# SENTINEL-1 TIME-SERIES ANALYSIS TO IDENTIFY PRE-COLUMBIAN MODIFIED ENVIRONMENTS

*Vinícius Peripato[1,2], Wesley A. Campanharo[1,2], Juliana M. F. S. Diniz[2], José Iriarte[3], and Luiz E. O. C. Aragão[1,2]*

[1]Tropical Ecosystems and Environmental Sciences Laboratory - TREES, Av. dos Astronautas 1758, Jardim da Granja, São José dos Campos, São Paulo, Brazil.
[2] Programa de Pós-Graduação em Sensoriamento Remoto - PGSER, Divisão de Observação da Terra e Geoinformática - DIOTG, Instituto Nacional de Pesquisas Espaciais - INPE.
[3]Department of Archaeology, College of Humanities, University of Exeter, North Park Road, Exeter, EX4 4RJ, UK.

## ABSTRACT

There has been increasing amount of evidences showing large pre-Columbian settlements and extensive domesticated landscapes across Amazonia. However, one of the challenges to understand the extent to which pre-Columbian societies transformed Amazonian landscapes is the fact that many settlements are covered by dense vegetation. Thus, the analysis of the forest structure though time can point out locations witch has better conditions to develop a more stabilized environment, caused by pre-Columbian management and cultivation practices. The identification of pre-Columbian sites within forested areas can enlighten uncertainties about the resistance and resilience of Amazonia ecosystems. Here, we compare SAR's vegetation indexes anomalies within pre-Columbian modified environments to other forested areas in their surroundings using a time-series (TS) from Sentinel-1A RADAR data (C-band). The results show that non-pre-Columbian areas have more anomaly events, and up to two time's higher recurrence. This approach can enable the search for new pre-Columbian archaeological sites hidden under the Amazonian forest canopy.

***Key words* —** SAR System, time-series, vegetation indexes, archeology, domesticated.

## 1. INTRODUCTION

During the pre-Columbian era, Amazonia was once home to dense and complex societies [1]. These societies had the ability to modify the terrain, soil and vegetation, domesticating the landscape of the Amazonia [2]–[5]. Several effects of past forest domestication can still be found hidden within the forest, such as earthworks and Amazonian Dark Earth are traces that persisted through time, and influence till today the Amazonian grounds [6], [7]. In addition, the lack of monumental architectural or metallurgical refinements – characteristic of Amazonian pre-Columbian societies [8] – highlights the need to explore the forest structure to quantify the scale and intensity of the pre-Columbian landscape domestication.

In this sense, Synthetic Aperture Radar (SAR) systems such as a Sentinel-1, which operates on microwaves (C-band, 5.4 cm) provides information of forest structure due the volumetric scattering. Moreover, an almost insensitive to cloud sensor can provide continuous information throughout a time-series (TS) without any pos-processing, filtering or interpolation technique.

The discovery of widespread human-modified forests throughout Amazonia is critical to develop a solid understanding of the interactions between human societies, and the Amazonian forest [9]. Indeed, rather than developing into a pristine climax ecosystem, if the widespread pre-Columbian occupation of the forest were to be proven, this ancient social-ecological system would age into an old secondary forest [10]. Moreover, this investigation can indicate if the Amazonian forest have long-term responses to climate change [11], improving greenhouse gas emissions and global climate change models.

The main goal of this study is to evaluate SAR's vegetation indexes anomalies within the Sentinel-1 time series to identify pre-Columbian modified environments.

## 2. MATERIAL AND METHODS

The study area comprises on locations with records of pre-Columbian modified environments located inside Conservation Units and Indigenous Lands, here denominated Protected Areas (PA). We considered two different types of pre-Columbian modified environments: (1) earthworks, which are man-made earth-building structures such as geoglyphs, fortified settlements, trenches and wells; and (2) Amazonian Dark Earth (ADE), which are anthropogenic soils with increased fertility developed during pre-Columbian management. The two types of presence data and the PA distribution throughout Amazonia can be seen in Figure 1. Pre-Columbian earthworks and ADE can be found in 28 different PA and 109 different PA, respectively, in the Amazon. Due to the lower likelihood of interference from anthropogenic disturbances (deforestation and fire), locations within PA were selected for the analysis of SAR's vegetation indexes anomalies.

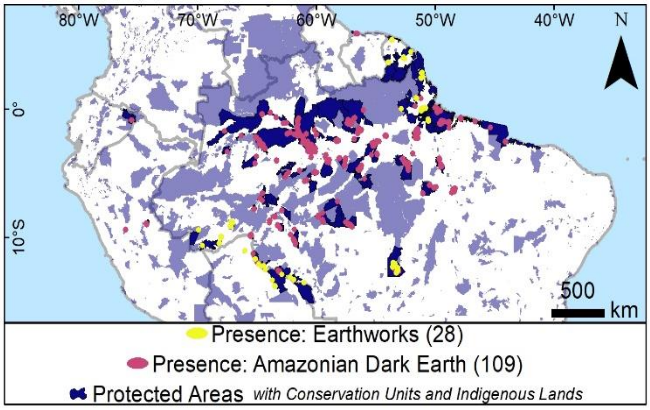


**Figure 1. Locations of Amazonian pre-Columbian modified environments inside Protected Areas.**

Pre-Columbian settlement records are composed of point data, compiled from four datasets provided by: (1) Amazonian Archaeological Sites Network (AmazonArch); (2) Brazilian National System of Archaeological Sites (CNSA); (3) Pre-Columbian Amazon Scale Transformations (PAST) project; and (4) Archaeological Research Center (NuPArq) Institute of Scientific and Technological Research of the State of Amapá (IPEA). We performed a 500 m buffer from each record centroid, and merged the same type of data – earthwork and ADE separately – located inside the same PA.

### 2.1 Sentinel Time-Series (TS) data

The collection of Sentinel-1A data on Google Earth Engine (GEE) platform provides dual-band polarization data in C-band (VV and VH). The data processed in Sentinel-1 Toolbox (S1TBX), developed by the European Space Agency (ESA) generates Ground Range Detected (GRD) scenes in to backscatter coefficient (σ°) with raw power values at 10 m nominal resolution. In order to standardize the Time Series (TS) composite, only Interferometric Wide (IW) swath mode data from a descending orbit pass were collected. Monthly average scenes were used to compare data acquired at different dates within pre-Columbian modified environments and the entire conservation unit. Moreover, due to this aggregation, monthly and within polygons, no speckle filter were apply. We collected all S1 monthly records from October 2014 until august 2022.

The raw power values of the backscatter coefficient (σ°) was used to calculate four index: (1) Cross-polarization Ratio (CR), as CR = VV/VH which highlights the relative biomass and water content [12]; (2) Normalization Ratio (NL), as NL = (VV*VH)/(VV+ VH) which highlights target discretization [13]; (3) Radar Gap Index (RGI), as RGI = (VV-VH)/(VV+VH) which highlights the difference on bare soil and forest areas [12]; and (4) Radar Vegetation Index (RVI), as RVI = (4*VH)/(VV+VH) which highlights the tree cover and biomass [14].

### 2.2 Anomaly detection

Anomaly detection were carried out in the pre-Columbian modified environments and in its PA surroundings using '*anomalize*' library in R version 4.0.2 [15]. The TS data were decomposed using Seasonal-Trend Decomposition for time series (STL) based on LOcally weighted polynomial regrESSion (LOESS) [16]. The regression-based approach detect patterns underlying the data, by distinguishing a trend and a seasonal component from the remainder. This method allows for a choice of the settings of the decomposition by the user, based on the objective of the decomposition and the knowledge of the processes affecting the data, giving STL and intrinsic flexibility [17]. Based on monthly data records, frequency and trend parameters were settled to auto.

Outliers on the decomposed data were identified using Generlized Extreme Studentized Deviate Test (GESD), which progressively eliminates outliers using a Student's T-Test comparing the test statistic to a critical value. Each time an outlier is removed, the test statistic is updated [18]. Once test statistic drops below the critical value, all outliers are considered removed. We performed these tests using a 90 % confidence interval.

The selection of the most appropriate index to evaluate vegetation indexes anomalies to identify pre-Columbian modified environments was performed by the higher ratio of anomalies identified throughout the TS.

## 3. RESULTS

In general, forested areas that held pre-Columbian modified environments have less anomalies across all explored indexes (Table 1). The highest anomalies ratio between the pre-Columbian presence and the overall forested area is highlighted by the VV polarization. The CR and NL indexes also presented a higher ratio for earthworks analysis. Lastly, RGI and RVI index showed the same pattern, however, with a higher ratio for earthworks analysis. All the SAR indexes on earthworks anomalies areas pointed a higher ratio in comparison to ADE anomalies areas.

| Index | Accumulated Anomalies | | | | | |
|---|---|---|---|---|---|---|
| | Earth works | PA | Ratio | ADE | PA | Ratio |
| VV | 121 | 161 | 1,33 | 448 | 549 | 1,23 |
| VH | 102 | 127 | 1,25 | 392 | 445 | 1,14 |
| CR | 78 | 98 | 1,26 | 399 | 459 | 1,15 |
| NL | 107 | 134 | 1,25 | 411 | 470 | 1,14 |
| RGI | 79 | 94 | 1,19 | 386 | 454 | 1,18 |
| RVI | 79 | 94 | 1,19 | 386 | 454 | 1,18 |

**Table 1. Accumulated anomalies across earthworks and Amazonian Dark Earth locations compared with their respective Protected Areas.**

The time difference between two detected anomaly events is longer in forested areas where pre-Columbian modified environments were present. The wider width on lower values on time minimum, maximum, mean and median on Figure 2 suggests that anomalies events are more recurrent on areas without pre-Columbian practices. In addition, the mean time recurrence of anomalies events can be up to two times higher in PA when comparing RGI and/or RVI indexes on earthworks locations.

## 4. DISCUSSION

Microwaves emitted by SAR sensors penetrates down into the canopy layer and the penetration depth depends on the wavelength. Even a C-band SAR sensor is able to penetrate in a well-developed canopy [19]. Thus, S1 data has a huge potential to identify changes, including anomalies, in the canopy structure within forested areas. Indeed, areas with pre-Columbian modified environments as earthworks have more water/moisture availability and biomass due their accumulation on the trenches, and ADE has higher water retention capacity. Therefore, forested areas within these modified environments will present a more constant biomass and water content throughout a TS of SAR indexes when compared to non-pre-Columbian areas. Moreover, the non-pre-Columbian areas will present more anomaly events with a higher recurrence intervals.

In this sense, the index with higher ratio between anomalies detected at PA in comparison to pre-Columbian modified environments was the single polarization VV channel. The VV intensity may be more sensitive to measurements of the canopy gap fraction [20], indicating that the conditions within modified environments defy their development. On the other hand, because both regions are essentially completely covered in trees, the vegetation scattering, which is stronger in VH polarization. In general, all the SAR indexes on earthworks areas show a higher ratio then in ADE areas. Thus, the higher biomass density within moist trenches could points out some information that should be further explored. Moreover, the same behavior through RGI and RVI indexes should be investigated.

In the future, a TS analysis with a pixel-by-pixel approach can be developed to identify possible locations of yet undiscovered pre-Columbian environments. Considering their site area, SAR data can be resampled to 500 m to circumvent some of the limitations of GEE cloud computing. This analysis can also be limited to regions free of degradation and deforestation brought on by contemporary human activity – aiming only on continuous forested areas – by using annual forests/vegetation masks from Global Forest Change (GFC) or MapBiomas. In addition, other TS analysis tools should be explored, such as identifying significant anomalies along the TS and even analyzing optical indexes.

Despite the fact that the volumetric scattering from SAR data provides so much information about the forest structure,

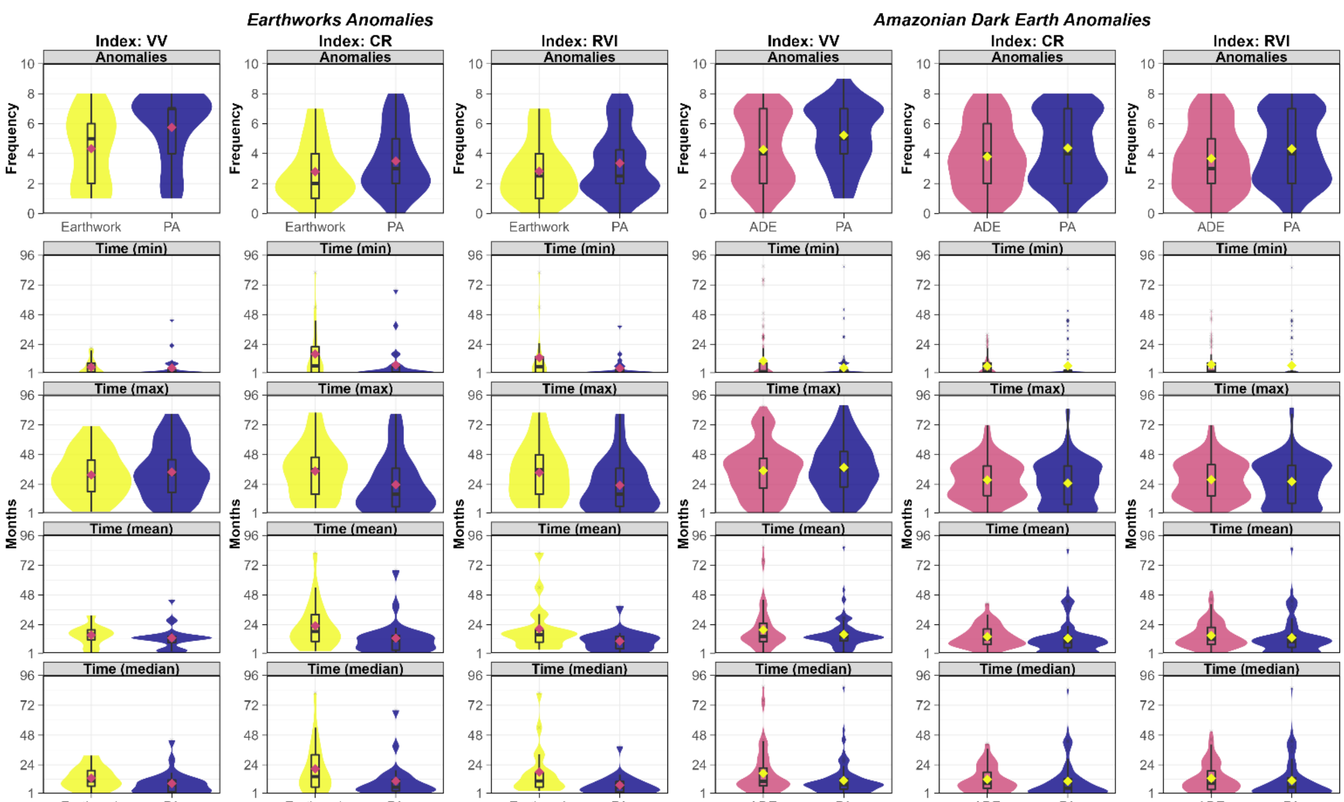


**Figure 2. Boxplots of SAR vegetation indexes anomalies frequency and time interval between anomalies events. The boxplot's shape, which is thicker for higher frequencies and thinner for lower frequencies, shows the frequency of the data. Mean values are represented by a diamond shape.**

data collection within S1 C-band only started in 2014. However, data acquisition (and availability) of optical sensors data have been continuously produced from 1972 through Landsat Collections, or 2000 by Moderate Resolution Imaging Spectroradiometer (MODIS). The data longevity from optical sensors can reveal forested dynamics, and should be explored to enhance the identification for possible locations of pre-Columbian modified environments.

## 5. CONCLUSIONS

This study arose from the desire – and the need – to investigate pre-Columbian Amazonia, observing how past human occupation can influence the forest today and in the coming centuries. Nowadays, 80% Amazonian forest is still on foot; therefore, we must explore remote sensing techniques to investigate this vast forested area spanning 6.7 million km$^2$. Pre-Columbian areas have less SAR indexes anomaly events, and longer recurrence intervals, than compared to its surrondings. This approach can enable the search for new pre-Columbian archaeological sites hidden under the Amazonian forest canopy. Moreover, with an increasing data availability and cloud computing to handle the processing of large image data sets, the need to develop researches that covers more ground is crucial.

**Acknowledgements**

We would like to thank members of the following projects/groups for providing data and support: Amazonian Archaeological Sites Network (AmazonArch); Archaeological Research Center (NuPArq) Institute of Scientific and Technological Research of the State of Amapá (IPEA); Pre-Columbian Amazon-Scale Transformations (PAST- European Research Council Consolidator Grant to J.I. - ERC_Cog 616179). This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) - Finance Code 001.

## 6. REFERENCES

[1] M. J. Heckenberger. Lost Cities of the Amazon, *Scientific American*, v. 301, pp. 64–71, Oct. 2009.

[2] C. de P. Moraes and E. G. Neves. Earthworks of the Amazon, in *Encyclopedia of Global Archaeology*, Springer International Publishing, Cham, 2019, pp. 1–13.

[3] C. Levis *et al.* Persistent effects of pre-Columbian plant domestication on Amazonian forest composition, *Science*, v. 355, pp. 925–931, Mar. 2017.

[4] C. Levis *et al.* How People Domesticated Amazonian Forests, *Frontiers in Ecology and Evolution*, v. 5, pp. 1–21, Jan. 2018.

[5] J. G. de Souza *et al.* Pre-Columbian earth-builders settled along the entire southern rim of the Amazon, *Nature Communications*, v. 9, p. 1125, Dec. 2018.

[6] H. Prümers *et al.* Lidar reveals pre-Hispanic low-density urbanism in the Bolivian Amazon, *Nature*, v. 606, pp. 325–328, Jun. 2022.

[7] J. Iriarte *et al.* Geometry by Design: Contribution of Lidar to the Understanding of Settlement Patterns of the Mound Villages in SW Amazonia, *Journal of Computer Applications in Archaeology*, v. 3, pp. 151–169, Apr. 2020.

[8] D. P. Schaan. *Sacred Geographies of Ancient Amazonia: Historical Ecology of Social Complexity*. Routledge, New York, 2012.

[9] C. N. H. McMichael *et al.* Ancient human disturbances may be skewing our understanding of Amazonian forests, *Proceedings of the National Academy of Sciences*, v. 114, pp. 522–527, Jan. 2017.

[10] M. J. Heckenberger. Amazonia 1492: Pristine Forest or Cultural Parkland?, *Science*, v. 301, pp. 1710–1714, Sep. 2003.

[11] R. J. W. Brienen *et al.* Long-term decline of the Amazon carbon sink, *Nature*, v. 519, pp. 344–348, Mar. 2015.

[12] A. F. Carneiro *et al.* Exploiting SENTINEL-1 SAR Time Series to Detect Grasslands in the Northern Brazilian Amazon, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, v. XLIII-B3-2, pp. 259–265, Aug. 2020.

[13] D. Lu *et al.* A Comparison of Multisensor Integration Methods for Land Cover Classification in the Brazilian Amazon, *GIScience & Remote Sensing*, v. 48, pp. 345–370, Jul. 2011.

[14] D. Mandal *et al.* Dual polarimetric radar vegetation index for crop growth monitoring using sentinel-1 SAR data, *Remote Sensing of Environment*, v. 247, p. 111954, Sep. 2020.

[15] R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria. Available online at https://www.R-project.org/.

[16] R. B. Cleveland *et al.* STL: A Seasonal-Trend Decomposition Procedure Based on Loess (with Discussion), *Journal of Official Statistics*, v. 6, pp. 3–73, 1990.

[17] E. Pavlidou *et al.* Finding a needle by removing the haystack: A spatio-temporal normalization method for geophysical data, *Computers & Geosciences*, v. 90, pp. 78–86, May 2016.

[18] A. R. Martel. The Detection of Outliers in Nondestructive Integrations with the Generalized Extreme Studentized Deviate Test, *Publications of the Astronomical Society of the Pacific*, v. 127, pp. 258–265, Mar. 2015.

[19] M. El Hajj *et al.* Penetration Analysis of SAR Signals in the C and L Bands for Wheat, Maize, and Grasslands, *Remote Sensing*, v. 11, p. 31, Dec. 2018.

[20] F. N. Numbisi and F. Van Coillie. Does Sentinel-1A Backscatter Capture the Spatial Variability in Canopy Gaps of Tropical Agroforests? A Proof-of-Concept in Cocoa Landscapes in Cameroon, *Remote Sensing*, v. 12, p. 4163, Dec. 2020.