MDPI

Article

Biomass Prediction Using Sentinel-2 Imagery and an Artificial Neural Network in the Amazon/Cerrado Transition Region

Luana Duarte de Faria [1,*], **Eraldo Aparecido Trondoli Matricardi** [1,*], **Beatriz Schwantes Marimon** [2],
Eder Pereira Miguel [1], **Ben Hur Marimon Junior** [2], **Edmar Almeida de Oliveira** [2],
Nayane Cristina Candido dos Santos Prestes [2] and **Osmar Luiz Ferreira de Carvalho** [3]

[1] Forestry Department, College of Technology, University of Brasília, Campus Darcy Ribeiro,
Brasilia 70910-900, DF, Brazil; edermiguel@unb.br
[2] Plant Ecology Laboratory, Mato Grosso State University, Campus Nova Xavantina, P.O. Box 08,
Nova Xavantina 78690-000, MT, Brazil; biamarimon@unemat.br (B.S.M.); bhmjunior@gmail.com (B.H.M.J.);
edmardinei@gmail.com (E.A.d.O.); nayaneprestesnx@gmail.com (N.C.C.d.S.P.)
[3] Department of Electrical Engineering, Campus Darcy Ribeiro, Brasilia 70910-900, DF, Brazil;
osmarcarvalho@ieee.org
* Correspondence: luanaduarte005@gmail.com (L.D.d.F.); ematricardi@unb.br (E.A.T.M.)

Abstract: The ecotone zone, located between the Cerrado and Amazon biomes, has been under intensive anthropogenic pressures due to the expansion of commodity agriculture and extensive cattle ranching. This has led to habitat loss, reducing biodiversity, depleting biomass, and increasing CO_2 emissions. In this study, we employed an artificial neural network, field data, and remote sensing techniques to develop a model for estimating biomass in the remaining native vegetation within an 18,864 km^2 ecotone region between the Amazon and Cerrado biomes in the state of Mato Grosso, Brazil. We utilized field data from a plant ecology laboratory and vegetation indices from Sentinel-2 satellite imagery and trained artificial neural networks to estimate aboveground biomass (AGB) in the study area. The optimal network was chosen based on graphical analysis, mean estimation errors, and correlation coefficients. We validated our chosen network using both a Student's *t*-test and the aggregated difference. Our results using an artificial neural network, in combination with vegetation indices such as AFRI (Aerosol Free Vegetation Index), EVI (Enhanced Vegetation Index), and GNDVI (Green Normalized Difference Vegetation Index), which show an accurate estimation of aboveground forest biomass (Root Mean Square Error (RMSE) of 15.92%), can bolster efforts to assess biomass and carbon stocks. Our study results can support the definition of environmental conservation priorities and help set parameters for payment for ecosystem services in environmentally sensitive tropical regions.

Keywords: biomass estimation; Amazon/Cerrado ecotone; remote sensing; artificial neural network; Google Earth Engine

1. Introduction

Brazilian biomes are recognized for their high biodiversity, with over 33,000 plant species, constituting a staggering 26.5% of all known species on Earth [1]. More specifically, the Amazonia/Cerrado ecotone is a unique transitional ecoregion covering more than 4000 km across the ecotone between the two greatest biomes of South America [2]. The region is dominated by a highly seasonal climate and a wide diversity of vegetation types. These vegetation types range from open savannas, which receive abundant solar radiation, to dense forest formations with denser canopy and higher air humidity and soil moisture levels.

Beyond its rich vegetation, this region of high ecological and biological significance harbors a large array of species. However, this biodiversity faces threats as pastures and crops expand into this ecotone, leading to massive deforestation [3,4]. The consequence

is a notable decrease in the native vegetation and biomass stock [5] due to the increasing deforestation rates in the Amazon and Cerrado biomes [6], a situation often worsened by forest fires [6,7].

Forest biomass is a critical factor in assessing the carbon sequestration and carbon balance capabilities of these ecosystems [8,9]. Accurately estimating aboveground biomass (AGB) is crucial to understanding the carbon cycle and its effects on climate changes and on terrestrial ecosystems and biodiversity [8–11], especially in tropical regions where reliable data are lacking [8,9].

Biomass estimation using remote sensing data has been widely applied at global, regional, and local scales. It has substantially improved in recent years [12], replacing conventional AGB estimation approaches. It enables temporal analysis of the environment and land cover [12] and, in the case of land use changes, contributes significantly to detecting, quantifying, and understanding vegetation behavior over time [13].

Several approaches have been developed and applied to accurately estimate carbon biomass. The authors of [14] accurately estimated aboveground biomass and stand volume in Hinton, the USA, by applying a methodological approach based on the relationship of forest structure attributes acquired in the field and Landsat ETM+ imagery. The authors of [15] successfully quantified live aboveground forest biomass in the states of Arizona and Minnesota using Landsat imagery and forest inventory data. The authors of [16] assessed Landsat 8 imagery to estimate aboveground biomass in the Umgeni catchment, South Africa. The authors of [17] applied boosted regression tree models, field data, and Sentinel-2 and Synthetic Aperture Radar (SAR) combined imagery acquired on different dates and were able to estimate aboveground biomass and forest cover.

Studies carried out by [18] combined vegetation indices retrieved from a Vegetation Sensor onboard the SPOT-4 satellite and Moderate Resolution Imaging Spectroradiometer (MODIS) and climate data to estimate primary production in Harvard Forest, Petersham, MA, the USA. The study by [19] observed a strong positive correlation between vegetation indices and biomass. Another study by [20] successfully estimated forest aboveground biomass (AGB) by combining Landsat and MODIS imagery.

New technologies based on machine learning and artificial intelligence have improved even more modeling approaches to predict biomass worldwide. Empirical modeling using deep learning algorithms has achieved highly accurate results in estimating AGB based on field sampling distributions with no assumptions. For example, [21] developed Sentinel-2 imagery and a machine learning model to estimate biomass in northern Anhui, China. Similarly, ref. [22] applied radar and optical imagery and a deep learning-based approach to estimate forest biomass in Tibet, China. The authors of Ref. [23] successfully combined an Artificial Neural Network (ANN) with vegetation indices retrieved from Landsat imagery to predict aboveground biomass for a study site in the Amazon region. However, they are more difficult to interpret and require accurate field data as the model input [24].

In this study, we developed and applied a model to estimate aboveground biomass in an Amazonia/Cerrado transition zone in the state of Mato Grosso, Brazil, using field data, remote sensing, and Artificial Neural Networks (ANNs). Our goal was to accurately estimate AGB using medium spatial resolution and freely available remotely sensed data (Sentinel-2 imagery) with an ANN, a method not previously applied to this large ecotone region. These study results are significant as they can facilitate further analyses of deforestation and forest fire impacts in this tropical region, which have profoundly affected forest structure by reducing tree cover and increasing herbaceous species. These herbaceous plants are more susceptible to water stress, making the region prone to recurrent and intense fire events [25].

Our model showed promising results for estimating and monitoring aboveground biomass and can play a pivotal role in supporting the implementation of payments for ecosystem services. This represents a technological advance in environmental preservation and conservation research, particularly in transitional zones that lack information on biomass stocks. From a critical perspective, conserving biomass in this study area, which is

near Brazil's largest indigenous territory (Xingu Indigenous Land), may have significant and positive impacts on the well-being and sustainable existence of traditional populations in their territories [26].

2. Materials and Methods

2.1. Regional Setting

Our study area encompassed a total of 18,863.6 Km^2 located in the Ecotone region between the Amazon and Cerrado biomes in Brazil. We selected permanent long-term measurement plots established and monitored by the Plant Ecology Laboratory of Mato Grosso State University (LABEV-UNEMAT) in the study region (Campus of Nova Xavantina, State of Mato Grosso, Brazil). The sample plots are in the municipalities of Gaúcha do Norte, Querência, and Ribeirão Cascalheira, state of Mato Grosso.

Field measurements were conducted in 12 sample plots, each measuring 100 m × 100 m and subdivided into 60 subplots of 100 m × 20 m (Figure 1). These measurements were carried out during the dry season (July to October) in 2014, 2018, 2020, and 2021. We selected this study area due to its environmental sensitivity and socioeconomic characteristics, as it is situated in the transition zone between the Cerrado and Amazonia biomes. The area is particularly notable for its proximity to indigenous lands and the significant deforestation activities reported in recent decades, especially in the region known as the "Arc of Deforestation" of the Brazilian Amazon [2].
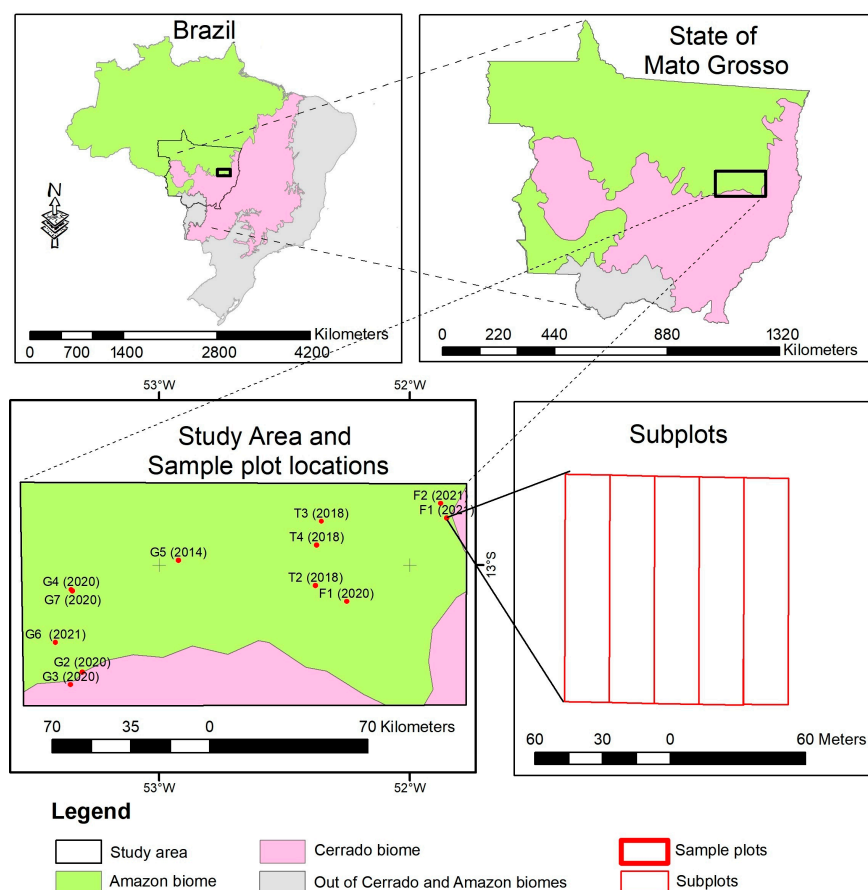

Figure 1. The study area is located within the ecotone zone of the Amazonia and Cerrado biomes in the state of Mato Grosso, Brazil. Field measurements were conducted in 12 sample plots, each measuring 10,000 m² and subdivided into 60 subplots of 2000 m² each, in the years 2014, 2018, 2020, and 2021. The year of sampling is indicated in black above each sample plot in the study area.

The study region features diverse soil types with distinct characteristics. These soils are characterized by low nutrient availability and elevated levels of aluminum toxicity. In

the interfluvial areas, medium-textured red-yellow latosols predominate, creating favorable conditions for forest establishment. Additionally, these latosols feature patches of anthropogenic soils created by ancient indigenous populations, known locally as 'terra preta de índio' or Amazonian Dark Earth (ADE). ADE is rich in pyrogenic carbon, leading to a higher concentration of organic matter on the surface and increased pH in deeper layers. In floodplains, clay-textured fluvic neosols are prevalent, containing higher potassium content but facing phosphorus restriction, poor drainage, and elevated aluminum and iron levels [7,27].

The permanent plots of this study are predominantly surrounded by Seasonal Forest (Fse) and Typical Cerrado (Sd), which are characteristic of the Central-West region of Brazil [28]. According to the Köppen climate classification, the study region is characterized by an Aw climate type, which is tropical seasonal [29], with two distinct seasons: a dry season from May to October and a rainy season from November to April [30]. As described by [7], the region's topography varies from flat to gently undulating. It includes plateaus and plains in the central area, mountains to the east, and residual depressions to the south [31].

2.2. Dendrometric Variables of the Inventory

The inventories were conducted in 2014 (one plot), 2018 (three plots), 2020 (five plots), and 2021 (three plots) by collaborators from the Forest Ecology Laboratory at the State University of Mato Grosso. The objective was to monitor vegetation within permanent plots across different strata, soil types, climatic zones, and regional groups. Sampling was randomized, with 12 sampling units, each measuring 100 × 100 m. Each unit was divided into five transects, resulting in a total of 60 subsamples measuring 100 × 20 m (Figure 1). There was only one sample unit showing different dimensions, covering an area of 180 × 60 m with transects measuring 36 × 60 m. We adopted the sampling protocol proposed by [7] to ensure data reliability.

We collected detailed information on species, families, tree diameters, and heights in each plot. To estimate basic wood density, we used the ForestPlots.net database, which includes data on over 2000 neotropical species [32,33]. Aboveground biomass was calculated using Microsoft Excel 2016, incorporating data on diameter at breast height, total height, and basic wood density. All data were analyzed following RAINFOR guidelines and the methodologies outlined by these authors.

Complementarily, we conducted a statistical analysis using field-collected data to examine variations in dendrometric characteristics within our study area. Descriptive table analysis allowed us to summarize and describe inventory variables, enabling comparisons with similar areas and contributing to the scientific understanding of this field.

2.3. Forest Biomass

To effectively develop methods for assessing Aboveground Biomass (AGB), it is crucial to acquire on-site estimates of this biomass, commonly referred to as "in situ" measurements. The in-situ estimates serve as essential data for the calibration and validation of algorithms designed to calculate biomass. Additionally, field-collected data provide valuable information to estimate various tree characteristics, including basal area and the total aboveground and/or belowground biomass. In our analysis, the forest inventory data were utilized to predict the aboveground biomass within the transitional area using remotely sensed data and an artificial neural network. This prediction considers the equation proposed by [34] for our field samples located within the Amazon biome:

$$AGB = 0.0673 \times \left(Wd \times Ht \times DBH^2 \right) 0.976 \qquad (1)$$

where:

AGB = Aboveground Biomass (kg);
Wd = Basic wood density for each tree species (g.cm^3);

Ht = Total height (m);

DBH = Tree diameter at 1.3 m from the ground (cm).

In addition, we calculated AGB for our field samples located within the Cerrado biome using a specific allometric equation developed for the Cerrado environment [35], as follows:

$$AGB = 0.4913 + 0.0291 \times DGH^2 \times Ht \tag{2}$$

where:

AGB = Aboveground Biomass (kg);

DGH = The diameter of trees at their base (ground), specifically for trees with a diameter equal to or greater than 5 cm;

Ht = tree height.

The biomass was estimated by applying allometric equations and utilizing tree-specific variables for each subplot within the sample plots. Subsequently, these values were normalized per unit area to calculate the results in Tons per hectare (ton·ha^{-1}).

2.4. Sentinel-2 Imagery

We utilized images acquired by the MultiSpectral Instrument (MSI) sensor aboard the Sentinel-2 satellite, which provided spectral information about vegetation. This sensor captures the red band, crucial for characterizing vegetation due to the presence of chlorophyll in plants [36]. The satellite's spatial resolution varies according to the spectral bands: 10 m for visible and near-infrared bands, 20 m for red edges and other infrared bands, and 60 m for water vapor and cirrus bands. Sentinel-2 features 13 spectral bands ranging from 0.442 μm to 2.202 μm, with a revisit frequency of every five days [37].

In this analysis, we used a total of five Sentinel-2 scenes acquired from 2016 to 2021, all during August of each year, to minimize seasonal effects on the remotely sensed products. All scenes, covering the entire study area, were level 1c orthorectified TOA (Top of Atmosphere) reflectance and were acquired in the same year as the forest inventory data for 2018, 2020, and 2021. The only exception was the image acquired in 2016, which was used to relate to field data collected in 2014 because there were no Sentinel images available for that year. Subsequently, we retrieved vegetation indices from the Sentinel-2 images using the Google Earth Engine (GEE) platform. The Sentinel-2 scenes' IDs and acquisition dates are listed in Table 1.

Table 1. Sentinel 2A sensor MSI (Multispectral Instrument) scenes acquired through Google Earth Engine (GEE) and used for retrieving the vegetation indices applied in this analysis.

ID Sentinel-2A, Sensor MSI	Data
20160807T135257_T22LBL	7 August 2016
20180802T135108_T22LCL	2 August 2018
20200801T135115_T22LBL	1 August 2020
20200803T134216_T22LCL	3 August 2020
20210813T134211_T22LDL	13 August 2021

2.5. Vegetation Indices

In this analysis, we included various vegetation indices based on different spectral band combinations to leverage their potential sensitivity in capturing diverse vegetation characteristics and enhancing the relationship between vegetation indices and forest AGB. The indices utilized were NDVI (Normalized Difference Vegetation Index), EVI (Enhanced Vegetation Index), GNDVI (Green Normalized Difference Vegetation Index), AFRI (Aerosol Free Vegetation Index), MSAVI (Modified Soil-Adjusted Vegetation Index), NDRE (Normalized Difference Red Edge Index), SAVI (Soil-Adjusted Vegetation Index), and MSAVIaf (Modified Soil-Adjusted Vegetation Index aerosol free), all described as follows.

Normalized Difference Vegetation Index (NDVI)

The NDVI, developed by [38], is one of the most widely used vegetation indices. It relies on the relationship between the difference in reflectance in the near-infrared and red spectral bands and the sum of the reflectance of these two bands. This index enables the assessment of the photosynthetic activity of vegetation, with values ranging from −1 to 1. In contrast, water surfaces or clouds typically exhibit values below 0 [39]. Its definition is as follows:

$$\text{NDVI} = \frac{\rho\text{NIR} - \rho\text{Red}}{\rho\text{NIR} + \rho\text{Red}} \tag{3}$$

where ρNIR is the reflectance in the near-infrared spectral band and ρRed is the reflectance in the red spectral band.

Enhanced Vegetation Index (EVI)

The Enhanced Vegetation Index (EVI), developed by [40], aims to minimize atmospheric effects and improve NDVI sensitivity. It is notable for its sensitivity in analyses of canopy structural variations and densely forested areas [41]. Its definition is as follows:

$$\text{EVI} = \text{G} * \frac{(\rho\text{NIR} - \rho\text{Red})}{\rho\text{NIR} + (\text{C1} \times \rho\text{Red}) - (\text{C2} \times \rho\text{Blue}) + \text{L}} \tag{4}$$

where ρNIR is the reflectance in the near-infrared spectral band, ρRed is the reflectance in the red spectral band, ρBlue is the reflectance in the blue spectral band, G is the gain factor (default value: 2.5), L is the canopy background adjustment factor (default value: 1.0), and C1 and C2 are coefficients to correct aerosol effects.

Enhanced Vegetation Index 2 (EVI 2)

The Enhanced Vegetation Index 2 (EVI2), developed by [42], aims to achieve results similar to its original version (EVI) but using only two spectral bands (excluding the blue band). It proves particularly useful when utilizing high-quality remote sensing data with minimal atmospheric effects. Its definition is as follows:

$$\text{EVI2} = \text{G} * \frac{(\rho\text{NIR} - \rho\text{Red})}{\rho\text{NIR} + 2.4 \times \rho\text{Red} + 1} \tag{5}$$

where ρNIR is the reflectance in the near-infrared spectral band, ρRed is the reflectance in the red spectral band and, G is the gain factor (default value: 2.5).

GNDV (Green Normalized Difference Vegetation Index)

The Green Normalized Difference Vegetation Index (GNDVI), a modification of the NDVI developed by [43], is used to estimate chlorophyll content in vegetation. This makes it valuable for distinguishing between senescent vegetation and vegetation experiencing various degrees of water stress. GNDVI replaces the red band with the green band from NDVI, aiming to mitigate vegetation saturation effects in denser conditions [43]. Its definition is as follows:

$$\text{GNDVI} = \frac{\rho\text{NIR} - \rho\text{Green}}{\rho\text{NIR} + \rho\text{Green}} \tag{6}$$

where ρNIR is the reflectance in the near-infrared spectral band and ρGreen is the reflectance in the green spectral band.

AFRI (Aerosol Free Vegetation Index)

The Aerosol Free Vegetation Index (AFRI) was developed by [44] with the aim of mitigating the effects of aerosols and atmospheric disturbances on vegetation index calculations. This index has the capability to penetrate the atmosphere more effectively, providing accurate information about vegetation and other soil characteristics, even under adverse conditions such as forest fire situations with the presence of smoke [44]. One of the main

advantages of AFRI is its resilience to smoke interference in data acquisition, distinguishing it from other conventional indices [44]. Its definition is as follows:

$$\text{AFRI} = \frac{\rho\text{NIR} - 0.5\rho\text{SWIR}}{\rho\text{NIR} + 0.5\rho\text{SWIR}} \tag{7}$$

where: ρNIR is the reflectance in the near-infrared spectral band and ρSWIR is the reflectance in the shortwave infrared 1 band.

SAVI (Soil-Adjusted Vegetation Index)

The Soil-Adjusted Vegetation Index (SAVI) was developed by [45] with the aim of minimizing soil interference in canopy spectral measurements. This index allows for calibration so that variations in soil substrate are normalized in vegetation estimates [45]. Its definition is as follows:

$$\text{SAVI} = \frac{\rho\text{NIR} - \rho\text{Red}}{\rho\text{NIR} + \rho\text{Red} + \text{L}} \times (1 + \text{L}) \tag{8}$$

where ρNIR is the reflectance in the near-infrared spectral band, ρRed is the reflectance in the red spectral band, and L is the soil adjustment factor (default value = 0.5).

MSAVI (Modified Soil-Adjusted Vegetation Index)

The Modified Soil-Adjusted Vegetation Index (MSAVI), developed by [46], was designed to enhance its original version, SAVI. Both MSAVI and SAVI utilize soil adjustment factors [46]. MSAVI proves to be a more effective option in terms of time and resources, particularly in areas where vegetation density is uncertain or varies significantly [46]. Its definition is as follows:

$$\text{MSAVI} = \frac{\text{NIR} - \rho\text{Red}}{\text{NIR} + \rho\text{Red} + \text{L}} \times (1 + \text{L}) \tag{9}$$

where ρNIR is the reflectance in the near-infrared spectral band, ρRed is the reflectance in the red spectral band, and L is the soil adjustment calculated using Equation (10):

$$\text{L} = [(\rho\text{NIR} - \rho\text{Red}) \times \text{s} + 1 + \rho\text{NIR} + \rho\text{Red}]^2 - 8.0 \times \text{s} \times (\rho\text{NIR} - \rho\text{Red}) \tag{10}$$

where s = 1.2 (slope of the soil line calculated from surface reflectance at non-forested areas).

MSAVIaf (Modified Soil-Adjusted Vegetation Index aerosol free

The MSAVIaf was developed by [12] with the aim of reducing atmospheric effects on vegetation index estimations. It has been demonstrated to be more sensitive to vegetation variations than the Aerosol Free Vegetation Index under anomalous atmospheric conditions in the Amazon region [12]. Its definition is as follows:

$$\text{MSAVIaf} = \frac{\rho\text{NIR} - 0.5\rho\text{SWIR}}{\rho\text{NIR} + 0.5\rho\text{SWIR} + \text{L}} \times (1 + \text{L}) \tag{11}$$

where ρNIR is the reflectance in the near-infrared spectral band, ρSWIR is the reflectance in the shortwave infrared spectral band (central wavelength: 1.6137 μm), and L is the soil adjustment factor, calculated as previously presented (Equation (10)).

NDRE (Normalized Difference Red Edge Index)

The Normalized Difference Red Edge Index (NDRE), developed by [47], was designed to measure plant physiological parameters, particularly those associated with chlorophyll content, nitrogen concentration, and canopy structure. It can be applied in identifying and classifying crops and land covers [48]. Its definition is as follows:

$$NDRE = \frac{\rho NIR - \rho Rededge}{\rho NIR + \rho Rededge} \tag{12}$$

where ρNIR is the reflectance in the near-infrared spectral band and $\rho Rededge$ is the reflectance in the red edge spectral band (central wavelength: 0.704 μm).

2.6. Correlation Analysis

The evaluation of vegetation indices for predicting biomass in our study area was performed by analyzing the correlation matrix between the nine indices retrieved from remotely sensed data and the field-measured biomass. To assess the normality of biomass and vegetation index datasets, we applied the Shapiro–Wilk test.

2.7. Modeling of the Artificial Neural Network (ANN)

In this study, we employed a Multilayer Perceptron (MLP) type of Artificial Neural Network (ANN), adjusted and trained using Statistica software (STATSOFT), version 12, to estimate forest biomass using the field-sampling data of LABEV-UNEMAT. The software utilizes the Intelligent Problem Solver (IPS) tool to optimize the network architecture, including the number of layers, neurons, and cycles to achieve more efficient results [49]. Training is conducted using the Broyden–Fletcher–Goldfarb–Shanno quasi-Newton algorithm by IPS for neural network processing, which has been shown to be highly capable of solving optimization and prediction problems, in addition to being the most popular quasi-Newton method [50–52].

In this analysis, the input layer of the neural network consisted of both categorical and numerical variables. The categorical variable pertained to the two types of strata in the study area: Perennial Seasonal Forest and typical Cerrado. The numerical variables included the vegetation indices NDVI, EVI, EVI2, GNDV, AFRI, MSAVI, NDRE, SAVI, and MSAVIaf. The hidden layer comprised 'n' neurons, while the output layer consisted of a single neuron responsible for estimating AGB.

To train the Artificial Neural Networks (ANNs), we selected 40 subsamples, representing 70% of the total 60 field-demarked subsamples during the inventories. The remaining 20 subsamples were used for result validation and testing. Multilayer Perceptron (MLP) ANNs calculate the weighted arithmetic mean of these inputs [53], and in this case, were activated by an exponential function. To assess the performance of the models developed using ANNs, we considered the parameters of the correlation coefficient (R) and root mean square error (RMSE). These coefficients have been utilized in other research involving ANNs to predict solar energy using weather data, as demonstrated by [54].

For the validation of the performance of the best ANNs, we conducted statistical analyses using Student's *t*-tests. To determine whether there was AGB underestimation or overestimation, we calculated the aggregate difference in percentage terms (AD%). The Aggregate Difference (AD%) corresponds to the difference between the sum of the observed values and the sum of the estimated values, in percentage, obtained by the following expression:

$$AD\% = \frac{\sum_{i=i}^{n} y_i - \sum_{1=1}^{n} \hat{y}_i}{\sum_{i=1}^{n} y_i} \times 100 \tag{13}$$

where AD% = Aggregate Difference; y_i = observed values; $\hat{y}_i$ = estimated values; and n = number of observations.

The statistical analyses were performed using Microsoft Excel software, Microsoft Office 365, Version 2408.

3. Results

3.1. Vegetation Inventory

The results in Table 2 show significant differences in the assessed variables, highlighting substantial variation in dendrometric characteristics between the Cerrado and the Amazon plots. Notably, trees in forest plots showed an average aboveground biomass

approximately eight times higher than those in Cerrado plots. This difference can be attributed to wider trunks (37% larger in the forest compared to Cerrado) and trees that were approximately three times taller in the forest. Interestingly, the average wood density was quite similar in both formations (Amazon and Cerrado) within the study area (Table 2).

Table 2. Dendrometric variables calculated from forest inventory for the areas of the Plant Laboratory (LABEV) Mato Grosso State University (UNEMAT) plots.

Cerrado Plots					Forest Plots				
Statistics	**DBH**	**Ht**	**WD**	**AGB**	**Statistics**	**DBH**	**Ht**	**WD**	**AGB**
Minimum	10.0	4.0	0.41	14.35	Minimum	10.0	10.0	0.20	66.56
Maximum	39.0	13.0	0.84	23.49	Maximum	93.2	30.0	1.09	331.38
Mean	13.81	6.64	0.66	18.38	Mean	19.12	13.99	0.67	146.86
Variance	15.38	1.65	0.01	10.97	Variance	105.19	17.5	0.019	2572.76
Deviation	3.92	1.28	0.10	3.31	Deviation	10.25	4.18	0.14	50.72
CV (%)	28.4	19.33	15.43	18.02	CV (%)	53.64	29.9	20.4	34.54

Where DBH = diameter at breast height (cm), Ht = total height (m), WD = average wood density (g·cm^3), AGB = aboveground biomass (Ton·ha^{-1}), and CV (%) = Coefficient of Variation (%).

The biomass measurements in the Forest samples showed themselves to be statistically consistent, showing an average value of 146.84 t·ha^{-1}. When examining a forest fragment located on the southern edge of the study area, we observed biomass variability ranging from 155 to 195 t·ha^{-1}.

3.2. Correlation Analysis of Biomass and Vegetation Indices

In this study, we created a mosaic of the Sentinel-2 images acquired in August 2019 to retrieve the vegetation indices for the study area (Table 3).

Table 3. Average of the independent variables in the study area. AFRI = Aerosol Free Vegetation Index; EVI = Enhanced Vegetation Index; GNDVI = Green Normalized Difference Index; EVI2 = Enhanced Vegetation Index–2; MSAVIaf = Modified Soil-Adjusted Vegetation Index aerosol free; MSAVI = Modified Soil-Adjusted Vegetation Index; NDVI = Normalized Difference Vegetation Index; NDRE = Normalized Difference Red Edge Index; SAVI = Soil-Adjusted Vegetation Index.

Vegetation Indices	Average
AFRI	0.564
EVI	0.571
GNDVI	0.569
EVI2	0.392
MSAVIaf	0.324
MSAVI	0.545
NDRE	0.515
NDVI	0.697
SAVI	0.408

The Shapiro–Wilk test indicated non-normality of the analyzed variables (vegetation indices and biomass). We then applied the Spearman correlation matrix, recommended for non-parametric data analysis. The Spearman correlation results indicated positive and significant correlations ($\alpha < 0.05$) among aboveground biomass and all vegetation indices, as well as among the vegetation indices themselves (Table 4).

Table 4. Spearman's correlation matrix was used to analyze the relationship between aboveground biomass and vegetation indices of the study area.

	AFRI	EVI	EVI2	GNDVI	MSAVIaf	MSAVI	NDRE	NDVI	SAVI	Biomass
AFRI	1									
EVI	0.887 **	1								
EVI2	0.875 **	0.963 **	1							
GNDVI	0.897 **	0.902 **	0.951 **	1						
MSAVIaf	0.948 **	0.963 **	0.969 **	0.950 **	1					
MSAVI	0.868 **	0.975 **	0.933 **	0.866 **	0.942 **	1				
NDRE	0.954 **	0.853 **	0.882 **	0.921 **	0.922 **	0.825 **	1			
NDVI	0.972 **	0.909 **	0.902 **	0.907 **	0.943 **	0.887 **	0.971 **	1		
SAVI	0.831 **	0.853 **	0.889 **	0.863 **	0.907 **	0.850 **	0.786 **	0.820 **	1	
Biomass	0.469 *	0.443 *	0.532 **	0.621 **	0.555 **	0.404 *	0.509 **	0.466 *	0.594 **	1

** Significant at $\alpha < 0.01$; * Significant at $\alpha < 0.05$. Where: AFRI = Aerosol Free Vegetation Index; EVI = Enhanced Vegetation Index; GNDVI = Green Normalized Difference Index; MSAVIaf = Modified Soil-Adjusted Vegetation Index aerosol resistant; MSAVI = Modified Soil-Adjusted Vegetation Index; NDVI = Normalized Difference Vegetation Index; NDRE = Normalized Difference Red Edge Index; SAVI = Soil-Adjusted Vegetation Index.

Based on the results of the correlation matrix, we subsequently proceeded with a stepwise regression analysis to select our predictive variables (vegetation indices). The stepwise technique involves adding or removing independent variables from the model one at a time, based on specific criteria such as the *p*-value. This procedure is implemented automatically to identify a subset of variables that are most relevant for predicting the dependent variable (in this case, aboveground biomass).

In contrast to the correlation matrix results, this complementary stepwise regression analysis found that the AFRI, EVI, and GNDVI indices (Figure 2) were the most suitable (highest statistical significance at $\alpha < 0.05$) vegetation indices to be used as input neurons for the ANN modeling. It is likely that retrieving vegetation indices from different spectral band combinations (near-infrared, middle infrared, red, and blue bands) greatly contributed to increasing their sensitivity and capturing aboveground biomass variation in the study area.

3.3. Biomass Modeling

After training the artificial neural networks (ANNs) with the most suitable independent variables (AFRI, EVI, and GNDVI) indicated by the stepwise regression analysis, we selected the top five performing ANNs based on correlation coefficients (r) exceeding 0.90 and validation errors less than 16%. The selected ANN showed low variation between training, selection, and evaluation indices, demonstrating stability during the training process [55]. An in-depth analysis of fit and accuracy statistics revealed that Neural Network 1 showed the strongest predictive capability for aboveground biomass, as indicated by the RMSE% values in Table 5.

Additionally, the results provided by Neural Network 1 indicated a satisfactory distribution of residuals (Figure 3—B1 training, B2 testing, and B3 validation) and accurate, consistent predictions of aboveground biomass (Figure 3—A1 training, A2 testing, and A3 validation) in the study area. The model showed a good fit, which indicates that it minimized the differences between observed and predicted values without significant bias.
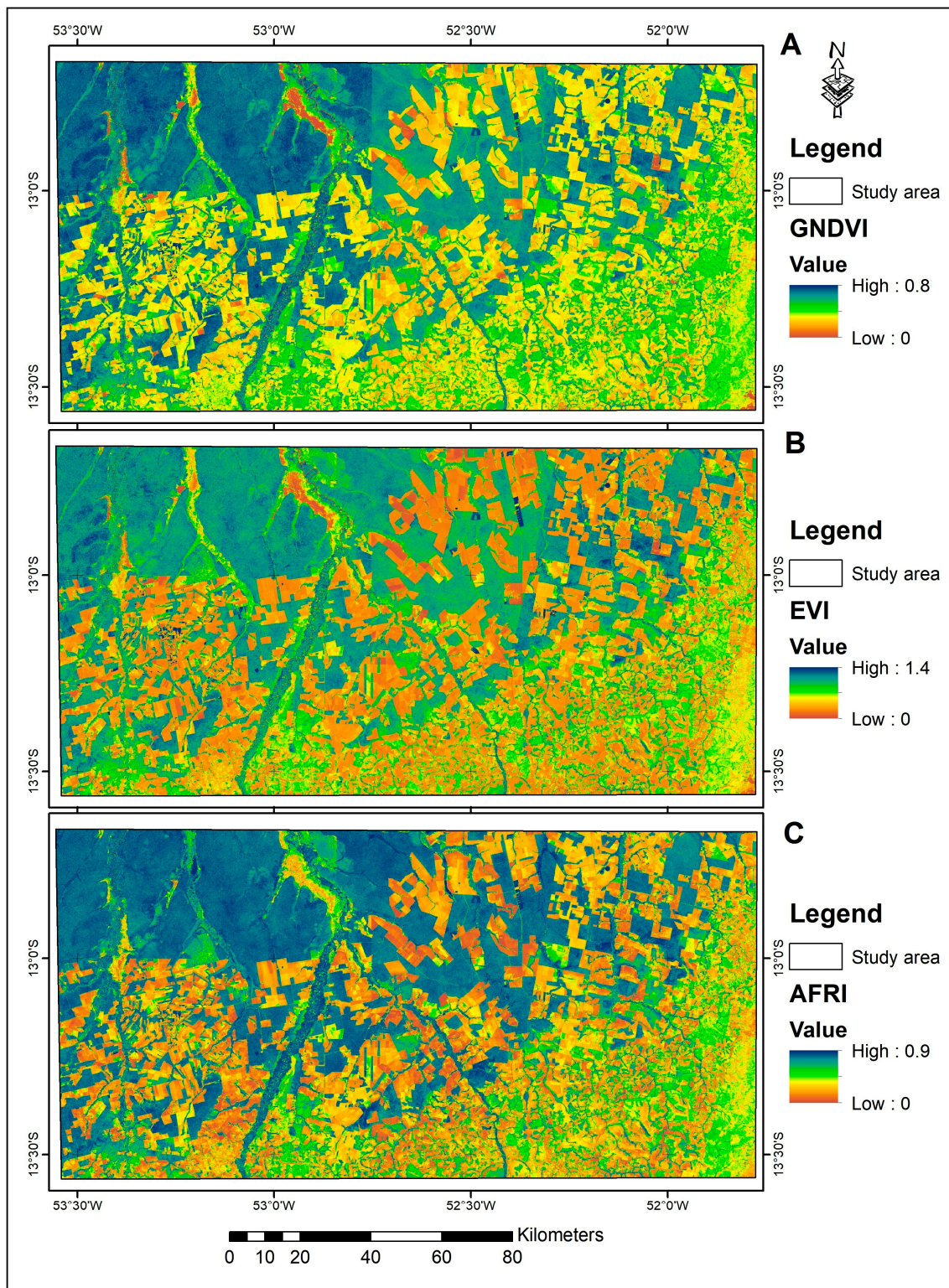
Figure 2. Vegetation indices ((**A**) = Green Normalized Vegetation Index—GNDV; (**B**) = Enhanced Vegetation Index—EVI; and (**C**) = Aerosol Free Vegetation Index—AFRI) retrieved from Sentinel-2 imagery acquired in August 2016, 2018, 2020, and 2021 covering the entire study region.

Table 5. Accuracy statistics of the selected artificial neural networks (ANNs) for prediction of aboveground biomass for the LABEV-UNEMAT plots located in the Cerrado/Amazon ecotone.

ANN	Architecture	N° of Cycles	Activation Hidden	Activation Output	Adjustment		Validation		Test	
					RMSE%	R	RMSE%	R	RMSE%	R
1	MLP 3-12-1	860	Tang	Tang	18.09	0.93	15.76	0.94	15.92	0.94
2	MLP 3-11-1	1630	Logistic	Exponential	19.44	0.92	16.09	0.93	16.18	0.93
3	MLP 3-8-1	910	Logistic	Identity	19.77	0.92	16.41	0.93	16.92	0.93
4	MLP 3-13-1	950	Tang	Exponential	19.53	0.92	16.62	0.93	16.91	0.93
5	MLP 3-11-1	670	Logistic	Identity	20.19	0.91	17.91	0.91	17.12	0.91

ANN	Predictor variables	Neurons per layer			Adjust			
		Input	Hidden	Output	TI	SI	AI	Algorithm
1	AFRI, EVI, GNDVI	3	12	1	0.08	0.08	0.09	BFGS
2	AFRI, EVI, GNDVI	3	9	1	0.10	0.11	0.12	BFGS
3	AFRI, EVI, GNDVI	3	5	1	0.10	0.12	0.13	BFGS
4	AFRI, EVI, GNDVI	3	13	1	0.11	0.13	0.10	BFGS
5	AFRI, EVI, GNDVI	3	7	1	0.13	0.15	0.17	BFGS

ANN = artificial neural network; MLP = Multilayer perceptron; RMSE% = Root Mean Square Error Percentage; R = correlation between observed and estimated values; TI= Training indices (network definition), SI = Selection Indices of training stop, AI = Assessment Indices (quality assessment of trained network); BFGS = Broyden–Fletcher–Goldfarb–Shannon.

The accuracy of aboveground biomass estimates is a crucial indicator of the model's effectiveness. The architecture of ANN-1 (Figure 4) comprises three layers: the input layer with three neurons representing predictor variables (EVI, AFRI, and GNDVI), a hidden layer of 12 neurons for data processing activated using a tangential function, and an output layer representing the variable of interest (AGB) activated with a logistic function.

3.4. Statistical Analysis

The Student's *t*-test is a statistical tool used to determine whether there is a significant difference between the means of two independent samples. In this test, we formulate a null hypothesis (H0) asserting that there is no difference between the means of the two samples, and an alternative hypothesis (H1) suggesting that there is a significant difference between them. Following the *t*-test, we compute a *p*-value. If the *p*-value falls below the chosen significance level (typically 0.05), it indicates statistical evidence to reject the null hypothesis in favor of the alternative hypothesis. In simpler terms, this means there is a significant difference between the means of the observed values compared to the estimated values. Conversely, if the *p*-value exceeds the significance level, there is not enough evidence to reject the null hypothesis, indicating no statistically significant difference between the observed mean values and the estimated values.

The *p*-value is a statistical measure that aids in interpreting the results of a hypothesis test in statistics. It indicates the probability of obtaining a result as extreme or more extreme than the one observed, assuming the null hypothesis is true. The null hypothesis typically states that there is no effect or difference between the compared groups, while the alternative hypothesis suggests the opposite. In short, the *p*-value provides a way to quantify how much the results support or refute the null hypothesis.

In this analysis, the application of the Student's *t*-test revealed that the calculated *p*-value for the selected neural network was greater than the established significance level ($\alpha = 0.05$), specifically $p = 0.952$. This indicates that there is insufficient statistical evidence to reject the null hypothesis, which shows no differences between the observed and predicted values by the neural network for the validation plots.
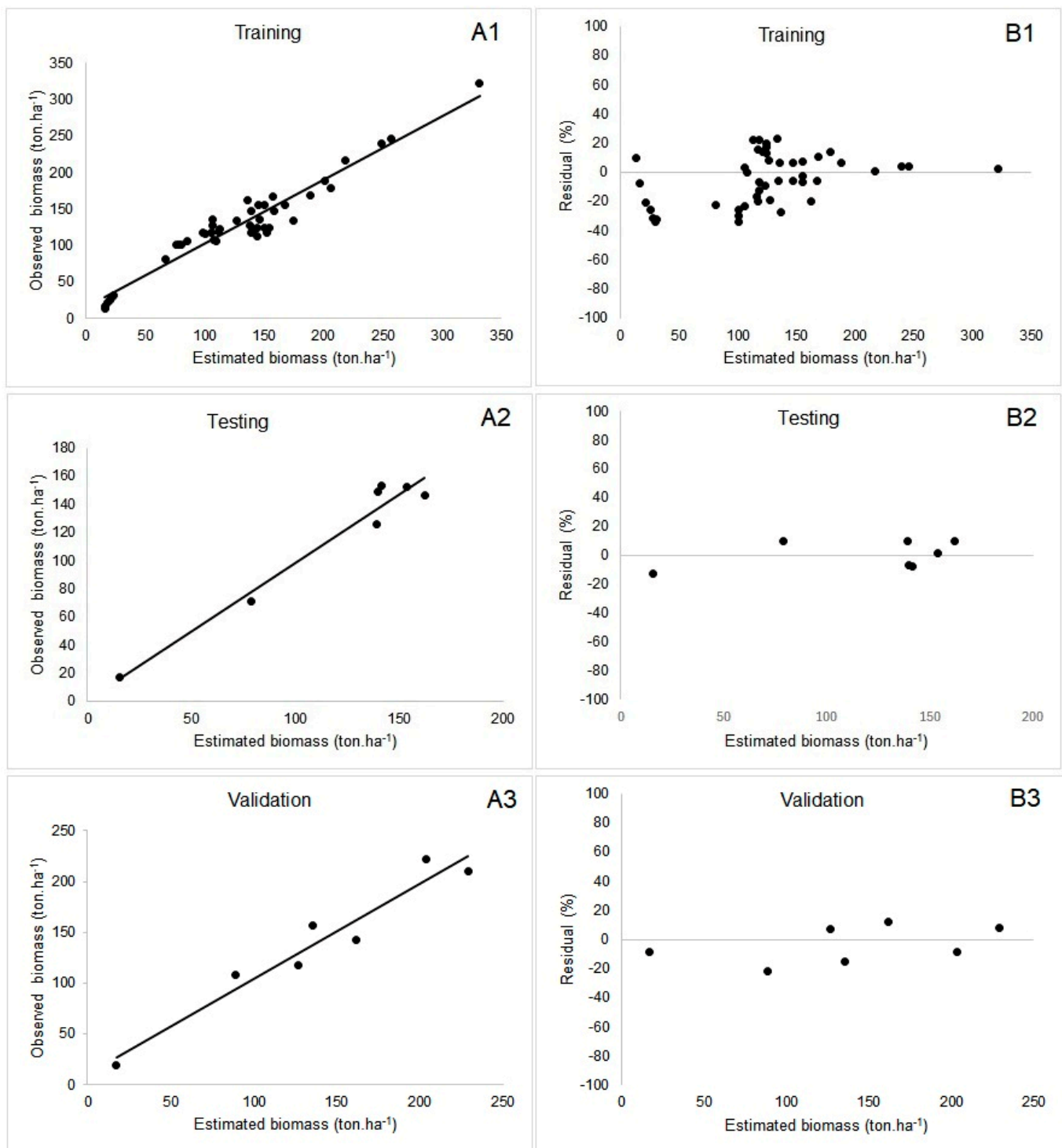
Figure 3. Observed and estimated aboveground biomass in the study area ((**A1**) = Training; (**A2**) = Testing; (**A3**) = Validation) and distribution of residuals ((**B1**) = Training; (**B2**) = Testing; (**B3**) = Validation) for Artificial Neural Network 1 (ANN-1).
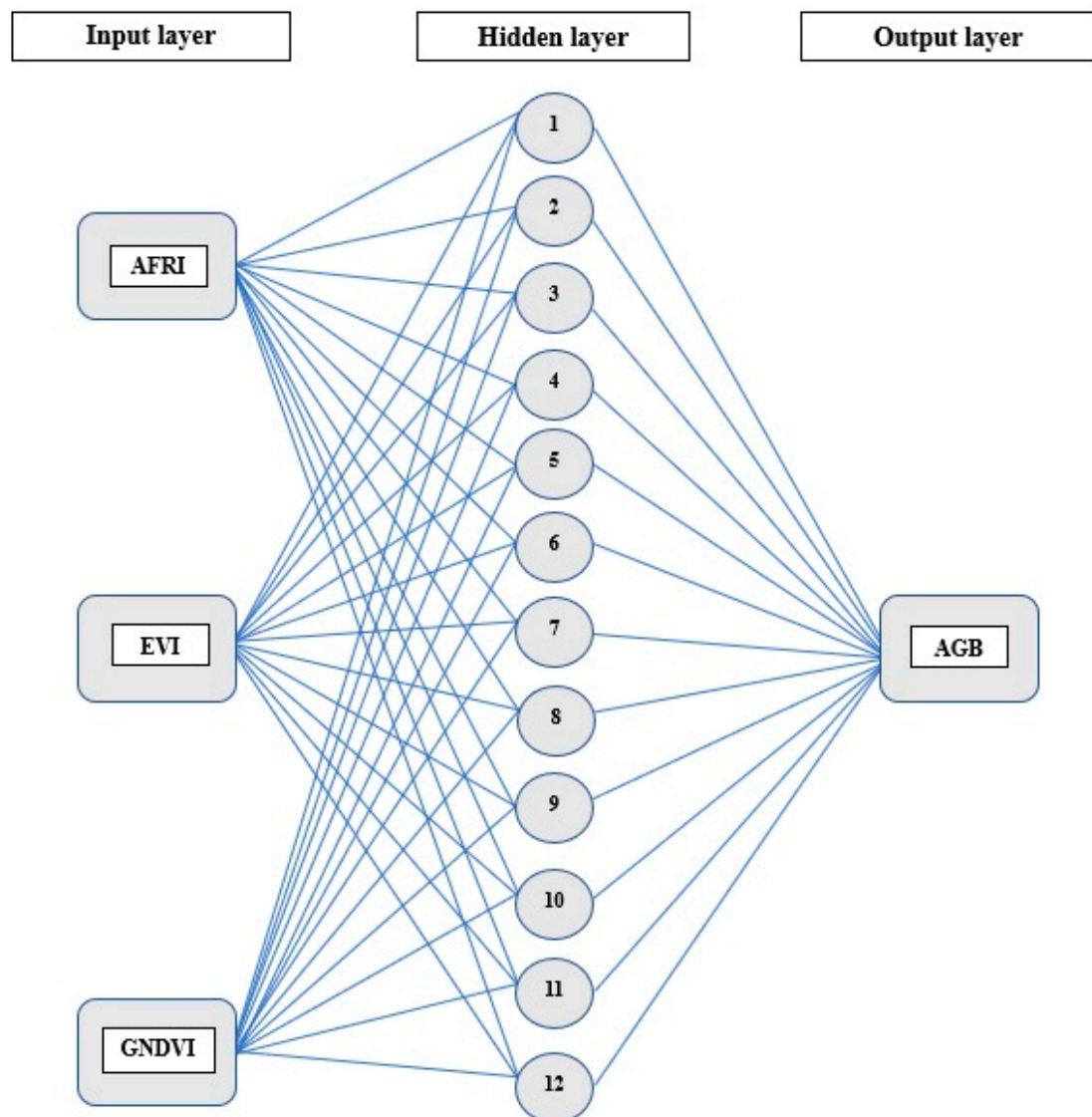
Figure 4. Architecture of ANN-1 selected for the prediction of aboveground biomass for the study area.

Additionally, the Aggregate Difference (AD) analysis indicated a slight tendency to overestimate the values predicted by the neural network, with a deviation of -0.1637%. Nevertheless, these results align with the accuracy of the information obtained during the ANN training process, confirming its proficiency in providing precise estimates for Aboveground Biomass (AGB). Consequently, these findings suggest that the ANN-generated estimates are both accurate and dependable for predicting AGB in areas of biome transition.

3.5. Analyzing the Spatial Distribution of Biomass

Based on the results obtained from the training of the neural networks, we were able to extend our estimates of AGB across the entire area covered by native vegetation in this study region. Consequently, the total biomass of the study area, considering the land use and land cover of native vegetation, was estimated at 109,118,121 tons. The most common AGB values in the study area were in the range of 0 to 50 t·ha^{-1}, followed by the range of 100 to 150 t·ha^{-1} (Figure 5).
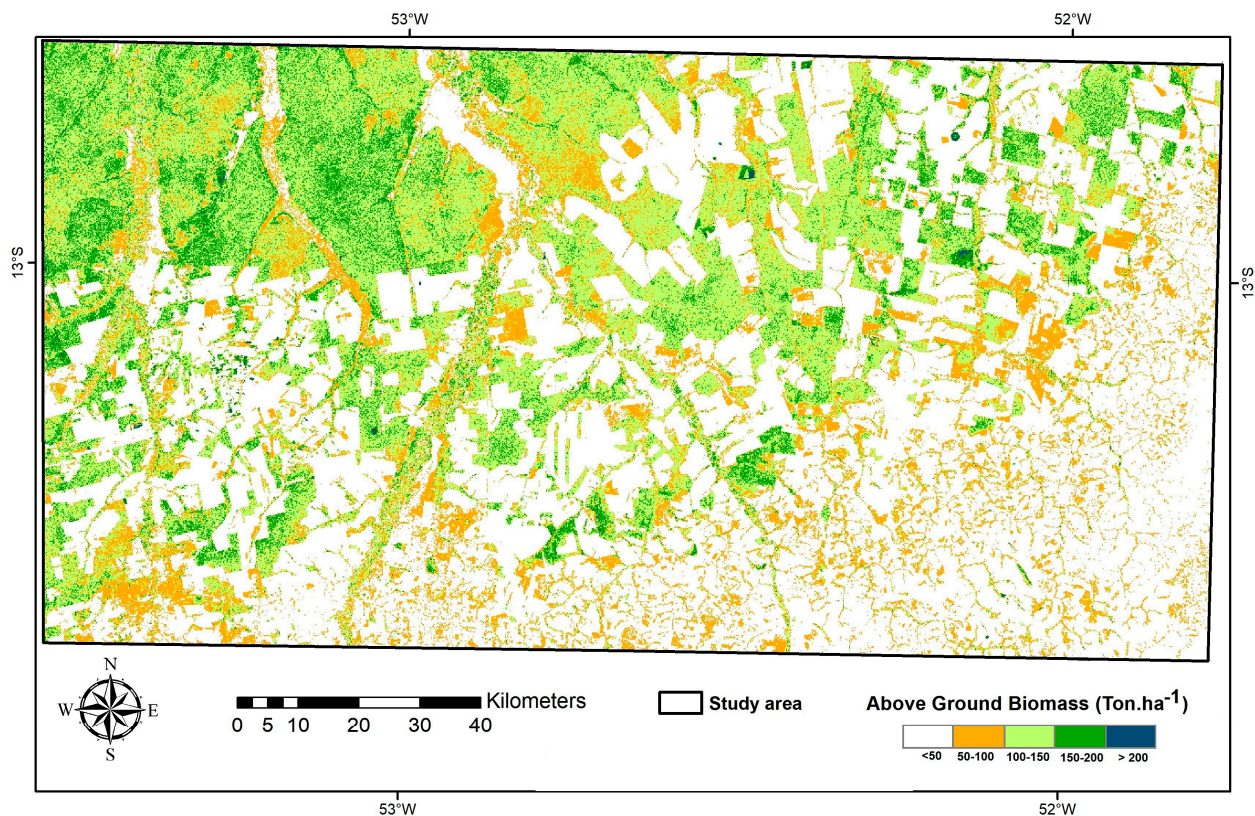
Figure 5. Spatial distribution of forest biomass is estimated for the Amazon–Cerrado ecotone zone. Darker areas indicate higher aboveground biomass, while lighter areas indicate lower biomass.

4. Discussion

4.1. Forest Biomass and Land Use and Land Cover

The amount of aboveground biomass (AGB) varies significantly within native forest formations in the study area, predominantly ranging between 100 and 150 t·ha^{-1}, followed by classes of 0–50 and 50–100 t·ha^{-1} occupied by savanna and transitional forest formations. This variation can be attributed to various factors, including climatic, geological, and soil conditions, as well as distinct previous vegetation disturbances and land use patterns in the study region [56].

The use of ANNs (Artificial Neural Networks) proved effective in estimating biomass per unit area while eliminating the basic assumptions of conventional mathematical modeling, such as normality and linearity of forest attributes [57]. These attributes often require various mathematical transformations for traditional modeling, which can result in a loss of quality and selection of models, leading to biased estimates of the variable of interest.

One hypothesis explaining this relatively low range of total AGB in the study region is the impact of anthropogenic activities, particularly agriculture, selective logging, fire, and livestock farming. These disturbances can increase edge effects and forest degradation, especially when caused by selective logging activities and forest fires [12,58]. Addressing this requires the definition and implementation of public policies to enforce sustainable land use management, conservation of natural ecosystems, environmental law enforcement, climate awareness, and fire prevention measures [59].

The increase in soybean cultivation over the last few decades has had severe impacts on natural ecosystems and the natural landscape in the study region. These impacts may directly lead to decreased rainfall, increased land surface temperatures, and soil and water contamination due to pesticides and chemical fertilizers. Additionally, pastures cover nearly 16 percent of the study area and can cause significant environmental impacts, including greenhouse gas emissions and soil and water degradation [60].

In summary, land use and land cover in the study area comprise a complex landscape mix of agricultural, livestock, and forestry activities. The potential environmental and social impacts associated with these activities add complexity to achieving a balance between economic development and the conservation of natural resources [59,61].

4.2. Selection of Independent Variables (Vegetation Indices)

The AFRI, EVI, and GNDVI indices showed the most significant correlation with aboveground biomass in our study area. The high observed correlations among vegetation indices and biomass are likely due to a combination of a broader range of spectral bands within the electromagnetic spectrum (middle- and near-infrared, red, green, and blue) required to retrieve these three vegetation indices from Sentinel-2 imagery, compared to other assessed indices in this analysis. The broader range of spectral bands increases their sensitivity to capture subtle vegetation variations and changes. Such sensitivity is crucial when using remotely sensed data, especially in ecotone regions that exhibit high vegetation variability and complexity.

The significant correlations observed between Green Reflectance (GREEN), as represented by the GNDVI (Green Normalized Difference Vegetation Index), and biomass in the study area can be attributed to variations in chlorophyll and anthocyanin content in the leaves [62]. These factors are closely related to vegetation development and maturity [62]. The Green Vegetation Index (GVI) and the Green Normalized Difference Vegetation Index (GNDVI), derived from reflectance equations, exhibit stronger correlations with nitrogen content in forest biomass leaves compared to the Ratio Vegetation Index (RVI) and the Red Normalized Difference Vegetation Index (RNDVI), indicating a greater sensitivity to variations in vegetation [63]. The combination of green and infrared bands plays an important role in aboveground biomass analysis, serving as critical descriptors in this index and providing dependable and precise information on biomass quantities at specific locations [64].

The use of the Near-Infrared (NIR) and Shortwave Spectrum (SWIR) bands in calculating the AFRI index has demonstrated efficiency in monitoring vegetation water content and dry biomass, particularly in regions with sparse vegetation [65]. Moreover, the AFRI index showed a stronger correlation with biomass in the study area located within an ecotone region between forest and savanna.

Commonly used vegetation indices such as NDVI and EVI have been applied worldwide to assess vegetation health. However, these indices are influenced by various factors, including terrain topography [66]. Our study showed that the soil adjustment factor "L" may heavily impact EVI results compared to NDVI, making EVI more sensitive to topographical conditions. This sensitivity is particularly critical in hilly terrain, where topographic effects can significantly affect vegetation indices with a simple band-ratio format, such as NDVI.

The choice of satellites for spectral data collection can influence the accuracy of biomass estimation. Nevertheless, our analysis found that Sentinel-2 satellite images were suitable for our study. The authors of [67] reported that the quality of MSI/Sentinel-2 sensor images, particularly in bands with a 10 m resolution, highlights the utility of this satellite for vegetation assessment research, especially when compared to aerial sensors with a spatial resolution of 0.13 m.

4.3. Training the Neural Networks

Our results indicate that the trained Artificial Neural Networks (ANNs) showed a satisfactory fit and high-accuracy statistics. The correlation coefficient (R) consistently equaled or exceeded 0.9, and the root mean square estimation (RMSE) errors remained below 14%. Among the five trained networks, Network 1 outperformed the others with an R^2 of 0.94 and an RMSE% of 10.76, making it a promising choice for the intended application. These findings underscore the feasibility of biomass estimation through remote sensing in natural forests.

The use of ANNs is effective in estimating biomass per unit area and does not require the basic assumptions of conventional mathematical modeling, such as normality and linearity of forest attributes [68]. These attributes often require various mathematical transformations for traditional modeling, which can result in lower model quality and selection capability, leading to biased and less accurate estimates of the variable of interest.

The authors of [68] also yielded positive results in estimating the components of total biomass, with an R^2 of 0.97 and an RMSE% of 25.04. Furthermore, the simulation of terrain elevation data along ICESat-2 and Landsat satellite profiles demonstrated significant potential for generating a forest biomass estimation product, achieving an R^2 of 0.66 [69].

These results align with the findings of our research, highlighting the robust performance of the models developed for tree biomass estimation. To assess the predictive capacity of the selected Artificial Neural Network, we examined the relationship between observed and predicted values. When analyzing the distribution of ANN errors, we observed that most errors fell within the −1.5% to −12% and 0% to 10% ranges. Additionally, errors exceeding the ±16% threshold were infrequent. Moreover, it was determined that a training dataset size of approximately 60 subplots or fewer was sufficient to achieve a good fit with the linear functional model.

5. Conclusions

Our research findings indicate that the combination of various vegetation indices integrating different spectral bands, such as EVI, AFRI, and GNDVI, with a Multilayer Perceptron Artificial Neural Network has led to more efficient and precise estimation of aboveground biomass in our study area. This approach facilitated the generation of high-resolution biomass distribution maps and provided a cost-effective and time-saving alternative to traditional forest inventories. Accurate estimates of forest biomass are crucial for understanding vegetation dynamics and ecological processes, as well as for formulating effective forest resource management policies. Additionally, our study results are valuable for forest biomass monitoring, including the assessment of environmental services and the formulation of conservation strategies for protected areas and indigenous territories. The advanced knowledge of forest biomass can also support sustainable forest management practices and enable the prediction of impacts of land use and land cover changes on forest biomass. Alternative approaches, such as deep learning and machine learning methods, could prove effective for estimating aboveground biomass in tropical regions and should be explored in future research endeavors.

References

1. Lima, S.K.R.; Coêlho, A.G.; Lucarini, M.; Durazzo, A.; Arcanjo, D.D.R. The *Platonia insignis* Mart. as the Promising Brazilian 'Amazon Gold': The State-of-the-Art and Prospects. *Agriculture* **2022**, *12*, 1827. [CrossRef]
2. Marques, E.Q.; Marimon-Junior, B.H.; Marimon, B.S.; Matricardi, E.A.T.; Mews, H.A.; Colli, G.R. Redefining the Cerrado–Amazonia transition: Implications for conservation. *Biodivers. Conserv.* **2020**, *29*, 1501–1517. [CrossRef]
3. Ratter, J.A.; Richards, P.W.; Argent, G.; Gifford, D.R. Observations on the vegetation of northeastern Mato Grosso: I. The woody vegetation types of the Xavantina-Cachimbo Expedition area. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* **1973**, *266*, 449–492.
4. Marimon, B.S.; Lima, E.D.S.; Duarte, T.G.; Chieregatto, L.C.; Ratter, J.A. Observations on the vegetation of northeastern Mato Grosso, Brazil. IV. An analysis of the Cerrado-Amazonian Forest ecotone. *Edinb. J. Bot.* **2006**, *63*, 323–341. [CrossRef]
5. Balch, J.K.; Nepstad, D.C.; Curran, L.M.; Brando, P.M.; Portela, O.; Guilherme, P.; Reuning-Scherer, J.D.; de Carvalho, O. Size, species, and fire behavior predict tree and liana mortality from experimental burns in the Brazilian Amazon. *For. Ecol. Manag.* **2011**, *261*, 68–77. [CrossRef]
6. MAPBIOMAS. News: In 38 Years, Brazil Has Lost 15% of Its Natural Forests. Available online: https://brasil.mapbiomas.org/en/noticias/ (accessed on 8 March 2024).
7. Nogueira, D.S.; Marimon, B.S.; Marimon-Junior, B.H.; Oliveira, E.A.; Morandi, P.; Reis, S.M.; Elias, F.; Neves, E.C.; Feldpausch, T.R.; Lloyd, J.; et al. Impacts of Fire on Forest Biomass Dynamics at the Southern Amazon Edge. *Environ. Conserv.* **2019**, *46*, 285–292. [CrossRef]
8. Brown, S.; Sathaye, J.; Cannell, M.; Kauppi, P.E. Mitigation of carbon emissions to the atmosphere by forest management. *Commonw. For. Rev.* **1996**, *75*, 80–91.
9. Houghton, R.A.; Hall, F.; Goetz, S.J. Importance of biomass in the global carbon cycle. *J. Geophys. Res. Biogeosci.* **2009**, *114*. [CrossRef]
10. Lu, D. Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon. *Int. J. Remote Sens.* **2005**, *26*, 2509–2525. [CrossRef]
11. Li, Y.; Li, M.; Li, C.; Liu, Z. Forest aboveground biomass estimation using Landsat 8 and Sentinel-1A data with machine learning algorithms. *Sci. Rep.* **2020**, *10*, 9952. [CrossRef]
12. Matricardi, E.A.T.; Skole, D.L.; Pedlowski, M.A.; Chomentowski, W.; Fernandes, L.C. Assessment of tropical forest degradation by selective logging and fire using Landsat imagery. *Remote Sens. Environ.* **2010**, *114*, 1117–1129. [CrossRef]
13. Silva, F. Sensoriamento Remoto para Detecção de Queimadas no Cerrado Maranhense: Uma Aplicação no Parque Estadual do Mirador. *Rev. Geogr. Acad.* **2019**, *13*, 90–105.
14. Hall, R.J.; Skakun, R.S.; Arsenault, E.J.; Case, B.S. Modeling forest stand structure attributes using Landsat ETM+ data: Application to mapping of aboveground biomass and stand volume. *For. Ecol. Manag.* **2006**, *225*, 378–390. [CrossRef]
15. Powell, S.L.; Cohen, W.B.; Healey, S.P.; Kennedy, R.E.; Moisen, G.G.; Pierce, K.B.; Ohmann, J.L. Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. *Remote Sens. Environ.* **2010**, *114*, 1053–1068. [CrossRef]
16. Dube, T.; Mutanga, O. Evaluating the utility of the medium-spatial resolution Landsat 8 multispectral sensor in quantifying aboveground biomass in uMgeni catchment, South Africa. *ISPRS J. Photogramm. Remote Sens.* **2015**, *101*, 36–46. [CrossRef]
17. Fremout, T.; De Vinatea, J.C.; Thomas, E.; Huaman-Zambrano, W.; Salazar-Villegas, M.; la Fuente, D.L.D.; Bernardino, P.N.; Atkinson, R.; Csaplovics, E.; Muys, B. Site-specific scaling of remote sensing-based estimates of woody cover and aboveground biomass for mapping long-term tropical dry forest degradation status. *Remote Sens. Environ.* **2022**, *276*, 113040. [CrossRef]
18. Xiao, X.; Zhang, Q.; Braswell, B.; Urbanski, S.; Boles, S.; Wofsy, S.; Moore, B., III; Ojima, D. Modeling gross primary production of temperate deciduous broadleaf forest using satellite images and climate data. *Remote Sens. Environ.* **2004**, *91*, 256–270. [CrossRef]
19. Patel, N.R.; Pandya, M.R.; Patel, N.K. Estimation of Biomass of Wheat Using Vegetation Indices. *Int. J. Curr. Microbiol. Appl. Sci.* **2016**, *5*, 288–296.
20. Song, X.; Li, L.; Zhuo, W.; Wu, C. Estimating forest aboveground biomass by combining Landsat and MODIS data: A case study for the Sierra National Forest, California, USA. *Remote Sens.* **2014**, *6*, 2107–2136.
21. Chen, X.; Yang, K.; Ma, J.; Jiang, K.; Gu, X.; Peng, L. Aboveground biomass inversion based on Object-Oriented Classification and Pearson-mRMR-Machine Learning Model. *Remote Sens.* **2024**, *16*, 1537. [CrossRef]
22. Lyu, G.; Wang, X.; Huang, X.; Xu, J.; Li, S.; Cui, G.; Huang, H. Toward a more robust estimation of forest biomass Carbon stock and Carbon Sink in mountainous region: A case study in Tibet, China. *Remote Sens.* **2024**, *16*, 1481. [CrossRef]
23. Costa, A.C.; Pinto, J.R.; Miguel, E.P.; Xavier, G.D.; Marimon, B.H.; Aparecido Trondoli Matricardi, E. Artificial intelligence tools and vegetation indices combined to estimate aboveground biomass in tropical forests. *J. Appl. Remote Sens.* **2023**, *17*, 024512. [CrossRef]
24. Tian, L.; Wu, X.; Tao, Y.; Li, M.; Qian, C.; Liao, L.; Fu, W. Review of remote sensing-based methods for forest aboveground biomass estimation: Progress, challenges, and prospects. *Forests* **2023**, *14*, 1086. [CrossRef]
25. Novo, E.; Ferreira, L.G.; Barbosa, C.; Carvalho, C.; Sano, E.E.; Shimabukuro, Y.E.; Miura, T. Advanced remote sensing techniques for global changes and Amazon ecosystem functioning studies. *Acta Amaz.* **2005**, *35*, 259–272. [CrossRef]
26. Bertier, F.; Silva, R.; Nora, G. Fire in the Woods, Danger for Real? Community Considerations About Using Fire in the Cerrado of Mato Grosso. *Rev. Rencima Edição Espec.* **2020**, *11*, 144–157.

27. Ivanauskas, M.M.; Ivanauskas, N.M.; Monteiro, R.; Rodrigues, R.R. Composição florística de trechos florestais na borda sul-amazônica. *Acta Amaz.* **2004**, *34*, 399–413. [CrossRef]

28. BRASIL Ministério das Minas e Energia, Secretaria Geral. *Projeto RADAMBRASIL—Levantamento dos Recursos Naturais*; Folha SB.21—Araguaia e Folha SC-22—Tocantins, geologia, geomorfologia, pedologia, vegetação e uso potencial da terra; BRASIL Ministério das Minas e Energia: Rio de Janeiro, RJ, Brazil, 1974; p. 552.

29. Peel, M.C.; Finlayson, B.L.; McMahon, T.A. Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.* **2007**, *11*, 1633–1644. [CrossRef]

30. Brasil, L.S.; Batista, J.D.; Giehl NF, D.S.; Valadão MB, X.; Santos JO, D.; Dias-Silva, K. Integridade ambiental e composição de espécies de libelinhas em riachos amazônicos na região do "arco do desmatamento", Mato Grosso, Brasil. *Acta Limnol. Bras.* **2014**, *26*, 278–287. [CrossRef]

31. IBGE—Brazilian Institute of Geography and Stastistic. Downloads—Geociências. Available online: https://www.ibge.gov.br/geociencias/downloads-geociencias.html (accessed on 13 April 2023).

32. Baker, T.R.; Phillips, O.L.; Malhi, Y.; Almeida, S.; Arroyo, L.; Di Fiore, A.; Erwin, T.; Killeen, T.J.; Laurance, S.G.; Laurance, W.F.; et al. Variation in wood density determines spatial patterns in Amazonian Forest biomass. *Glob. Change Biol.* **2004**, *10*, 545–562. [CrossRef]

33. Peacock, J.; Baker, T.R.; Lewis, S.L.; Lopez-Gonzalez, G.; Phillips, O.L. The RAINFOR database: Monitoring forest biomass and dynamics. *J. Veg. Sci.* **2007**, *18*, 535–542. Available online: https://www.jstor.org/stable/4499259 (accessed on 5 February 2024). [CrossRef]

34. Chave, J.; Muller-Landau, H.C.; Baker, T.R.; Easdale, T.A.; Ter Steege, H.; Webb, C.O. Regional and phylogenetic variation of wood density across 2456 neotropical tree species. *Ecol. Appl.* **2006**, *16*, 2356–2367. [CrossRef] [PubMed]

35. Rezende, A.; Vale, A.D.; Sanquetta, C.R.; Figueiredo Filho, A.; Felfili, J.M. Comparação de modelos matemáticos para estimativa do volume biomassa e estoque de carbono da vegetação lenhosa de um cerrado sensu stricto em Brasília, D.F. *Sci. For.* **2006**, *71*, 65–76.

36. Curran, P.J.; Dungan, J.L.; Gholz, H.L. Exploring the Relationship Between Reflectance Red Edge and Chlorophyll Content in Slash Pine. *Tree Physiol.* **1990**, *7*, 33–48. [CrossRef] [PubMed]

37. ESA—European Space Agency. Sentinel Overview. Available online: https://sentinel.esa.int/web/sentinel/missions (accessed on 3 January 2023).

38. Rouse, J.W., Jr.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring vegetation systems in the Great Plains with ETRS. In Proceedings of the Third Earth Resources Technology Satellite-1 Symposium., Washington, DC, USA, 10–14 December 1973; Paper A-20, pp. 309–317.

39. Jensen, J.R. *Remote Sensing of the Environment: An Earth Resource Perspective 2/e*; Pearson Education India: Bangalore, India, 2009.

40. 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 Sens. Environ.* **2002**, *83*, 195–213. [CrossRef]

41. Yan, E.; Wang, G.; Lin, H.; Xia, C.; Sun, H. Phenology-based classification of vegetation cover types in Northeast China using MODIS NDVI and EVI time series. Rev. *Int. J. Remote Sens.* **2015**, *36*, 489–512. [CrossRef]

42. Jiang, Z.; Huete, A.; Didan, K.; Miura, T. Development of a two-band enhanced vegetation index without a blue band. *Remote Sens. Environ.* **2008**, *112*, 3833–3845. [CrossRef]

43. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, *58*, 289–298. [CrossRef]

44. Karnieli, A.; Bayasgalan, M.; Bayarjargal, Y.; Agam, N.; Khudulmur, S.; Tucker, C.J.; Zhang, X. Use of NDVI and land surface temperature for drought assessment: Merits and limitations. *J. Clim.* **2001**, *14*, 2833–2845. [CrossRef]

45. Huete, A.R. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [CrossRef]

46. Qi, J.; Chehbouni, A.; Huete, A.R.; Kerr, Y.H.; Sorooshian, S. A modified soil-adjusted vegetation index. *Remote Sens. Environ.* **1994**, *48*, 119–126. [CrossRef]

47. Barnes, E.M.; Clarke, T.R.; Richards, S.E.; Colaizzi, P.D.; Haberland, J.; Kostrzewski, M.; Waller, P.; Choi, C.; Riley, E.; Thompson, T.; et al. Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data. In Proceedings of the Fifth International Conference on Precision Agriculture, Bloomington, MN, USA, 16–19 July 2000; Robert, P.C., Ed.; ASA: Madison, WI, USA, 2000.

48. Kang, Y.; Hu, X.; Meng, Q.; Zou, Y.; Zhang, L.; Liu, M.; Zhao, M. Land cover and crop classification based on red edge indices features of GF-6 WFV time series data. *Remote Sens.* **2021**, *13*, 4522. [CrossRef]

49. STATSOFT. Comparativo de Versões do Software Statistica. Available online: http://www.statsoft.com.br/ftp/COMP_VERS_STATISTICA.pdf (accessed on 20 March 2023).

50. Carrijo, J.V.N.; Miguel, E.; Vale, A.T.D.; Matricardi, E.; Monteiro, T.; Rezende, A.; Inkotte, J. Artificial intelligence associated with satellite data in predicting energy potential in the Brazilian savanna woodland area. *Iforest-Biogeosci. For.* **2020**, *13*, 48. [CrossRef]

51. Guerrout, E.H.; Ait-Aoudia, S.; Michelucci, D.; Mahiou, R. Hidden Markov random field model and Broyden–Fletcher–Goldfarb–Shanno algorithm for brain image segmentation. *J. Exp. Theor. Artif. Intell.* **2018**, *30*, n415–n427. [CrossRef]

52. Borsato, D.; Moreira, I.; Nobrega, M.M.; Moreira, M.B.; Dias, G.H.; Ferreira da Silva, R.S.D.S.; Bona, E. Aplicação de redes neurais artificiais na identificação de gasolinas adulteradas comercializadas na região de Londrina–Paraná. *Química Nova* **2009**, *32*, 2328–2332. [CrossRef]

53. Shiblee, M.; Chandra, B.; Kalra, P. Learning of geometric mean neuron model using resilient propagation algorithm. *Expert Syst. Appl.* **2010**, *37*, 7449–7455. [CrossRef]

54. Fischer, D.R.; Paixão, J.L.; Sausen, J.P.; Abaide, A.R. Previsão de Curto Prazo para Geração Fotovoltaica a partir de Dados Meteorológicos via RNA. In Proceedings of the Congresso Brasileiro de Automática—CBA, Porto Alegre, Brazil, 23–26 November 2020; Volume 2.

55. Binoti, M.L.M.S.; Binoti, D.H.B.; Leite, H.G. Aplicação de redes neurais artificiais para estimação da altura de povoamentos equiâneos de eucalipto. *Rev. Árvore* **2013**, *37*, 639–645. [CrossRef]

56. Bustamante, M.C.; Roitman, I.; Aide, T.M.; Alencar, A.; Anderson, L.O.; Aragão, L.; Asner, G.P.; Barlow, J.; Berenguer, E.; Chambers, J.; et al. Toward an integrated monitoring framework to assess the effects of tropical forest degradation and recovery on carbon stocks and biodiversity. *Glob. Change Biol.* **2016**, *22*, 92–109. [CrossRef]

57. Egrioglu, E.A.; Ufuk, Y.B.; Cagdas, H.A.; Eren, B. Recurrent Multiplicative Neuron Model Artificial Neural Network for Non-Linear Time Series Forecasting. *Procedia Soc. Behav. Sci.* **2014**, *109*, 1094–1100. [CrossRef]

58. Matricardi, E.A.T.; Skole, D.L.; Pedlowski, M.A.; Chomentowski, W. Assessment of forest disturbances by seletive logging and forest fires in the Brazilian Amazon using Landsat data. *Int. J. Remote Sens.* **2013**, *34*, 1057–1086. [CrossRef]

59. Santos, T. Os impactos do desmatamento e queimadas de origem antrópica sobre o clima da Amazônia brasileira: Um estudo de revisão. *Rev. Geográfica Acadêmica* **2017**, *11*, 157–181.

60. Scremin, A.P.; Kemerich, P.D.C. Impactos ambientais em propriedade rural de atividade mista. *Disc. Sci. Sér Ciências Nat. Tecnol. Santa Maria* **2010**, *11*, 126–148.

61. Silva, F.L.; Oliveira, F.D.A.; Amin, M.M.; Beltrão, N.E.S.; Sales de Andrade, V.M. Dimensões do Uso e Cobertura da Terra nas Mesorregiões do Estado do Pará. *Espacios* **2016**, *37*, 5.

62. Merzlyak, M.N.; Chivkunova, O.B.; Solovchenko, A.E.; Naqvi, K.R. Absorção de luz por antocianinas em folhas juvenis, estressadas e senescentes. *J. Exp. Bot.* **2008**, *59*, 3903–3911. [CrossRef] [PubMed]

63. Bronson, K.F.; Chua, T.T.; Booker, J.D.; Keeling, J.W.; Lascano, R.J. In-season nitrogen status sensing in irrigated cotton: II. *Leaf nitrogen and biomass. Soil Sci. Soc. Am. J.* **2003**, *67*, 1439–1448. [CrossRef]

64. Barrachina, M.; Cristóbal, J.; Tulla, A. Estimating above-ground biomass on mountain meadows and pastures through remote sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *38*, 184–192. [CrossRef]

65. Huang, J.; Chen, D.; Cosh, M.H. Sub-pixel reflectance unmixing in estimating vegetation water content and dry biomass of corn and soybeans cropland using normalized difference water index (NDWI) from satellites. *Int. J. Remote Sens.* **2009**, *30*, 2075–2104. [CrossRef]

66. Matsushita, B.; Yang, W.; Chen, J.; Onda, Y.; Qiu, G. Sensitivity of the enhanced vegetation index (EVI) and normalized difference vegetation index (NDVI) to topographic effects: A case study in high-density cypress forest. *Sensors* **2007**, *7*, 2636–2651. [CrossRef]

67. Bollas, N.; Kokinou, E.; Polychronos, V. Comparison of sentinel-2 and UAV multispectral data for use in precision agriculture: An application from northern Greece. *Drones* **2021**, *5*, 35. [CrossRef]

68. Güner, Ş.T.; Diamantopoulou, M.J.; Poudel, K.P.; Çömez, A.; Özçelik, R. Employing artificial neural networks for effective biomass prediction: An alternative approach. *Comput. Electron. Agric.* **2022**, *192*, 106596. [CrossRef]

69. Narine, L.L.; Popescu, S.C.; Malambo, L. Synergy of ICESat-2 and Landsat for mapping forest aboveground biomass with deep learning. *Remote Sens.* **2019**, *11*, 1503. [CrossRef]