Article
Mapping Above-Ground Carbon Stocks at the Landscape Scale to Support a Carbon Compensation Mechanism: The Chocó Andino Case Study

Francisco Cuesta 1,*, Marco Calderón-Loor 2,*, Paulina Rosero 1, Noam Miron 2, Andrei Sharf 2, Carolina Proaño-Castro 3 and Felipe Andrade 3

1 Grupo de Investigación en Biodiversidad Medio Ambiente y Salud-BIOMAS, Universidad de las Américas, Vía a Nayón, Quito 170124, Ecuador; poliroseroa.pr@gmail.com
2 Albo Climate, Ahad Haam, 9, Tel Aviv 6525101, Israel; noam.miron@albosys.com (N.M.); andrei.sharf@albosys.com (A.S.)
3 Fundación Futuro, Av. De Las Palmeras N45-74 y de las Orquídeas, Quito 170502, Ecuador; cproano@futuro.com.ec (C.P.-C.); fandrade@futuro.com.ec (F.A.)

* Correspondence: francisco.cuesta@udla.edu.ec (F.C.); marco.calderon@albosys.com (M.C.-L.)

Abstract: (1) Background: Tropical Mountain forests (TMF) constitute a threatened major carbon sink due to deforestation. Carbon compensation projects could significantly aid in preserving these ecosystems. Consequently, we need a better understanding of the above-ground carbon (AGC) spatial distribution in TMFs to provide project developers with accurate estimations of their mitigation potential; (2) Methods: integrating field measurements and remote sensing data into a random forest (RF) modelling framework, we present the first high-resolution estimates of AGC density (Mg C ha$^{-1}$) over the western Ecuadorian Andes to inform an ongoing carbon compensation mechanism; (3) Results: In 2021, the total landscape carbon storage was 13.65 Tg in 194,795 ha. We found a broad regional partitioning of AGC density mediated primarily by elevation. We report RF-estimated AGC density errors of 15% (RMSE = 23.8 Mg C ha$^{-1}$) on any 10 m pixel along 3000 m of elevation gradient covering a wide range of ecological conditions; (4) Conclusions: Our approach showed that AGC high-resolution maps displaying carbon stocks on a per-pixel level with high accuracy (85%) could be obtained with a minimum of 14 ground-truth plots enriched with AGC density data from published regional studies. Likewise, our maps increased precision and reduced uncertainty concerning current methodologies used by international standards in the Voluntary Carbon Market.

Keywords: remote sensing; permanent plots; Andes; Sentinel; artificial intelligence; mountain forests; biomass

1. Introduction

Human-induced land use change is one of the main contributors to the current atmospheric concentration of greenhouse gases [GHG; [1]]. Reducing GHG emissions is of transcendent importance, as curving them will directly impact the global climatic system [2] and can help achieve the goals of the Paris Agreement [3]. To meet the ambitious 1.5 °C goal set by the Paris Agreement, substantial reductions in deforestation and associated CO$_2$ emissions are required, with targets of approximately 70% by 2030 and 95% by 2050 [4,5]. Moreover, forest restoration and regrowth, leading to CO$_2$ sequestration, must continue as a primary mitigation strategy [6].

However, the extent and distribution of climate mitigation opportunities available from preserving, restoring, or enhancing tropical mountain forest cover have been insufficiently explored [7,8]. One major challenge in designing carbon compensation projects at subnational and national levels is the uncertainty associated with the spatial distribution of carbon stored in above-ground biomass (AGB) [9–11]. This uncertainty is particularly
pronounced in tropical mountain forests (TMFs) due to sharp environmental gradients, limited data availability, and knowledge gaps [12,13]. Likewise, in many cases, tropical mountain forests' remnants are embedded in human-dominated landscapes at different successional stages of recovery or degradation [11,14], hence increasing the spatial variation of above-ground carbon (AGC) [13].

Accurate estimation of AGC density (Mg C ha\(^{-1}\)) at the landscape scale depends mainly on the availability and quality of ground-truth and remote sensing (RS) data together with the modeling techniques [15–17]. While local forest plot measurements provide accurate information at the plot level, their scarcity in tropical mountainous regions poses a challenge when extrapolating their values to the landscape scale [18]. Integrating local plot data with existing high-quality ground-truth AGC datasets offers a promising approach to improve AGC stock predictions at a pixel scale and to estimate model uncertainty [19–21].

The integration of RS data is crucial for extrapolating plot-based AGC density estimates and deriving pixel-level estimates of AGC stocks and changes [9,22]. Recent advances in RS technology have provided comprehensive coverage, high spatial resolution, and short return cycles. The Sentinel-1 C-band Synthetic Aperture Radar (SAR) and the Sentinel-2 multispectral instrument by the European Space Agency (ESA), are examples of these advances [23]. The combination of different sensors has the potential to overcome limitations associated with single remote sensing techniques for AGC density estimates [24].

The choice of a modelling framework is also critical to improving the accuracy of AGC predictions beyond field data [25,26]. While parametric models such as linear regressions have been widely used [27–29], non-parametric algorithms, such as machine learning models, offer a better ability to capture complex relationships and deal with skewed data [30]. Machine learning algorithms have demonstrated high accuracy in producing spatially explicit AGC estimates at the pixel level [15,31]. However, previous AGC mapping exercises in tropical regions have often been limited by coarse spatial resolution (~500 m to 1000 m) and high uncertainties, rendering them less suitable for local-scale carbon compensation projects [32–35].

Natural landscapes in Ecuador have changed rapidly in the last four decades [36–39]. The lowland and mountain forests of north-western Ecuador, the Chocó-Andes region, is one of the three regions of the country where most of the deforestation has occurred [40,41]. One of the largest remaining forest tracts is embedded in the UNESCO Chocó Andino Biosphere Reserve—RBCA hereafter [42]. From 1991 to 2017, forest loss within the RBCA was estimated at an annual gross loss rate of 0.66%. However, forest regrowth was observed for the same period, equal to 8.5% of the forest coverage in 1991 [43]. Moreover, the RBCA has witnessed a high density of forest conservation and restoration efforts in the last decade, offering a unique opportunity to establish a compensation mechanism to reduce deforestation and enhance carbon sequestration.

This research presents an innovative approach to map AGC forest stocks and their temporal variation using RS coupled with advanced statistical methods at high spatial resolution (10 m-pixel resolution) where local spatiotemporal ground-truth data is limited. Our approach leverages machine learning techniques to establish complex non-linear relationships between observed AGC values, satellite imagery channels, and ancillary environmental data. Specifically, we developed a set of high-resolution maps of AGC, with geospatially explicit uncertainty estimates for the RBCA to inform an ongoing local carbon compensation mechanism, the Zero Carbon program (https://nftree.com.ec/, accessed on 2 June 2023). This study focused exclusively on the AGC of standing trees ≥ 5 cm in diameter, not on the belowground or necromass carbon pools.

2. Materials and Methods

2.1. Study Area

This study was conducted in the western versant of the Ecuadorian Andes (0°11’17.8656″ N, 0°7’36.21″ S; −78°34’8.8566″ W, −78°54’46.242″ W), which spans over 286,800 ha from
the outer foothills of the western cordillera of the Andes to its inner foothills, along 4000 m of elevation gradient (Figure S1). Forests in the study area are humid, evergreen ecosystems with limited or no seasonality characteristic of tropical mountain systems. Forest composition gradually changes along the elevation gradient driven by temperature conditions [44,45], defining four ecosystems with different structure, composition, and above-ground biomass [46]. Mean annual air temperature diminishes from 21.6 °C at 653 m a.s.l. to 7.2 °C at 3507 m a.s.l., decreasing ca. 0.56 °C per 100 m a.s.l., reaching values of 0.7 °C at the upper end (3507 m a.s.l.). By 2020, native forest covered nearly 68% (i.e., 195,000 ha) of the study area, 3% (i.e., 10,000 ha) was mountain grasslands located above the forest line (above 3900 m a.s.l.), and 9% was mountain shrublands (i.e., 25,000 ha). The remaining 20% (i.e., 57,000 ha) constituted agricultural lands. Agriculture and cattle raising constitute an essential livelihood, with over 80% of the Biosphere Reserve’s productive land use being dedicated to extensive cattle raising [43].

2.2. Ground-Truth Data

The ground-truth data inputs for this study come from an in situ AGB dataset from the Pichincha Forest plot network, established in 2015 [45,47], enriched with a published regional AGB dataset: https://doi.org/10.5061/dryad.59zw3r26f, accessed on 8 November 2022 [48] coming from a regional Andean forest plot network—Red Bosques Andinos—RBA (https://redbosques.condesan.org/, accessed on 8 November 2022) and a subset of an extensive AGB pantropical dataset [19]. Both the local and regional databases provided AGB estimations based on permanent plots to study forest dynamics. From the larger pantropical dataset [19], we generated a subset based on the following criteria: (1) AGB records with tree cover < 70%, based on [49], were removed from the analysis. Albeit, removing these plots might limit the applicability of our AGC model to less mature forests (i.e., forests between 30% and 70% of tree cover), more than 90% of the forest area inside the study area has a tree cover > 70% (Figure S2); (2) the geographic extent included only AGB data from the north-western Amazonia, the Andes (Perú and Bolivia), and the Chocó forests (Colombia, Ecuador, and Panama); (3) latitudinal and longitudinal coordinates were not degraded or incomplete; (4) AGB field measurements were sampled from 2000 onwards; (5) extreme AGB values falling below the 5th percentile or exceeding the 95th percentile were removed.

The resulting dataset with 494 AGB plots was randomly split into a training and validation dataset, 80% and 20%, respectively (Table 1), covering the tropical lowland and montane forests of Colombia, Ecuador, Perú, Bolivia, and Panamá (Figure 1). The dataset covers an elevation gradient of 3240 m a.s.l. with a mean value of 431 m a.s.l. (±531 m a.s.l.), with a skewed distribution towards the lower section of the elevation range (i.e., 75% of the plots correspond to an altitudinal range bounded between 100 and 560 m a.s.l.; Figure 2a). All AGB plot locations had a mean forest cover of 97% ± Sd = 4.4 (Figure 2b). The mean plot size was 0.63 ha (±0.32 ha). Lastly, to obtain AGC values, we multiplied the individual plot AGB value by a conversion constant of 0.465 [50].

Table 1. Training and validation dataset’s composition.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Training</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pichincha Forest plot network—RBCA</td>
<td>15</td>
<td>8</td>
</tr>
<tr>
<td>Regional</td>
<td>379</td>
<td>92</td>
</tr>
<tr>
<td>Total</td>
<td>394</td>
<td>100</td>
</tr>
</tbody>
</table>
Figure 1. Above-ground biomass (AGB) reference dataset for the tropical lowland and montane forests of Colombia, Ecuador, Perú, Bolivia, and Panamá.

Figure 2. Above-ground biomass data distribution along (a) elevation gradient and (b) tree cover values for the year 2000 based on [49].

2.3. Remote Sensing Data Assembling

Multispectral Sentinel-2 (S-2, hereafter) and Sentinel-1 Synthetic Aperture Radar (S-1, hereafter) imagery from the European Space Agency were accessed from the Google Earth Engine (GEE) platform [51]. We used atmospherically corrected S-2 scenes (bands B2-B8, B11, and B12) to predict AGC in the RBCA. We created a stack of images for the target year, i.e., 2018–2021 ± one year (e.g., images from 2017 to 2019 were used for producing the year 2018 mosaic) to minimize missing data and cloud interference [52]. We then...
applied the s2cloudless algorithm implemented in GEE for masking clouds and cloud-shadows to every stack. Each multiyear stack contained around 5 to 20 valid observations at pixel-level (year and location dependent). In addition, we created yearly S-1 mosaics. S-1 images were retrieved from the ‘Descending orbit’ (to eliminate potential discrepancies in the backscatter) and pre-processed by applying an additional border noise correction, a multi-temporal boxcar speckle filtering and a radiometric terrain normalization ([53]; Table S1).

Different spectral indices were calculated and appended to the cloud-free stacks (Table S1). Likewise, elevation and slope data (30-m) from the Shuttle Radar Topography Mission (SRTM) dataset [54] were included as predictor variables. Lastly, X (longitude) and Y (latitude) coordinates were generated for each pixel of the study area and integrated with the RS dataset. The geographical coordinates were included to capture the environmental gradients that influence our study region, e.g., harsher climatic conditions towards the east and a marked carbon stocks variation from north to south. All the aligned and registered layers were resampled to a 10 m resolution if necessary. The final composites were produced by computing the median value of every variable/band at the pixel level in the 2017–2019 period.

2.4. Above-Ground Carbon (AGC) Regression Models

We used a random forest (RF) algorithm [55] to develop a regression-based machine learning model. RF is a computationally efficient tree-based machine learning model robust to random and systematic noise [56,57]. RF builds several trees, previously specified by the user, iteratively from random training data samples (n = 394). The ensemble of the predictions of all trees constitutes the final model. We parametrized the model to generate 1000 trees with 15 variables per split; a grid search algorithm was used to find the optimum number of trees. We used the scikit-learn library for fine tuning the RF model in Python [58]. A total of 23 variables were used as predictors (Table S2).

To increase the robustness of our regression and to provide a measure of the model’s estimation uncertainty at the pixel level we used the mean of ten distinct RF models as the AGC value for every pixel and every time step. Our dataset, n = 494, was divided into ten folds [59]; for each fold, an 80% random sample (i.e., n = 394) was used to train an RF model. For every time step, ten AGC maps were generated and averaged to produce AGC density estimates at the pixel level.

2.5. Error and Uncertainty Estimation

We assessed the error estimation and precision for the predicted AGC models by comparing the predicted AGC (i.e., mean of the 10 maps) against the observed ground-truth AGC. Error and precision were calculated for the entire training dataset (n = 393), validation dataset (n = 100) and the Pichincha Forest plot network exclusively (n = 23). To evaluate the performance of the AGC models, we used the adjusted $R^2$ metric, which considers the model’s goodness of fit while adjusting for model complexity to avoid overfitting. Additionally, we employed three widely used performance indicators: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the Mean Absolute Percentage Error (MAPE) [23,60]. These metrics allowed us to quantify the magnitude of differences between the observed and predicted AGC values and provided a comprehensive evaluation of model accuracy.

Lastly, we assessed the generalization capabilities of our AGC model by comparing the outputs of independent random forest (RF) models against localized ground-truth data. Four distinct areas of interest (AOI) were defined, each containing a different number of local plots (Table 2). In an iterative process, 600 different RF models were trained for each AOI. For constructing the training and validation datasets for each AOI, the following procedure was followed: (1) plots belonging to the specific AOI under consideration were removed from the entire dataset; (2) the remaining dataset was randomly split into training and validation subsets, with proportions of 80% and 20%, respectively, in each iteration;
(3) local plots were then randomly selected and incrementally added to both the training and validation datasets; (4) the maximum number of local plots that were added to the training subset was limited to 80% of the total local plots available within the AOI. This iterative process allowed us to progressively evaluate the impact of adding more localized data on the model’s performance. The validation dataset, consisting of 20% of the data, including local validation points, was used to estimate the model’s performance for each AOI. This approach ensured a comprehensive assessment of the model’s ability to generalize across different areas with varying degrees of local data availability.

Table 2. Metadata of four areas of interest (AOI) where model’s generalization capabilities were assessed.

<table>
<thead>
<tr>
<th>AOI</th>
<th>AOI Area (km²)</th>
<th>Coordinates Extent (°)</th>
<th>Median Elevation (Interquartile Range) (m a.s.l.)</th>
<th>Local Plots Size (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBCA</td>
<td>2868</td>
<td>79.381, 78.36, −0.643, −0.395</td>
<td>1267 (1380)</td>
<td>23</td>
</tr>
<tr>
<td>Bolivia</td>
<td>1,165,573</td>
<td>−69.5, −68, −15.5, −14</td>
<td>1494 (1462)</td>
<td>23</td>
</tr>
<tr>
<td>Peru</td>
<td>1,804,273</td>
<td>72.146, 70.986, −13.368, −11.781</td>
<td>1519 (2188)</td>
<td>28</td>
</tr>
<tr>
<td>Bolivia and Peru</td>
<td>7,761,764</td>
<td>−72.206, −70.86, −15.5, −11.67</td>
<td>1200 (1973)</td>
<td>67</td>
</tr>
</tbody>
</table>

2.6. Carbon Maps’ Post-Processing

A post-processing was applied to all the generated AGC maps. First, non-forest areas were masked-out from the results using the Global Forest Change V1.10 [49]. Then, a temporal coefficient of variation (CV) at the pixel-level was calculated and used for improving the temporal consistency of the AGC maps. Pixels with a temporal CV > 1.5, i.e., high temporal uncertainty, and high slope (>25°) were smoothed by using a focal window. We used a circular kernel of 80 m radius to calculate the average AGC inside the moving window. In each time-step the AGC in the selected pixels was replaced with the smoothed AGC value. Then, a temporal correction was applied to all the previously smoothed pixels. AGC values in these pixels were replaced by their minimum temporal AGC value to reduce temporal inconsistencies [49].

2.7. Carbon Stocks Spatial Patterns, Gains, and Losses

As elevation is the single most important factor controlling AGC stocks in tropical mountain ecosystems [61–63], AGC maps were decomposed into six elevation ranks to assess the spatial distribution of the above-ground carbon across the study area. Elevation ranks were set based on natural breaks in the histogram of the elevation grid. We also generated a profile along the elevation gradient and summarised the AGC value per pixel for each elevation rank and generated AGC averages per each rank. Lastly, we compared our AGC model against pre-existing AGC global datasets [19,64,65] across the elevation gradient in the RBCA.

We estimated carbon stock change between 2018 and 2021 at the pixel scale to assess carbon gains and losses. We defined the AGC net change (Mg ha⁻¹ yr⁻¹) as the difference between AGC stock in 2021 and AGC stock in 2018 divided by the elapsed time (t; in years) \[\text{AGC net change} = (\text{AGC}_{2021} - \text{AGC}_{2018})/4\]. Positive values were defined as forest carbon stock gain, whereas negative values as carbon loss.

3. Results

In 2018, the forests of the RBCA stored a mean of 70.26 (±8.1) Mg C ha⁻¹, ranging from 55.6 to 102.8 Mg C ha⁻¹ (Table 3, Figure 3). In 2021 a slight increase in AGC density was observed at 70.64 (±8.1) Mg C ha⁻¹ at the pixel scale. In 2021, total carbon...
storage at the landscape scale was 13.66 Tg (million megagrams) in its above-ground compartment, evidencing an annual net gain of 71,886 Mg C, equalling a productivity of 0.37 Mg C ha\(^{-1}\) yr\(^{-1}\) (Table 4). This net increase resulted from a greater increase in pixels that gained biomass (mean = 0.60 ± 0.66 Mg C ha\(^{-1}\) yr\(^{-1}\)) than pixels that lost biomass (mean = −0.49 ± 0.61 Mg C ha\(^{-1}\) yr\(^{-1}\)) between 2018 and 2021.

Table 3. Estimated Above-Ground Carbon density (AGC, Mg C ha\(^{-1}\)) and AGC stocks (landscape) in the Chocó Andino Biosphere Reserve (RBCA) for the period of 2018–2021.

<table>
<thead>
<tr>
<th>Year</th>
<th>AGC Pixel-Level Mean (±Sd)</th>
<th>AGC Pixel-Level Range (Min–Max)</th>
<th>Landscape Total (Tg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018</td>
<td>70.26 (±8.1)</td>
<td>55.6–102.8</td>
<td>13.45</td>
</tr>
<tr>
<td>2019</td>
<td>70.65 (±8.2)</td>
<td>55.6–102.8</td>
<td>13.71</td>
</tr>
<tr>
<td>2020</td>
<td>70.64 (±8.2)</td>
<td>55.7–102.8</td>
<td>13.76</td>
</tr>
<tr>
<td>2021</td>
<td>70.64 (±8.1)</td>
<td>55.6–101.9</td>
<td>13.66</td>
</tr>
</tbody>
</table>

Figure 3. Variation in above-ground carbon density (AGC, Mg C ha\(^{-1}\)) at 10 m pixel resolution throughout the Chocó Andino Biosphere Reserve (RBCA for its Spanish acronym) in the western Ecuadorian Andes, derived from an integrated model of remote sensing, local, regional, and global field plot datasets integrated in a random forest algorithm.

Table 4. Forest cover, AGC stocks, and AGC net change between 2018 and 2021 at the pixel and landscape scale in the Chocó Andino Biosphere Reserve (RBCA).

<table>
<thead>
<tr>
<th>Variable</th>
<th>2018</th>
<th>2021</th>
<th>Productivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGC (Mg)</td>
<td>13,446,451</td>
<td>13,662,109</td>
<td></td>
</tr>
<tr>
<td>Forest cover (ha)</td>
<td>198,727</td>
<td>194,795</td>
<td></td>
</tr>
<tr>
<td>Landscape (Mg C yr(^{-1}))</td>
<td></td>
<td></td>
<td>71,886</td>
</tr>
<tr>
<td>Pixel (Mg C ha(^{-1}) yr(^{-1}))</td>
<td></td>
<td></td>
<td>0.37</td>
</tr>
</tbody>
</table>

We found a broad regional partitioning of standing carbon stocks mediated primarily by elevation, longitude, and latitude (Figure 4). The mean AGC of the piedmont forest
(340–1200 m a.s.l.) was 80 (±6.2) Mg C ha\(^{-1}\) and decreased monotonically along the gradient to as low as 59.8 (±3.8) Mg C ha\(^{-1}\) in the upper montane forests (>3500 m a.s.l.; Figure 5).

![Figure 4](image)

**Figure 4.** Variable importance in a random forest model based on the Gini impurity index for the Chocó Andino Biosphere Reserve (RBCA for its Spanish acronym).

![Figure 5](image)

**Figure 5.** The variation with elevation of the predicted above-ground carbon density (AGC, Mg C ha\(^{-1}\)) at a pixel size of 10 m in the remnant forests of the Chocó Andino Biosphere Reserve (RBCA for its Spanish acronym).

Based on the analysis of the entire validation dataset (\(n=100\)), we successfully predicted above-ground carbon (AGC) values across the study region with high precision (\(R^2 = 0.6; p < 0.0001\)) and accuracy at a fine resolution of 0.01 ha (RMSE = 44.3 Mg C ha\(^{-1}\) in 2021; Table 5). Subsequently, we focused solely on the Pichincha plots (\(n=23\)), which resulted in improved precision (adj. \(R^2 = 0.7\)); however, it also led to an increase in the error, accounting for nearly 28% of the mean AGC density (RMSE = 23.8 Mg C ha\(^{-1}\); Table 5). Likewise, the RF models with different training subsets for the four AOs depict that our AGC model can achieve an overall mean accuracy of ~85%, when at least 14 local sampling
plots are combined with regional plots to reach a relatively high mapping accuracy (Table 6). Tables 5 and 6 report differences in RMSE, MAE, and MAPE at the RBCA level. Main reasons for these differences are the number of RF models that were employed in each analysis. While we used the average of 10 maps for calculating the accuracy and error metrics in Table 5, the tabular output of 600 RF models was employed for calculating and assessing the generalization capabilities of our approach in Table 6.

Table 5. Models’ error assessment based on different datasets for the year 2021 based on the linear fit between the observed above-ground carbon values against the random forest modelled estimates at 10 m pixel resolution, where: AGC: above-ground carbon; RMSE: Root Mean Square Error; MAE: Mean Absolute Error; and MAPE: Mean Absolute Percentage Error.

<table>
<thead>
<tr>
<th>Total Plots (n)</th>
<th>Type</th>
<th>Observed AGC (Mg ha(^{-1}))</th>
<th>Adj. (R^2)</th>
<th>RMSE (Mg C ha(^{-1}))</th>
<th>MAE (Mg C ha(^{-1}))</th>
<th>MAPE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>494</td>
<td>Entire dataset</td>
<td>85.7</td>
<td>0.87</td>
<td>12.1</td>
<td>8.29</td>
<td>0.12</td>
</tr>
<tr>
<td>393</td>
<td>Training dataset</td>
<td>85.8</td>
<td>0.95</td>
<td>8.3</td>
<td>6.07</td>
<td>0.09</td>
</tr>
<tr>
<td>100</td>
<td>Validation dataset</td>
<td>85.1</td>
<td>0.6</td>
<td>20.6</td>
<td>15.71</td>
<td>0.24</td>
</tr>
<tr>
<td>23</td>
<td>RBCA entire dataset</td>
<td>84.5</td>
<td>0.7</td>
<td>23.8</td>
<td>17.39</td>
<td>0.23</td>
</tr>
<tr>
<td>8</td>
<td>RBCA validation dataset</td>
<td>81</td>
<td>0.4</td>
<td>34.6</td>
<td>28.62</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Table 6. Models’ error assessment in four areas of interests (AOIs) with different number of local plots, where: AGC: above-ground carbon; RMSE: Root Mean Square Error; NRMSE: Normalized RMSE; and MAPE: Mean Absolute Percentage Error. Minimum plots refer to the minimal number of local plots added to the dataset that ensured a mean accuracy of \(\geq 85\%\).

<table>
<thead>
<tr>
<th>AOI</th>
<th>Total Plots (n)</th>
<th>Minimum Plots (n)</th>
<th>Observed AGC (Mg ha(^{-1}))</th>
<th>RMSE (Mg C ha(^{-1}))</th>
<th>NRMSE</th>
<th>MAPE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBCA</td>
<td>23</td>
<td>23</td>
<td>84.5 (40.7)</td>
<td>18.4 (3.7)</td>
<td>0.22 (0.044)</td>
<td>16.4 (0.01)</td>
</tr>
<tr>
<td>Bolivia</td>
<td>23</td>
<td>19</td>
<td>82.9 (44)</td>
<td>19.9 (4.7)</td>
<td>0.24 (0.06)</td>
<td>14.8 (0.02)</td>
</tr>
<tr>
<td>Peru</td>
<td>28</td>
<td>14</td>
<td>87.6 (34.3)</td>
<td>16.3 (2.7)</td>
<td>0.19 (0.03)</td>
<td>15.4 (0.03)</td>
</tr>
<tr>
<td>Bolivia and Peru</td>
<td>67</td>
<td>35</td>
<td>90 (36.3)</td>
<td>21.5 (2.6)</td>
<td>0.24 (0.03)</td>
<td>15.2 (0.01)</td>
</tr>
</tbody>
</table>

When comparing our AGC model with the global maps available for the tropics [19,22,65], we found that all global maps tended to overestimate AGC density across the study area, and the variation reported by all global maps was at least two times higher than our results (Table 7). Further, the three global datasets failed to represent the AGC/elevation trend reported for our study area. All global AGC maps predicted the montane forests between 1800 and 2500 m a.s.l. to contain the higher AGC density followed by the range between 1200 and 1800 m a.s.l. (Figure S3, Table 7).

Table 7. Estimated above-ground carbon (AGC) stocks in the Chocó Andino Biosphere Reserve along the elevation gradient by the three global datasets and this study, using 2019 as a reference midpoint of the period studied (2018–2021). The six elevation bins used resemble the ecosystems with distinct carbon contents (based on [46]).

<table>
<thead>
<tr>
<th>Elevation (m a.s.l.)</th>
<th>[22] AGC (±Sd) Mg ha(^{-1})</th>
<th>[19] AGC (±Sd) Mg ha(^{-1})</th>
<th>[65] Mean (±Sd)</th>
<th>This Study AGC (±Sd) Mg ha(^{-1})</th>
<th>Landcape (Mg C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>300–1200</td>
<td>79.05 (±12.9)</td>
<td>88.57 (±23.2)</td>
<td>91.7 (±36.9)</td>
<td>80.01 (±6.19)</td>
<td>23,954,412</td>
</tr>
<tr>
<td>1200–1800</td>
<td>94.02 (±25.3)</td>
<td>97.33 (±32.4)</td>
<td>100.2 (±42.4)</td>
<td>72.98 (±6.27)</td>
<td>4,514,733</td>
</tr>
<tr>
<td>1800–2500</td>
<td>122.23 (±30.7)</td>
<td>125.74 (±41.9)</td>
<td>113 (±42.2)</td>
<td>65.87 (±4.18)</td>
<td>3,273,434</td>
</tr>
<tr>
<td>2500–3000</td>
<td>106.3 (±35.9)</td>
<td>106.93 (±52.9)</td>
<td>99.1 (±49.5)</td>
<td>63.82 (±4.31)</td>
<td>1,641,143</td>
</tr>
<tr>
<td>3000–3500</td>
<td>94.6 (±34.1)</td>
<td>81.42 (±45.3)</td>
<td>82.5 (±50.2)</td>
<td>62.45 (±4.16)</td>
<td>856,181</td>
</tr>
<tr>
<td>3500–4777</td>
<td>71.5 (±27)</td>
<td>45.87 (±29.1)</td>
<td>43.51 (±39.2)</td>
<td>59.78 (±3.75)</td>
<td>59,682</td>
</tr>
<tr>
<td>All</td>
<td>100.7 (±31.7)</td>
<td>89.13 (±47.2)</td>
<td>99.87 (±44.1)</td>
<td>70.65 (±8.2)</td>
<td>13,706,235</td>
</tr>
</tbody>
</table>
4. Discussion

Our predictive random forest (RF) models for above-ground carbon (AGC) revealed significant carbon storage in the montane forest of the RBCA, with a total of 13,662,109 Mg C stored in the above-ground compartment in 2021. The piedmont (600–1200 m a.s.l.) and lower montane (1200–1800 m a.s.l.) forests on the northwestern versant of the study area were identified as the primary sinks, accounting for nearly 60% of the total AGC stored in the RBCA. Above this elevation range, the AGC stocks showed a steady decrease (refer to Table 7 and Figure 5). This observation aligns with previous studies that reported a decline in AGC with increasing elevation along elevation gradients in the RBCA and the tropical Andes [48,61,62,66], reinforcing the capability of our AGC models to accurately capture the spatial variation of AGC influenced by elevation.

The observed increase in AGC over time was found to be lower than that reported for mountain forests in the Andes using permanent plots (0.67 ± 0.08 Mg C ha$^{-1}$ yr$^{-1}$; [48]). Additionally, between 2015 and 2019, 16 out of the 21 local plots in our study area exhibited an AGC increase at an annual rate of 1.20 ± 0.89 Mg ha$^{-1}$ yr$^{-1}$ [46]. However, at larger spatial scales environmental heterogeneity increases, leading to varied AGC trends. As a result, many pixels across the RBCA indicated a decrease in AGC between 2018 and 2021. These pixels showed a mean AGC loss of $-0.49$ (±0.61) AGC ha$^{-1}$ yr$^{-1}$ (Table S3). On the other hand, pixels reporting a consistent increase in carbon stocks showed a mean AGC gain of 0.60 (±0.66) AGC ha$^{-1}$ yr$^{-1}$. To better understand the trajectory of these pixels over time, further analyses are needed. For example, conducting a time series analysis for a subset of pixels (e.g., 15% of all pixels exhibiting AGC loss and gain) surrounding field plot locations using spectral–temporal segmentation algorithms (e.g., LandTrendr https://github.com/eMapR/LT-GEE, accessed on 2 June 2023), can provide valuable insights into forest cover changes based on the pixel’s spectral history, utilizing moderate or high-resolution imagery [67]. Such an approach would complement our understanding of AGC pixel trends over shorter timeframes.

4.1. Uncertainty and Its Impact on the Calculation of Carbon Credits

Our approach has proven highly effective in estimating above-ground carbon (AGC) densities at the landscape scale, providing predictive values (Mg ha$^{-1}$) at the pixel level that accurately reflect AGC spatial variation within complex landscape mosaics, such as the western versant of the Ecuadorian Andes. The increased AGC resolution achieved through our methodology has significant implications for the design of local Payment for Ecosystem Services (PES) schemes, particularly in determining the size of carbon benefits that can be generated for any project area. Notably, our high-resolution maps offer enhanced precision and reduced uncertainty compared to current methodologies used by international standards in the Voluntary Carbon Market (e.g., Plan Vivo, Verra), where average AGC density per hectare is estimated for each forest stratum [68,69].

Significant differences emerged when comparing our approach with available global AGC datasets (Table 7). Global datasets consistently estimated higher AGC contents at the pixel scale, with considerable dispersion around the means, ranging from 31% to 53% variation. In contrast, our model demonstrated a considerably lower dispersion, with only 12% variation. On the landscape scale, on average, the three datasets estimated 7.4 million Mg more carbon stored in the RBCA than our estimates, potentially leading to an overestimation of the mitigation potential (refer to Table 6). Several factors contribute to these discrepancies: (1) The inability of the three global datasets to capture the AGC/elevation trend reported for tropical mountain ecosystems (e.g., [12,48,61]). All global AGC maps predicted that montane forests between 1200 and 2500 m a.s.l. contained the highest carbon density in our study area (Table 7, Figure S3). As this elevation band constitutes 60% of the forest cover in the RBCA, it significantly increases the estimated total carbon at the landscape scale. (2) There is limited plot coverage in tropical mountain regions of the Andes, which forms the basis for global products. Plot distribution tends to be skewed toward lowland areas [18,70] leading to a lack of representation in higher elevation regions.
This limitation hinders the applicability of global datasets and increases the uncertainty surrounding the quantified AGC estimates. (3) Globally available forest masks at a 30 m resolution (i.e., [49]), may overestimate forest cover remnants in agricultural mosaics. A 30 m pixel resolution product might inaccurately depict forest cover remnants in complex topographic regions, where forest fragments exist in areas of high slope and are embedded within diverse land use mosaics (e.g., palmetto plantations and cacao agroforestry systems). This can lead to overestimation of carbon stocks in such regions.

4.2. Robustness of the AGC Model

Our approach demonstrated a remarkable reduction in the number of local plots required to achieve a statistically valid sample for obtaining highly accurate AGC estimates. Current methodologies for REDD+ projects typically recommend a minimum of 45 local plots for training (70%) and validating remote-sensing-based AGC models [68]. However, our approach showcased that AGC high-resolution maps, displaying carbon density at a per-pixel level with high accuracy (≥85%), can be obtained with as few as 15 local plots enriched with AGC data from published regional studies and existing forest and carbon datasets. This significant reduction in the number of required local plots can lead to substantial cost savings and improved efficiency in REDD+ project designs. Furthermore, our proposed approach allows for near real-time monitoring of AGC dynamics across the entire study region. Although we utilized S-1 and S-2 imagery, our approach is sensor-agnostic and can be applied to different satellite imagery sources, whether public or commercial, at various spatiotemporal resolutions. These findings can have a meaningful impact on REDD+ project implementation, particularly for projects covering extensive areas with diverse land-use/land-cover types and limited accessibility, where reaching a statistically valid sampling size using traditional plot-based biomass measurements can be challenging.

The inclusion of environmental covariates in the RF model significantly improved the accuracy of our AGC estimates. Elevation, longitude, and latitude were identified as the most influential factors in the RF model (refer to Figure 4). This aligns with previous research that highlights the strong influence of environmental variables on AGC spatial distribution, especially in complex topographic regions [61,63,71]. Interestingly, an average pixel-level error of 20.6 Mg C ha$^{-1}$ compares favourably with past studies that reported uncertainties ranging from 21 Mg C ha$^{-1}$ [15] to 40 Mg C ha$^{-1}$ in tropical forests [60,72–74], see Figure S4. The accuracy achieved by our AGC model is a crucial aspect for the success of carbon sequestration schemes and can support more informed decision-making in climate mitigation efforts.

4.3. Implications for Forest Conservation and Restoration in the RBCA

The AGC maps for the RBCA were generated to be integrated into the Zero Carbon Compensation Program. The Zero Carbon Program emerges as a mechanism to promote, on the one hand, the decarbonization of Grupo Futuro (GF) companies (https://www.futuro.com.ec/, accessed on 2 June 2023) and, on the other, as a financing mechanism for the conservation of biodiversity. The Zero Carbon Program is structured into three components. The first focuses on quantifying greenhouse gas (GHG) emissions from GF companies, also known as carbon footprint (based on international quality standards, i.e., ISO 14064 [75], 14067 [76] standards; IPCC Protocol for the AFOLU sector). The second component centres on implementing strategies to reduce GF companies’ carbon footprint. The third component focuses on designing and implementing compensation strategies through a trust fund, the Carbon Neutral Fund, formed in 2019. Funds from this trust are invested in the conservation and restoration of forests via conservation agreements between GF and landowners in the RBCA.

The third component, related to the compensation mechanism, is implemented on a technological platform called NFTrees (https://nftree.com.ec/, accessed on 2 June 2023). Our carbon maps were uploaded in blockchain technology to allow interested buyers to purchase digital tokens to compensate for their emissions. The purchase is made through a
credit card or wire transfer directly to the Carbon Neutral Trust Fund. The NFTrees platform allows transactional traceability and avoids double accounting (i.e., different parties cannot claim the same carbon removal or reduction credit). Likewise, blockchain technology generates a unique and unalterable record (a Token). The token guarantees that any hectare of forest in which the carbon has been compensated cannot be sold again. Currently, the GF’s Zero Carbon Program model generates USD 250,000 per year, equivalent to the carbon offset of around 13,000 Mg of CO$_2$-eq of 12 companies of the GF, which translates into the conservation of 3000 ha of forests within the RBCA. With the AGC maps integrated with blockchain technology and digital contracts to gain transparency and traceability, the Zero Program envisions generating USD 1.9 million in the next five years to finance the conservation of 22,600 hectares of forest and benefit at least 700 families within the RBCA.

4.4. Limitations and Further Research

RS mapping of above-ground carbon (AGC) stocks relies on various factors that can influence the quality of the outputs. While we introduced a method to reduce the need for extensive local ground-truth data, the quantity and quality of regional/global AGC data remain critical for accurately representing AGC in areas with high environmental heterogeneity, such as complex topography and diverse climates. The adequate representation of such heterogeneity requires a sufficient number of well-distributed regional records, obtained from forest plots in regional/global datasets. In the case of the RBCA, AGC stocks were partly underestimated due to the scarcity of up-to-date AGC data from mountainous areas, particularly above 1000 m a.s.l. (Figure 2a). The lack of updated regional ground-truth data from regions with similar environmental conditions can increase the Root Mean Square Error (RMSE) of models and reduce their prediction capabilities. To address this, initiatives releasing global AGB/AGC datasets that can be used for training and validating RS-based models must be properly supported with fair compensation mechanisms, ensuring the maintenance of long-term plot monitoring programs [77]. Additionally, the incorporation of ground-truth data from the Global Ecosystem Dynamics Investigation (GEDI) Light Detection and Ranging (LiDAR) sensor [78] as a source of ground-truth data can increase the prediction capabilities of RS-based AGC models (e.g., [79]).

Accurate carbon credits and potential mitigation estimates need the combination of removals and emissions (i.e., AGC stocks) and activity data (i.e., land cover change). Locally generated land change information needs to be developed alongside AGC data for improving wall-to-wall AGC estimates while ensuring the generated data’s transparency. Here we have used a previously generated global forest cover dataset at 30 m [49] for generating a forest vs. non-forest mask for each of the four AGC maps (2018–2021). Although the Global Forest Change [49] product has been validated and extensively used worldwide, it presents limitations when used at local scales, particularly in landscape mosaics of forest remnants embedded in agricultural lands.

Similarly, further research is needed to produce temporally consistent AGC stock maps, especially in tropical mountainous regions. The atmospheric and topographic characteristics of the regions add an extra layer of complexity to the consolidation of quality multitemporal satellite images. Here, we implemented and applied different preprocessing algorithms to the multitemporal stacks. However, some remnant noise can influence the AGC stocks at the pixel level. Time series analyses to decompose the raw data can be beneficial for removing seasonality and noise signals from the stacks [80]. Furthermore, adding the AGC time series in the training dataset can also improve the temporal consistency of AGC estimates. Although different authors have used statistical approaches for generating historical AGC maps (e.g., validating these products requires updated and multitemporal local and regional AGB/AGC data (e.g., [81])).

5. Conclusions

In this study, we have addressed a critical knowledge gap by providing comprehensive insights into the spatial distribution of above-ground carbon (AGC) density in the TMFs
of the western Ecuadorian Andes. Our findings hold significant implications for carbon compensation projects aimed at preserving these threatened ecosystems. By integrating field measurements and remote sensing data within an RF modeling framework, we have achieved a pioneering advancement in accurately estimating AGC density at high spatial resolutions. By employing this approach, we have successfully produced the first high-resolution estimates of AGC density (Mg C ha$^{-1}$) across the western Ecuadorian Andes. Notably, our approach yields a remarkable accuracy of 85%, with RF-estimated AGC density errors of only 15% (RMSE = 23.8 Mg C ha$^{-1}$) on any 10 m pixel along a 3000 m elevation gradient, encompassing diverse ecological conditions. The resulting AGC density maps provide a detailed overview of carbon stocks at a per-pixel level, offering insights into the factors controlling the spatial distribution of AGC density at the landscape scale. The outcomes of our study revealed a pronounced regional partitioning of AGC density predominantly driven by elevation variations. This nuanced understanding of AGC distribution enriches the foundation for carbon compensation mechanisms, enabling project developers to make more precise estimations of the mitigation potential inherent in TMFs. Furthermore, our investigation established the feasibility of obtaining highly accurate AGC density maps by utilizing a minimum of 14 ground-truth plots augmented with AGC density data from published regional studies. This approach not only enhances the precision of carbon stock assessments but also contributes to reducing the uncertainty associated with existing methodologies employed within the Voluntary Carbon Market and international standards.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/f14091903/s1, Figure S1: Location of the 23 permanent plots across an elevation gradient on the Choco Andino Biosphere Reserve (RBCA for its Spanish acronym); Figure S2: Forest cover density (>70%) in 2019 on the Chocó Andino Biosphere Reserve. Forest cover density derived from [49]; Figure S3: AGC stock estimates by four different models. Maps A-C correspond to global carbon models for the tropics at different spatial resolutions; Figure S4: Variation in AGC uncertainty at pixel scale based on the coefficient of variation of 10 random forest models. White areas within the RBCA represent no-forest cover (i.e., pixels with less than 70% of forest cover, based on [49] for the year 2021). Polygons inside de RBCA delimit Quito municipality conservation and sustainable use areas (ACUs for its Spanish acronym); Table S1: Description of satellite bands used to build the Chocó Andino Biosphere Reserve AGC model; Table S2: Variable relative importance of the random forest model performed to predict AGC in the RBCA for the reference year 2019. Number of Trees: 1000; Table S3: Pixels that increase and decrease in above-ground carbon for the period of 2018–2021 in the Chocó Andino Biosphere Reserve.

**Author Contributions:** Conceptualization, F.C., M.C.-L., C.P.-C. and F.A.; methodology, M.C.-L., F.C., P.R. and A.S.; formal analysis, M.C.-L., F.C. and P.R.; writing—original draft preparation, F.C., M.C.-L. and P.R.; writing—review and editing, F.C., M.C.-L., P.R., N.M., A.S., C.P.-C. and F.A.; project administration, F.C. and C.P.-C.; funding acquisition, F.C. and C.P.-C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Universidad de las Américas, Ecuador (Grant No. DIVFCC.22.02), and Fundación Futuro (Grant No. UDLA-002).

**Data Availability Statement:** Not Applicable.

**Acknowledgments:** We want to thank Hedi Zisling for his help with retrieving and preparing the RS data. We thank Reut Keller, Huanhuan Huang, and Ira Bar for their inputs on an earlier version of the manuscript. We express our gratitude to Esteban Pinto, Antonella Bernardi, Andrea TeráinValdez, Laura Cifuentes, and Emilia Ayala for their invaluable work in leading and collecting the AGB in the Pichincha long-term monitoring transect. Likewise, we would like to express our gratitude to the owners and administrators of the private reserves where the forest dynamics permanent plots are established: Alejandro Solano, María Emilia Arcos, and Oliver Torres from Mashpi Shungo and Pambiliño; Mateo Robleda from Mashpi Lodge/Fundación Futuro, Sonia Saltos from Rio Bravo, Juan Diego Ortiz, and Sol Acosta from Sacha Urcu, Inti Arcos and Nina Duarte from Intyllacta, Richard Parson and Favió Fernández from Bellavista Cloud Forest, Germán Toasa from El Cedral Ecolodge,
Maria Elisa Vizuete and Enrique Maldonado from Verdecoccha and the Jocotoco foundation from the Yanacocha Reserve. We thank Universidad de las Americas for financially supporting this research (Grant No. DIV.FCC.22.02) and Fundación Futuro (Grant No. UDLA-002). We acknowledge ALBO Climate for their continuous support throughout this project.

Conflicts of Interest: The authors declare no conflict of interest.

References


73. Mascaro, J.; Asner, G.P.; Muller-Landau, H.C.; van Breugel, M.; Hall, J.; Dahlin, K. Controls over Aboveground Forest Carbon Density on Barro Colorado Island, Panama. *Biogeosciences* 2011, 8, 1615–1629. [CrossRef]


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.