

Article

Not peer-reviewed version

Isoscape of Oxygen Stable Isotopes in Woods of the Amazon

[Ana Claudia Gama Batista](#)*, Maria Gabriella Araújo, Isabela Maria Souza Silva, [Deoclécio Jardim Amorim](#), Fabiana Cristina Fracassi Adorno, [Gabriela Bielefeld Nardoto](#), [Vladimir Eliodoro Costa](#), Mario Tomazello Filho, [Niro Higuchi](#), [Perseu da Silva Aparicio](#), Yasmin Lara Bezerra Vieira da Silva, Marta Silvana Volpato Scoti, [Ana Carolina Barbosa](#), [Fabio José Viana Costa](#), João Paulo Sena Souza, Gabriel J. Bowen, [Luiz Antonio Martinelli](#)

Posted Date: 12 March 2026

doi: 10.20944/preprints202603.0937.v1

Keywords: stable isotope; $\delta^{18}\text{O}$; isoscapes; random forest; multiple regression; Amazon; forensic tracking



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

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

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.

Article

Isoscape of Oxygen Stable Isotopes in Woods of the Amazon

Ana Claudia Gama Batista ^{1,*}, Maria Gabriella Araujo ¹, Isabela Maria Souza Silva ¹, Deoclécio Jardim Amorim ¹, Fabiana Cristina Fracassi Adorno ¹, Gabriela Bielefeld Nardoto ², Vladimir Eliodoro Costa ³, Mario Tomazello Filho ⁴, Niro Higuchi ⁵, Perseu da Silva Aparicio ⁶, Yasmin Lara Bezerra Vieira da Silva ⁶, Marta Silvana Volpato Scoti ⁷, Ana Carolina Barbosa ⁸, Fabio José Viana Costa ⁹, João Paulo Sena Souza ¹⁰, Gabriel J. Bowen ¹¹ and Luiz Antonio Martinelli ¹

¹ Center for Nuclear Energy in Agriculture, Av. Centenario 303, 13416-000 Piracicaba-SP, Brazil

² Department of Ecology, Institute of Biological Sciences, University of Brasilia, Campus Universitário Darcy Ribeiro Bloco E s/n, CEP 70910-900 Brasília, Distrito Federal, Brazil

³ UNESP, Institute of Biosciences, Stable Isotopes Center, 18618689, Botucatu-SP, Brasil

⁴ Department Forest Sciences, Luiz de Queiroz College of Agriculture, University of São Paulo, Av. Pádua Dias, 11, Piracicaba, São Paulo, 13418-900, Brazil

⁵ National Institute for Amazon Research, Av. André Araújo, 2936, Petrópolis, CEP 69067-375, Manaus, Amazonas, Brasil

⁶ State University of Amapá, Av. Pres. Vargas, 650 - Central, CEP 68901-258 Macapá - AP, Brazil

⁷ Federal University of Santa Maria, Av. Roraima nº 1000 Cidade Universitária Bairro - Camobi, CEP 97105-900 Santa Maria - RS, Brazil

⁸ Department of Forest Sciences, University of Lavras, P.O.Box 3037, CEP 37203-202, Lavras, Minas Gerais, Brazil

⁹ National Institute of Criminalistic, Federal Police, Setor Policial Sul, Lote 7, Asa Sul, 70610-902, Brasília, DF, Brazil

¹⁰ State University of Montes Claros, Melo, CEP 39580-000, Francisco Sá - MG, Brazil

¹¹ Geology and Geophysics Department, University of Utah, Salt Lake City, UT, USA

* Correspondence: anaclaudiabatista89@gmail.com

Abstract

Stable oxygen isotopes ($\delta^{18}\text{O}$) in wood provide integrative records of plant water use and regional hydroclimatic processes, offering a powerful framework for spatial ecological analysis in tropical forests. Here, we present the first regional-scale $\delta^{18}\text{O}$ isoscapes for Amazonian wood based on 387 trees sampled across 25 sites. After α -cellulose extraction, $\delta^{18}\text{O}$ values were modeled using multiple linear regression (MLR) and Random Forest (RF) approaches. A Moran's I test revealed no significant spatial autocorrelation ($p = 0.73$), indicating that geostatistical interpolation methods such as kriging were not appropriate for this dataset. The MLR model based on site-average data achieved an R^2 of 0.70, with a mean absolute error (MAE) of 0.56‰ and root mean square error (RMSE) of 0.68‰. The RF model showed comparable performance ($R^2 = 0.67$; MAE = 0.64‰; RMSE = 0.77‰). Both approaches reproduced a coherent southeast-to-northwest gradient, with lower $\delta^{18}\text{O}$ values in the western Amazon and higher values in the east, consistent with regional patterns in precipitation isotopic composition and evapotranspiration. These findings demonstrate that climate-driven statistical modeling effectively captures large-scale isotopic structure across the Amazon basin, providing a robust spatial representation of $\delta^{18}\text{O}$ variability in tropical forest wood.

Keywords: stable isotope; $\delta^{18}\text{O}$; isoscapes; random forest; multiple regression; Amazon; forensic tracking

1. Introduction

Stable isotopes have been extensively utilized in both ecological and forensic sciences [7,32,33]. The isotopic composition of elements such as carbon, hydrogen, nitrogen, oxygen, and sulfur varies based on the climatic, geographic, and ecological characteristics of the environment. These spatial variations, referred to as isoscapes, can be modeled using statistical and machine learning techniques such as kriging and Random Forest [6,15,18,22,24,37]. Isoscapes have proven valuable in tracing the geographic origin of materials ranging from timber and wildlife to human remains and narcotics [19,26,30,31].

These environmental signals are absorbed by plants and incorporated into organic tissues through well-established physiological fractionation processes [1,36]. In wood, cellulose is preferentially analyzed because it preserves the isotopic signal formed during biosynthesis and provides a more consistent representation of environmental conditions than bulk wood. Among stable isotopes, $\delta^{18}\text{O}$ in tree-ring cellulose is particularly informative because it integrates both the isotopic composition of source water and leaf-level evaporative enrichment during transpiration. Consequently, $\delta^{18}\text{O}$ values in cellulose provide a robust proxy for local hydroclimatic conditions and, ultimately, geographic origin [17].

Although the field of isotopic spatial modeling is well established, isoscapes remain scarce in tropical regions, particularly in the Brazilian Amazon, where applications are urgently needed. In particular, forensic approaches to timber tracking could benefit substantially from spatial isotopic tools. Despite existing regulatory frameworks, traceability challenges persist in parts of the Amazon timber supply chain, highlighting the need for complementary origin-verification methods based on chemical or genetic signatures. Recent studies have demonstrated the potential of isotopic models for this purpose [10,29,33,39], yet efforts remain fragmented and geographically limited.

In this study, we develop a regional $\delta^{18}\text{O}$ isoscape for Amazonian tree-ring cellulose to examine the extent to which large-scale hydroclimatic gradients structure isotopic variability across the basin. We hypothesize that $\delta^{18}\text{O}$ patterns in tropical wood are primarily governed by broad atmospheric and moisture-related controls rather than by local spatial dependence, reflecting the integration of source-water isotopic composition and evaporative enrichment processes at the leaf level. By evaluating how different statistical modeling frameworks represent these environmental controls, we seek to clarify the mechanisms shaping $\delta^{18}\text{O}$ spatial structure in highly heterogeneous tropical forests. Through this approach, we aim to advance the conceptual and methodological foundation for constructing robust isotopic surfaces in the Amazon, where complex hydroclimatic dynamics and ecological variability challenge conventional spatial modeling assumptions.

2. Material and Methods

2.1. Isotopic Data Collection

To construct the $\delta^{18}\text{O}$ -isoscape, we sampled 387 trees across 25 locations throughout the Amazon Basin (Figure 1). Species of commercial importance were initially targeted. However, due to the challenges of sampling the same species across the Amazon, expanded the sampling strategy to include mature trees regardless of their commercial significance.

Representatives from 58 genera distributed across 24 botanical families were collected. These forest species represent different ecological groups, including pioneers, secondary, and climax or later-successional, occurring in mixed or dense ombrophilous forests and reflecting the phytosociology diversity of the Amazon region.

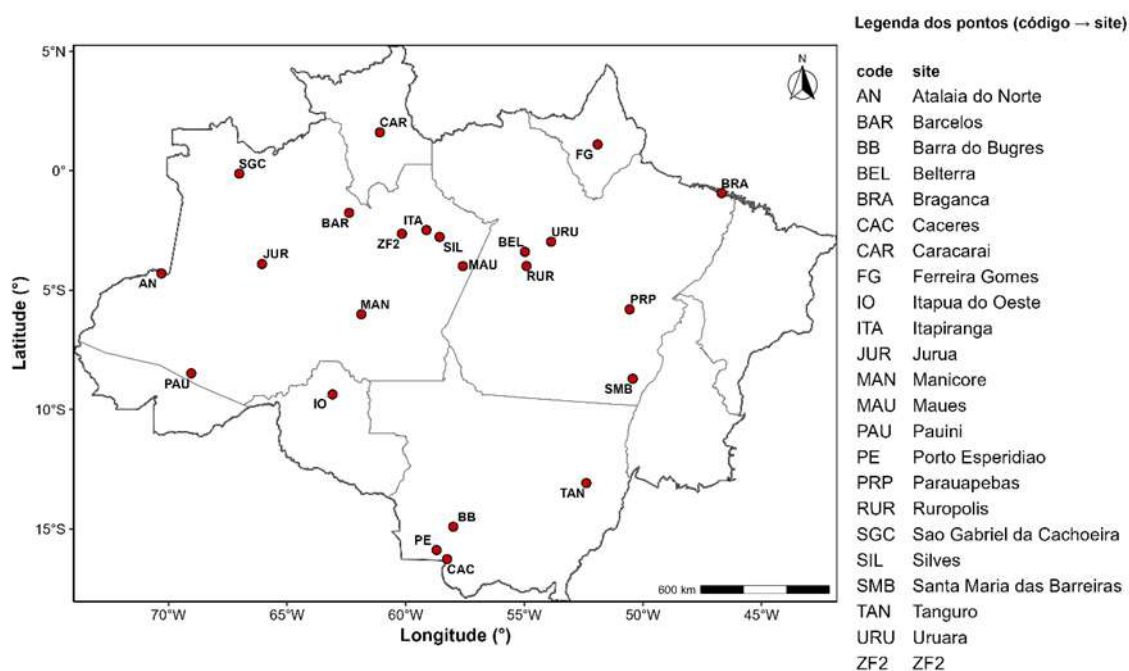


Figure 1. Map of sample distribution in the Amazon basin. The red dots indicate the sampling locations, covering various regions of the Legal Amazon, including National Forest (Flonas) and strategic areas for isotopic wood analysis.

Samples were collected from the base of trees harvested under authorized forest management plans within the Amazon region, in accordance with CONAMA Resolution N^o 406 of February 2, 2009[14]. The minimum cutting diameter (MCD) was ≥ 50 cm, and discs with a minimum thickness of 5 cm were obtained. Each wood disc was processed radially from the pith to the bark, and 2 mm thick laths were prepared, with their length varying according to the diameter of the collected disc.

The laths were proportionally segmented along the radial at (0% (pith), 25%, and 50% (heartwood), 75% (heartwood-sapwood transition), 100% (sapwood up to the bark)) (Figure 2). Each segment approximately 5 to 10 growth rings, enabling characterization of the radial variation in isotopic signal.

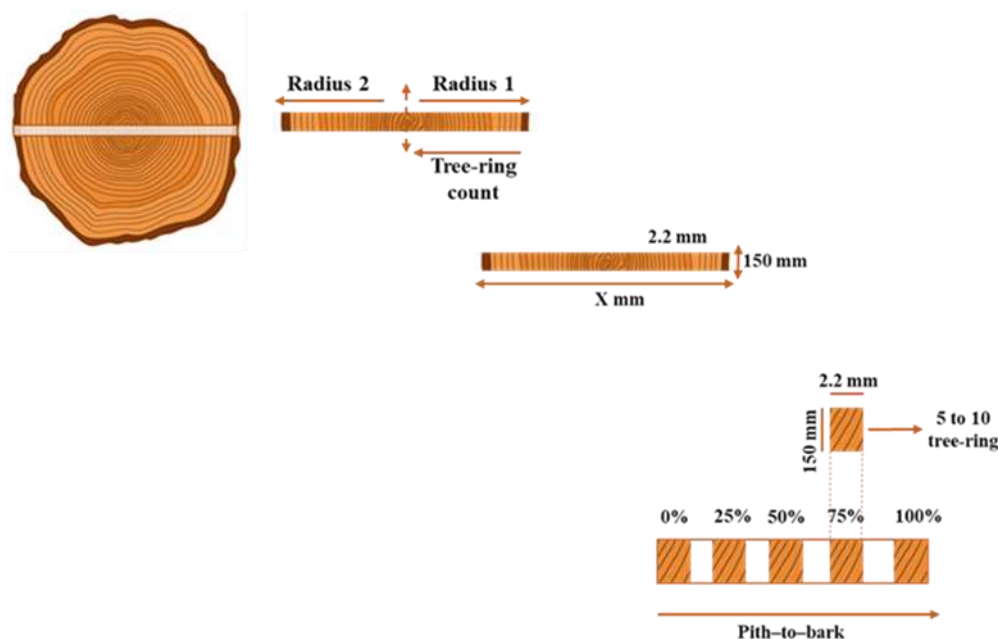


Figure 2. Preparation of the tree-ring laths. An entire radial section from pith to bark was removed from both sides of the stem disc. As both sides were considered symmetrical, the radius was divided in half, and ring counting started at the pith. The length of the laths was not standardized and depended on the disc diameter; however, all laths were sawn to a uniform thickness of 2 mm. The laths were marked at 0% (pith), 25% and 50% (heartwood), 75% (heartwood–sapwood transition), and 100% (sapwood–bark transition). These segment positions were defined to maximize the number of growth rings per section. The segments were cut into 2 cm pieces, each containing approximately 5–10 growth rings.

The tree-ring laths were chemically treated to obtain α -cellulose following the protocols established by the Isotopic Ecology Laboratory – CENA/USP [4], which are based on previously published and widely applied methodologies [23,28]. Isotopic analyses were partially conducted at the Stable Isotope Laboratory of the University of California, Davis (USA), and at the Stable Isotope Center at São Paulo State University (UNESP), Botucatu, Brazil. An intercalibration procedure was performed between laboratories to ensure analytical consistency and standardization of the results.

For oxygen isotope analyses, internationally recognized USGS reference materials (USGS54, USGS55, USGS56, USGS90, and USGS91) were used for calibration and quality control. The samples were analyzed for $^{18}\text{O}/^{16}\text{O}$ ratios using an isotope ratio mass spectrometer (IRMS) coupled with an automated pyrolysis device. The δ values were expressed as the relative difference between the isotopic ratio of the sample and that of the international standard (VSMOW for ^{18}O) [35], as shown in Equation 1.

$$\delta = ((R_{\text{sample}} / R_{\text{standard}}) - 1) \times 1000 \quad (1)$$

2.2. Environmental Data

Environmental variables such as Mean Annual Temperature (MAT), Mean Annual Precipitation (MAP), Potential Evapotranspiration (PET), and Vapor Pressure Deficit (VPD) were obtained from the WorldClim platform, version 2.1 (2020).

The data were extracted at a spatial resolution of 10 minutes (~340 km²). Each dataset contains 12 GeoTiff (.tif) files, one for each month of the year, numbered from January (1) to December (12). The extraction of these files was performed using QGIS software.

The Digital Elevation Model (DEM) was also obtained from the WorldClim platform, with elevation data derived from the Shuttle Radar Topography Mission (SRTM). These files were extracted at the same spatial resolution of 10 minutes, with an elevation precision of 10 meters. The rasters were acquired using QGIS software. Atmospheric Pressure (PA) was calculated according to Equation 2.

$$P(h) = P_0 \times (1 - h / 44330)^{5.255} \quad (2)$$

where $P(h)$ is the atmospheric pressure at different elevations, P_0 is the sea-level pressure (1013.25 hPa), and h is the altitude in meters from the DEM.

Relative Humidity (RH) was obtained using Equation 3. For the calculation of RH, water vapor pressure (kPa) raster files were first extracted from the WorldClim platform, with a spatial resolution of 10 minutes. The extraction of these rasters was performed using QGIS software.

$$\text{RH} = (C_a / e_s(T)) \times 100 \quad (3)$$

where RH is the relative humidity (%), C_a is the vapor pressure obtained from the raster, and $e_s(T)$ is the saturation vapor pressure, calculated using Equation 4.

$$e_s(T) = 0.6108 \times \exp((17.24 \times T) / (237.3 + T)) \quad (4)$$

where $e_s(T)$ is the saturation vapor pressure (kPa), 0.6108 is a constant expressed in kPa, \exp is the exponential function, 17.24 represents the linear dependence on temperature, $237.3 + T$ accounts for the increase in vapor pressure with temperature, and T is temperature in degrees Celsius (°C).

The $\delta^{18}\text{O}$ values of precipitation ($\delta^{18}\text{O}_{\text{ppt}}$) were determined by interpolating precipitation data from The Global Network of Isotopes in Precipitation (GNIP) database using IsoriX, an R package for constructing isoscapes using Generalized Linear Mixed Models (GLMMs) [16]. This process enables the creation of precipitation isoscapes using annual mean $\delta^{18}\text{O}$, generating raster data and extracting $\delta^{18}\text{O}$ values from precipitation for each sampling point based on latitude and longitude.

The GLMM model in IsoriX accounts for spatial and temporal variations in isotopes and allows the response variable to follow a distribution other than normal, incorporating both fixed and random effects (Equation 5).

$$\delta^{18}\text{O}(S_i) = \beta_0 + \beta_1\text{DEM} + \beta_2\text{T} + \mu(S_i) \quad (5)$$

where $\delta^{18}\text{O}(S_i)$ is the isotopic value at spatial location S_i , β_0 is the intercept, β_1 and β_2 are coefficients for the covariates DEM and temperature, and $\mu(S_i)$ is the spatially autocorrelated error term, modeled using the Matérn function.

All environmental variables were extracted for each sampling location using the WorldClim database (Table 1). In addition, the isotopic composition of precipitation ($\delta^{18}\text{O}_p$) was interpolated for each site using the IsoriX package [16], based on the geographic coordinates and WorldClim environmental predictors. These interpolated $\delta^{18}\text{O}_p$ values were later used as covariates in the modeling of $\delta^{18}\text{O}$ in wood cellulose. The spatial distribution of modeled $\delta^{18}\text{O}_p$ values reflect large-scale atmospheric processes, such as Rayleigh distillation and temperature-driven fractionation along air-mass trajectories [20,38].

Table 1. Average data of environmental variables extracted from WorldClim (2020).

Site	State	Lat DD	Lon DD	VPD Kpa	RH%	PET mm	DEM m	MAT °C	MAP mm	$\delta^{18}\text{O}_{\text{ppt}}$ %
Maués	Amazonas	-57.589	-3.996	0.95	72.16	1390	51	27.00	2209.36	-3.81
São Gabriel da Cachoeira	Amazonas	-67.013	-0.121	0.90	72.45	1256	97	25.97	2902.27	-3.93
Rurópolis	Pará	-54.908	-3.992	1.09	67.93	1320	120	25.94	1830.03	-3.55
Juruá	Amazonas	-66.055	-3.907	0.93	71.87	1235	81	26.45	2849.93	-4.57
Atalaia do Norte	Amazonas	-70.291	-4.304	0.89	72.81	1205	107	26.45	2702.18	-5.12
Manicoré	Amazonas	-61.869	-6.010	0.97	71.57	1276	64	26.84	2751.36	-4.45
ZF2	Amazonas	-60.150	-2.638	1.08	68.02	1328	97	26.39	2228.53	-4.06
Itapiranga	Amazonas	-59.121	-2.493	0.96	70.90	1361	113	26.47	2274.51	-3.99
Tanguro	Mato Grosso	-52.377	-13.081	1.26	60.52	1521	381	24.76	1621.02	-3.71
Itapiranga	Amazonas	-59.121	-2.497	0.96	70.90	1361	113	26.47	2274.51	-3.99
Belterra	Pará	-54.974	-3.398	1.09	67.64	1313	144	25.73	1847.88	-3.54
Rurópolis	Pará	-54.908	-3.992	1.09	67.93	1320	120	25.94	1830.03	-3.55
Uruará	Pará	-53.871	-2.974	1.08	67.43	1309	196	25.27	1897.40	-3.31
Barcelos	Amazonas	-62.382	-1.764	0.91	72.65	1299	34	26.38	2576.45	-4.03
Pauini	Amazonas	-69.036	-8.484	0.99	69.01	1182	207	24.62	2352.68	-5.39
Itapuã do Oeste	Rondônia	-63.083	-9.366	1.10	67.08	1257	138	25.23	2249.53	-4.91
Barra do Bugres	Mato Grosso	-57.993	-14.906	1.24	61.26	1481	282	24.79	1695.15	-4.72
Santa Maria das Barreiras	Pará	-50.429	-8.709	1.31	61.42	1467	216	25.83	1744.29	-3.06
Ferreira Gomes	Amapá	-51.909	1.100	1.07	67.06	1530	248	25.48	2506.87	-2.84
Cáceres	Mato Grosso	-58.253	-16.264	1.29	60.20	1623	233	26.11	1378.89	-4.66
Porto Esperidião	Mato Grosso	-58.697	-15.882	1.38	58.47	1640	198	26.24	1466.49	-4.70
Caracarái	Roraima	-61.088	1.604	1.08	68.66	1506	60	27.33	1809.22	-3.84
Silves	Amazonas	-58.569	-2.773	1.09	67.82	1365	77	26.74	2190.13	-3.90
Parauapebas	Pará	-50.560	-5.805	1.18	63.98	1383	233	25.22	1898.79	-3.04

Bragança	Pará	-46.681	-0.940	1.16	66.21	1562	14	26.34	2369.48	-1.85
----------	------	---------	--------	------	-------	------	----	-------	---------	-------

Lat DD Latitude in decimal degrees, *Long DD* Longitude in decimal degrees, *VPD* Vapor Pressure Deficit, *RH* Relative Humidity, *PET* Potential Evapotranspiration, *DEM* Digital Elevation Model, *MAT* Mean Annual Temperature, *MAP* Mean Annual Precipitation, $\delta^{18}\text{O}_{\text{ppt}}$ $\delta^{18}\text{O}$ of Precipitation.

These parameters (Table 1) were considered due to the global relationship between latitude, longitude, and $\delta^{18}\text{O}_{\text{ppt}}$, allowing for the prediction of local $\delta^{18}\text{O}$ values [11,12,40,43]. Additionally, large-scale patterns in atmospheric vapor transport also influence $\delta^{18}\text{O}$ in precipitation.

2.3. Construction of the Isotopic Model

Multiple Linear Regression Model

The isotopic model was developed using the caret package in R [25]. First, the predictor variables were standardized: the mean was subtracted from individual observations, and the resulting values were divided by the standard deviation, allowing for the observation of effect size. Subsequently, a maximal model (Equation 6) was constructed, incorporating potential predictor variables for $\delta^{18}\text{O}$ in the Amazon region.

$$\delta^{18}\text{O} = \beta_0 + \beta_1\text{VPD} + \beta_2\text{RH} + \beta_3\text{PET} + \beta_4\text{DEM} + \beta_5\text{PA} + \beta_6\text{MAT} + \beta_7\text{MAP} + \beta_8\delta^{18}\text{O}_{\text{ppt}} \quad (6)$$

From this maximal model, all possible combinations of candidate variables were tested using the dredge function from the MuMIn package [3]. This process resulted in 256 unique models, from which the best was selected based on Information Theory (IT), considering those with $\Delta\text{AICc} < 4$ [9,21]. As a result, five models were retained.

To evaluate model performance, Leave-One-Out Cross-Validation (LOOCV) was employed [8]. This method involves excluding one test data point, building the model with the remaining data, and calculating error metrics such as R^2 , RMSE, and MAE. With model selection complete, the model equations were applied to variable values from the geospatial raster grids to construct the isoscape using the simplest (parsimonious) model.

We tested the regression model using both individual tree-level $\delta^{18}\text{O}$ values and site-averaged $\delta^{18}\text{O}$ values to evaluate its performance and applicability for timber provenance. Our main objective in making this comparison was to address a persistent question in timber forensics. In real-world applications, one typically has only a single wood sample to work with, rather than multiple samples from the same site that can be averaged to reduce variability. Therefore, while site-average models are valuable for identifying regional isotopic patterns, individual-level models are more realistic and directly applicable for forensic provenance, where predictions must be made from single samples. However, this approach introduces greater uncertainty due to natural intra-site variability, which in turn tends to reduce the model's coefficient of determination. It is thus important to address the trade-off between model precision and practical applicability, balancing statistical performance with the realities of forensic casework.

Finally, we estimated Moran's I statistic for the residuals of the regression model to assess spatial autocorrelation. The result ($p = 0.73$) indicated no significant spatial structure, justifying the exclusion of kriging from the modeling approach.

2.4. Random Forest

Random Forest (RF) is a machine learning (ML) technique applied in modeling[27,41]. We predicted $\delta^{18}\text{O}$ of cellulose using RF and tested different covariates to select the most accurate model. Recursive feature elimination (RFE) was applied to select representative variables using the caret package. RFE constructs a model with all variables, ranks their importance, and iteratively eliminates the least important. Model performance was evaluated using repeated 5-fold cross-validation (10 repetitions) via the trainControl function. Metrics such as RMSE and R^2 were obtained from this process.

Like for the multiple regression model, we trained Random Forest models using individual tree-level $\delta^{18}\text{O}$ values and site-averaged $\delta^{18}\text{O}$ values to evaluate model performance and applicability to timber provenance. For provenance applications specifically, the model trained on site-mean values was selected for isoscape prediction and geographic assignment due to several critical methodological considerations. However, we also trained a Random Forest model using individual tree-level $\delta^{18}\text{O}$ values in order to evaluate the full isotopic variability, including within-site variation. This allowed us to assess how the inclusion of biological noise arising from factors such as species differences, microclimatic conditions, and tree age impacts model performance.

The RF model was trained using *caret package* [25], with a grid search for tuning hyperparameters (Table 2). Variable importance was assessed, and predictions were compared to observed values. Environmental rasters were used to predict $\delta^{18}\text{O}$ values across the Amazon Basin, and an isoscape was created and plotted using *ggplot2 package* [42], with the Legal Amazon boundaries serving as a reference. To better understand the structure of isotopic variability, we trained two Random Forest models: one using individual $\delta^{18}\text{O}$ values from each tree and another using site-level means.

Table 2. Hyperparameter for variable selection in the random forest model.

Hyperparameter	Description	Unit
Longitude	Geographic coordinate (X)	Decimal degrees ($^{\circ}$)
Latitude	Geographic coordinate (Y)	Decimal degrees ($^{\circ}$)
VPD	Vapor pressure deficit	KPa
RH	Relative humidity	%
PET	Potential evapotranspiration	mm
DEM	Digital elevation model	Meters (m)
MAT	Annual average temperature	$^{\circ}\text{C}$
MAP	Annual average precipitation	mm
$\delta^{18}\text{O}_{ppt}$	Predictive model of $\delta^{18}\text{O}$ in precipitation	‰
$\delta^{18}\text{O}$	Isotopic ratio of ^{18}O	‰

Table 3. Models selected by information theory, $\Delta\text{AICc} < 4$.

	Models	$\Delta\text{AICc} < 4$
1	$\delta^{18}\text{O} = \beta_1\text{DEM} + \beta_2\text{Isorix} + \beta_3\text{MAP} + \beta_4\text{MAT} + \beta_5\text{PET} + \beta_6\text{RH} + \beta_7\text{VDP}$	4456
2	$\delta^{18}\text{O} = \beta_1\text{DEM} + \beta_2\text{Isorix} + \beta_3\text{MAP} + \beta_4\text{MAT} + \beta_5\text{RH} + \beta_6\text{VPD}$	4458
3	$\delta^{18}\text{O} = \beta_1\text{DEM} + \beta_2\text{Isorix} + \beta_3\text{MAP} + \beta_4\text{PET} + \beta_5\text{RH} + \beta_6\text{VPD}$	4459

3. Results

3.1. Multiple Linear Regression

The model with the lowest AIC included the following climatic variables: *mat*, *vpd*, *map*, *rh*, and *pet*. The coefficients produced by this model are shown in Table 1. The model built on the site-average data achieved an R^2 of 0.70, indicating that these predictors account for a significant portion of the variance in $\delta^{18}\text{O}$ of the cellulose. The MAE was 0.56 ‰ , and the RMSE was 0.68 ‰ (Table 4). Although not shown, the residuals followed a normal distribution, with no discernible pattern between the predicted values and the residuals. For the entire data set, there was a decrease in the R^2 to 0.42, and an increase of the MAE and RMSE to 0.92 ‰ and 1.10 ‰ , respectively (Table 4).

Table 4. Quality metrics for multiple linear regression (MLR) and random forest (RF) models used to predict the $\delta^{18}\text{O}$ of tree-ring cellulose in the Amazon region. Models were trained using either individual tree-level data or site-averaged isotopic values. MAE is the mean absolute error; RMSE is the root mean square error; R^2 is the coefficient of determination. For the MLR models, metrics were calculated from observed versus predicted values. For the RF models, metrics were estimated using internal leave-one-out cross-validation via the *caret package* in R.

	MLR-individual	MLR-site average	RF-individual	RF-site average
MAE ()	0.92	0.56	0.90	0.64
RMSE ()	1.10	0.68	1.12	0.77
R²	0.42	0.70	0.44	0.67
Std.dev (‰)				
Min.	1.15	0.72	0.01	0.17
1st Q	1.16	0.75	0.95	0.70
Median	1.17	0.78	1.25	0.87
3rd Q	1.19	0.83	1.48	1.10
Max.	2.00	1.86	1.97	2.00

The isoscape prediction grid spans longitudes between -74.00 and -43.50, and latitudes between -16.67 and 4.50, with a total of 12,392 points. The spatial grid has 183 cells on the x-axis (longitude) and 127 on the y-axis (latitude), with each cell measuring approximately 0.16 degrees, corresponding to spatial resolution of about 18 km² per cell. Model predictions based on site-average data ranged from 22.7‰ to 35.0‰, with a median value of 26.1‰ and interquartile range from 25.1‰ to 27.7‰. Predictions using individual tree $\delta^{18}\text{O}$ values yielded a similar distribution, ranging from 23.8‰ to 33.2‰, with the same median (26.1‰) and interquartile range (25.1‰ to 27.7‰). These predicted values align closely with the observed $\delta^{18}\text{O}$ measurements, which had a median of 25.8‰, interquartile range (24.8‰ to 26.7‰) and ranged from 22.1‰ to 30.0‰.

A marked spatial gradient in $\delta^{18}\text{O}$ values is evident across the Amazon region, with values decreasing from east to west, particularly along a southeast-to-northwest axis (Figure 3A). Although Figure 3 presents the isoscape generated from site-average $\delta^{18}\text{O}$ values, a remarkably similar spatial pattern was observed in the isoscape derived from individual tree-level data (Figure S2), reinforcing the robustness of the observed isotopic gradient across the region.

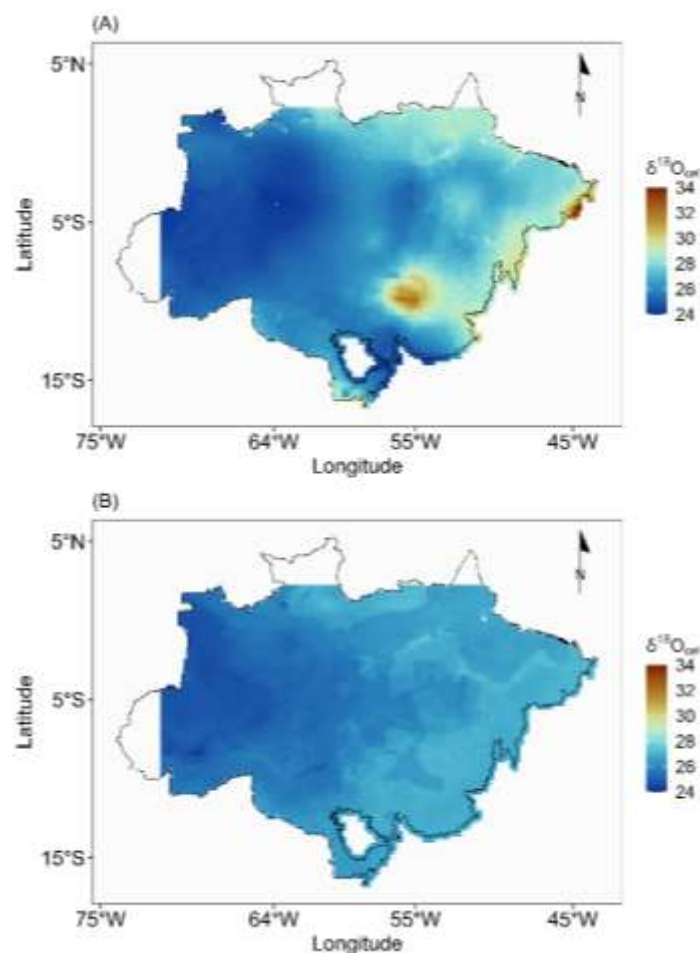


Figure 3. Average isoscapes of $\delta^{18}\text{O}$ in cellulose across the Amazon region, modeled using (A) multiple linear regression (MLR) and (B) random forest (RF), based on site-averaged $\delta^{18}\text{O}$ values. Black lines indicate the boundaries of the Amazon region. Warmer colors represent higher $\delta^{18}\text{O}$ values, while cooler colors indicate lower values.

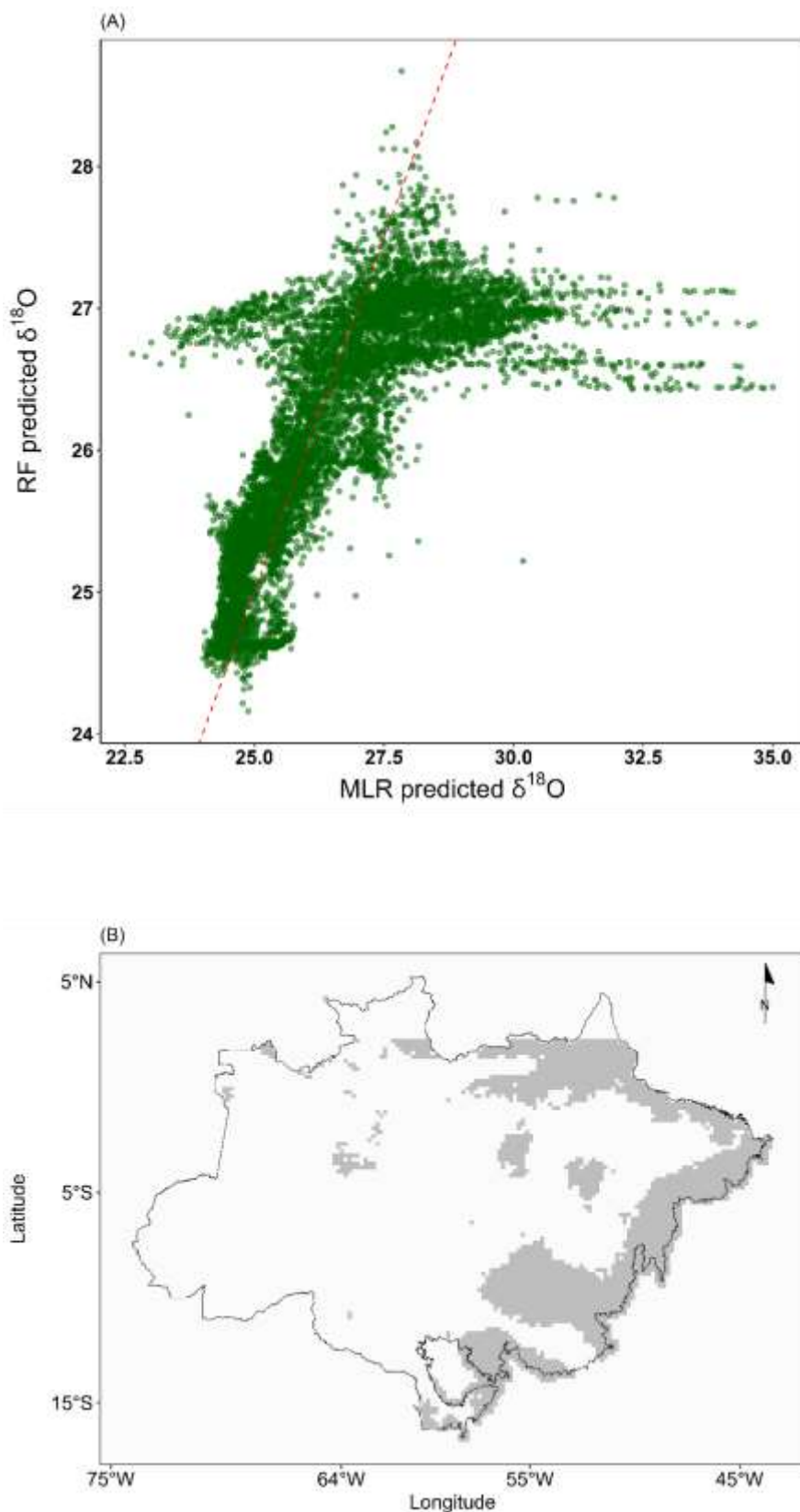


Figure 4. Comparison of $\delta^{18}\text{O}$ predictions in tree-ring cellulose across the Amazon region from the multiple linear regression (MLR) and random forest (RF) models. (A) Scatterplot of $\delta^{18}\text{O}$ predictions from both models.

The dashed red line represents the 1:1 relationship. (B) Spatial map highlighting areas where the difference between MLR and RF predictions falls within the $\pm 1\%$ interval (white), and areas with differences exceeding $\pm 1\%$ (gray).

The standard deviation of the predictions using site-average $\delta^{18}\text{O}$ values ranges from 0.72‰ to 1.9‰. Descriptive statistics indicate an interquartile range of 0.75‰, to 0.83‰, with a median of 0.78‰ and mean of 0.83 ‰. The standard deviation of the predictions using individual trees $\delta^{18}\text{O}$ values were higher, as expected. The median was 1.17‰, the interquartile range from 1.16‰ to 1.19‰, and the range from 1.15‰ to 2.05‰. Areas of low prediction uncertainty were distributed across the area, indicating relatively high confidence across much of the Amazon region (Figures S3 and S4). However, the areas with higher standard deviations tend to coincide with regions where sampling density is lower (Figure 1). Notably, many of these higher-uncertainty areas fall within the so-called "arc of deforestation", a zone of intense land-use changes and ongoing illegal logging activity.

3.2. Random Forest

The random forest model trained on the site-average data, yielding the lowest MAE of 0.64‰, and RMSE of 0.77‰ and a coefficient of determination (R^2) of 0.67, indicating that the model explains approximately 67% of the variance in the data (Table 4). Thus, such quality metrics are comparable with those produced by the multiple linear regression model. These metrics deteriorate when the whole dataset is used to train the model. The MAE and RMSE increased to 0.90‰ and 1.12‰, with a R^2 of 0.44 (Table 4). Likewise, these metrics were also equivalent to those produced by the multiple linear regression model when using individual tree data.

Both the site-average and individual-tree isoscapes produced by the RF model showed similar spatial patterns, characterized by a southeast-to-northwest gradient in $\delta^{18}\text{O}$ values. The site-average RF isoscape is presented in Figure 3B, while the individual-tree RF isoscape appears separately in the Supplementary Material (Figure S3). Despite these similarities, an important distinction emerged when comparing the RF and MLR model outputs: the range of predicted values was considerably narrower in the RF isoscape than in the MLR-based one (Figure 3A). Notably, the greatest spatial differences between the two models were found in the eastern Amazon, in the so called "arc of deforestation" where the models diverged most markedly (Figure 3B).

The standard deviation of the predictions using QRF and the site average dataset had slightly higher values than the standard deviations estimated using the MLR model. The median value was 0.87‰, the Q1 and Q3 were 0.70‰ and 1.10‰, respectively; and the range was from 0.2‰ to 2.0‰. However, the spatial distribution of standard deviation values across the Amazon was different than the one estimated by the MLR models. While the standard deviation isoscape by MLR produced a smooth surface with gradual transition (Figures S3 and S4), the standard deviation isoscape by QRF produced patchy surfaces due to the differences in model structure between the two methods (Figures S5 and S6). The standard deviation isoscape produced by the QRF model exhibits a patchy spatial pattern due to its data-driven estimation of local prediction uncertainty, while the smoother standard deviation surface from the MLR model reflects its assumption of homoscedasticity and linear relationships between predictors and response.

4. Discussion

This study represents the first effort to construct a $\delta^{18}\text{O}$ isoscape from tree-ring cellulose across the Brazilian Amazon, aiming to contribute to the development of a multi-isotopic model to trace the origin of timber and combat illegal logging in the region. We compared the performance of two spatial interpolation approaches: multiple linear regression (MLR), and random forest (RF), a machine learning technique capable of capturing complex, nonlinear interactions between environmental variables. We also tested two data handling approaches: one using the $\delta^{18}\text{O}$ values from individual trees, and the other using site-averaged values.

Our main findings show that the multiple linear regression (MLR) model performed slightly better than the Random Forest (RF) model, suggesting that the $\delta^{18}\text{O}$ signal is primarily linear (Table 4). The comparison also highlights a central conceptual distinction in isotopic modeling for timber provenance. Models trained on site-averaged values capture regional or population-level isotopic patterns, whereas models based on individual samples explicitly account for within-site variability and are theoretically more appropriate for forensic provenance, where predictions must be made for single trees. Understanding this theoretical distinction is essential when designing models for enforcement or ecological applications. Additionally, it is crucial to evaluate how the use of individual versus site-averaged data affects the resulting $\delta^{18}\text{O}$ isoscapes, as this may influence their spatial resolution and interpretability.

Both methods yielded isoscapes that reveal similar broad spatial patterns, with lower $\delta^{18}\text{O}$ values in the western Amazon, higher values toward the east, and intermediate values in the central region. A latitudinal trend was also apparent, with elevated values at northern and southern extremes and lower values in the central Amazon. These patterns suggest that, despite differences in their modeling approaches, both methods can capture some of the large-scale environmental gradients influencing $\delta^{18}\text{O}$ distribution.

The east-to-west decrease in $\delta^{18}\text{O}$ values of cellulose mirrors the isotopic pattern observed in rainfall across the Amazon Basin [1,37]. This trend likely reflects Rayleigh-type fractionation, whereby heavy oxygen isotopes (^{18}O) are progressively removed from atmospheric moisture during condensation and precipitation events along the prevailing Easterly moisture pathway [2,13]. As air masses travel inland from the Atlantic Ocean, successive rainfall events lead to the gradual depletion of ^{18}O in water vapor, a signal ultimately captured in plant cellulose. Higher $\delta^{18}\text{O}$ values in the east suggest shorter vapor transport distances and limited moisture recycling, whereas lower values in the west may indicate more extensive continental recycling and greater isotopic distillation. This east-west gradient underscores the dominant influence of Atlantic-derived moisture and its progressive isotopic modification during inland transport.

This example illustrates how synoptic-scale isotope effects on precipitation become imprinted in sessile organisms such as trees. The consistency between rainfall isotopes and cellulose $\delta^{18}\text{O}$ reinforces the value of tree-ring isotopes as a powerful tool for deciphering complex climatic and hydrological patterns across large spatial scales.

Although the general spatial trends captured by the MLR and RF models were similar, the RF model predicted a noticeably narrower range of $\delta^{18}\text{O}$ values compared to the MLR model. This pattern likely reflects the limited ability of Random Forest to extrapolate relationships between predictor and response variables beyond the calibration range, resulting in more conservative predictions within the bounds of the training data. In contrast, the MLR model can extend predictions linearly outside this range, potentially capturing broader gradients. Expanding the dataset through future sampling particularly in underrepresented regions will provide an opportunity to test the strength and stability of the linear relationships estimated by the MLR model and to evaluate whether these patterns persist as spatial coverage improves.

Most predicted standard deviations in the isoscapes were below 1‰; however, certain localized areas exhibited notably higher values, as shown in the standard deviation isoscapes generated by the MLR model (Figures S2 and S3). Many of these high-uncertainty regions overlap with the arc of deforestation. This convergence of sparse sampling, elevated model uncertainty, and ongoing environmental disturbance highlights a critical gap in current monitoring efforts. Future sampling strategies and model improvements should focus on reducing uncertainties in these vulnerable areas to improve the accuracy of provenance assessments and reinforce the application of isotopic tools in environmental enforcement.

Considerable variation in $\delta^{18}\text{O}$ was observed among individual trees within the same plots, likely reflecting species-specific physiological traits and localized micro-environmental differences [5]. This highlights the importance of extensive within-site sampling to capture this variability and improve model resolution. While the current $\delta^{18}\text{O}$ isoscapes are effective at distinguishing trees from

isotopically distinct regions such as the far eastern and far western Amazon they are less reliable in isotopically homogeneous areas, particularly in the eastern Amazon, where logging pressure is most intense.

The predicted spatial range of $\delta^{18}\text{O}$ values across the basin is approximately 3‰, while the model uncertainty, expressed as RMSE, is around 0.7‰. This suggests that regional differentiation is robust only where isotopic gradients exceed model uncertainty, for example, between western and eastern sites, but becomes ambiguous within low-contrast areas of central and eastern Amazonia. These results highlight both the potential and limitations of using $\delta^{18}\text{O}$ isoscapes for timber provenance: while broad-scale distinctions are feasible, fine-scale geographic assignments remain uncertain. Expanding the spatial coverage of sampling and integrating complementary isotopic tracers (e.g., $\delta^2\text{H}$ or $^{87}\text{Sr}/^{86}\text{Sr}$) will be essential to enhance discriminative power in isotopically homogeneous regions where enforcement needs are greatest.

5. Conclusions

This study provides the first regional $\delta^{18}\text{O}$ isoscapes for Amazonian wood and offers important methodological insights for isotope-based timber provenance in tropical forests. Although spatial autocorrelation was not statistically detected in our dataset, the coherent east–west gradients observed in the predictive maps indicate that isotopic structure likely exists but was not fully captured at the current sampling scale. Both multiple linear regression and random forest models generated broadly similar spatial patterns and comparable predictive errors; however, differences in predicted isotopic range were more evident in regions undergoing rapid environmental transformation, such as the arc of deforestation, highlighting that model selection may influence spatial interpretation and assignment sensitivity. By explicitly comparing site-averaged and individual-tree approaches, we demonstrate a fundamental trade-off between statistical robustness and forensic realism: while site means improve model fit and correlation strength, individual-level data more accurately reflect real-world enforcement scenarios, despite greater within-site variability. These findings underscore that isoscape development in highly heterogeneous tropical forests must balance generalizability with operational applicability. Future advances will depend on expanding spatial sampling coverage and integrating ecophysiological relevant covariates, including traits such as stomatal conductance, successional status, and wood density, which influence plant water use and isotopic discrimination processes. In this sense, Amazonian timber isoscapes should no longer be regarded as static climatic surfaces, but as dynamic eco-physiological representations of forest structure, function, and environmental change.

Supplementary information: Supplementary data associated with this article can be found in the online version at <https://doi.org/10.5281/zenodo.18852894>.

Authorship contribution: Ana C.G. Batista: Writing, data analysis, review, and editing. Luiz A. Martinelli: Writing, data analysis, review. Isabela M.S. Silva: Review. Maria G. Araújo: Review. Mario Tomazello-Filho: Review and sample collection donation. Niro Higuchi: Review and sample collection donation. Gabriela B. Nardoto: Review. Deoclecio J. Amorim: Data analysis and review. Ana C. Barbosa: Review and sample collection donation. Vladmir E. Costa: Review. Perseu S. Aparicio: Review. Yasmin L.B.V. Silva: Collected and prepared sample and review. Marta S.V. Scoti: Review. Fabio J.V.C.: Review. João P. Sena-Souza: Review. Gabriel J. Bowen: Review. Fabiana C.F.A.: Methodology process and prepared samples.

Funding: This work was funded by the CAPES funding agency (Academic Cooperation Program in Public Security and Forensic Sciences – PROCAD), by INCT – CNPq (Forensic Metrology and Traceability in Agro-Environmental Quality – MRFor), and by The Nature Conservancy Brazil (TNC), in partnership with Google.

Declarations: Conflict of interest We have nothing to declare.

Acknowledgments: Ana Claudia Gama Batista acknowledges the Coordination for the Improvement of Higher Education Personnel – Brazil (CAPES) for the doctoral scholarship granted through the Academic Cooperation

Program in Public Security and Forensic Sciences (Process nº 88887.626727/2021-00), which supported the development of this research. This work was supported by Forest Management – INPA, the Dendrochronology Laboratory – UFLA, and the Wood Anatomy and Identification Laboratory – ESALQ, to whom we express our sincere gratitude for the donation of wood samples. We also acknowledge Forest Ark for logistical support and assistance with wood sample collection in the Jamari National Forest (Floresta Nacional do Jamari – FLONA Jamari), Rondônia, Brazil, and we sincerely thank Fernanda Almeida for her support during field activities. We further thank RRX Timber, a forest concessionaire operating in the Amapá National Forest (Floresta Nacional do Amapá – FLONA Amapá), under the responsibility of Forest Engineer Guilherme Stucchi, for supporting field activities and facilitating wood sample collection. We especially thank Aparecido Candido Siqueira, technician at the Wood Anatomy and Identification Laboratory – ESALQ/USP, for his valuable technical support in sample preparation. We are also grateful to the technicians of the Isotopic Ecology Laboratory – CENA/USP Fabiana C. Fracassi Adorno, Gustavo Gobert Baldi, and Isadora S. Ottani for their assistance in the analytical processing of the samples. We further acknowledge Sarah Lima, secretary of the Isotopic Ecology Laboratory – CENA/USP, for her administrative support and assistance throughout the project. Isabela Maria Souza Silva (co-author) received a scholarship from the São Paulo Research Foundation (FAPESP, grant #2023/13568-7). Ana Carolina Barbosa was Supported by CNPq (grant no. PQ 313129/2022/3). Finally, we extend our appreciation to all collaborators at the Isotopic Ecology Laboratory who contributed in any way to the completion of this project.

Data availability: Data will be made available on request.

References

1. AMPUERO, A.; STRÍKIS, N. M.; APAÉSTEGUI, J.; VUILLE, M.; NOVELLO, V. F.; ESPINOZA, J. C.; CRUZ, F. W.; VONHOF, H.; MAYTA, V. C.; MARTINS, V. T. S.; CORDEIRO, R. C.; AZEVEDO, V.; SIFEDDINE, A. The forest effects on the isotopic composition of rainfall in the northwestern Amazon Basin. *Journal of Geophysical Research: Atmospheres*, v. 125, n. 4, 2020.
2. BARTON, K. **MuMIn**: Multi-Model Inference. Vienna: R Project, 2024.
3. BATAILLE, C. P.; CROWLEY, B. E.; WOOLLER, M. J.; BOWEN, G. J. Advances in global bioavailable strontium isoscapes. *Palaeogeography, Palaeoclimatology, Palaeoecology*, v. 555, 1 out. 2020.
4. BATISTA, A. C. G. (2025). Tracing Amazon timber provenance using stable oxygen isotope (PhD Thesis, Centro de Energia Nuclear na Agricultura, Universidade de São Paulo).
5. BATISTA, A. C. G.; SILVA, I. M. S.; ARAÚJO, M. G. S.; AMORIM, D. J.; NARDOTO, G. B.; COSTA, F. J. V.; HIGUCHI, N.; TOMAZELLO-FILHO, M.; BARBOSA, A. C.; COSTA, V. E.; PONTON, S.; MARTINELLI, L. A. Within- and between-site variability of $\delta^{18}\text{O}$, $\delta^{13}\text{C}$, and $\delta^{15}\text{N}$ in Amazonian tree rings: Climatic drivers and implications for geographic traceability. *Forest Ecology and Management*, v. 597, 1 dez. 2025.
6. BOWEN, G. J. Isoscapes: Spatial pattern in isotopic biogeochemistry. *Annual Review of Earth and Planetary Sciences*, v. 38, p. 161–187, 2010.
7. BOWEN, G. J.; WILKINSON, B. Spatial distribution of ^{18}O in meteoric precipitation. *Geology*, 2002. Disponível em: <<http://pubs.geoscienceworld.org/gsa/geology/article-pdf/30/4/315/3523533/i0091-7613-30-4-315.pdf>>.
8. BRUNELLO, A. T.; NARDOTO, G. B.; SANTOS, F. L. S.; SENA-SOUZA, J. P.; QUESADA, C. A. N.; LLOYD, J. J.; DOMINGUES, T. F. Soil $\delta^{15}\text{N}$ spatial distribution is primarily shaped by climatic patterns in the semiarid Caatinga, Northeast Brazil. *Science of the Total Environment*, v. 908, 15 jan. 2024.
9. BURNHAM, K. P.; ANDERSON, D. R.; HUYVAERT, K. P. AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology*, v. 65, n. 1, p. 23–35, 18 jan. 2011.
10. CAMARGO, P. B. de; BORGES, L. M.; GONTIJO, A. B. Isótopos forenses na proteção dos recursos florestais. In: NARDOTO, G. B.; MAYRINK, R. R.; BARBIERI, C. B.; FÁBIO JOSÉ VIANA. **Isótopos Forenses**. 1. ed. Campinas: Millenium Editora, 2022. cap. 6.
11. CAMERA, C.; BRUGGEMAN, A.; HADJINICOLAOU, P.; PASHIARDIS, S.; LANGE, M. A. Evaluation of interpolation techniques for the creation of gridded daily precipitation (1 × 1 km²); Cyprus, 1980–2010. *Journal of Geophysical Research: Atmospheres*, v. 119, n. 2, p. 693–712, 27 jan. 2014.

12. CHAPMAN, L.; THORNES, J. E.; BRADLEY V, A. Modelling of road surface temperature from a geographical parameter database. Part I: Statistical. *Meteorological Applications*, v. 8, n. 4, p. 409–419, 2001.
13. CINTRA, B. B. L.; GLOOR, M.; BOOM, A.; SCHÖNGART, J.; BAKER, J. C. A.; CRUZ, F. W.; CLERICI, S.; BRIENEN, R. J. W. Tree-ring oxygen isotopes record a decrease in Amazon dry season rainfall over the past 40 years. *Climate Dynamics*, v. 59, p. 1401–1414, 2022.
14. CONAMA – Conselho Nacional do Meio Ambiente. Resolução nº 406, de 2 de fevereiro de 2009. Diário Oficial da União, nº 26, 6 fev. 2009, p. 100. Brasília, DF, 2009.
15. COTRONEO, S.; KANG, M.; CLARK, I. D.; BATAILLE, C. P. Applying Machine Learning to investigate metal isotope variations at the watershed scale: A case study with lithium isotopes across the Yukon River Basin. *Science of the Total Environment*, v. 896, 20 out. 2023.
16. COURTIOL, A.; ROUSSET, F.; ROHWAEDE, M.-S.; KRAMER-SCHADT, S. **Isoscape Computation and Inference of Spatial Origins using Mixed Models**. 14 nov. 2023.
17. CRAIG, H.; GORDON, L. I.; CRAIG, H. K. Deuterium and oxygen-18 variations in the ocean and marine atmosphere. Spoleto Meeting on Nuclear Geology, *Stable Isotopes in Oceanography Studies and Paleo Temperature*, 1965.
18. DEMIR, Y.; ERSOY MIRICI, M. Effect of land use and topographic factors on soil organic carbon content and mapping of organic carbon distribution using regression kriging method. *Carpathian Journal of Earth and Environmental Sciences*, v. 15, n. 2, p. 311–322, 2020.
19. DUDLEY, B. D.; YANG, J.; SHANKAR, U.; GRAHAM, S. L. A method for predicting hydrogen and oxygen isotope distributions across a region's river network using reach-scale environmental attributes. *Hydrology and Earth System Sciences*, v. 26, n. 19, p. 4933–4951, 7 out. 2022.
20. GALEWSKY, J.; STEEN-LARSEN, H. C.; FIELD, R. D.; WORDEN, J.; RISI, C.; SCHNEIDER, M. Stable isotopes in atmospheric water vapor and applications to the hydrologic cycle. *Reviews of Geophysics*, 1 dez. 2016.
21. HARRISON, X. A.; DONALDSON, L.; CORREA-CANO, M. E.; EVANS, J.; FISHER, D. N.; GOODWIN, C. E. D.; ROBINSON, B. S.; HODGSON, D. J.; INGER, R. A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, v. 6, p. e4794, 23 maio 2018.
22. HIEMSTRA, P. H.; PEBESMA, E. J.; TWENHÖFEL, C. J. W.; HEUVELINK, G. B. M. **Package 'automap'- Automatic Interpolation Package Computers and Geosciences**. Vienna: R Project, 2009.
23. KAGAWA, A.; SANO, M.; NAKATSUKA, T.; IKEDA, T.; KUBO, S. An optimized method for stable isotope analysis of tree rings by extracting cellulose directly from cross-sectional laths. *Chemical Geology*, v. 393–394, p. 16–25, 30 jan. 2015.
24. KESKIN, H.; GRUNWALD, S. Regression kriging as a workhorse in the digital soil mapper's toolbox. *Geoderma*. Elsevier B.V., 15 set. 2018.
25. KUHN, M. **The caret Package**. Vienna: R Project, 2011.
26. KUMAR, A.; SANYAL, P.; AGRAWAL, S. Spatial distribution of $\delta^{18}\text{O}$ values of water in the Ganga river basin: Insight into the hydrological processes. *Journal of Hydrology*, v. 571, p. 225–234, 1 abr. 2019.
27. LAI, K.; TWINE, N.; O'BRIEN, A.; GUO, Y.; BAUER, D. Artificial Intelligence and Machine Learning in Bioinformatics. In: **Encyclopedia of Bioinformatics and Computational Biology**. Amsterdam: Elsevier, 2019. p. 272-286.
28. LOADER 'I, N. J.; ROBERTSON, I.; BARKER, A. C.; SWITSUR, V. R.; WATERHOUSE, J. S. An improved technique for the batch processing of small whole wood samples to a-cellulose ISOTOPE GEOSCIENCE. *Chemical Geology*, v 136,313-317,1997.
29. MARTINELLI, L. A.; BATAILLE, C. P.; BATISTA, A. C.; SOUZA-SILVA, I. M.; ARAÚJO, M. G.; ABDALLA FILHO, A. L.; BRUNELLO, A.; TOMMASIELLO FILHO, M.; HIGUCHI, N.; BARBOSA, A. C.; COSTA, F.; NARDOTO, G. B. Bioavailable strontium isoscape for the Amazon region using tree wood. *Forest Ecology and Management*, v. 594, 2025.
30. MARTINO, J. C.; TRUEMAN, C. N.; MAZUMDER, D.; CRAWFORD, J.; DOUBLEDAY, Z. A. The universal imprint of oxygen isotopes can track the origins of seafood. *Fish and Fisheries*, v. 23, n. 6, p. 1455–1468, 1 nov. 2022.

31. MASIOL, M.; ZANNONI, D.; STENNI, B.; DREOSSI, G.; ZINI, L.; CALLIGARIS, C.; KARLICEK, D.; MICHELINI, M.; FLORA, O.; CUCCHI, F.; TREU, F. Spatial distribution and interannual trends of $\delta^{18}\text{O}$, $\delta^2\text{H}$, and deuterium excess in precipitation across North-Eastern Italy. *Journal of Hydrology*, v. 598, 1 jul. 2021.
32. MATOS, M. P. V.; JACKSON, G. P. Isotope ratio mass spectrometry in forensic science applications. *Forensic Chemistry Elsevier B.V.*, 1 maio 2019.
33. MISHRA, R.; KAUR, E.; VITA, D.; AVINASH, P.; SINGH, B.; SINGH, A. P.; SINGH, J. Hydrogen isotope: An ideal tracer for forensic examination. *International Journal of Health Sciences*, 2022.
34. NOWAKOWSKI, S. **Hierarchical Clustering of Univariate**. 1d Vienna: Data CRAN, 2023.
35. PAUL, D.; SKRZYPEK, G.; FÓRIZ, I. Normalization of measured stable isotopic compositions to isotope reference scales - A review. *Rapid Communications in Mass Spectrometry*. John Wiley and Sons Ltd, 2007.
36. RODEN, J. S.; LIN, G.; EHLERINGER, J. R. A mechanistic model for interpretation of hydrogen and oxygen isotope ratios in tree-ring cellulose. *Geochimica et Cosmochimica Acta*, v. 64, n. 1, p. 21–35, 2000.
37. SALATI, E.; DALL'OLIO, A.; MATSUI, E.; GAT, J. R. Recycling of water in the Amazon Basin: An isotopic study. *Water Resources Research*, v. 15, n. 5, p. 1250–1258, 9 out. 1979
38. SENA-SOUZA, J. P.; HOULTON, B. Z.; MARTINELLI, L. A.; BIELEFELD NARDOTO, G. Reconstructing continental-scale variation in soil $\delta^{15}\text{N}$: a machine learning approach in South America. *Ecosphere*, v. 11, n. 8, 1 ago. 2020.
39. SOUZA-SILVA, I. M.; MARTINELLI, L. A.; HOLMES, B.; BATISTA, A. C. G.; ARAÚJO, M. G. S.; GARÇÃO, A. L.; PONTON, S.; GROENENDIJK, P.; LOCOSSELLI, G. M.; ORTEGA-RODRIGUEZ, D. R.; AMORIM, D. J.; COSTA, F. J. V.; NARDOTO, G. B.; BRUNELLO, A. T.; COSTA, V. E.; ASSIS-PEREIRA, G.; TOMAZELLO-FILHO, M.; HIGUCHI, N.; BARBOSA, A. C.; SENA-SOUZA, J. P.; BATAILLE, C. P. Machine-learning models of $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ isoscapes in Amazonian wood. *Biogeosciences*, 23, 881–904, 2026.
40. ST. JOHN GLEW, K.; GRAHAM, L. J.; MCGILL, R. A. R.; TRUEMAN, C. N. Spatial models of carbon, nitrogen and sulphur stable isotope distributions (isoscapes) across a shelf sea: An INLA approach. *Methods in Ecology and Evolution*, v. 10, n. 4, p. 518–531, 1 abr. 2019.
41. WALSH, E. S.; KREAKIE, B. J.; CANTWELL, M. G.; NACCI, D. A Random Forest approach to predict the spatial distribution of sediment pollution in an estuarine system. *Plos ONE*, v. 12, n. 7, e0179473, 24 jul. 2017.
42. WICKHAM, H.; CHANG, W.; HENRY, L.; TAKAHASHI, K.; WILKE, C.; WOO, K.; YUTANI, H.; DUNNINGTON, D.; BRAND, T. van den. **Create Elegant Data Visualizations Using the Grammar of Graphics**. New York, NY. Springer New York, 2024.
43. ZHAO, S.; HU, H.; TIAN, F.; TIE, Q.; WANG, L.; LIU, Y.; SHI, C. Divergence of stable isotopes in tap water across China. *Scientific Reports*, v. 7, 2 mar. 2017.

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.