Supplementary Material

Appendix 1 - Exploratory Data Analysis for Stationarity and Anisotropy.

In the geostatistical framework, the intrinsic hypothesis of stationarity was assumed, whereby variance is considered a function solely of the distance between pairs of points, rather than their absolute spatial location. This assumption was evaluated through exploratory data analysis, which included: (i) detecting and assessing the influence of outliers on variance; (ii) examining the variation of values as a function of spatial coordinates (longitude and latitude) to identify potential first-order spatial trends; and (iii) investigating anisotropy, which is fundamental to understanding how spatial continuity varies with direction.

Anisotropy assessment was conducted by calculating experimental semivariograms in four principal directions (0°, 45°, 90°, and 135°), corresponding respectively to the North–South, Northeast–Southwest, East–West, and Northwest–Southeast orientations. These directions capture spatial continuity both along the primary sampling axes and the diagonals. Anisotropy was diagnosed by comparing the ranges (the distance at which the semivariogram reaches the sill) and the sills themselves across different directions. Differences in range indicate geometric anisotropy, while differences in sill values suggest zonal anisotropy, often associated with underlying trends or distinct spatial processes. The anisotropy ratio, obtained from the proportion between the maximum and minimum ranges, was then used to calibrate the mathematical semivariogram model to the observed spatial variability.

Appendix 2 - Regression kriging (RK).

The mathematical model of RK used is presented in matrix form as follows:

$$\begin{bmatrix} y1 \\ y2 \\ \vdots \\ yn \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ & \vdots & \vdots & \vdots & \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

where: $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$

Where: **X** is the matrix of the independent variables, starting with a column of 1's for the intercept estimation. **β** is the vector of parameters estimated by least squares, e **ε** is the vector of errors (residuals) from the multiple linear regression model estimation, which encompasses the usual formula below:

$$y=\beta_0+\beta_1 x_1+\beta_x x_2+\ldots+\beta_k x_{ik}+\varepsilon_i$$

Where, α: Intercept; β: Regression coefficients of independent variables, **X**: Independent variables; ε: Random error. Following the recommendation of Draper & Smith (1980), a correlation matrix was constructed using AGB from GEDI and covariate data (NDVI, NDRE, SAVI, EVI, Altitude, Slope, Roughness and aspect).

Tab. S1 - Theoretical semivariogram models fitted to assess the spatial dependence of residuals and values regressed by the multiple linear model.

Models	
Gaussian	$\gamma(h)=\beta_0+\beta_1[1-\exp\left(-3\left(\frac{h}{\beta_2}\right)\right)]$
Spheric	$\gamma(h)=\beta_0+\beta_1[1{,}5\left(\frac{h}{\beta_2}\right)-0{,}5\left(\frac{h}{\beta_2}\right)]$
Exponential	$\gamma(h)=\beta_0+\beta_1[1-\exp\left(\frac{-h}{\beta_2}\right)]$

Where: y(h) = semivariance; β_0 is the nugget effect; β_1 = sill; β_2 = range; h = distance between sampled points.

Validation and test data

Tab. S2 - Model performance statistics using training data.

Training			
Coefficients (β)	**Fitted Values**	**Standard Error**	**t-valor**
B_0	375.9935	9.3510	2^{-16*}
β_1	6.2080	0.2004	2^{-16*}
β_2	-828.3882	26.2875	2^{-16*}
R^2_{adj}		0.42	
RMSE		59.6456	
MAE		46.063	

Tab. S3 - Model performance statistics using test data.

Test			
Coefficients (β)	**Fitted Values**	**Standard Error**	**t-valor**
B_0	418.0097	17.5254	2^{-16*}
β_1	2.4999	0.2067	2^{-16*}
β_2	-876.0725	49.1821	2^{-16*}
R^2_{adj}		0.42	
RMSE		61.8344	
MAE		47.9651	

Appendix 3 - Kriging and cokriging results.

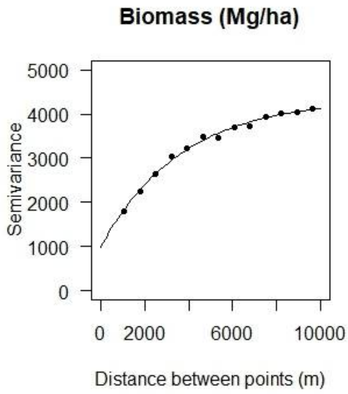

Fig. S1 - Modeling of the isotropic experimental semivariogram of biomass measured by LiDAR in Mg ha^{-1} for protected areas in the Carajás region, southeastern State of Pará.

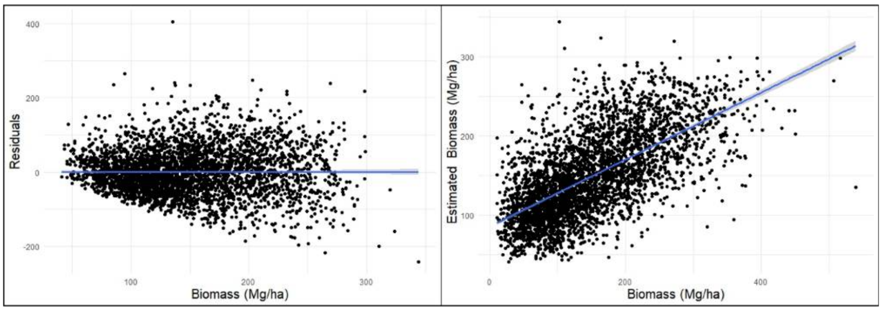

Fig. S2 - Modeling using multiple linear regression presenting the trend line of residuals and the Equation of the line between AGB values measured by LiDAR and those estimated by the model.

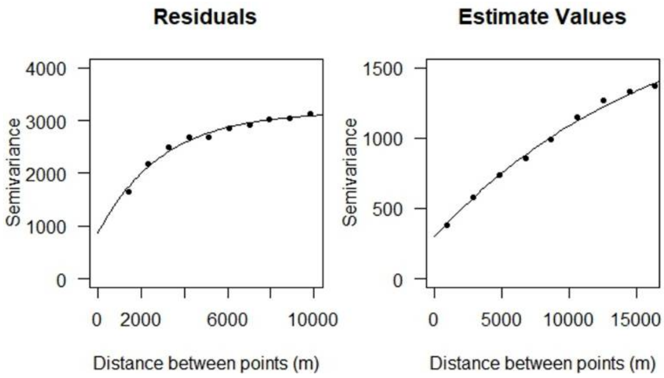

Fig. S3 - Modeling of the experimental semivariogram for residuals and regressed values using the exponential model. The residuals exhibited more significant variability compared to the estimated values, indicating that a classical statistical model may introduce high errors by not considering spatial continuity.

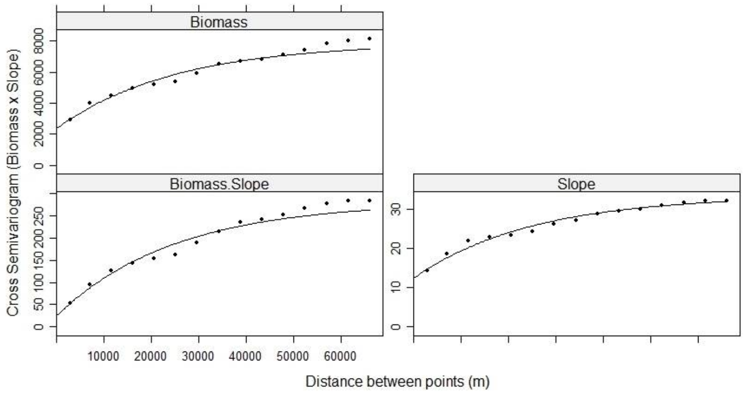

Fig. S4 - Modeling of the cross-experimental semivariogram between AGB and slope, resulting in a scaled semivariogram that captures the variability of this interaction.