

Article

Not peer-reviewed version

---

# Establishing Models for Predicting Above Ground Carbon Stock Based on Sentinel 2 Imagery for Evergreen Broadleaf Forests in South Central Coastal Ecoregion, Vietnam

---

Tam Huu Nguyen , [Loi Van Nguyen](#) , [Tuan Huy Hoang](#) \*

Posted Date: 27 February 2025

doi: 10.20944/preprints202502.2136.v1

Keywords: above ground carbon; cross validation; principal component analysis; sentinel 2; Vietnam



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

*Article*

# Establishing Models for Predicting Above Ground Carbon Stock Based on Sentinel 2 Imagery for Evergreen Broadleaf Forests in South Central Coastal Ecoregion, Vietnam

Nguyen Huu Tam <sup>1,2</sup>, Nguyen Van Loi <sup>1</sup> and Hoang Huy Tuan <sup>1,\*</sup>

<sup>1</sup> University of Agriculture and Forestry, Hue University, 102 Phung Hung Street, Hue City, Vietnam; tamnh.dhnl22@hueuni.edu.vn (N.H. Tam), nvanloi@hueuni.edu.vn (N.V. Loi), hhtuan@hueuni.edu.vn (H.H. Tuan)

<sup>2</sup> Department of Agriculture and Environment of Da Nang City, 24 Tran Phu Street, Da Nang City, Vietnam

\* Correspondence: hhtuan@hueuni.edu.vn

**Abstract:** In Vietnam, the models for estimating above ground biomass (AGB) for converting to carbon stocks prediction mostly based on diameter at breast height (DBH), tree height (H), wood density (WD) meanwhile the remote sensing application has considered as suitable method since improving accuracy and reducing cost. With this context, this study was conducted with aim to develop correlation equations among total above ground carbon (TAGC) and indices of Sentinel 2 images to directly predict carbon stock for assessing carbon emission and removal. In this study, remote sensing indices great influencing TAGC were determined by principal component analysis (PCA) and forest inventory factors from 115 sample plot was used to calculate the TAGC. Regression models were established by Ordinary Least Squares and Maximum Likelihood methods and validated by Monte Carlo cross-validation method. The study found out that NDVI, SAVI, NIR and three variable combination (NAVI, ARVI), (SAVI, SIPI), (NIR, EVI) have strongly influenced on TAGC. Total 36 linear and non-linear with weight models basing on above selected variables were established, in which quadratic models used NIR and variable combination (NIR, EVI) with AIC of 756.924, 752.493,  $R^2$  value of 0.86, 0.87 and MPSE of 22,04%, 21,63% respectively, were found as optimal models. Therefore, the study these models have recommended for predicting carbon stocks for Evergreen Broadleaf Forests in South Central Coastal Ecoregion, Vietnam

**Keywords:** above ground carbon; cross validation; principal component analysis; sentinel 2; Vietnam

## 1. Introduction

Many vegetation indices play a critical role in estimating biomass and monitoring plant growth, yet each index employs a distinct approach and has certain limitations. Soil Adjusted Vegetation Index (SAVI) was developed to mitigate the influence of soil conditions, thereby enhancing accuracy in areas with sparse vegetation, though it is less effective in regions with dense vegetation [1]. Chlorophyll Vegetation Index (CVI) measures chlorophyll concentration and is highly useful for assessing plant health and development; however, it may be limited in areas with low chlorophyll levels [2]. Green Leaf Index (GLI) reflects the greenness of leaves and is particularly accurate during periods of rapid growth, but its effectiveness diminishes in areas with dense vegetation cover [3]. Normalized Difference Vegetation Index (NDVI) stands out for its stability when applied on a large scale and its suitability for regions with abundant vegetation. It is calculated based on the difference between near-infrared and red light reflectance, which helps measure the level of photosynthetic activity and vegetation growth, thereby effectively estimating biomass and stored carbon [4]. Furthermore, NDVI provides precise information regarding vegetation density and health, key

factors in assessing carbon stocks [5]. Additionally, NDVI has stability over time, making it particularly advantageous for use in areas with dense and abundant vegetation, thus facilitating effective large-scale assessments [6]. Therefore, some previous studies have developed regression equations for estimating above ground biomass (AGB) of natural forests based on vegetation indices from Landsat, Sentinel imageries, such as NDVI, EVI, SAVI [7–10].

Currently, estimation of forest carbon stock and flows have applied by various methods, from the simple forest biomass inventory to complex, sophisticated experiments and models, which are forest inventory, remote sensing, eddy covariance and the inverse method [11]. Among these methods, remote sensing technology has become the main tool to overcome some limitations of ground data collection from field sample plots in forest monitoring and inventory at a landscape scale by improving accuracy and reducing costs [12–14]. It also has considered as the cheapest method if there are suitable resolution and spatial scale since the developing countries with limited capacities in collecting and managing data need to apply the low cost methodologies but spatial and temporal resolutions are acceptable and ratio of sample plots are appropriate [11]. Therefore, over the past decade, Sentinel-2 satellite data has been widely applied for estimating the biomass of natural forests due to its high resolution. Investigating the relationship between Sentinel-2 image indices (NDVI, EVI, NDI45, etc.) and above-ground biomass (AGB) in privately managed tropical forests in Indonesia found that the NDI45 exhibited strong correlation with AGB in comparison to other indices ( $r = 0.89$ ;  $R^2 = 0.79$ ) [10]. Moreover, Sentinel-2 imagery has been integrated with EnMAP to map and monitor environmental changes [8] or with PlanetScope to develop biomass mapping models [15]. Estimating biomass and carbon accumulation capacity of tropical rainforest in Kon Ha Nung plateau based on EVI index of Sentinel-2 images, Dang, H.N. et al. demonstrated that lin-log models established to estimate biomass from EVI index of Sentinel-2 images in 2016 and 2021 both had the highest  $R^2$  value, 0.76 and 0.765, respectively [8].

In Vietnam, the UN-REDD (Reducing Emissions from Deforestation and forest Degradation) program has developed allometric equations for each forest type in each ecological region and a generic equation for a nation by using the destructive method to estimate biomass and carbon stocks [16]. However, the destructive method is both time-consuming and costly, making it impractical for application in large, inaccessible forest areas [17]. Although the relationship between remote sensing indices and forest inventory factors has been determined, biomass is a decisive condition for the effectiveness of remote sensing in forest resource monitoring. However, it appears that very few studies have examined the correlation between biomass, carbon stocks, and remote sensing indices for monitoring forest carbon at national, regional, or global scales [18]. Meanwhile, the Sentinel 2 image has quite high resolution and suitable for analyzing the relationship of forest vegetation reflectance indices with forest biomass, furthermore the Sentinel 2 images are free of charge so it contributes to reducing costs in estimating natural forest carbon stocks. In this context, this study was conducted with aim to develop correlation equations between natural forest carbon stocks and indices of Sentinel 2 images through linear and non-linear models to directly estimate stored carbon for assessing carbon emission and removal in the South Central Coastal Ecoregion, Vietnam.

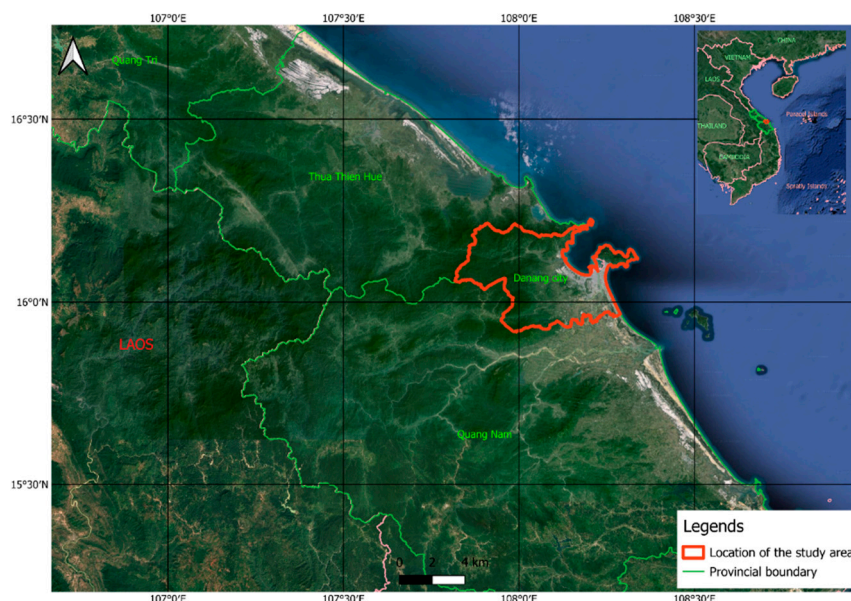
## 2. Materials and Methods

### 2.1. Study Area

This study was conducted in Da Nang city that belongs to South Central Coastal (SCC) ecoregion of Viet Nam (Figure 1), is one of eight agricultural ecoregion divided following climate, altitude and soil conditions [19,20]. Da Nang city is located in a typical monsoon climate zone with two distinct seasons: dry season from January to August and rainy season from September to December. Average annual temperature is 25.8°C, average annual precipitation is 2,153 mm, and average humidity is 83.4%. Its terrain has both coastal delta and mountains, in which the mountainous area has evaluation of 700-1,500 m and slopes over 40° with the watershed forests occupies a large area [21]. Natural forest



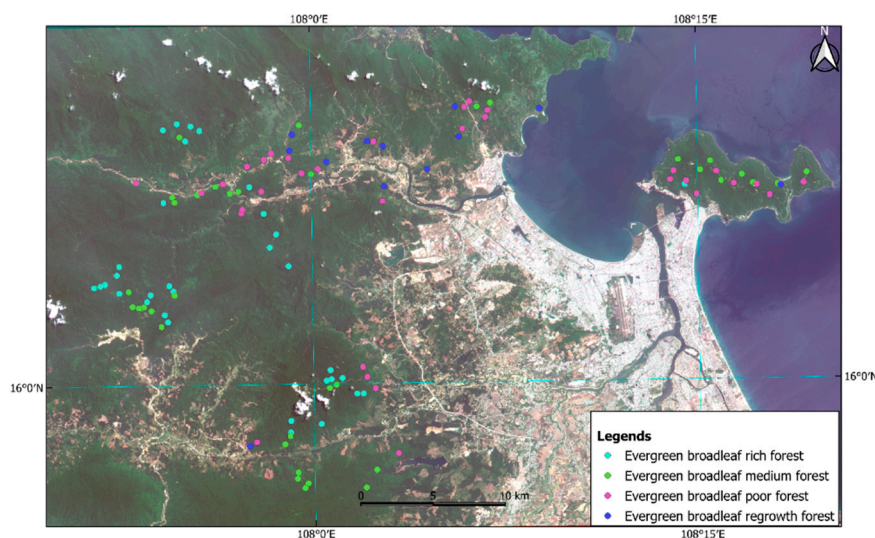
in Da Nang city is Evergreen Broadleaf Forest (EBF) type with total of 43,061.90 hectares [22], including four forest types: rich forest, medium forest, poor forest, and regrowth forest.



**Figure 1.** Location of the Da Nang city.

## 2.2. Sample Plots and Estimation of Total Above Ground Carbon

Total 115 sample plots of 1,000 m<sup>2</sup> (40 m x 25 m) were established that are representative of four forest classes, which are 33 sample plots representative of rich forest, 37 sample plots representative of medium forest, 33 sample plots representative of poor forest, and 12 sample plots representative of regrowth forest (Figure 2). All standing trees with DBH  $\geq$  6cm within sample plots were defined species name (both local and scientific names) and were measured DBH (cm), H (m), meanwhile WD were derived from the data that used for establishing biomass equations for EBF in South Central Coastal Ecoregion of Viet Nam [23] or from the ICRAF's WD data (<http://db.worldagroforestry.org/wd>) in case of these trees are outside of above species.



**Figure 2.** Location of sample plots representing forest classes in study area.

For this study, AGB of individual tree was estimated by applying the AGB equation developed by Huy et al (2016) for EBF in SCC ecoregion of Vietnam. Due to the equation with combination of two covariates (DHB<sup>2</sup>H) and of three covariates (DHB<sup>2</sup>HWD) are more appropriate than with single variable [23], this study, therefore was applied the equation below for estimating AGB

$$AGB = 0.598313 \times (DBH^2 \times HWD)^{0.959790}$$

(1)

Where: AGB is expressed in kg; DBH is expressed in centimeters cm; H is expressed in m; WD is expressed in g/cm<sup>3</sup>; and 0.598313 and 0.959790 are constants.

Total above ground biomass (TAGB) of sample plot was calculated from AGB of individual tree, then converted to per hectare (Mgha<sup>-1</sup>). The TAGC was converted from TAGB by using the carbon fraction (CF) of IPCC (2006), specific equation for calculating TAGC is [24]:

$$TAGC = TAGB \times 0.47$$

(2)

Where: TAGC is expressed in Mgha-1; TAGB is expressed in Mgha-1; and 0.47 is default value for CF.

2.3. Sentinel-2 Image and Identification of Key Indices

In Da Nang city, the dry season extends from January to August each year; therefore, remote sensing images taken during this season should be selected as they will be less affected by clouds, ensuring accuracy in the image interpretation process. This study used Sentinel-2 images with a panchromatic resolution of 10 m × 10 m and a multispectral resolution of 20 m × 20 m to interpret and calculate vegetation indices (VIs) based on four typical multispectral bands: BLUE, GREEN, RED, and NIR. The scene “S2B\_MSIL1C\_20230522T030529\_N0509\_R075,” captured on May 22, 2024 and downloaded from <https://scihub.copernicus.eu/>, was interpreted and calculated VIs related to vegetation covers.

Using VIs obtained from remote sensing analysis of foliage to quantify and qualitatively assess vegetation cover, vitality and growth is a fairly simple and effective method. However, different VIs reflect different vegetation, so each VI is often suitable for a few specific uses under certain conditions [25]. Therefore, this study has identified 5 VIs closely related to forest vegetation to establish a TAGC estimation model (Table 1)

Table 1. Vegetation Indices used for estimating TAGC.

VIs	Definition	Sources (References)
ARVI	$(NIR - (2 \times RED) + BLUE)/(NIR + (2 \times RED) + BLUE)$	[26]
EVI	$2.5 \times (NIR - RED)/(NIR + 6 \times RED - 7.5 \times BLUE + 1)$	[27]
NDVI	$(NIR - RED)/(NIR + RED)$	[2]
SAVI	$1.428 \times (NIR - RED)/(NIR + RED + 0.428)$	[1]
SIPI	$(NIR - BLUE)/(NIR - RED)$	[28]

2.4. Development of Regression Models

The VIs and multispectral bands that have a significant influence on TAGC prediction were identified using Principal Component Analysis (PCA). Total 10 variables were conducted through PCA with package ‘ggplot2’, ‘factoextra’, ‘dplyr’, ‘ggfortify’, ‘pracma’ [29]. In which, TAGC that was calculated from 115 sample plots, and 5 VIs (ARVI, EVI, NDVI, SAVI, SIPI) and 4 multispectral bands (BLUE, GREEN, RED, NIR) that were identified from Sentinel-2 image “S2B\_MSIL1C\_20230522T030529\_N0509\_R075”. Based on the PCA results, the first two principal components (PC) are selected to identify the VIs and multispectral bands that have a significant influence on TAGC if the overall variability of the dataset reaches 80% or more; if not, additional PCs

will be included until this threshold is met, since PCA provides an approximation of the original dataset by using some PCs [30]

PC1 is used to determine the weight of each variable, and this is combined with the component weight relationship graph between PC1 and PC2 to select those VIs and multispectral bands with high absolute weight values and a significant influence on TAGC. Based on these findings, the linear and non-linear with weight models (both single variable and variable combination) will be determined. In addition to applying powers models as many previous studies have applied to establish correlation equations to estimate AGB based on forest inventory factors such as DHB, H, WD, etc. [23] or based on VIs [7–10], exponential and quadratic models are also established to select the fit models.

## 2.5. Cross Validation

Correlation models were established and validated for model comparison and selection using Monte Carlo cross-validation method [23,31,32]. The dataset was randomly divided into two parts for each iteration, in which 80% of the sample was used for model establishment and the remaining 20% was used for model validation. The cross-validation process was repeated 100 times, the statistics and errors of the model were averaged for 100 iterations.

Linear models were constructed and cross validated using the ordinary least squares method with the 'lm' package [29]. The weighting scheme was defined as  $W = 1/X^\alpha$ , where  $X$  represents the vegetation indices (VIs) and multispectral bands that significantly influence TAGC, and  $\alpha = \pm 2$  [33]. Non-linear models were developed and cross validated using the Maximum Likelihood method via the 'nlme' package [29,31], with the weighting scheme defined as  $W = 1/X^\delta$ , where  $\delta$  is the variance function coefficient [23,31,32,34].

The Akaike Information Criterion (AIC) developed by [35] and applied by Huy et al. [23,31,32,34] was used to validate, compare and select fit models. In addition, the  $R^2$  and the main errors such as Average Systematic Error (ASE), Root Mean Square Error (RMSE), and Mean Percent Standard Error (MPSE) [23,31,36] are also used along with the AIC for model selection. AIC,  $R^2$ , ASE, RMSE, and MPSE are calculated from 20% of the dataset used to validate the model for each iteration and are averaged over 100 iterations.

$$R^2 = \frac{1}{R} \sum_1^R \left( 1 - \frac{\sum_{j=1}^m (Y_j - \hat{Y}_j)^2}{\sum_{j=1}^m (Y_j - \bar{Y})^2} \right) \quad (3)$$

$$ASE (\%) = \frac{1}{R} \sum_1^R \frac{100}{m} \sum_{j=1}^m \frac{Y_j - \hat{Y}_j}{\hat{Y}_j} \quad (4)$$

$$RMSE (Mg ha^{-1}) = \frac{1}{R} \sum_1^R \sqrt{\frac{1}{m} \sum_{j=1}^m (Y_j - \hat{Y}_j)^2} \quad (5)$$

$$MPSE (\%) = \frac{1}{R} \sum_1^R \frac{100}{m} \sum_{j=1}^m \left| \frac{Y_j - \hat{Y}_j}{\hat{Y}_j} \right| \quad (6)$$

Where  $R$  is the number of iterations (100) in cross-validation;  $m$  is the number of sample plots in the validation dataset (20% of randomly selected data); and  $Y_j$ ,  $\hat{Y}_j$  and  $\bar{Y}$  represent the TAGC ( $Mg ha^{-1}$ ) of the observed, predicted, and mean values, respectively, for the  $j$ -th sample plot during the  $R$  iterations of cross-validation.

The selected models are those with the best indices (i.e., the lowest AIC, the highest  $R^2$ , and the smallest errors in ASE, RMSE, and MPSE). However, in practice, it is very challenging for a model to optimize all five indices simultaneously. Therefore, models with an AIC lower than the average AIC of all models and an  $R^2$  greater than 0.85 (with higher values being preferable) are chosen for further parameter identification and validation.

The statistical software R [29] is employed using 100% of the data and 100 iterations to determine and verify the existence of the model parameters. Models that exist all parameters (with p-value < 0.05) and have an MPSE of less than 30% are considered acceptable.

3. Results

3.1. Vegetation Indices and Multispectral Bands Influencing on Above Ground Carbon

From the PCA results of 10 variables from the data of 115 sample plots (Figure 3), the first two PCs were selected to analyze and determine the VIs and multispectral bands that have a large influence on TAGC to establish correlation models, because these two PCs account for 96.35% of the variation of the original dataset.

Importance of components:										
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Standard deviation	2.864	1.1976	0.45214	0.33361	0.2024	0.09233	0.01477	0.005087	0.002567	0.001149
Proportion of Variance	0.820	0.1434	0.02044	0.01113	0.0041	0.00085	0.00002	0.000000	0.000000	0.000000
Cumulative Proportion	0.820	0.9635	0.98390	0.99503	0.9991	0.99997	1.00000	1.000000	1.000000	1.000000

Figure 3. Result summary of PCA from 115 dataset of 10 variables.

Since PC1 accounts for the largest share of variance in the original dataset (82%), it was used to determine the weight of each of the 10 variables (Figure 4) and to analyze the relationship matrix of these variables in the context of the correlation between PC1 and PC2 (Figure 5).

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
TAGC	0.3243373	-0.17075275	0.090292535	-0.917404924	0.05011524	-0.115448857	0.003170630	-0.003704769	-0.0001262039	0.0007073273
NDVI	-0.3351920	0.22878575	-0.057174241	-0.148897859	-0.06333439	-0.169613042	-0.441165068	-0.315237907	-0.6856373927	0.1139469269
EVI	-0.3443245	0.07620415	-0.277828821	-0.163619214	-0.12205892	-0.016265880	0.548890422	-0.626225824	0.1721095817	-0.1828916489
SAVI	-0.3420016	0.16060869	-0.006559612	-0.176027982	-0.05987280	0.167999012	-0.178914963	0.014719087	0.4556482813	0.7446877826
ARVI	-0.3329627	0.23548760	-0.188396941	-0.186915805	-0.06149948	0.003411275	-0.446058115	0.243754815	0.3779666440	-0.5985095426
SIPI	-0.3270426	0.18088792	0.596264678	-0.017311043	-0.05533496	-0.606273110	0.280821583	0.230471730	0.0398576836	-0.0171981440
NIR	-0.3469365	0.05102502	0.080904407	-0.203819754	-0.05471612	0.604781278	0.359617881	0.441306274	-0.3643421208	-0.0580331384
RED	-0.2269835	-0.60042064	0.531802434	0.039599324	-0.21411055	0.260753858	-0.250681447	-0.310775837	0.0926152208	-0.1474967594
GREEN	-0.3204391	-0.29405600	-0.074317802	-0.008242110	0.89464527	-0.069698105	-0.002242768	0.001590014	0.0017754437	-0.0003498977
BLUE	-0.2335349	-0.59088584	-0.474804192	0.004981321	-0.34485157	-0.351034493	0.045521747	0.323058660	-0.0830117713	0.1239904768

Figure 4. Weight of variability by principal components .

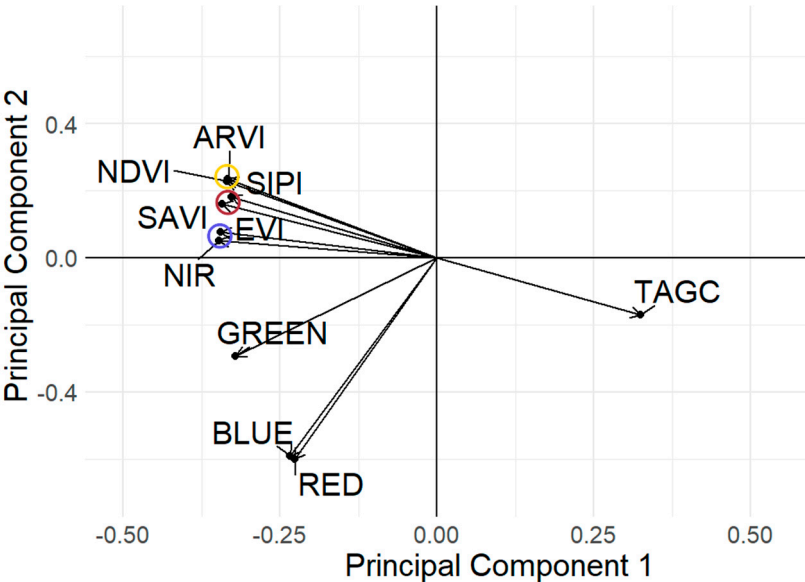


Figure 5. Biplot of variable weights of principal component 1 versus principal component 2.

The equation indicating weight of variables following PC1 is presented as follows:



$$\begin{aligned} \text{PC1} = & 0.3243373 \times \text{TAGC} - 0.3351920 \times \text{NDVI} - 0.3443245 \times \text{EVI} \\ & - 0.3420016 \times \text{SAVI} - 0.3329627 \times \text{ARVI} - 0.3270426 \times \text{SIPI} - \\ & 0.3469365 \times \text{NIR} - 0.2269835 \times \text{RED} - 0.3204391 \times \text{GREEN} - \\ & 0.2335349 \times \text{BLUE}. \end{aligned}$$

(7)

The variables BLUE, RED and GREEN have small weights and are separate, with unclear effects on TAGC (Figure 5), so these variables are excluded from establishing correlation models. TAGC has a close and inverse relationship with the variables NAVI, EVI, SAVI, ARVI, SIPI, NIR; and these variables have a close relationship with each other in pairs, so they form variable groups in the regression model with TAGC, which are the groups (NAVI, ARVI), (SAVI, SIPI), (NIR, EVI), respectively represented in the yellow, red, and blue circles in Figure 5. In addition, in each of the above variable combinations, the variables NAVI, SAVI, NIR have higher weights than the remaining variable (Figure 4), so they are also selected to establish a regression model with TAGC in single variable form.

3.2. Establishment of Above Ground Carbon Estimation Models

The PCA results selected three single variables and three variable combinations that have a significant influence on TAGC. An analysis of TAGC variations with respect to NDVI, SAVI, and NIR showed that TAGC differentiates strongly when these variables are small (Figure 6). Consequently, it is necessary to develop models of single variable or variable combination with weight to improve the differentiation of the model’s predicted values (Table 2).

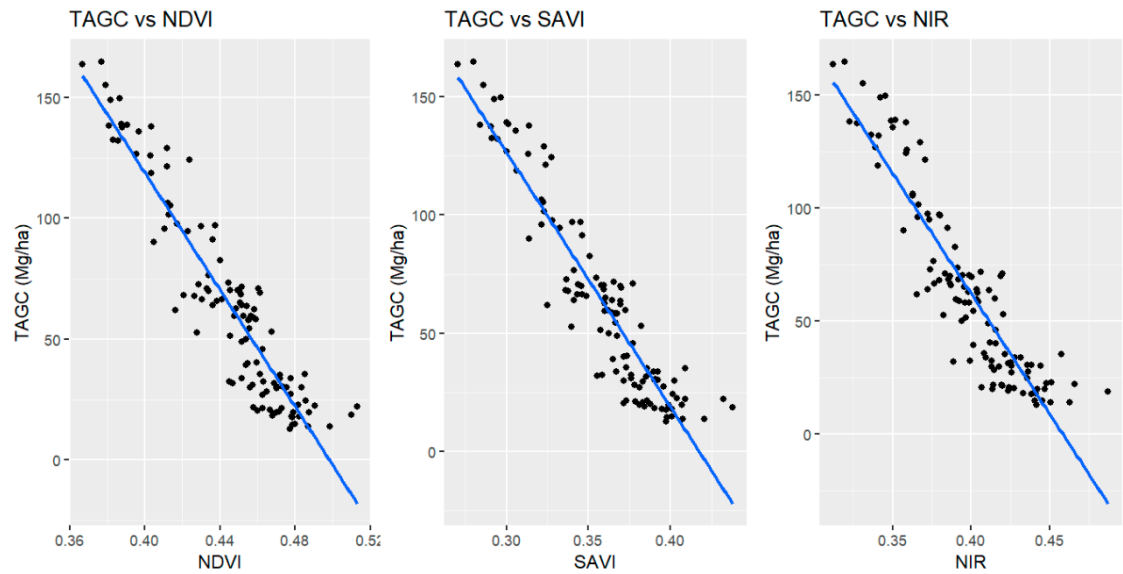


Figure 6. TAGC volatility by NDVI, SAVI and NIR.

Table 2. Established models forms with weights to select the optimal model.

Model	Correlation equation form	Weight
Linear	$\text{TAGC} = f(\text{NDVI})$	$1/\text{NDVI}^2$
	$\text{TAGC} = f(\text{SAVI})$	$1/\text{SAVI}^2$
	$\text{TAGC} = f(\text{NIR})$	$1/\text{NIR}^2$
	$\text{TAGC} = f(\text{NDVI}, \text{ARVI})$	$1/\text{NDVI}^2$
	$\text{TAGC} = f(\text{SAVI}, \text{SIPI})$	$1/\text{SAVI}^2$



Model		Correlation equation form	Weight
Non-linear (Power, Exponential, Quadratic)		$TAGC = f(NIR, EVI)$	$1/NIR^{-2}$
		$TAGC = f(NDVI)$	$1/NDVI^b$
		$TAGC = f(SAVI)$	$1/SAVI^b$
		$TAGC = f(NIR)$	$1/NIR^b$
		$TAGC = f(NDVI, ARVI)$	$1/NDVI^b$
		$TAGC = f(SAVI, SIPI)$	$1/SAVI^b$
		$TAGC = f(NIR, EVI)$	$1/NIR^b$

From the equations presented in Table 2, 36 equations were establish and compared for selecting fit models. The regression equations were established with an average AIC value of 766.691 (the smallest value was 748.717, the largest value was 785.717) and R<sup>2</sup> values ranging from 76.3% to 89.1%. Among these, 16 regression equations with an AIC < 766.691 (ranging from 748.983 to 765.790) and R<sup>2</sup> > 0.85 were chosen for estimating and testing the existence of parameters, including 2 linear equations of single variable, 1 linear equation of variable combination, 3 non-linear equations of single variable, and 10 non-linear equations of variable combination (Table 3).

Table 3. Result of model development with fit statistics.

ID	Equation form	AIC	R <sup>2</sup>	ASE (%)	RMSE (Mg ha <sup>-1</sup> )	MPSE (%)
1	$TAGC = a + b \times NDVI$	762.452	0.87215	0.75	15.08	34.47
2	$TAGC = a \times e^{(b \times NDVI)}$	769.343	0.85502	-0.96	16.15	24.67
3	$TAGC = a + b \times NDVI + c \times NDVI^2$	754.401	0.88567	4.76	14.34	33.30
4	$TAGC = a \times NDVI^b$	776.901	0.83989	-2.30	16.95	25.37
5	$TAGC = a + b \times (NDVI \times ARVI)$	773.041	0.85401	-8.30	16.32	35.60
6	$TAGC = a \times e^{(b \times NDVI \times ARVI)}$	762.113	0.87105	-3.02	15.28	23.96
7	$TAGC = a + b \times (NDVI \times ARVI) + c \times (NDVI \times ARVI)^2$	752.085	0.88764	-59.26	14.36	90.02
8	$TAGC = a \times (NDVI \times ARVI)^b$	774.778	0.84658	-2.11	16.70	25.88
9	$TAGC = a + b \times NDVI + c \times ARVI$	762.164	0.87367	-3.99	14.81	33.62
10	$TAGC = a \times e^{(b \times NDVI + c \times ARVI)}$	769.258	0.86312	-2.83	15.89	25.34
11	$TAGC = a + b \times NDVI + c \times NDVI^2 + d \times ARVI + e \times ARVI^2$	756.739	0.88757	1.76	14.19	32.49
12	$TAGC = a \times NDVI^b \times ARVI^c$	775.078	0.85231	-3.17	16.40	25.57
13	$TAGC = a + b \times SAVI$	765.790	0.86697	-5.12	15.38	36.42
14	$TAGC = a \times e^{(b \times SAVI)}$	762.800	0.85757	-2.84	16.35	23.62
15	$TAGC = a + b \times SAVI + c \times SAVI^2$	748.983	0.88897	1.27	14.14	42.69
16	$TAGC = a \times SAVI^b$	772.618	0.82260	-1.70	17.68	24.74
17	$TAGC = a + b \times (SAVI \times SIPI)$	771.465	0.85792	-10.79	15.83	46.47
18	$TAGC = a \times e^{(b \times SAVI \times SIPI)}$	765.410	0.85607	-2.23	15.73	23.22
19	$TAGC = a + b \times (SAVI \times SIPI) + c \times (SAVI \times SIPI)^2$	753.226	0.88436	4.28	13.88	26.51

ID	Equation form	AIC	R <sup>2</sup>	ASE (%)	RMSE (Mg ha <sup>-1</sup> )	MPSE (%)
20	TAGC = a x (SAVI x SIPI) <sup>a</sup>	777.733	0.82186	-1.70	17.54	25.55
21	TAGC = a + b x SAVI + c x SIPI	766.979	0.86804	0.03	15.44	61.61
22	<b>TAGC = a x e<sup>(b x SAVI + c x SIPI)</sup></b>	<b>762.561</b>	<b>0.86136</b>	-2.72	15.74	23.62
23	<b>TAGC = a + b x SAVI + c x SAVI<sup>2</sup> + d x SIPI + e x SIPI<sup>2</sup></b>	<b>751.194</b>	<b>0.89076</b>	3.02	14.39	29.86
24	TAGC = a x SAVI <sup>b</sup> x SIPI <sup>c</sup>	773.127	0.82525	-1.30	17.58	24.99
25	TAGC = a + b x NIR	785.321	0.83064	24.52	17.03	95.31
26	TAGC = a x e <sup>(b x NIR)</sup>	768.493	0.82896	-0.57	16.92	23.82
27	<b>TAGC = a + b x NIR + c x NIR<sup>2</sup></b>	<b>756.924</b>	<b>0.86649</b>	0.70	15.50	23.17
28	TAGC = a x NIR <sup>b</sup>	775.476	0.78901	-0.84	18.75	24.77
29	TAGC = a + b x (NIR x EVI)	785.717	0.82968	-9.43	17.54	41.36
30	TAGC = a x e <sup>(b x NIRxEVI)</sup>	761.065	0.84910	-2.65	16.20	23.33
31	<b>TAGC = a + b x (NIR x EVI) + c x (NIR x EVI)<sup>2</sup></b>	<b>752.493</b>	<b>0.87647</b>	-0.16	14.65	22.54
32	TAGC = a x (NIR x EVI) <sup>b</sup>	777.524	0.76316	-0.18	20.00	24.77
33	TAGC = a + b x NIR + c x EVI	776.007	0.85124	-7.96	15.90	51.13
34	TAGC = a x e <sup>(b x NIR + c x EVI)</sup>	768.716	0.82342	-0.27	17.77	24.71
35	<b>TAGC = a + b x NIR + c x NIR<sup>2</sup> + d x EVI + e x EVI<sup>2</sup></b>	<b>756.972</b>	<b>0.87294</b>	2.35	15.23	24.13
36	TAGC = a x NIR <sup>b</sup> x EVI <sup>c</sup>	775.920	0.78908	-0.02	19.85	25.11

Note: **Bold**: Selected models based on AIC and R<sup>2</sup> for estimating parameters.

3.3. Determination of Above Ground Carbon Estimation Models

The fit models were estimated and tested for the existence of parameters with data of 115 samples (100% of the dataset). Among the 16 selected fit models, 9 models have parameters with p-value <0.001 and MPSE values ranging from 21.63% to 26.32% (lower than 30%), including one single variable exponential model (Equation 14), two exponential models of variable combination (Equations 6 and 18), three single variable quadratic models (Equations 3, 15, and 27), and three quadratic models of variable combination (Equations 7, 19, and 31) (Table 4).

**Table 4.** Result of fit models selection with parameter estimation and fit statistics.

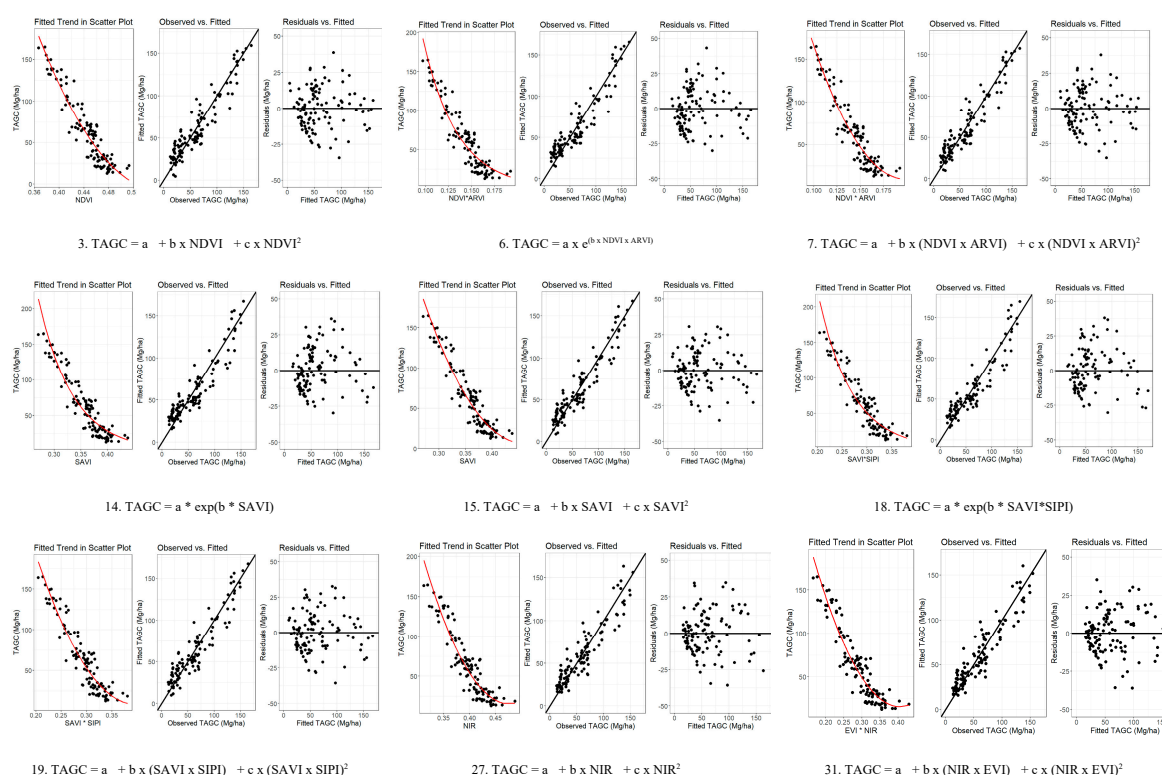
ID	Equation form	Parameters	P-value	Std. Error	R <sup>2</sup>	MPSE (%)
1	TAGC = a + b x NDVI	a b	590 -1181,4	<0.001 <0.001	20.8 46.1	0.87064 35.99
3	<b>TAGC = a + b x NDVI + c x NDVI<sup>2</sup></b>	a b c	<b>1523,206</b> <b>-5441,707</b> <b>4837,304</b>	<b>&lt;0.001</b> <b>&lt;0.001</b> <b>&lt;0.001</b>	217.6896 989.7512 1121.616	0.88581 <b>26.32</b>
6	<b>TAGC = a x e<sup>(b x NDVI x ARVI)</sup></b>	a b	<b>2617,904</b> <b>-26,8626</b>	<b>&lt;0.001</b> <b>&lt;0.001</b>	356.5325 1.0567	0.87025 <b>23.62</b>

ID	Equation form	Parameters	P-value	Std. Error	R <sup>2</sup>	MPSE (%)	
7	TAGC = a + b x (NDVI x ARVI) + c x (NDVI x ARVI) <sup>2</sup>	a	648,275	<0.001	53.1819	0.88712	23.03
		b	-6450,795	<0.001	743.7517		
		c	16281,53	<0.001	2576.696		
9	TAGC = a + b x NDVI + c x ARVI	a	607,43	<0.001	22.71	0.87415	31.13
		b	-2406,51	<0.001	674.97		
		c*	1644,34	0.071	903.84		
11	TAGC = a + b x NDVI + c x NDVI <sup>2</sup> + d x ARVI + e x ARVI <sup>2</sup>	a	1348,19	<0.001	256.874	0.88784	23.47
		b*	10401,88	0.348	11037.17		
		c*	-12747,94	0.291	12032.41		
		d*	-20840,87	0.154	14536.67		
		e*	31994,84	0.146	21860.44		
13	TAGC = a + b x SAVI	a	432,86	<0.001	15.47	0.86581	48.71
		b	-1031,04	<0.001	42.15		
14	TAGC = a x e <sup>(b x SAVI)</sup>	a	14563,32	<0.001	3089.371	0.85639	23.38
		b	-15,606	<0.001	0.6266		
15	TAGC = a + b x SAVI + c x SAVI <sup>2</sup>	a	1070,272	<0.001	109.0798	0.88933	22.62
		b	-4640,056	<0.001	608.0028		
		c	5060,527	<0.001	843.6183		
18	TAGC = a x e <sup>(b x SAVI x SIPI)</sup>	a	4288,168	<0.001	701.5295	0.85790	23.42
		b	-14,784	<0.001	0.5977		
19	TAGC = a + b x (SAVI x SIPI) + c x (SAVI x SIPI) <sup>2</sup>	a	754,192	<0.001	67.7088	0.88606	22.85
		b	-3758,794	<0.001	458.1902		
		c	4737,115	<0.001	769.794		
22	TAGC = a x e <sup>(b x SAVI + c x SIPI)</sup>	a*	387,9781	0.610	759.3648	0.86151	23.23
		b	-20,1558	<0.001	2.5505		
		c*	6,3853	0.065	3.4379		
23	TAGC = a + b x SAVI + c x SAVI <sup>2</sup> + d x SIPI + e x SIPI <sup>2</sup>	a*	-762,477	0.795	2932.537	0.88978	22.24
		b	-5846,867	<0.001	1706.388		
		c	6568,468	0.004	2266.891		
		d*	4858,904	0.532	7756.204		
		e*	-2845,5	0.543	4663.326		
27	TAGC = a + b x NIR + c x NIR <sup>2</sup>	a	1537,576	<0.001	143.5515	0.86646	22.04
		b	-6398,241	<0.001	700.553		
		c	6723,375	<0.001	852.3433		
31	TAGC = a + b x (NIR x EVI) + c x (NIR x EVI) <sup>2</sup>	a	505,7588	<0.001	33.2636	0.87646	21.63
		b	-2411,523	<0.001	214.8227		
		c	2967,038	<0.001	343.2117		
35	TAGC = a + b x NIR + c x NIR <sup>2</sup> + d x EVI + e x EVI <sup>2</sup>	a	1513,702	<0.001	205.311	0.87259	21.76
		b	-7727,276	0.018	3225.838		
		c	8721,605	0.022	3780.838		

ID	Equation form	Parameters	P-value	Std. Error	R <sup>2</sup>	MPSE (%)
		d*	790,201	0.563	1364.302	
		e	-638,933	0.470	881.386	

Note: **Bold**: Selected models based on cross validation with MPSE < 30%; \*: Parameter with p-value > 0.05.

Among the 9 models identified as fit models, the quadratic models showed the best AIC, R<sup>2</sup>, and MPSE values. In particular, the variable combination model (NIR, EVI) (Equation 31) and the single variable NIR model (Equation 27) had the best MPSE (lowest) and equivalent, at 21.63% and 22.04% respectively (Table 4). In addition, these two models have fitted trend in the middle of the data cloud (left) or closely follow the observed value on the diagonal (middle) and the residuals (right) have a narrow distribution, and are all according to the fitted values (Figure 7). From these results, two mentioned-above models determined as optimal equations, in which, one equation is representative of single variable quadratic model ( $TAGC = a + b \times NIR + c \times NIR^2$ ) and one is representative of quadratic model with variable combination ( $TAGC = a + b \times (NIR \times EVI) + c \times (NIR \times EVI)^2$ ).



**Figure 7.** Comparison of Fitted Trend in Scatter Plot, Observed vs. Fitted, and Residuals vs. Fitted (left to right) of 9 selected equations.

## 4. Discussion

### 4.1. Determination of Indices of Sentinel 2 Imagery Influencing TAGC Prediction

Many studies have shown that satellite images are commonly used to build biomass estimation models through their VIs, but most only determine the correlation between AGB and each individual VI [7–10], [37,38]). Meanwhile, VIs are determined from the multispectral bands BLUE, GREEN, RED, NIR [25], so there is a complex relationship between TAGC, VIs and multispectral bands. Consequently, it is necessary to identify multiple VIs and multispectral bands that are strongly associated with TAGC in order to develop regression equations with sing variable form or variable combination form.



From the PCA results, it was determined that NDVI, SAVI, and NIR have a great influence on TAGC. This result is similar to the study of Poudel et al. (2023), reporting that among the 12 VIs of Sentinel 2 images, NDVI and SAVI are the 2 VIs that have the closest relationship with AGB in both linear and quadratic models [38]. Unlike previous studies, which often only used single indices to establish models, this study identified additionally 3 variable combinations (NDVI, ARVI), (SAVI, SIPI), (NIR, EVI) that have a great influence on TAGC. This limits the loss of information in the original dataset, because PCA reduces the multidimensionality of large datasets, increases the ability to interpret/explain, and minimizes the loss of information in the dataset [39].

As PCA results of previous studies such as PCA results of the Dalat pine's diameter growth and 4 ecological environmental factors with 66.85%, and 9 climate factors with 77.90% [31], this study conducted PCA with assumption using the first two PCs if they account for 80% total variability in the original dataset. With this assumption, VIs and multispectral band were determined for TAGC prediction with high reliability, since cumulative variance proportion of first two PCs account for 96.35% of total variability in the original dataset (Figure 3)

#### 4.2. Establishment and Validation of Models for Predicting TAGC

Most publications related to the relationship model between AGB and Sentinel 2 image indices (ARVI, EVI, NDVI, SAVI, SIPI) are set up in a linear form without weight, so  $R^2$  only reaches from 0.57 to 0.75 [38], NDVI of Landsat 8 image only gives  $R^2 = 0.43$  [40] or linearized log-log, log-lin function with EVI gives  $R^2 = 0.60-0.76$  [8]. With the weighted linear equations between TAGC and single variables (NDVI, SAVI, NIR) or combination of variables (NDVI, ARVI), (SAVI, SIPI), (NIR, EVI) by the Ordinary Least Squares method in this study,  $R^2 = 0.83 - 0.87$  (Table 3), should be superior to the linear model or log-log, log-lin linearization. Similarly, with the weighted nonlinear model established by the Maximum Likelihood method,  $R^2 = 0.76 - 0.89$  (Table 3), which is also much higher than the research results of Poudel, A. et al (2023),  $R^2 = 0.72 - 0.78$  with the quadratic form,  $R^2 = 0.55 - 0.61$  with the power form and  $R^2 = 0.53 - 0.62$  with the exponential form [38]. This result has shown that establishing regression models, including both linear with weight and non-linear with weight models, improves error variability.

The quadratic, exponential and power models all gave higher  $R^2$  than the linear model, respectively  $R^2 = 0.87 - 0.89$ ,  $R^2 = 0.82 - 0.87$ ,  $R^2 = 0.76 - 0.85$  (Table 3). This result is similar to the study of Poudel, A. et al (2023), the quadratic models gave the highest  $R^2$ , followed by the exponential model and finally the power model [38]. Therefore, non-linear models should be established to increase reliability compared to the linear and linearized models, because there is a complex, multivariate and non-linear relationship between AGB/TAGC and VIs and multispectral bands.

Regarding model validation, most previous studies have not applied the cross validation method but mainly divided the dataset into two independent parts, 50% of the data used to establish the model and 50% of the data for model validation [8,38,40]. This study applied the cross validation method with 80% of the data used to establish the model and 20% of the data for validation. The validation process was carried out with 100 iterations, so the models all had  $R^2 = 0.76 - 0.89$  (Table 3), which was superior to previous studies that all had  $R^2 < 0.78$  [8,38,40].

## 5. Conclusion

Establishing TAGC estimation models based on VIs and multispectral bands of Sentinel 2 images plays an increasingly important role in estimating forest carbon stocks, because this method increases reliability and reduces costs compared to the destructive method and can be applied at large scales such as national, regional and global.

The variables NDVI, SAVI, NIR and the combination of variables (NDVI, ARVI), (SAVI, SIPI), (NIR, EVI) have a close relationship with TAGC. This study selected 9 suitable models with  $R^2 = 0.85 - 0.89$  and MPSE = 21.63% - 26.32%, including both single variable and variable combination non-linear models. Among them, two optimal models were determined to estimate TAGC in the South Central Coastal Ecoregion, Vietnam:

$$\text{TAGC} = 1537.576 - 6398.241 \times \text{NIR} + 6723.375 \times \text{NIR}^2$$

$$\text{TAGC} = 505.7588 - 2411.523 \times (\text{NIR} \times \text{EVI}) + 2967.038 \times (\text{NIR} \times \text{EVI})^2$$

## References

1. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment* **1988**, 25(3), 295–309. [https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/10.1016/0034-4257(88)90106-X)
2. Gitelson, A.; Merzlyak, M.N. Remote estimation of chlorophyll content in higher plant leaves. *International Journal of Remote Sensing* **1997**, 18(12), 2691–2697. <https://doi.org/10.1080/014311697217558>
3. Wu, J.; Chen, B.; Reynolds, G.; Xie, J.; Liang, S.; O'Brien, M.J.; Hector, A. Monitoring tropical forest degradation and restoration with satellite remote sensing: A test using Sabah Biodiversity Experiment. In *Advances in Ecological Research*; Academic Press Inc., **2020**; 62, pp. 117–146. <https://doi.org/10.1016/bs.aecr.2020.01.005>
4. Tucker, C.J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment* **1979**, 8(2), 127–150. [https://doi.org/10.1016/0034-4257\(79\)90013-0](https://doi.org/10.1016/0034-4257(79)90013-0)
5. Baccini, A.; Friedl, M.A.; Woodcock, C.E.; Warbington, R. Forest biomass estimation over regional scales using multisource data. *Geophysical Research Letters* **2004**, 31(10). <https://doi.org/10.1029/2004GL019782>
6. Pettorelli, N.; Vik, J.O.; Mysterud, A.; Gaillard, J.M.; Tucker, C.J.; Stenseth, N.C. Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology & Evolution* **2005**, 20(9), 503–510. <https://doi.org/10.1016/j.tree.2005.05.011>
7. Priatama, A.R.; Setiawan, Y.; Mansur, I.; Masyhuri, M. Regression Models for Estimating Aboveground Biomass and Stand Volume Using Landsat-Based Indices in Post-Mining Area. *Jurnal Manajemen Hutan Tropika* **2022**, 28(1), 1–14. <https://doi.org/10.7226/jtfm.28.1.1>
8. Dang, H.N.; Ba, D.D.; Trung, D.N.; Viet, H.N.H. A Novel Method for Estimating Biomass and Carbon Sequestration in Tropical Rainforest Areas Based on Remote Sensing Imagery: A Case Study in the Kon Ha Nung Plateau, Vietnam. *Sustainability* **2022**, 14, 16857. <https://doi.org/10.3390/su142416857>
9. Khan, K.; Iqbal, J.; Ali, A.; Khan, S.N. Assessment of Sentinel-2-Derived Vegetation Indices for the Estimation of Above-Ground Biomass/Carbon Stock, Temporal Deforestation, and Carbon Emissions Estimation in the Moist Temperate Forests of Pakistan. *Applied Ecology and Environmental Research* **2020**, 18(1), 783–815. [https://doi.org/10.15666/aeer/1801\\_783815](https://doi.org/10.15666/aeer/1801_783815)
10. Askar, N.; Nuthammachot, N.; Phairuang, W.; Wicaksono, P.; Sayektiningsih, T. Estimating Aboveground Biomass on Private Forest Using Sentinel-2 Imagery. *Journal of Sensors* **2018**, Article ID 6745629, 11 pages. <https://doi.org/10.1155/2018/6745629>
11. Zhang, X.; Zhao, Y.; Ashton, M.S.; Lee, X. Measuring Carbon in Forests. In *Managing Forest Carbon in a Changing Climate*; Ashton, M., Tyrrell, M., Spalding, D., Gentry, B., Eds.; Springer: Dordrecht, The Netherlands, **2012**. [https://doi.org/10.1007/978-94-007-2232-3\\_7](https://doi.org/10.1007/978-94-007-2232-3_7)
12. Naesset, E.; Gobakken, T.; Solberg, S.; Gregoire, T.G.; Ståhl, G.; Lange, H.; Dick, O.; Gobakken, T.; Astrup, R. Mapping and estimating forest area and aboveground biomass in miombo woodlands in Tanzania using data from airborne laser scanning, TanDEM-X, RapidEye, and global forest maps: A comparison of estimated precision. *Remote Sensing of Environment* **2016**, 175, 282–300. <https://doi.org/10.1016/j.rse.2016.01.006>
13. McRoberts, R.E.; Næsset, E.; Gobakken, T. Estimation for inaccessible and non-sampled forest areas using model-based inference and remotely sensed auxiliary information. *Remote Sensing of Environment* **2014**, 154, 226–233. <https://doi.org/10.1016/j.rse.2014.08.028>
14. Esteban, J.R.E.; Montealegre, A.L.; Miranda, D.; Segura, A.S.; Ruiz, M.M. A model-based volume estimator that accounts for both land cover misclassification and model prediction uncertainty. *Remote Sensing* **2020**, 12(20), 3360. <https://doi.org/10.3390/rs12203360>
15. Jędrych, M.; Zagajewski, B.; Marcinkowska-Ochtyra, A. Application of Sentinel-2 and EnMAP new satellite data to the mapping of environmental changes. *Polish Cartographical Review* **2017**, 49(3), 107–119. <https://doi.org/10.1515/pcr-2017-0011>

16. Phuong, V.T.; Inoguchi, A.; Birigazzi, L.; Henry, M.; Sola, G. Introduction and Background of the Study. In *Tree Allometric Equation Development for Estimation of Forest Above-Ground Biomass in Viet Nam (Part A)*; Inoguchi, A., Henry, M., Birigazzi, L., Sola, G., Eds.; UN-REDD Programme: Hanoi, Vietnam, **2012**.
17. Moradi, F.; Darvishsefat, A.A.; Pourrahmati, M.R.; Deljouei, A.; Borz, S.A. Estimating Aboveground Biomass in Dense Hyrcanian Forests by the Use of Sentinel-2 Data. *Forests* **2022**, *13*(1), 104. <https://doi.org/10.3390/f13010104>
18. Huy, B. Allometric Model and Remote Sensing-GIS to Estimate Carbon Removal of Evergreen Broadleaf Forests in the Central Highland Region; Publication House of Science and Technique: Hanoi, Vietnam, **2013**.
19. Phuong, V.T.; Linh, N.T.M. Final Report on Forest Ecological Stratification in Vietnam; UN-REDD Programme: Ha Noi, Vietnam, **2011**.
20. Sola, G.; Inoguchi, A.; Garcia-Perez, J.; Donegan, E.; Birigazzi, L.; Henry, M. Allometric Equations at National Scale for Tree Biomass Assessment in Viet Nam: Context, Methodology and Summary of the Results; UN-REDD Programme: Ha Noi, Vietnam, **2014**.
21. Wikipedia. Da Nang City. Available online: [https://vi.wikipedia.org/wiki/%C4%90%C3%A0\\_N%E1%BA%B5ng](https://vi.wikipedia.org/wiki/%C4%90%C3%A0_N%E1%BA%B5ng) (accessed on 24 June 2024).
22. Da Nang Province People's Committee. Decision 430/QĐ-UBND Dated 04/03/2024 on Approval of Forest Status and Land Use Planning for Forest Development in Da Nang City in 2023. **2024**.
23. Huy, B.; Poudel, K.P.; Temesgen, H. Aboveground biomass equations for evergreen broadleaf forests in South Central Coastal ecoregion of Viet Nam: Selection of eco-regional or pantropical models. *Forest Ecology and Management* **2016**, *376*, 276–283. <https://doi.org/10.1016/j.foreco.2016.06.031>
24. IPCC. *Guidelines for National Greenhouse Gas Inventories*; Eggleston, H.S.; Buendia, L.; Miwa, K.; Ngara, T.; Tanabe, K., Eds.; IGES: Japan, **2006**.
25. Xue, J.; Su, B. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. *Journal of Sensors* **2017**, 2017, Article ID 1353691, 17 pages. <https://doi.org/10.1155/2017/1353691>
26. Kaufman, Y.J.; Tanre, D. Atmospherically Resistant Vegetation Index (ARVI) for EOS-MODIS. *IEEE Transactions on Geoscience and Remote Sensing* **1992**, *30*(2), 261–270.
27. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the Radiometric and Biophysical Performance of the MODIS Vegetation Indices. *Remote Sensing of Environment* **2002**, *83*(1–2), 195–213. [https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2)
28. Peñuelas, J.; Filella, I.; Lloret, P.; Muñoz, F.; Vilajeliu, M. Reflectance Assessment of Mite Effects on Apple Trees. *International Journal of Remote Sensing* **1995**, *16*(14), 2727–2733.
29. R Core Team. A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria, **2023**. Available online: <http://www.r-project.org/index.html>
30. Greenacre, M.; Groenen, P.J.F.; Hastie, T.; et al. Principal Component Analysis. *Nature Reviews Methods Primers* **2022**, *2*, 100. <https://doi.org/10.1038/s43586-022-00184-w>
31. Huy, B.; Nam, L.C.; Poudel, K.P.; Temesgen, H. Individual Tree Diameter Growth Modeling System for Dalat Pine (*Pinus dalatensis* Ferré) of the Upland Mixed Tropical Forests. *Forest Ecology and Management* **2021**, *480*, 118612: 1–15. <https://doi.org/10.1016/j.foreco.2020.118612>
32. Huy, B.; Truong, N.Q.; Khiem, N.Q.; Poudel, K.P.; Temesgen, H. Stand Growth Modeling System for Planted Teak (*Tectona grandis* L.f.) in Tropical Highlands. *Trees For People* **2022**, *9*, 100308. <https://doi.org/10.1016/j.tfp.2022.100308>
33. Picard, N.; Saint-André, L.; Henry, M. *Manual for Building Tree Volume and Biomass Allometric Equations: From Field Measurement to Prediction*; FAO: Rome, Italy; Centre de Coopération Internationale en Recherche Agronomique pour le Développement: Montpellier, France, **2012**; 215 pp.
34. Huy, B.; Khiem, N.Q.; Truong, N.Q.; Poudel, K.P.; Temesgen, H. Additive Modeling Systems to Simultaneously Predict Aboveground Biomass and Carbon for *Litsea glutinosa* of Agroforestry Model in Tropical Highlands. *Forest Systems* **2023**, *32*(1), e006. <https://doi.org/10.5424/fs/2023321-19780>
35. Akaike, H. Information Theory and an Extension of the Maximum Likelihood Principle. In *Proceedings of the 2nd International Symposium on Information Theory*; Akademiai Kiado: Budapest, Hungary, **1973**; pp. 267–281.

36. Zeng, W.; Zhang, L.; Chen, X.; Cheng, Z.; Ma, K.; Li, Z. Construction of Compatible and Additive Individual-Tree Biomass Models for *Pinus tabulaeformis* in China. *Canadian Journal of Forest Research* **2017**, *47*, 467–475. <https://doi.org/10.1139/cjfr-2016-0342>
37. Pandit, S.; Tsuyuki, S.; Dube, T. Estimating Above-Ground Biomass in Sub-Tropical Buffer Zone Community Forests, Nepal, Using Sentinel-2 Data. *Remote Sensing* **2018**, *10*(4), 601. <https://doi.org/10.3390/rs10040601>
38. Poudel, A.; Shrestha, H.L.; Mahat, N.; Sharma, G.; Aryal, S.; Kalakheti, R.; Lamsal, B. Modeling and Mapping of Aboveground Biomass and Carbon Stock Using Sentinel-2 Imagery in Chure Region, Nepal. *International Journal of Forestry Research* **2023**, 2023, Article ID 5553957. <https://doi.org/10.1155/2023/5553957>
39. Jolliffe, I.T.; Cadima, J. Principal Component Analysis: A Review and Recent Developments. *Philosophical Transactions of the Royal Society A* **2016**, *374*, 20150202. <https://doi.org/10.1098/rsta.2015.0202>
40. Luong, V.N.; Tateishi, R.; Kondoh, A.; Sharma, R.C.; Hoan, T.N.; Tu, T.T.; Minh, D.H.T. Mapping Tropical Forest Biomass by Combining ALOS-2, Landsat 8, and Field Plots Data. *Land* **2016**, *5*(4), 31. <https://doi.org/10.3390/land5040031>

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.