the aspect of chronosequence rehabilitated tropical forest stands in Malaysia. It compares the carbon stock of different age classes and forest types, and evaluates the effectiveness of forest rehabilitation. The rehabilitated forests have tree carbon ranging from 0.1 - 54.0 Mg C ha^{-1}. In contrast to the natural regenerating secondary forest, tree carbon was at 61.0 Mg C ha-1 [71].

AGC Map Accuracy

The AGC map was validated by using separate sample plots that were allocated for validation purposes. The predicted AGC values were fitted against the actual values measured at validation plots. The validation scatter plot is a common tool used to measure the performance of a model. It is used to visualise the relationship between the predicted values and the actual values of a model. The scatter plot shows how well the model is able to predict the actual values, and how much variation there is between the predicted and actual values. The closer the points are to the line of perfect prediction, the better the model's performance. This plot is particularly useful when evaluating regression models, as it allows to measure the performance of the models developed

to predict continuous variables. The validation scatterplots are shown in Figure 12. The accuracy of the model's performance was also assessed by determining the RMSE and SMAPE.

The study found that the AGC predicted on mangrove forest attained the best accuracy at 84.85% with ±22.51 Mg C ha^{-1}. Lower accuracies obtained for peat swamp and dry inland forests, with the attainable accuracies at 77.14% and 77.34%, respectively. Table 8 summarises the overall accuracies of the predictions resulted from the models.

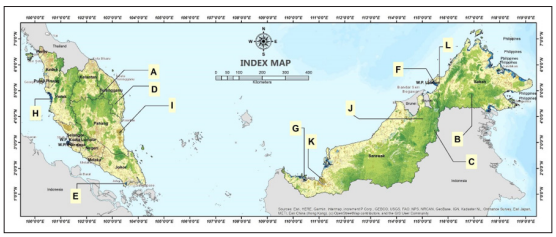


Figure 11: Map showing locations of the selected areas.

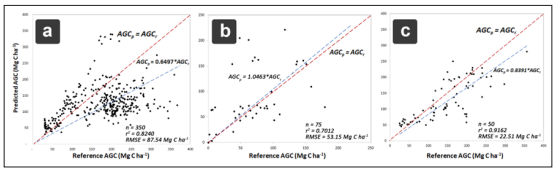


Figure 12: Validation scatterplots for the assessment of models' performance.

AGC Maps Comparisons

Comparisons demonstrate that the maps that the AGC predicted from previous studies are in agreement with the one produced from this study (Figure 13). The reference AGC maps were produced at different pixel resolutions, which Baccini produced at 30-m resolution and Hensen produced at 300-m resolution. Therefore, the predicted AGC map pixel resolution for the year 2012, as generated by this study was also resampled accordingly to match the spatial resolutions of the reference maps. The comparisons were made for all pixels within the AGC maps, regardless the land use /land cover classes, depicted in linear correlations as shown in Figure 14. The scatterplots show that there are two major domains in the AGC values, one at higher ends belongs to the forests and the lower parts is for the other land use/land cover classes. The comparisons are satisfying with r^2 of 0.8251 and 0.8624 for Hensen and Baccini, respectively. However, it was notable that the predicted AGC in this study was underestimated compared to Hensen and slightly overestimated compared to Baccini these occur due to several factors, which are the different dates of data sources, the map scales, and the resampled pixel size. Baccini used Landsat satellite data between years 2010 and 2012 to produce the final AGC maps, while Hensen et al. used data from years 2011 to 2012 to come out with the final maps. Both have produced maps at a global scale.

This demonstrates that modelling AGC from a single source of remotely sensed data subject to uncertainties. Satellite images do not measure forest biomass carbon directly, but indirectly by integrating the spectral information offered by the sensors. Therefore, the uncertainties sourced from the radiometric characteristics itself, and all optical imageries only deal with the reflectance from the forest canopies, which does not directly represent the forest structure and biomass. Other sources of uncertainties include misplaced sample plots, spatial resolution that does not match plot size, forest sampling time that does not match image acquisition dates, tree measurement protocols, tree species, variations within different forest types and ecosystem habitat, wood density, topography/terrain features, and management regimes. These are among issues that are very complex to address and normally lead to uncertainties in forest carbon biomass estimations. Eventually, different AGC map products will always have different sets of values and not consistent over times.

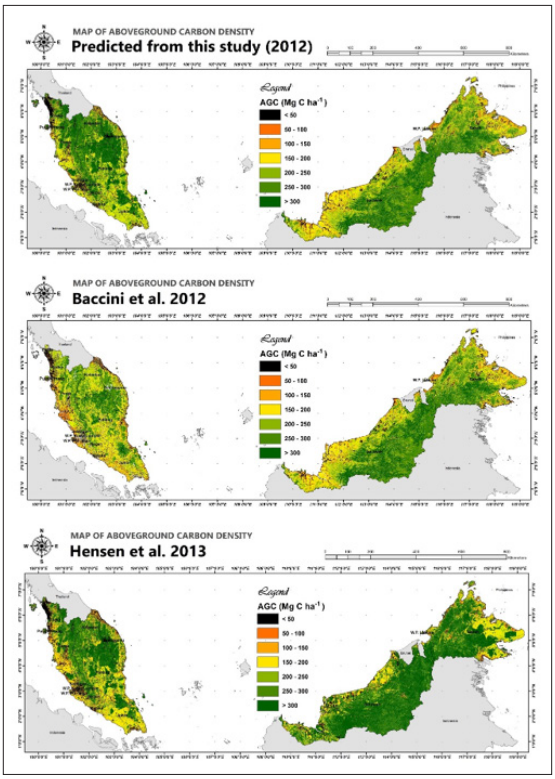


Figure 13: Comparison of different AGC maps.

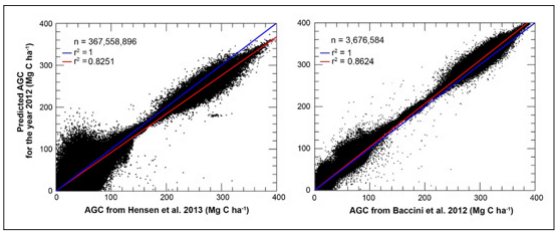


Figure 14: Scatterplots represent the comparison between AGC maps.

Table 8: Accuracies of the AGC predictions

Forest Type	RMSE (±Mg C ha^{-1})	SMAPE (%)	Absolute accuracy (%)	Overall performance
Dry inland forest	87.54	22.66	77.34	Underestimate
Mangrove Forest	53.15	22.86	77.14	Overestimate
Peat swamp forest	22.51	15.15	84.85	Underestimate

Table 9: Summary of AGC in all states in Malaysia for the year 2023.

State	Forest Types										Total AGC (Mg C)
	Dry inland forest		Mangrove forest		Peat Swamp forest		Forest plantation				
Johor	374,290	189.72	71,009,966	22,903	99.71	2,283,719	3,682	137.82	507,394	201	91.78	18,484	73,819,564
Kedah	316,978	200.28	63,484,726	8,312	144.77	1,203,284	n.a	n.a	n.a	n.a	n.a	n.a	64,688,010
Kelantan	756,532	150.37	113,757,308	966	94.61	168,613	n.a	n.a	n.a	n.a	n.a	n.a	113,848,671
Melaka	5,747	141.30	812,087	833	70.53	58,754	n.a	n.a	n.a	n.a	n.a	n.a	870,841
Negeri Sembilan	168,201	172.22	28,968,338	1,107	113.98	126,148	n.a	n.a	n.a	2,026	96.71	195,962	29,290,448
Pahang	1,781,989	168.27	299,846,698	4,185	131.25	549,305	116,289	123.28	14,336,526	25,268	73.32	1,852,771	316,585,299
Perak	1,007,343	189.68	191,077,300	41,088	87.27	3,585,867	n.a	n.a	n.a	3,207	69.79	223,840	194,887,008
Perlis	11,294	106.42	1,201,931	18	72.61	1,338	n.a	n.a	n.a	n.a	n.a	n.a	1,203,270
Pulau Pinang	8,460	190.97	1,615,736	695	94.79	65,896	n.a	n.a	n.a	n.a	n.a	n.a	1,681,632
Selangor	140,036	204.25	28,602,907	18,329	89.29	1,636,696	86,226	105.71	9,114,966	6,311	72.82	459,531	39,814,100
Terengganu	649,541	186.74	121,296,366	3,114	95.10	545,323	31,743	103.22	3,276,418			-	124,868,965
Federal Territory Kuala Lumpur & Putrajaya	425	137.33	58,386	n.a	n.a	n.a	n.a	n.a	n.a	n.a	n.a	n.a	58,386
Sabah	4,273,536	171.45	732,678,679	318,553	112.68	35,893,753	97,276	83.53	8,125,079	n.a*	n.a	n.a	776,697,511
Sarawak	7,328,029	165.79	1,214,899,320	127,460	86.84	11,068,877	320,207	114.09	36,532,736	n.a*	n.a	n.a	1,262,500,934
Total	16,822,403	-	2,869,309,750	547,564	-	56,861,181	655,422	-	71,893,119	37,014	-	2,750,588	3,000,814,638

n.a = Not available, which is not exist in certain states

n.a* = Insufficient information available. Forest plantations in Sabah and Sarawak are included in dry inland forest.

Table 10: Summary of AGC in all states in selected area, representing various conditions and types of forests in Malaysia.

Ref.	Name	State	Forest Function	Forest Type	Area (ha)	AGC (Mg C ha^{-1})				Total AGC (Mil. Mg C)
						Min	Mean	Max	Std. Dev	
A	Greater National Park	Pahang, Terengganu & Kelantan	Totally Protected Area	Inland forest	457,673.99	2.81	203.61	469.42	47.95	93.19
B	Maliau Basin	Sabah	Conservation Area		64,362.33	2.79	215.72	455.81	41.58	13.88
C	Pulong Tau	Sarawak	National Park		77,995.24	2.80	229.10	432.28	43.39	17.87
D	Jengai & Sungai Nipah Forest Reserve	Terengganu	Timber production area		78,475.49	5.54	192.55	321.37	42.06	15.11
E	Sungai Pulai	Johor	Protection and Production areas; Ramsar Site	Mangrove forest	7,320.47	0.00	82.62	365.54	67.05	0.60
F	Menumbok	Sabah	Totally Protected Area, Wildlife Conservation Area		18,370.02	0.00	102.64	336.59	82.24	1.89
G	Kuching Wetlands	Sarawak	Totally Protected Area, Ramsar Site		5,949.55	0.00	111.92	313.76	90.67	0.67
H	Matang Mangrove Reserves	Perak	Production area with some Protection Forest		38,231.83	0.00	86.69	328.20	63.13	3.31
I	Pekan Forest Reserve	Pahang	Protected area	Peat swamp forest	55,263.83	2.13	146.19	364.71	61.28	8.08
J	Loagan Bunut	Sarawak	National Park		11,785.40	2.35	122.76	341.32	74.51	1.45
K	Maludam	Sarawak	National Park		44,367.18	15.05	148.02	319.96	47.59	6.57
L	Klias	Sabah	Totally Protected Area, Wildlife Conservation Area		3,977.90	10.86	177.15	297.47	37.87	0.70

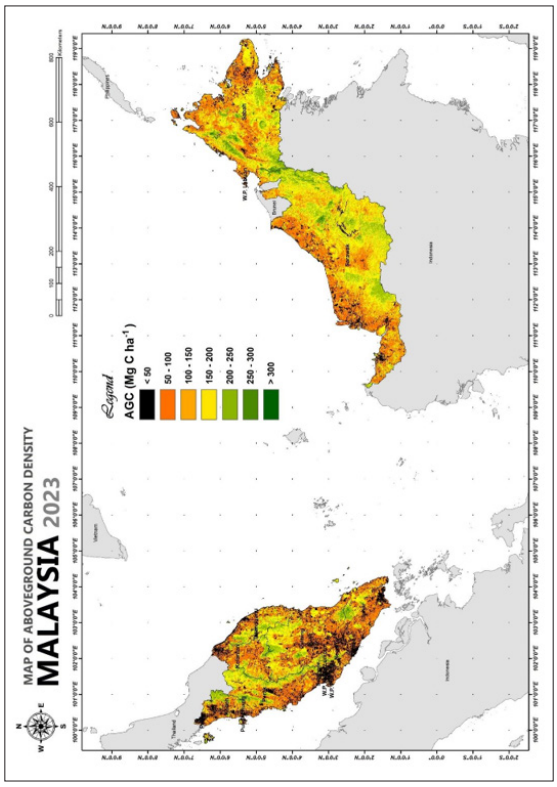


Figure 15: Map showing spatial distribution of AGC over Malaysia for the year 2023.

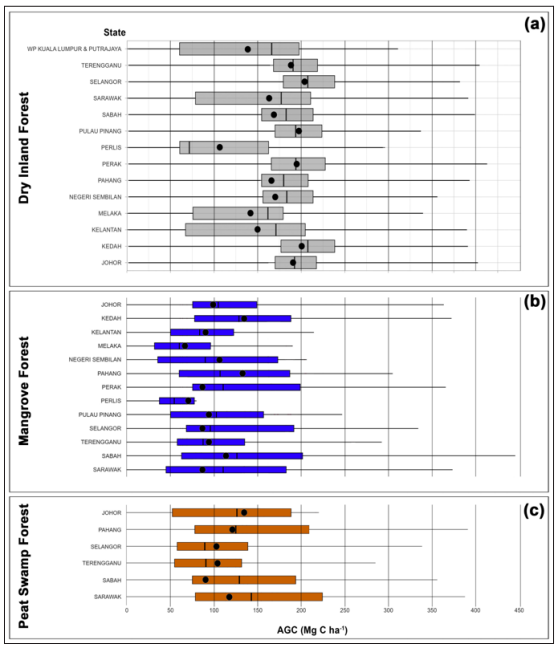


Figure 16: Summary of AGC in (a) dry inland forest, (b) mangrove forest, and (c) peat swamp forest within all states in Malaysia.

Conclusions

Landsat data have proven to be a valuable resource for forest biomass prediction, offering insights into forest ecosystems and their response to environmental changes. The combination of Landsat data with advanced modelling techniques, the use of cloud-based platforms such as GEE and other advanced technologies has enhanced the ability to estimate biomass accurately. Further research is needed to address challenges, refine methodologies, and improve the accuracy of forest biomass predictions using Landsat data.

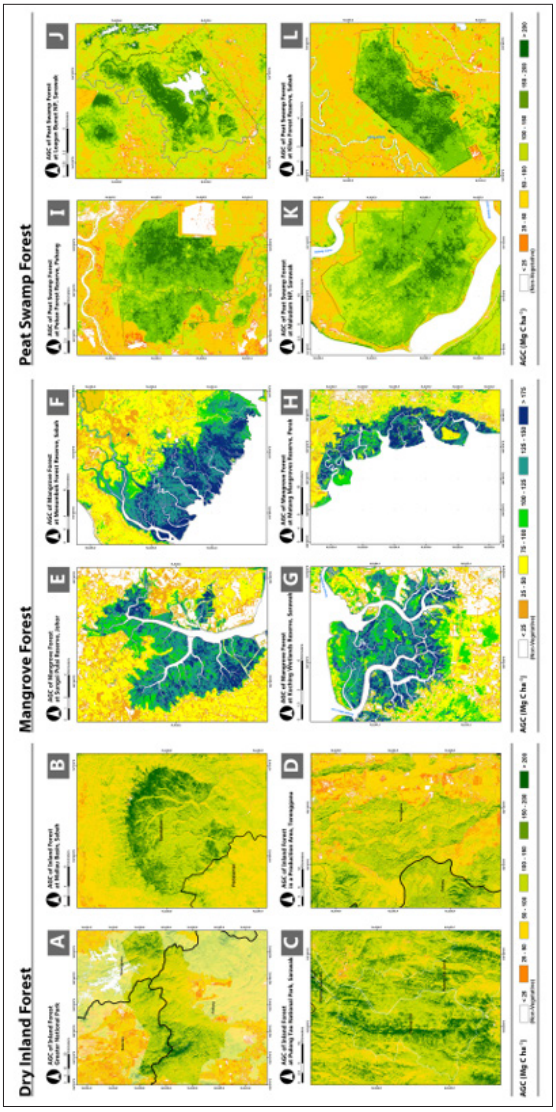


Figure 17: Maps showing spatial distribution of AGC over the selected landscapes of dry inland forest (A-D), peat swamp forest (E-H), and mangrove forest (I-L).

The AGC map will be useful in the development of baseline for carbon-related, nature-based solution approaches. The scrutiny against carbon project in the international voluntary markets, in recent years, demand for more accuracy and rigorous assessment of data to (i) support evidence of additionality through documented forest loss or degradation; (ii) support robustness and quantification of GHG emission where data is use to estimates the deforestation or degradation rates at project, subnational and national level; (iii) assess non-permanence risks including site susceptibility to natural hazards; and (iv) support evidence of co-benefits.

This study can be expanded for generation of time-series assessment over at least a 2-year interval [72,73]. This data will also facilitate the subsequent carbon verification process and ensures the validity and accountability of emissions data, the success of emissions reduction projects, confirming that the

emissions reductions are permanent and genuine. The reported results can be used for the national/subnational mitigation efforts including the REDD+ implementation. REDD+ is constructed on the principles of additionality against a baseline to ensure no leakages and avoid double counting. The generation of subnational/jurisdictional level FREL will enable the Government to develop more effective mitigation measures in achieving the Malaysian Nationally Determined Contribution (NDC) and offer the potential to scale up emissions reductions more rapidly with greater environmental integrity. More than 73 countries have implemented their carbon pricing instrument (CPI) in emission trading scheme and/or carbon tax as a means of bringing down emissions and driving investment into cleaner options [74]. The foundation of how allocation is determined under these instruments are based on historical intensity of emission from the targeted sectors. This study can be used as a basis to determine allocation for the forestry sector, if CPI is implemented in Malaysia.

Although the study provided a comprehensive map of AGC for the entire Malaysia, there are some limitations that are foreseen to have potentially be addressed in the future. Spatial resolution of Landsat data, which currently offers at 30-m resolution images can affect the accuracy of biomass predictions, particularly in heterogeneous landscapes. Integration with other data sources e.g., LiDAR and SAR, can improve the accuracy of biomass predictions. Continuous calibration and validation of biomass prediction models are also crucial to ensure their accuracy and reliability and these processes are expected to become a requirement in the future, especially when dealing with carbon projects at a state- or project-level.

In conclusion, the availability of comprehensive inventory data is instrumental in unveiling the intricate correlation patterns between aboveground carbon levels and the image variables extracted from Landsat data [75]. This symbiotic relationship between ground-based measurements and remote sensing imagery enables better comprehension of the dynamics of terrestrial carbon sequestration. With a wealth of inventory data at the disposal, more holistic understanding is gained of how various ecological and environmental factors influence aboveground carbon stocks. This knowledge not only enriches our understanding of our planet's carbon balance but also empowers us to make informed decisions for sustainable land management and climate change mitigation.

Appendix A: The Sampling Plot Design

Appendix A summarises the designs of the sampling plots that were adopted for sampling work in the dry inland, peat swamp and mangrove forests.

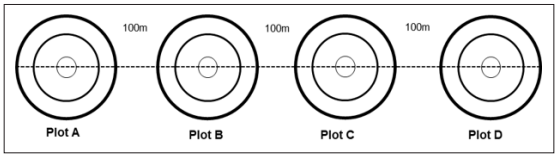


Figure 18: Layout of a cluster for dry inland forest.

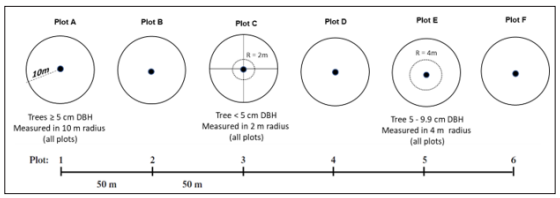


Figure 19: Layout of a cluster for peat swamp forest.

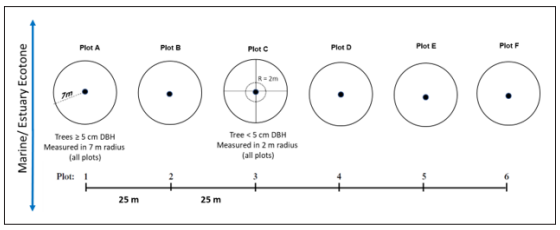
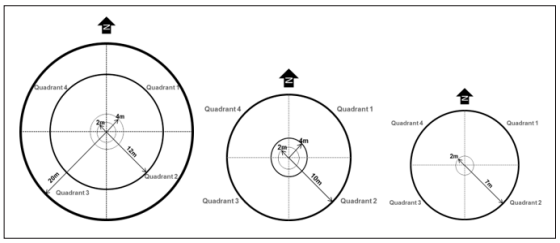


Figure 20: Layout of a cluster for mangrove forest.



*All figures are not true to scale.

Figure 21: Layout of a sampling plot for inland forest (left), peat swamp forest (centre), and mangrove forest (right).

Table 11: Summary living trees measurement in a sampling plot.

Forest Type	Nest radius (m)	Size	Tree size
All forest types	2	Sapling	< 5 cm & ≥ 1.3 m in height
Dry Inland Forest	4	Small	5 – 14.9 cm, dbh
	12	Medium	15 – 29.9 cm, dbh
	20	Large	≥ 30 cm, dbh
Peat Swamp Forest	4	Small - Medium	5 – 9.9 cm, dbh
	10	Large	≥ 10 cm, dbh
Mangrove Forest	7	Small - Large	≥ 5 cm, dbh

Appendix B: Number of Sampling Plots

Table 11 lists the locations and the respective number of sampling plots that have been collected between years 2011 and 2023. Data from these sampling plots we used in this study for the construction of AGC estimation model.

Table 12: Summarised number of sample plots data collected from years 2011 to 2023

State	Year												
	2011	2012	2013	2014	2016	2017	2018	2019	2020	2021	2022	2023	Total
Peninsular Malaysia													
Johor	4	0	3	2	0	0	0	0	0	0	54	0	63
Kedah	0	0	0	0	0	0	0	0	0	10	39	0	49
Kelantan	18	0	0	4	0	0	0	0	0	0	0	0	22
Melaka	0	0	0	0	0	0	0	0	0	12	0	0	12
Negeri Sembilan	58	4	12	0	40	0	0	0	0	0	15	0	129
Pahang	76	25	0	41	0	0	0	0	0	0	0	4	146
Perak	144	0	0	0	0	46	0	0	0	0	4	0	194
Perlis	10	0	0	0	0	0	0	0	0	0	0	0	10
Selangor	71	21	0	0	36	0	0	0	0	0	0	0	128
Terengganu	27	6	51	0	0	0	0	0	0	0	47	8	139
Sub-Total													892
East Malaysia													
Sabah	0	0	0	0	0	0	0	0	632	1380	1886	460	4358
Sarawak	0	0	0	0	0	127	275	67	126	0	0	0	595
Sub-Total													4953
Total													5845

Disclosures

All authors have no relevant financial interests in the manuscript and no other potential conflicts of interest to disclose.

Code, Data, and Materials Availability

Data sharing is not applicable to this article, as no new data were created or analysed.

Acknowledgments

Thanks to: The Ministry of Natural Resources, Environment and Climate Change (NRECC), 10th (2011-2015), 11th (2016-2020) and 12th (2021-2025) Malaysia Plans, Forest Research Institute Malaysia (FRIM), Mangrove's Technical, Research and Development Committee (JTRD), Forestry Department Peninsular Malaysia, States Forestry Department, Sabah Forestry Departments, Forest Department of Sarawak, Kumpulan Pengurusan Kayu Kayan Terengganu (KPKKT), Forestry and Forest Products Research Institute of Japan (FFPRI), International Tropical Timber Organization, Thematic Program Reducing Deforestation and Forest Degradation and Enhancing Environmental Services in Tropical Forests (ITTO-REDDES), WWF-Malaysia, and Malaysia Forest Fund (MFF). Thanks also to the USGS (https://earthexplorer.usgs.gov) that provides free-access Landsat images for this study.

References

1. Henson IE. An Assessment of Changes in Biomass Carbon Stocks in Tree Crops and Forests in Malaysia. J Trop for Sci. 2005. 17: 279-296.
2. Moktshim N. Forest management in Malaysia: The strategies undertaken towards achieving Sustainable Development Goals. IOP Conf. Ser.: Earth Environ. Sci. 2020. 561: 012041.
3. Ministry of Natural Resources, Environment and Climate Change Malaysia (NRECC). 2023.
4. Houghton RA, Hall F, Goetz SJ. Importance of biomass in the global carbon cycle. J. Geophys. Res. Biogeosciences. 2009. 114: 1-13.
5. Guillén F, Orozco R, Santaella JA. Measuring Climate Change: The importance of geospatial information with an application to carbon sequestration and storage in the System of Environmental-Economic Accounting-Ecosystem Accounting (SEEA EA)-9th IMF Statistical Forum, United Nations, Rome. 2021.
6. U.S. Geological Survey. Landsat-Earth Observation Satellites. Fact Sheet; U.S. Geological Survey: Reston, VA, USA. 2015. 4.
7. Wulder MA, Masek JG, Cohen WB, Loveland TR, Woodcock CE. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. Remote Sens. Environ. 2012. 122: 2-10.
8. Potapov P, Hansen MC, Kommareddy I, Kommareddy A, Turubanova S, et al. Landsat Analysis Ready Data for Global Land Cover and Land Cover Change Mapping. Remote Sens. 2020. 12: 426.
9. Lu D. The potential and challenge of remote sensing-based biomass estimation, Int. J. Remote Sens. 2006. 27: 1297-1328.
10. Hamdan O, Thirupathi RN, Norsheilla MJC, Nur Atikah AB, Muhamad Afizzul M. Utilization of Remote Sensing Technology for Carbon Offset Identification in Malaysian Forests. IntechOpen, 2021.
11. Shao Z, Zhang L. Estimating Forest Aboveground Biomass by Combining Optical and SAR Data: A Case Study in Genhe, Inner Mongolia, China. Sensors. 2016. 16: 834.

12. Li X, Zhang M, Long J, Lin H. Novel Method for Estimating Spatial Distribution of Forest Above-Ground Biomass Based on Multispectral Fusion Data and Ensemble Learning Algorithm. Remote Sens. 2021. 13: 3910.
13. Puliti S, Breidenbach J, Schumacher J, Hauglin M, Klingenberg TF, et al. Above-ground biomass change estimation using national forest inventory data with Sentinel-2 and Landsat. Remote sensing of environment. 2021. 265: 112644.
14. Lourenço P. Biomass Estimation Using Satellite-Based Data, Forest Biomass - From Trees to Energy. 2021.
15. Foody GM, Boyd DS, Cutler ME. Cutler. Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions. Remote Sens. Environ. 2003. 85: 463-474.
16. Tavasoli N, Arefi H. Comparison of Capability of SAR and Optical Data in Mapping Forest above Ground Biomass Based on Machine Learning. 2021. 5: 13.
17. Breiman L. Random Forest. Mach. Learn. 2001. 45: 5-32.
18. Zhang X, Li L, Liu Y, Wu Y, Tang J, et al. Improving the accuracy of forest aboveground biomass using Landsat 8 OLI images by quantile regression neural network for Pinus densata forests in southwestern China. Frontiers in Forests and Global Change. 2023. 6: 1162291.
19. Gizachew B, Solberg S, Næsset E, Gobakken T, Bollandsås OM, et al. Mapping and estimating the total living biomass and carbon in low-biomass woodlands using Landsat 8 CDR data. Carbon balance and management. 2016. 11: 13.
20. Shao Z, Zhang L. Estimating Forest Aboveground Biomass by Combining Optical and SAR Data: A Case Study in Genhe, Inner Mongolia, China. 2016. 16: 834.
21. Mette T, Papathanassiou KP, Hajnsek I, Zimmermann R. Forest biomass estimation using polarimetric SAR interferometry. InIEEE International Geoscience and Remote Sensing Symposium. 2002. 2: 817-819.
22. Purohit S, Aggarwal SP, Patel NR. Estimation of forest aboveground biomass using combination of Landsat 8 and Sentinel-1A data with random forest regression algorithm in Himalayan Foothills. Tropical Ecology. 2021. 62: 288-300.
23. Luo P, Ye H, Huang W, Liao J, Jiao Q, et al. Enabling Deep-Neural-Network-Integrated Optical and SAR Data to Estimate the Maize Leaf Area Index and Biomass with Limited in Situ Data. Remote Sensing. 2022. 14: 5624.
24. Lu D, Chen Q, Wang G, Moran E, Batistella M, et al. Aboveground Forest biomass estimation with Landsat and LiDAR data and uncertainty analysis of the estimates. International Journal of Forestry Research. 2012. 2012: 436537.
25. Li Y, Li M, Li C, Liu Z. Forest aboveground biomass estimation using Landsat 8 and Sentinel-1A data with machine learning algorithms. Scientific reports. 2020. 10: 9952.
26. Ministry of Natural Resources, Environment and Climate Change (NRECC). 2023.
27. Michinaka T. Approximating Forest Resource Dynamics in Peninsular Malaysia Using Parametric and Nonparametric Models, and Its Implications for Establishing Forest Reference (Emission) Levels under REDD+. Land. 2018. 7: 70.
28. Sato T, Niiyama K, Toriyama J, Kiyono Y. How to Estimate Forest Carbon Stocks? Application to Ground-Based Inventory. In Proceedings Workshop on REDD+ Research Project in Peninsular Malaysia, Forest Research Institute Malaysia. 2013.
29. Walker SM, Pearson TR, Casarim FM, Harris N, Petrova S, et al. Standard Operating Procedures for Terrestrial Carbon Measurement: Version. 2012.
30. IPCC. Guidelines for National Greenhouse Gas Inventories-Volume 4: Agriculture, Land Use and Forestry (GL-AFOLU). 2023.
31. Hamdan O, Valeria L, Muhamad Afizzul M. Guide to the Development of Forest Resources Inventory of Sabah. In FRIM Technical Handbook No. 52. Forest Research Institute Malaysia. 2021.
32. Hamdan O, Muhamad Afizzul M. Manual Kerja Lapangan Survei Karbon Hutan. In FRIM Technical Information Handbook No. 59. Forest Research Institute Malaysia. 2023.
33. Reyes G, Brown S, Chapman J, Lugo AE. Eds. In Wood densities of tropical tree species. General Technical Report SO-88, New Orleans, Louisiana.1992.
34. Ashton PS. Dipterocarpaceae. Flora Malesiana 1982. 9: 237-552.
35. Dwiyono A, Rachman S. Management and conservation of tropical peat forest of Indonesia. In Proceedings of a Workshop on Integrated Planning and Management of Tropical Lowland Peatlands, Cisarua, Indonesia. 2006.
36. Rieley JO, Page SE. Eds. In Wise Use Guidelines for Tropical Peatlands. Wageningen, The Netherlands, Alterra. 2005. 237.
37. Kauffman JB, Arifanti VB, Basuki I, Kurnianto S, Novita N, et al. Eds. In Protocols for the measurement, monitoring, and reporting of structure, biomass, carbon stocks and greenhouse gas emissions in tropical peat swamp forests. CIFOR, Bogor, Indonesia. 2016. 221.
38. Mitsch WJ, Gosselink JG. Eds, Wetlands (Fourth edition). John Wiley and Sons, Inc., New York, USA. 2007. 582.
39. Kauffman JB, Donato DC. Eds. In Protocols for the measurement, monitoring and reporting of structure, biomass and carbon stocks in mangrove forests. CIFOR, Bogor, Indonesia. 2012. 86.
40. Chave J. Improved allometric models to estimate the aboveground biomass of tropical trees. Glob Change Biol. 2014. 12629: 14.
41. Symington CF. Foresters' Manual of Dipterocarps. Malayan Forest Records No.16, Penerbit Universiti Malaya, Kuala Lumpur. 1943. 244.
42. Brown S. Measuring carbon in forests: current status and future challenges. Environ. Pollut 2002. 116: 363-372.
43. Rikimaru A, Roy PS, Miyatake S. Tropical Forest cover density mapping. Trop Ecol. 2002. 43: 39-47.
44. Azizi Z, Najafi A, Sohrabi H. Forest Canopy Density estimating using satellite images, In the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2008. 1127-1130.
45. Huang S, Tang L, Hupy JP, Wang Y, Shao G. A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. Journal of forestry research. 2021. 32: 1-6.
46. Li C, Li M, Li Y. Improving estimation of forest aboveground biomass using Landsat 8 imagery by incorporating forest crown density as a dummy variable. Can J For Res. 2020. 50: 390-398.
47. Arisanty D, Saputra AN, Rahman AM, Hastuti KP, Rosadi D. The Estimation of Iron Oxide Content in Soil based on Landsat 8 OLI TIRS Imagery in Wetland Areas. Pertanika Journal of Science & Technology. 2021. 29.

48. Nathalie P. NDVI from A to Z, The Normalized Difference Vegetation Index (Oxford, 2013; online ed, Oxford Academic, 8 May 2015).
49. García MJL, Caselles V. Mapping burns and natural reforestation using thematic mapper data. Geocarto. 1991. 6: 31-37.
50. Zhu Z, Woodcock CE. Object-based cloud and cloud shadow detection in Landsat imagery. Remote sensing of environment. 2012. 118: 83-94.
51. Huete AR. A soil-adjusted vegetation index (SAVI). Remote sensing of environment. 1988. 25: 295-309.
52. Liu JG, Mason PJ. Essential image processing and GIS for remote sensing. John Wiley & Sons. 2013.
53. Xu H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. International journal of remote sensing. 2006. 27: 3025-3033.
54. Huete A, Didan K, Miura T, Rodriguez EP, Gao X, et al. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote sensing of environment. 2002. 83: 195-213.
55. Tofallis C. A Better Measure of Relative Prediction Accuracy for Model Selection and Model Estimation. J Operational Research Society. 2015. 66: 1352-1362.
56. Baccini AG, Goetz SJ, Walker WS, Laporte NT, Sun M, et al. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. Nature climate change. 2012. 2: 182-185.
57. Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SA, et al. High-resolution global maps of 21st-century forest cover change. science. 2013. 342: 850-853.
58. Xie S, Liu L, Zhang X, Yang J, Chen X, et al. Automatic land-cover mapping using landsat time-series data based on google earth engine. Remote sensing. 2019. 11: 3023.
59. Tran TV, Reef R, Zhu X. A review of spectral indices for mangrove remote sensing. Remote Sensing. 2022. 14: 4868.
60. Hamdan O, Khairunnisa MR, Ammar AA, Hasmadi IM, Aziz HK. Mangrove carbon stock assessment by optical satellite imagery. Journal of Tropical Forest Science. 2013. 1: 554-565.
61. Rannestad M, Eid T, Bollandsås OM, Gobakken T, Tetemke B. Aboveground Biomass Prediction Model Using Landsat 8 Data: A Test on Possible Approaches for Seasonally Dry Forests of Northern Ethiopia. InConference of the Arabian Journal of Geosciences. 2019. 383-386.
62. Kho LK, Jepsen MR. Carbon stock of oil palm plantations and tropical forests in M alaysia: a review. Singapore Journal of Tropical Geography. 2015. 36: 249-266.
63. Hamdan O, Aziz HK, Hasmadi IM. L-band ALOS PALSAR for biomass estimation of Matang Mangroves, Malaysia. Remote Sensing of Environment. 2014. 155: 69-78.
64. Raihan A, Begum RA, Mohd Said MN, Pereira JJ. Assessment of carbon stock in forest biomass and emission reduction potential in Malaysia. Forests. 2021. 12: 1294.
65. Hamdan O, Hasmadi IM, Aziz HK, Norizah K, Zulhaidi MH. L-band saturation level for aboveground biomass of dipterocarp forests in peninsular Malaysia. Journal of Tropical Forest Science. 2015. 1: 388-399.
66. Omar H, Misman MA, Kassim AR. Synergetic of PALSAR-2 and Sentinel-1A SAR polarimetry for retrieving aboveground biomass in dipterocarp forest of Malaysia. Applied Sciences. 2017. 7: 675.
67. Hamdan Omar HO, Muhamad Afizzul Misman MA. Time-series maps of aboveground biomass in dipterocarps forests of Malaysia from PALSAR and PALSAR-2 polarimetric data. 2018.
68. Asner GP, Brodrick PG, Philipson C, Vaughn NR, Martin RE, et al. Mapped aboveground carbon stocks to advance forest conservation and recovery in Malaysian Borneo. Biological Conservation. 2018. 217: 289-310.
69. Matthew NK, Shuib A, Muhammad IS, Eusop ME, Ramachandran S, et al. Carbon stock and sequestration valuation in a mixed dipterocarp forest of Malaysia. Sains Malaysiana. 2018. 47: 447-455.
70. Hamdan O, Mohd Hasmadi I, Khali Aziz H, Helmi Zulhaidi MA, Norizah K. Estimating Biomass in Logged Tropical Forest Using L-Band SAR (PALSAR) Data and GIS. Sains Malaysiana. 2015. 44: 1085-1093.
71. Kueh RJ, Majid NM, Ahmed OH, Gandaseca S. Assessment of carbon stock in chronosequence rehabilitated tropical forest stands in Malaysia. Journal of forest and environmental science. 2016. 32: 302-310.
72. Nguyen TH, Jones SD, Soto-Berelov M, Haywood A, Hislop S. Monitoring aboveground forest biomass dynamics over three decades using Landsat time-series and single-date inventory data. International Journal of Applied Earth Observation and Geoinformation. 2020. 84: 101952.
73. Nguyen HT, Jones S, Soto-Berelov M, Haywood A, Hislop S. Landsat time-series for estimating forest aboveground biomass and its dynamics across space and time: A review. Remote Sensing. 2019. 12: 98.
74. Weltbank. State and trends of carbon pricing. World Bank. 2023.
75. Shobairi SO, Usoltsev VA, Chasovskikh VP, LI M. Exploring Forest aboveground biomass estimation using landsat, forest inventory and analysis data base. Analysis. 2018. 4: 632-641.

Copyright: © 2025 Hamdan Omar, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.