

# Ancient human occupation alters modern forest traits, structure and biomass in neotropical forests.

Christopher E. Doughty<sup>\*1</sup>, Benjamin C. Wiebe<sup>1</sup>, Megan M. Seeley<sup>2</sup>, Jesus Aguirre-Gutierrez<sup>3</sup>, Jason Nesbitt<sup>4</sup>, Francisco Estrada-Belli<sup>5</sup>, Marcello A. Canuto<sup>5</sup>, Luke Auld-Thomas<sup>6</sup>

<sup>1</sup>School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, AZ, USA

<sup>2</sup>Center for Global Discovery and Conservation Science, Arizona State University, Hilo, HI 96706, USA

<sup>3</sup>Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford, UK

<sup>4</sup>Department of Anthropology, Tulane University, New Orleans, LA, USA

<sup>5</sup>Middle American Research Institute, Tulane University, New Orleans, LA, USA

<sup>6</sup>Department of Anthropology, Northern Arizona University, Flagstaff, AZ, USA

\*To whom correspondence should be addressed [chris.doughty@nau.edu](mailto:chris.doughty@nau.edu)

**Keywords** – Biomass, Amazon, Peten, ADE, geoglyphs, traits, GEDI

**Abstract** – There is considerable debate as to how extensively neotropical forests were shaped by ancient people. This uncertainty has implications for the global carbon cycle because forest plots may be disproportionately located in formerly occupied areas, potentially biasing estimates of tropical forest growth and carbon uptake. Ground measurements within the formerly occupied Maya city of Xmakabatun (400BCE to 800CE), now covered by forest, revealed significantly higher forest biomass, modified leaf traits, and distinct species composition on archaeological structures (i.e. ancient buildings) relative to control and former agricultural areas. We then used GEDI and trait-based remote sensing across the broader region to show that forest regions with greater quantities of known archaeological settlement also had significantly higher biomass and modified leaf traits. We next extend this approach to the Amazon basin, where ancient land-use histories are less certain. Using similar remote sensing techniques, we found that forests near Amazonian Dark Earth (terra preta) sites (N=503) were shorter but had slightly more foliage (~6%), whereas forests near geoglyphs (N=961) were taller but had much less bulk (20%) than the control forest plots. From these relationships, we estimate that ancient human occupation has increased modern Amazonian biomass by, on average,  $2.8 \pm 1.6\%$ , an amount equivalent to ~1% of global above-ground biomass. Together, these results demonstrate that ancient land use has left detectable legacies in the traits and structure of modern forests and suggest that remotely sensed biomass and/or leaf trait datasets may also provide a tool for regional-scale archaeological mapping.

**Plain language summary** – To better understand global climate change, we need accurate estimates of how much carbon is stored in tropical forests, which holds the world's largest aboveground carbon stock. However, the forest plot networks used to estimate these carbon stocks might be disproportionately located in regions formerly occupied and modified by people, potentially biasing our understanding of tropical forest carbon storage. Here, we show, using ground transects and remote sensing, that areas formerly occupied by the Maya hundreds of years ago in Guatemala still differ in forest biomass, tree species, and leaf traits. We then apply similar remote sensing techniques in the Amazon Basin, where ancient land-use histories are less well understood, and show that forests with modified soils (i.e., Amazon Dark Earth; ADE) or with mounds (i.e., geoglyphs) also differ in their traits and biomass. We further use these forest signatures to identify areas that may contain previously undocumented geoglyphs or ADE sites. Finally, we estimate that pre-colonial populations increased present-day Amazonian aboveground carbon stocks by approximately 3%, equivalent to about 1% of global aboveground carbon.

#### Main points

1. Ground and remote sensing data suggest that ancient Maya land use increased forest biomass and modified species and forest traits.
2. Remote sensing data suggest that ancient Amazonian land use increased forest biomass and modified forest traits.
3. Ancient human land use may have increased present-day Amazonian aboveground carbon stocks by ~3%, or 1% globally.

## Introduction -

Neotropical forests are often hypothesized to contribute to the “missing carbon sink,” or the global terrestrial region where much of humanity's CO<sub>2</sub> emissions have recently been sequestered. The magnitude and persistence of this sink have direct implications for global climate mitigation strategies and for identifying priority regions for conservation. Several lines of evidence support this hypothesis, including atmospheric inversion (Basso *et al.*, 2023) and Earth system modeling studies (although the sink in models is about half of that predicted by field results; Koch, Hubau and Lewis, 2021). The strongest evidence for the Amazon as the missing carbon sink has been large forest plot networks like RAINFOR (Hubau *et al.*, 2020). However, the Amazonian forest inventory plot locations might be spatially biased toward areas that were previously settled by human populations (McMichael *et al.*, 2017). For instance, forest inventory plots are located disproportionately near archaeological evidence, and these plots may still be affected by past disturbances, including modified species, biomass, and trait compositions, which could impact our understanding of how these forests are responding to global change (McMichael *et al.*, 2017). Therefore, forest network plots in the neotropics that feed into determining carbon stocks/global climate monitoring might be inaccurate because they are based in areas where humans have modified the landscapes. However, it is difficult under thick tree cover to know exactly where humans have modified the landscape.

Once, it was difficult to know what percentage of the neotropical landscape had been occupied and with what intensity, but modern advances in remote sensing (e.g., lidar) have greatly expanded our ability to locate former settlements, particularly in areas with dense forest cover. For example, high-resolution lidar has been used to identify ~100,000 ancient structures in northern Guatemala and Mexico (Canuto *et al.*, 2018, Estrada-Belli *et al.* 2025). Extrapolation of this settlement density to the entire 95,000 km<sup>2</sup> of the Central Maya Lowlands has led to an estimated human population of 9.5 to 16 million (Estrada-Belli *et al.* 2025). Similarly, lidar initially collected to measure forest biomass in Mexico has also helped identify large, previously unknown urban centers and revealed a landscape crowded with archaeological settlement remains (Auld-Thomas *et al.*, 2024, Schroder *et al.* 2020), contributing to our expanding map of ancient human settlements (Estrada-Belli *et al.*, 2025).

Identifying pre-colonial occupation in the Amazon Basin has been more challenging, largely because wooden and earthen constructions that were more typical in Amazonia have lower archaeological visibility than the masonry constructions of Mesoamerica or the Andes. Even so, lidar and other remote sensing techniques have started to reveal the extent of pre-Columbian settlement systems in different parts of the Amazon (Heckenberger *et al.* 2008; Iriarte *et al.* 2020; Pärsinnen *et al.* 2026; Prümers *et al.*, 2022; Rostain *et al.*, 2024). In Brazil, many recent independently collected lidar flights were aggregated and used to identify 24 previously undetected pre-Columbian earthworks beneath the forest canopy. Peripato *et al.* (2023) estimated that between 10,272 and 23,648 Amazonian geoglyph sites remain to be discovered, as compared with the 20,000 in SW Amazonia alone (Pärsinnen *et al.* 2026).

Another indication of ancient human settlement in Amazonia is the presence of Amazonian Dark Earth (ADE), or terra preta, which most believe were artificially enriched soils to improve soil

productivity probably created by pre-colonial populations (Lombardo *et al.* 2022; Schmidt *et al.*, 2025), although there are some dissenters to this view (Silva *et al.*, 2021). McMichael *et al.* (2014) predicted that ~154,063 km<sup>2</sup>, or 3.2%, of forests are likely to be ADE. While McMichael *et al.* (2014) used climatic, topographic, hydrologic, and geologic data layers to develop these estimates, there is potential for remote sensing to locate ADE directly. High-spatial resolution (6m<sup>2</sup>) SPOT VNIR imagery was used to identify ADE over a 256 km<sup>2</sup> area near Santarem, Brazil, and estimated that ADEs would cover 3% of the upland Amazon (Söderström *et al.*, 2016). Another study using MODIS (250m<sup>2</sup> spatial resolution) found that ADE sites were characterized by shorter trees with lower biomass, normalized difference vegetation index, and canopy water content than nearby (<50 km) non-ADE forested (Palace *et al.*, 2017).

Archaeological sites discovered in a region of the Amazon not previously thought to be inhabited could increase Amazonian population estimates and thus our estimate of how much people impact the modern forest. The error brackets around pre-Columbian population estimates for Amazonia often vary by an order of magnitude (Denevan 2014; Hirst *et al.*, 2025). It was estimated that at least 5-6 million, and perhaps as many as 10 million people inhabited Amazonia before European colonization in the 1500s (Denevan, 2014). Though recent archaeology suggest that the population could have been significantly greater (Pärsinnen *et al.* 2026). However, new technologies are needed because large swaths of Amazonia have limited archaeological fieldwork.

Throughout the Neotropics, there are other ways for researchers to tell that people once occupied an area other than directly finding archaeological sites. For instance, there is increasing evidence that trees growing on formerly occupied areas have different traits, structure, and biomass. The evidence for enduring legacies of ancient peoples in modern forests continues to expand (Campbell *et al.*, 2006). For instance, Campbell *et al.* (2006) showed how ancient Maya people modified forest structure in settled areas. In expanding our understanding of ancient forest occupancy, the recent lidar campaigns support the assertion that contemporary forests of the Maya region are anthropogenic (Brokaw *et al.*, 2025, Eshleman *et al.* 2026). At geoglyph sites in Amazonia, species composition was often altered (Watling *et al.* 2017). Further, ADEs have been associated with altered species composition, notably having more palm species (Junqueira, Shepard and Clement, 2010; McMichael *et al.*, 2015). Overall, ADE plots tend to have higher soil nutrients (e.g., phosphorus) and net primary production but lower total biomass, possibly due to a community-level shift in survival strategy toward acquisitive trait syndromes (Doughty *et al.*, 2014). These modifications may affect modern carbon dynamics as evidenced by CO<sub>2</sub> enrichment experiments, which show initial increases in photosynthesis and plant growth, but plant growth quickly decreases as the forests become nutrient-limited (Norby and Zak, 2011). New remote sensing datasets may allow us to address these important questions.

New satellite data enables us to test our knowledge of the extent of former human settlements in the Neotropics. For instance, remote sensing can now accurately predict tropical forest canopy traits, including chemical composition, and forest structure (Asner *et al.*, 2014). The Global Ecosystem Dynamics Investigation (GEDI) has taken over 7 billion lidar measurements that reveal forest structure, biomass, and height (Dubayah *et al.*, 2023; Burns, Hakkenberg and Goetz,

2024). Analysis of tropical forest structure using GEDI has recently challenged long-held beliefs on how these forests are structured (Doughty *et al.*, 2023). More recently, remotely-sensed traits have been produced across all tropical forests using spectral sensors, allowing for continental-scale trait comparisons (Aguirre-Gutiérrez *et al.*, 2025). Here, we propose that trait-based remote sensing and lidar can detect changes in forest traits and structure linked to ADEs and geoglyphs through their effects on soil properties and species composition. We further propose that these vegetation impacts may reveal previously undetected ADE and geoglyph sites.

We investigated the legacies of ancient peoples on modern forest structure, biomass, and traits in both the central Maya lowlands and the Amazon. We conducted field transects in Guatemala, where land use history is known to the meter scale, and then upscaled our results to the broader Guatemalan landscape using satellite datasets. Next, we applied these satellite analyses to the Amazon, where the land history is less well known. This allowed us to ask the following questions:

**Q1 – Do archeological sites covered in forests show consistent impacts across the neotropics on forest structure and plant traits?**

**Q2 – Is tropical forest biomass higher in areas with evidence for ancient human settlement?**

## Methods

### Field campaign

In June 2025, in the start of the dry season, we measured leaf traits (leaf mass area (LMA) and elements (like phosphorus (P) and Calcium (Ca)), biomass, species composition and leaf spectra near Xmakabatun (89.2359498°W 17.5107969°N), in northeastern Guatemala, a small ancient city with a peak structure density of 264 structures per km<sup>2</sup> in its urban core and an average density of 40-60 structures per km<sup>2</sup> in its periurban zones extending over approximately 30 km<sup>2</sup> of hilly terrain. The city and its surroundings were continuously occupied from the Late Preclassic (400 BCE-250 CE) to Late Classic period (650 to 800 CE) and are now completely covered in forest (Figure 1 top). The site was mapped with high-resolution airborne lidar as part of the Pacunam Lidar Initiative. We specifically used the 309 km<sup>2</sup> Holmul survey block (Canuto *et al.*, 2018), which allowed us to identify individual ancient buildings and agricultural terraces with high confidence and precise spatial detail (Figure 1 top). Control areas that lacked evidence of archaeological settlement or agricultural terraces were selected based on field and lidar survey data (Estrada-Belli, 2017; Canuto *et al.*, 2018).

We made 29, 30m transects: 9 on archaeological buildings, 11 in terraced agriculture areas, and 9 in control regions. All transects were within 5 km of each other with similar underlying (not counting structures) aspect, slope and soil type. On each transect, at 0, 10, 20 and 30m, we collected soil from the O-horizon after scraping away the litter layer and ~5 shade leaves from the biggest tree by trunk diameter within a 2m radius with a 3m leaf clip pole. We measured in situ chemistry with a Bruker Tracer five XRF pointed at dried samples for one minute, using a soil calibration for soils and a leaf calibration for leaves. Samples were dried in a small desiccator oven for at least 24 hours. Leaf mass area (LMA) was measured by cutting 5, ~1.6 cm diameter leaf disks per sample that were dried and then weighed. A local expert (Don Israel) identified all individuals of five species within 5m of the transect known to have one or more uses for past or modern people living in the area: e.g. Copal (*Protium copal*), Corozo (*Attalea cohune*), Ramón (*Brosimum alicastrum*), Chico zapote (*Manilkara zapota*), Guano/Botán (*Sabal mauritiiiformis*). We estimated biomass along the transect with allometric equations from Chave *et al* (2005) using measurements of diameter at breast height (DBH) for all trees >10 cm DBH within the 4m-wide transect and approximate tree height determined using a laser range finder (Chave *et al.*, 2005).

**Remote sensing and Archeology databases** – In the Maya region of Guatemala, we compared three regions with low past human population densities (Figure 1d; Canuto *et al.*, 2018): E1 (with overall average density of 2 structures per km<sup>2</sup>), E2 (2/km<sup>2</sup>) and E4 (4/km<sup>2</sup>) which taken all together average 3 structures per km<sup>2</sup> and compare them to seven regions with high past human population densities Holmul (23/km<sup>2</sup>), Naachtun (89/km<sup>2</sup>), Tintal (43/km<sup>2</sup>), Peru (45/km<sup>2</sup>), Xultun (44/km<sup>2</sup>), Zotz-Tikal-Uaxactun (52/km<sup>2</sup>) and Corona (8/km<sup>2</sup>) which average 51 structures per km<sup>2</sup> (Figure 1; Canuto *et al.* 2018: Table 4). In the Amazon basin, we identified formerly occupied areas with ADE coordinates (N=503) from WinklerPrins and Aldrich (2010) and geoglyph coordinates (N=961) from Peripato *et al.* (2023) (Figure 1g and e).

**GEDI data** - We used the vertical forest structure (L2A and L2B, Version 2) and biomass (L4A) products from the GEDI instrument (Dubayah *et al.*, 2020) based on the ISS between 2019.04.18 and 2021.02.17 for Amazonia and the Maya region. The L2A product has been ground validated in tropical forests (Marselis *et al.*, 2018, 2020; Liu, Cheng and Chen, 2021; Cobb *et al.*, 2023). The specific L2A product used was Plant Area Volume Density (PAVD) profile to get vertical vegetation structure. PAVD, which incorporates both leaf and wood, is separated into 5-meter vertical bins of 0-5, 5-10, 10-15, 15-20, 20-25, 25-30 m heights. Both biomass and PAVD was previously quality checked into gridded rasters where the 25m diameter GEDI footprints were aggregated into 300 x 300 m pixels (Burns, Hakkenberg and Goetz, 2024). In the Maya regions, we compared all available GEDI data for the three survey blocks with low population density (E1, E2, E4) to seven blocks with high population densities (Holmul, Naachtun, Tintal, Peru, Xultun, Zotz-Tikal-Uaxactun and Corona) with areas that averaged 150-450 km<sup>2</sup> (Figure 1d).

For Amazonia, using the centroid coordinates from each of the ADE sites (N=503) and geoglyphs (N=961) plots, we found the 300 x 300 m averaged GEDI pixel that encompassed the plot. We then searched a wider area by incrementally averaging a gradually increasing area of 3 by 3, 7 by 7, 11 by 11, and 21 by 21 pixels. We compared the archeology coordinates to 3,000 coordinates randomly selected to be within a 750 km radius of the archeological coordinates (we also tested coordinates within areas of 375, 263 and 112 km and show the small impact of these changes in Figure S3). Not all coordinates were primary forest, so we screened biomass with three biomass thresholds (100, 150, or 200 Mg ha biomass) and forest height thresholds of 10, 15, or 20m to ensure forest cover. This screening typically removed between 50-60% of the data. Anything below the threshold is given a NaN. We then averaged forest biomass and height in the surrounding areas. We estimated a “small” region of 3x3 pixels around the centroid, resulting in 900x900m<sup>2</sup> pixels, and a “large” region of 21x21 pixels to obtain 6300x 6300m<sup>2</sup> (size modified in a sensitivity study) in order to understand the impact of distance from the centroid archeological site or control coordinates. For structure, we used similar methodology with PAVD and used the same forest height thresholds as above to screen out non-forest regions for both the Maya and Amazonian regions.

**Trait data** – We used forest trait data from Aguirre-Gutierrez *et al.* (2025), which calculated 10 by 10m pixels by combining field-collected data from more than 1,800 vegetation plots and tree traits with Sentinel-2 satellite remote-sensing, terrain, climate and soil data. These datasets predict variation across 13 morphological, structural and chemical functional traits of trees. For this project, we focused on the traits LMA, P and wood density (WD) (Aguirre-Gutiérrez *et al.*, 2025). Specifically, we gathered this trait data for the 10 survey regions (red densely settled and blue sparsely settled) shown in Fig 1D and the geoglyph (Fig 1F), the ADE (Fig 1G), and the 3000 control points. We did not use a biomass or forest height threshold to screen out non-forest regions due to the spatial mismatch of the trait and biomass data. Since E1 has little remaining forest cover, for the trait analysis, we remove the entire block of E1 in Figure S4 and show very similar results to Figure 3.

**Predict archeology and biomass** – To assess whether remote sensed and environmental data can predict the presence of pre-colonial settlement (geoglyphs, ADE or both) in Amazonia we use

simultaneous autoregressive spatial error models (SARerr) in R (Dormann *et al.*, 2007) using the R library ‘spdep’ (Bivand, Hauke, & Kossowski, 2013). SAR<sup>err</sup> models were chosen to account for general spatial autocorrelation and we tested different neighborhood distances from 10 km to 300 km. The corresponding correlogram showed reduced spatial autocorrelation, as measured with Akaike Information Criterion (AIC), at 40 km. We then ranked the best a SARerr models with a 40km neighborhood by AIC and used the VIF command in the CAR package (Fox J and S, 2019) to test for multicollinearity between variables. Using only coordinates where there was at least one GEDI footprint within the 300 by 300m area resulted in a sample size of (N=85 TP, Control=1020) for Terra Preta and (GG=52, Control = 583) for geoglyphs. Control points were chosen in the same manner as described above.

In addition to forest trait data from Aguirre-Gutierrez et al. (2025) and GEDI biomass and structure we used several environmental datasets. We used total cation exchange capacity (CEC) for the top 0-5 cm soil layer (units = mmol(c) kg<sup>-1</sup>) from soil grids (Batjes, Ribeiro and van Oostrum 2020) as a proxy for soil fertility and because it is a primary driver of variation in tropical forest wood production and biomass (Quesada et al 2012). We then averaged TerraClimate (Abatzoglou *et al.*, 2018) data between 2000 and 2018 to obtain Vapor Pressure Deficit (VPD in kPa), Mean Monthly Precipitation (MMP) (mm/month) and maximum temperature (Tmax (°C)). Thus, all data layers were: climate (MMP, Tmax, VPD), soils (CEC), traits (P, LMA, WD), and structure (biomass, height, PAVD 0-5).

We first used the variables described above to predict the binary presence of either geoglyphs, ADE or both. In a second model, we used the variables described above plus archeology (geoglyphs, ADE or both) as a binary (1=present and 0=not present) to predict GEDI biomass. Then, based on the variables from the best models, we first predicted archaeology and then biomass with archaeology as a possible predictor.

**Statistics** - For our field data in Guatemala, we use a two-sided t-test in matlab to test for significant differences between groups. If multiple groups were compared, we used matlab’s anova1 and multi compare with Tukey’s honestly significant difference procedure.

## Results

**Field results** - Biomass and leaf traits at Xmakabatun, a forested archeological zone in Guatemala, were significantly different among structures, agricultural terraces, and control sites. We observed higher ( $P < 0.01$ ) forest biomass on structures than on agricultural terraces or control sites ( $402 \pm 264$  Mg/ha  $^{***}$ ,  $N=9$ , vs  $163 \pm 111$ ,  $N=11$  and  $122 \pm 51$ ,  $N=9$ ) (Figure 2 and Table 1). Forest biomass in terraced former agricultural areas was higher than in control areas, though not significantly ( $P=0.12$ ). At both the structures and terraced agricultural areas, species composition differed significantly ( $P < 0.001$ ) from control areas, as these sites had more individuals of the five focal tree species (structures =  $8.9 \text{ ind} \pm 4.3$ ; terraced =  $7.6 \pm 4.6$ ; control =  $2.2 \pm 1.4$ ).

Soil and leaf chemistry similarly differed among the ancient land use categories. Soil Ca and P were significantly higher on structures ( $10.42 \text{ ppt} \pm 10$ ,  $N=21$  and  $0.081 \text{ ppt} \pm 0.08$ ,  $N=21$ ), than at control sites ( $2.62 \pm 2.5$ ,  $N=10$ , and  $0.067 \pm 0.08$ ,  $N=10$ , respectively). We found significantly more leaf Ca on structures ( $3.90 \text{ ppt} \pm 2.6$ ,  $N=44$ ) than at control sites ( $2.37 \pm 2.5$ ,  $N=39$ ) and non-significantly more P ( $0.081 \pm 0.08$ ,  $N=21$  vs  $0.067 \pm 0.08$ ,  $N=10$ ). We observed no significant differences in soil and leaf P and Ca between terraced agricultural and control sites. LMA at structures ( $81.80 \text{ g m}^{-2} \pm 28$ ,  $N=43$ ,  $P=0.06$ ), agricultural terraces ( $81.07 \pm 23$ ,  $N=41$ ,  $P=0.11$ ), and the control sites ( $88.87 \pm 23$ ,  $N=36$ ) did not differ significantly.

### Remote sensing – Petén

We next upscaled our field analyses from Xmakabatun and compared them to remotely sensed biomass and trait estimates across Petén (Northern Guatemala) for agricultural (the seven blocks with red numbers in Fig 1D) or the control areas (the three blocks with blue numbers in Fig 1D). Specifically, we compared GEDI-estimated biomass and Sentinel-2-estimated leaf traits across sites that were either more ( $8\text{-}89 \text{ structures/km}^2$ ) or less ( $2\text{-}3 \text{ structures/km}^2$ ) heavily structured. (Figure 1). We observed a reduction in LMA ( $131 \pm 7$  vs  $135 \pm 2$ ,  $P < 0.0001$ , – Cohen's  $D=0.74$ ),  $P$  ( $0.118 \pm 0.004$  vs  $0.120 \pm 0.004$ ,  $P < 0.0001$  – Cohen's  $D=0.54$ ), and WD ( $0.577 \pm 0.008$  vs  $0.576 \pm 0.006$ ,  $P < 0.001$  – Cohen's  $D=-0.11$ ) at more densely settled archeological sites compared to the control sites (Cohen's  $D$  indicating moderate to high effect size for LMA and %  $P$ ; Figure 3 and Figure S1). We compared the satellite trait values to field-measured trait data from agricultural terraces, which occupy more of the land area in our ground study site of Xmakabatun than do structures. Our field data comparing agricultural to control regions were in the same direction and had similar magnitude for biomass, LMA, and % $P$  as the RS data (Table 1). GEDI-estimated biomass was significantly ( $P < 0.05$ ) higher ( $\sim 6 \text{ Mg ha}^{-1}$ ) at the more heavily settled survey blocks ( $148 \pm 89$  for the high population versus  $142 \pm 88 \text{ Mg ha}^{-1}$  for the low population), broadly aligning with field results which showed higher biomass in both where structures were ( $402 \pm 264 \text{ Mg ha}^{-1}$ ) and in agriculture ( $163 \pm 111 \text{ Mg ha}^{-1}$ ) as compared to control areas ( $122 \pm 51 \text{ Mg ha}^{-1}$ ).

### Amazonia

Building on this work in Guatemala, we transferred our remote sensing approach to regions of the Amazon, where our confidence regarding the precise distribution of pre-Columbian

settlements was lower. Using the ADE (N=503) and geoglyph (N=961) archeological datasets, we compared biomass and leaf traits in settled regions to 3000 randomly chosen control sites (noting that there is a small chance that some of these random points were undiscovered settlements). Overall, geoglyph plots tended to have increased biomass while ADE sites had reduced biomass compared to controls, but this result varied by sampling strategy. At ADE sites, biomass estimated from GEDI was significantly lower (by 17 Mg ha<sup>-1</sup>; Figure 4) compared to control sites for both small (20 Mg ha<sup>-1</sup> - 900 by 900m) and large areas (14 Mg ha<sup>-1</sup> - 6300 by 6300m) regardless of biomass threshold (100 Mg ha<sup>-1</sup> biomass threshold (th) - 27 Mg ha<sup>-1</sup>, 150 th - 2 Mg ha<sup>-1</sup> and 200 th - 10 Mg ha<sup>-1</sup>). We further observed a linear increase in biomass positively associated with distance from the geoglyph sites (Figure S2). In other words, when comparing geoglyph to control sites, distance matters, and averaging over a larger region (21 by 21 pixels) showed a larger increase in biomass (10-15 Mg ha<sup>-1</sup>) as compared to averaging over a smaller region (3 by 3 pixels) with a ~3 Mg ha<sup>-1</sup> increase. This indicates spatial autocorrelation in our data, which we correct for in the next section.

In ADE forests, when averaging all data, biomass was higher (5 Mg ha<sup>-1</sup>) in geoglyph sites compared to control sites. However, this result diverged when we averaged over large areas (12 Mg ha<sup>-1</sup>) versus small areas (-1.5 Mg ha<sup>-1</sup>) and across several forest size thresholds (100 th -2 Mg ha<sup>-1</sup>, 150 th 11 Mg ha<sup>-1</sup> and 200th 7 Mg ha<sup>-1</sup>). Thus, these results again suggest there is spatial autocorrelation in the results presented above. This is why we account for spatial autocorrelation using the SARerr model. For comparison, in other studies, field methods often show lower biomass in ADE. For example, one study showed a total biomass of 110.9 Mg C ha<sup>-1</sup> at the ADE site versus 179.4 Mg C ha<sup>-1</sup> at a control site (Doughty *et al.*, 2014).

We found that forest height quantified with GEDI was slightly reduced (average -0.19m Figure 4) in ADE sites compared to control sites, with differing results depending on distance (small areas -.36m) and large areas 0.76m) surrounding the coordinates and under several forest size thresholds (100th -0.22m, 150th 0.37m and 200th 0.44m). However, forest height was significantly higher (1.8m) in geoglyph sites compared to control sites for small (1.2m) and large areas (2.4 m) surrounding the coordinates and under several forest size thresholds (100th 0.86m, 150th 2.3m and 200th 2.2m). For comparison, field methods often show lower forest height in ADE sites, with mean tall tree (>40 cm DBH) height of 25.3 ± 3.2 m for the ADE plot and 27.2 ± 5.5 m for the control plot (Doughty *et al.*, 2014), with the change likely due to a combination of different species, different soil nutrients or different land use histories (former farms).

Forest structure likewise changed significantly at most height bins between the archeological sites and control sites. On average, the geoglyph plots lost PAVD in the mid and bottom height bins (Figure 5). Thus, forests at geoglyph sites were taller but had lower PAVD (20%) than the control plots. On average, the ADE plots gained PAVD except at the very top height bins. In contrast, ADE sites were slightly shorter but with more PAVD (~6%). These differences were significant (P<0.05) at almost all height bins. For comparison, field methods often show increased leaf area index and stem density in ADE sites with, for example, leaf area index (LAI) (m<sup>2</sup> m<sup>-2</sup>) 5.0 ± 0.19 for control and 5.3 ± 0.21 for ADE sites and stem density (stems ha<sup>-1</sup>) 448 control and 547 for terra preta sites (Doughty *et al.*, 2014). Thus, ADE sites resembled a

disturbed, early successional forest with higher stem density and faster-growing trees (higher NPP). Figure S5 shows that both geoglyph and ADE sites increased in phosphorus using Sentinel 2 trait maps, relative to the control plots, though the effect size was small (~0.2%) and not significant. There was also spatial structure to the trait data that suggests spatial autocorrelation which we correct for in the next section.

### **Spatial models accounting for spatial autocorrelation**

Lastly, we combined all the climate (MMP, Tmax, VPD), soils (CEC), traits (P, LMA, WD), and structure (biomass, height, PAVD 0-5) data to both predict the presence of ADE and geoglyph sites and develop a spatially explicit understanding of the enduring impacts of ancient peoples on forests. We controlled for the spatial autocorrelation of nearby plots by using SARerr models and grouping all ADE and geoglyph plots in 40 km neighborhoods. Table 2 shows the weighting for the significant variables in the SARerr model as chosen by minimizing AIC. Leaf traits like LMA and % P were significant predictors of ADE sites but not geoglyphs, while wood density was a predictor for both. Forest structure was a strong predictor of geoglyphs; along with forest height, PAVD at 0-5m, and biomass were all important, while only PAVD 0-5m was important for ADE. Soil fertility was a predictor for both site types, but in opposite directions, with cation exchange capacity (CEC) reducing the likelihood of ADE while increasing it for geoglyphs. VPD is negatively associated with geoglyphs, while precipitation is positively associated with ADE. The results of our SARerr model broadly match but also differ from our results from Figs 2-4, possibly because when we group the data to account for spatial autocorrelation we remove some of the scaling issues noted above.

We then applied our models to the broader Amazon region to create occurrence probability maps of geoglyphs and ADE sites. Geoglyph probability peaks in the southwest Amazon with an average % likelihood of occupation of  $18 \pm 8\%$  per pixel, with several regions having up to a 50% likelihood, especially in the central SW Amazon. As this matches the spatial distribution of known geoglyphs, which are concentrated in SW Amazonia (Iriarte *et al.* 2020; Pärsinnen *et al.* 2026; Peripato *et al.* (2023)). In contrast, the distribution of our predictions is normal with a mean at 20% and has a low ( $R^2$  of 19%) prediction accuracy (although  $R^2$  is also a function of the number of control points chosen). In our ADE prediction maps, we identified areas near rivers with known ADE sites as hot spots. This matches many known ADE sites, although many are also found on terra firme sites (e.g., Parsinnen *et al.* 2026; Iriarte 2024; Balee 2013). We also note a hotspot in the NW Amazonia, but think this is likely a false signal due to natural forest species and trait changes in that region (ter Steege *et al.*, 2006). We have an average % likelihood of occupation of  $5.4 \pm 8\%$  (versus 3.2% of the forest in McMichael *et al.* (2014) with several regions having up to a 30% likelihood. Our ADE prediction accuracy is very low with an  $R^2$  of 10%.

We next predict forest biomass by creating a model using archaeology (ADE or geoglyph or control) as a predictor in addition to more traditional variables like forest structure, soils, and climate. We found that the presence of geoglyphs was a significant predictor of forest biomass, resulting in increased forest biomass by  $33 \text{ Mg ha}^{-1}$ , which was larger than what we found in Figure 3. Overall, forest biomass predictability was low at  $r^2=0.2$ , and removing archaeology

reduced it to 0.18, or by 10%. Including the ADE data layer did not improve predictions of forest biomass at ADE sites, which had a higher overall predictability at  $r^2=0.45$ . Combining both the ADE and geoglyphs as one archaeology-based predictor is also highly significant but improves the  $r^2$  by less (3%). Other significant predictors were forest structure (forest height, PAVD, WD), as shown in previous work (Doughty *et al.*, 2024), and Tmax. Leaf traits were not significant predictors for biomass.

Finally, we predicted the impact of past peoples on the current biomass of the Amazon region. To do this, we use the geoglyph coefficients from Table 3 and the data layers to estimate total biomass with (Figure 6c) archeology. We then subtract a map where archeology prediction is set to zero across the basin from Fig 6c to estimate only the carbon attributed to pre-colonial human activity (Figure 6d). The resulting biomass map displays a spatial trend aligning with many Amazonian biomass estimates, as biomass peaked in the Northeastern Amazon and was lower in the Northwestern Amazon (Saatchi *et al.*, 2011). Forest biomass in the Amazon with archeology (a “peopled Amazonia”) subtracted from forest biomass without archeology (an artificially “unpeopled Amazonia”) has similar spatial trends to our predicted geoglyph maps (Figure 6a). We divide the total carbon in Fig 6d by the total carbon in Fig 6c to estimate that past people increased Amazonian carbon by  $2.8 \pm 1.6\%$  for the entire Amazon basin (this estimate is based only on the geoglyph coefficient). Given that we estimated that a given pixel would have a 16% chance of being peopled on average (Figure 6a), there was an increase of 33 Mg to the mean “unpeopled” forest biomass of  $238 \text{ Mg ha}^{-1}$ , or  $\sim 5 \text{ Mg ha}^{-1}$  on average. Ancient human settlement in Amazonia appears to have directly increased forest carbon, albeit by a relatively small amount.

**Discussion** – In the Amazon basin, we connected remote sensing to archeological data to 1) assess whether the enduring impacts of ancient peoples on modern Amazonian forests could be detected using remote sensing and 2) refine our global carbon estimations according to ancient land use. Our remote sensing analysis suggests that ancient Amazonian societies may have increased current tropical forest biomass by  $2.8 \pm 1.6\%$  of Amazonian above-ground biomass (AGB) or  $0.8\%$  of global AGB, based on estimates that the Amazon contains  $\sim 28\%$  of global aboveground biomass carbon (Fendrich et al., 2025) (Saatchi et al., 2007). Both our field campaign in Guatemala and several other ground studies (McMichael et al., 2023) contribute to a triangulation that suggests ancient land use can modify both biomass and leaf traits in contemporary neotropical forests. For instance, changes in species composition in formerly occupied areas have been shown in both the Maya regions (Campbell *et al.*, 2006; Brokaw *et al.*, 2025) and the Amazon basin (Balée 2013; Junqueira, Shepard and Clement, 2010; McMichael et al., 2015).

We do not have in situ biomass or trait data for the Amazon, but we do have these for the Maya regions, which show similar trends to those in the Amazon. Our transects in Xmakabatur demonstrated enduring human impacts on species composition (both on structures and agricultural terraces), which could drive both the measured changes in biomass and trait values. Biomass on ancient structures was much higher than in control regions ( $402 \pm 264$  \*\*\* versus  $122 \pm 51$  Mg ha<sup>-1</sup>), largely due to the presence of large Ramón trees, which were valuable to the Maya, and have long been recognized to predominate on ruins (Lambert and Arnason, 1982). In terraced agricultural areas, palms were more abundant, which could change forest structure and leaf traits. However, palms tend towards increased LMA (i.e. in our limited dataset palm LMA was  $93 \pm 16$  versus  $83 \pm 25$  g m<sup>-2</sup>), thus opposite of the landscape RS scale trait change, so palms are not driving the landscape trait changes because they tend to be shorter than other vegetation but may contribute towards the observed structural changes.

In addition to species composition shifts, modern forest traits and biomass differences on ancient structures in the Maya region could differ because ancient peoples modified soil substrates by building structures with unweathered limestone blocks. These building practices resulted in higher soil calcium and phosphorus on structures than control plots (i.e. soil Ca  $10.42 \pm 10$  vs  $2.62 \pm 2.5$ ). The increased soil calcium was transferred to the trees growing on structures, as leaf-level elemental compositions of trees on structures differed from the control (i.e. Ca  $3.90 \pm 2.6$  versus  $2.37 \pm 2.5$ ). Ancient structures also affected LMA ( $81.80 \pm 28$  vs.  $88.87 \pm 23$ ), which may reflect both changes in species composition and trees being more water-limited during the dry season due to growing on unweathered rock or shallower soils. At this time, it is difficult to disentangle how much the biomass and trait changes can be attributed to species composition versus other factors (e.g. soil nutrients). Prior studies have found that Maya sites tend to have taller, denser vegetation with a fuller canopy during the dry season because of changes in soil chemistry and hydrology (LaRocque, Leblon and Ek, 2019).

The remotely sensed patterns showed similar magnitude and direction change to our ground terraced agriculture transects, which makes sense because improved agricultural areas (terraces, field walls, berms, etc.) comprise a larger area than do archaeological buildings *per se*.

Structures may also contribute to the increased biomass signal since they cover ~1-2% of land area and have ~3x more biomass. There is also a scale issue because there are two different scales (25m diameter for GEDI and 10 by 10m for trait maps), but even the smallest scale cannot detect individual trees. Our transects show that the focal species were dispersed among other trees and not always grouped, thus even 10 by 10m pixels would contain mixed pixels. Most of our ground terraced agriculture results were not statistically significant due to small sample sizes and larger future field campaigns should better verify these numbers. Additionally, we do not control for selective logging at any of the sites, which is another aspect that should be controlled for in future studies. Leaves on archaeological features at wet sites in Mexico are smaller, drier and denser with different spectral features than leaves growing in the surrounding forest and they could be classified (grown on vs off structures) with 76% accuracy (B. Wiebe personal communication). Others have shown archeological features such as ADE can impact forest spectral features (Söderström *et al.*, 2016; Palace *et al.*, 2017b) likely through leaf trait changes. Thus, in theory, optical satellite remote sensing could detect spectral changes in forests associated with archeological features—and this is what we found, for both Amazonia and northern Guatemala. In addition, active remote sensing like GEDI lidar could detect changes in forest structure due to ancient peoples because our ground transects showed very large forest structural changes from species changes. In Guatemala, remote sensing data showed that LMA and P are significantly reduced, and biomass was significantly increased by ~6 Mg ha<sup>-1</sup> in regions with more structures/km<sup>2</sup> compared to places with fewer.

We propose that maps of archaeological settlement in Amazonia stand to be improved by incorporating remote sensing data that capture the historical-ecological effects of ancient settlement on the contemporary forest (Howey *et al.*, 2020; C N H McMichael *et al.*, 2023). The ground transects in Guatemala and remotely sensed Amazonian results suggest that foliar traits and forest structure may contain additional information identifying previously inhabited areas. The precise spatial distribution(s) of ancient peoples is much more uncertain in the Amazon than in the Maya region. Recent estimates suggest at least 5-6 million, perhaps as many as 10 million people inhabited pre-Columbian Amazonia. 1.25 to 3 million may have lived in the modern region of Acre alone and as many as 6 million along the southern rim of the Amazon (Denevan 2014). However, with little data from many remote regions, even this wide bracket remains highly uncertain (Hirst *et al.* 2025; Pärsinnen *et al.* 2026). Since high-resolution airborne lidar is unlikely to be available across the entire basin, satellite data may provide useful information in data-poor regions. Much recent effort has gone into predicting both ADE and geoglyph sites across the Amazon basin beyond localized aircraft lidar (Peripato *et al.*, 2023) ; McMichael *et al.*, 2014; Walker *et al.*, 2023). However, these studies mainly use spatially extrapolated datasets (such as elevation, distance to rivers, soils and climate), and not remotely sensed datasets, so there is potential room for improvement by incorporating satellite multispectral and lidar data.

We investigated the consequences of two different ancient land-use categories on modern forests—geoglyphs, which are large soil designs, and ADE, where soil was enriched in nutrients. We found that forest structure (biomass, height, PAVD and WD) was most affected at geoglyph sites, while leaf traits such as LMA and P were most affected at ADE sites. Forests at the geoglyph sites were significantly taller but with much less volume (PAVD <20%) than the

control plots (Figures 4 and 5). We hypothesize that forest structure is different near geoglyph sites because ancient people planted palms, which reduced forest bulk due to their low LAI, and fruit trees, which would increase overall forest height (Montoya *et al.*, 2020). ADE forests often have different species compositions and typically have increased phosphorus and cation exchange capacity (Quesada *et al.*, 2010), which influence leaf chemistry and structure (Fyllas, Quesada and Lloyd, 2012). Here, leaf traits were significant predictors of ADE sites only after spatial autocorrelation was removed, suggesting that local variation in the trait data caused statistical noise. Our ADE results largely reflect those of a prior study using lower-resolution optical imagery (MODIS 250m resolution) that found ADE forests were shorter with lower biomass. However, they found that canopies were less green (lower NDVI), which we might not expect with increased P and LMA (Palace *et al.*, 2017), although others found increased greenness (EVI) (Choi, Wright and Lima, 2020). At both types of archeological sites, wood density was positively associated with ancient human activity after accounting for spatial autocorrelation, potentially as a result of ancient peoples planting large Amazonian fruit trees, which tend to be larger with higher wood density (Doughty *et al.*, 2016).

In predicting the location of these ancient sites using remotely sensed forest characteristics, the geoglyph plots were predicted much more strongly than the ADE plots with our linear models ( $r^2 = 0.2$  vs  $0.1$ ), although prediction for both was weak. However, with future “wall-to-wall” missions like BIOMASS, ESA’s P-band Synthetic Aperture Radar (SAR) instrument with a resolution of  $50 \times 50 \text{ m}^2$  or Surface Biology and Geology (SBG, an upcoming wall-to-wall hyperspectral satellite) (Schimel & Poulter, 2022), these predictions will likely improve.

Our predictions of new locations of geoglyphs differ slightly from prior predictions (Peripato *et al.*, 2023) that did not use remotely sensed data as predictor variables. Our predictions expand more into central Amazonia although neither prediction shows many sites in Western Amazonia. It is unknown whether this indicates a lack of people in these regions or just a cultural shift to non-geoglyph building people. There are natural variations to forest structure and traits across Amazonia (Doughty *et al.*, 2023) that are not currently accounted for in the results and predictability could be improved in the future by accounting for this. Our predictions of ADE differ more substantially from prior predictions (McMichael *et al.*, 2014). The large, predicted signal in the Northwest may not be real because it is far from the river basins traditionally thought to contain ADE and may reflect forest trait changes associated with species composition changes that follow a south-west/north-east soil fertility gradient and a north-west/south-east precipitation gradient (ter Steege *et al.*, 2006). The predictive power is low ( $r^2 < 0.1$ ), and we do not suggest our predictions should be taken at face value, but rather that they represent an early test case of methods that will improve rapidly in the future.

The Amazonian forest inventory plot locations might be spatially biased toward areas that had archeological evidence (both ADE and geoglyphs) suggesting human occupation (McMichael *et al.*, 2017). Our work suggests anthropogenic forests might have higher biomass and, thus, we might be overestimating total Amazonian carbon. We use traits and forest structure to predict that the past presence of people may have increased basin-wide biomass by  $2.8 \pm 1.6\%$ . Our estimated 2.8% increase in Amazonian aboveground biomass corresponds to approximately 0.8%

of global aboveground biomass carbon, based on estimates that the Amazon contains ~28% of global aboveground biomass carbon (Fendrich et al., 2025) (Saatchi et al., 2007). Spatially, we predict that the Eastern Amazon increased its forest biomass as the result of prior human inhabitation more than the western Amazon. Broadly, the increase in biomass from ancient settlement is likely due to humans increasing the presence of larger seed dispersed fruit trees versus smaller wind dispersed trees in non-anthropogenic areas, but this needs to be better verified with ground data. A prior study found Late Pleistocene megafaunal extinctions reduced carbon content in the Amazon by ~1.5% because of reduced spread of high wood density large fruit trees (Doughty *et al.*, 2016). Here we suggest that early Amazonian people may have continued the ecological roles of the extinct megafauna by continuing to plant similar, larger, higher wood density tree types and thus partially reversed the potential carbon loss from the extinctions. Early people would have likely had a different role on nutrient dynamics by concentrating (e.g., ADE) versus dispersing nutrients across the landscape as animals tend to do (Doughty, Wolf and Malhi, 2013).

To specifically answer our questions:

**Q1 – Do archeological sites covered in forests show consistent impacts across the neotropics on forest structure and plant traits?**

There are trends that former human occupation modified both traits and structure, but the signal varies depending on the type of human occupation (i.e. geoglyphs vs ADE) and also on structures versus on farming.

**Q2 – Is tropical forest biomass higher in areas with evidence for ancient human settlement?**

In most areas (Maya regions and Geoglyphs, but not ADE), biomass is higher where there was human archeological evidence.

**Conclusions**

Overall, we used ground transects to show how ancient peoples affect biomass and leaf traits in the Maya region, where occupancy is well known. These trends were also present at the landscape scale using remotely sensed trait and structure data. We then moved to the Amazon where occupancy is less well known and showed that the presence of geoglyphs and ADE sites impacted both traits and forest structure in ways that were largely consistent with the Maya region. We then predicted that ancient peoples would have a small increase in Amazonian carbon. However, although our prediction power was weak, it will likely improve in the future as more powerful active and passive remote sensors come online.

**Conflict of interest statement** – The authors have no conflict of interest to disclose.

**Acknowledgements** – We would like to thank the staff near Xmakabatun who made the field work possible.

577 **Open research - Code and data availability** - Description of the type(s) of data and/or  
578 software -  
579 Data - Data and its descriptions to create all figures and tables in this paper are available  
580 (Doughty 2026).  
581 Software – All code and its descriptions to create all figures and tables in this paper are available  
582 (Doughty 2026).  
583

## Tables

**Table 1** – Field results from Guatemala from 29, 30m transects where we measured leaf and soil phosphorus and calcium with an XRF, LMA, plot biomass and number of species of interest (Copál, Corozo, Ramón, Chico zapote, Guano/Botán). At the bottom, we compare these values to remotely sensed values for the polygons in Figure 1c for agriculture (red text) and control sites (blue text).

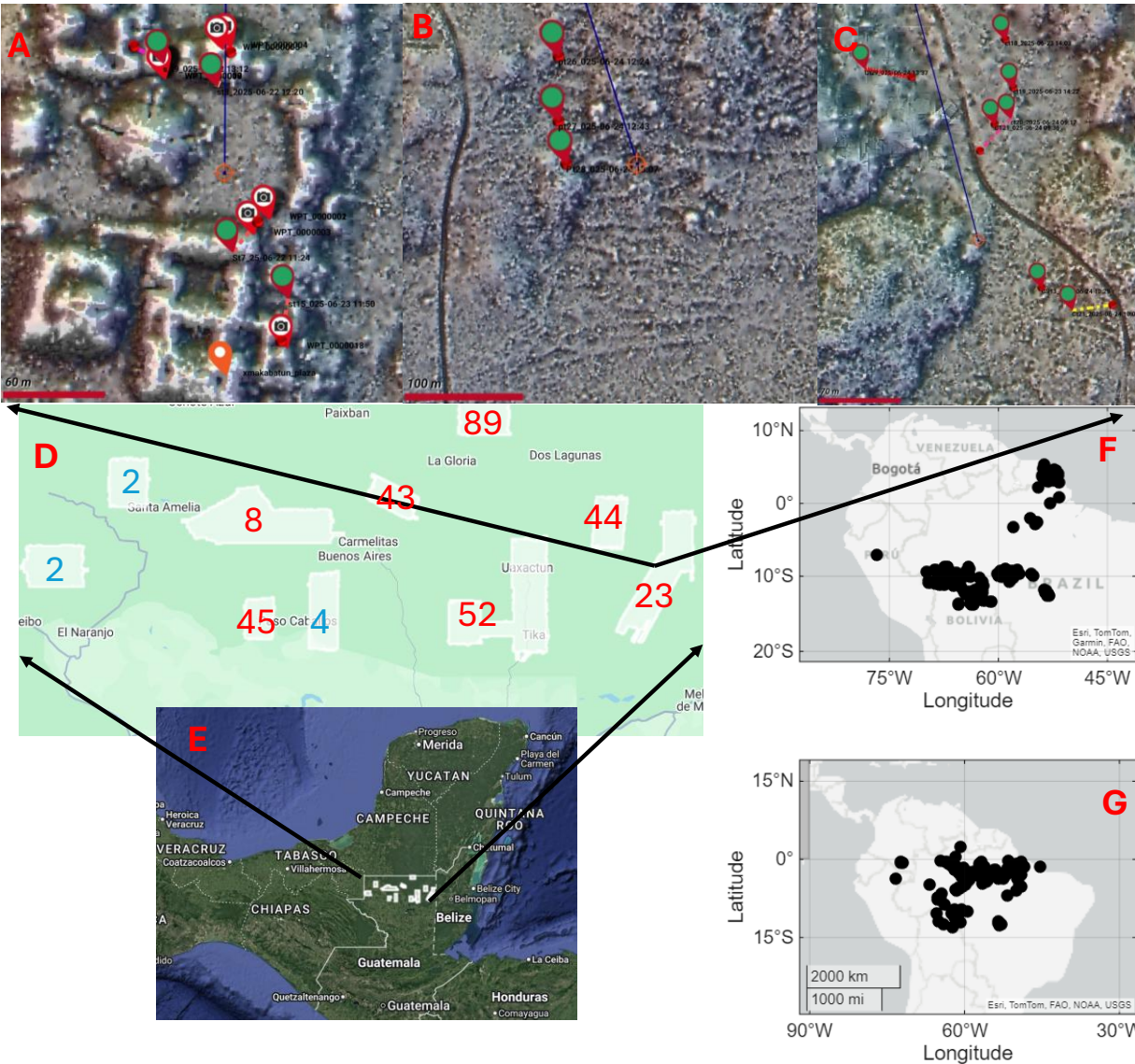
|                                      | <b>Control</b>     | <b>Structure</b>   | <b>Agriculture</b>  |
|--------------------------------------|--------------------|--------------------|---------------------|
| Phos ppt leaves                      | 0.065 ± 0.04, N=39 | 0.075 ± 0.04, N=43 | 0.061 ± 0.04, N=40  |
| Ca ppt leaves                        | 2.50 ± 1.5, N=39   | 3.90 ± 2.5, N=43   | 3.16 ± 2.7, N=40    |
| Phos ppt soil                        | 0.067 ± 0.08, N=10 | 0.081 ± 0.08, N=21 | 0.0178 ± 0.02, N=12 |
| Ca ppt soil                          | 2.62 ± 2.5, N=10   | 10.42 ± 10, N=21   | 1.82 ± 2.1, N=12    |
| LMA g m <sup>-2</sup>                | 88.87 ± 23, N=36   | 81.80 ± 28, N=43   | 81.07 ± 23, N=41    |
| Biomass Mg ha                        | 122 ± 51, N=9      | 402 ± 264, N=9     | 163 ± 111, N=11     |
| Individuals of focal species (count) | 2.1 ± 1.4, N=9     | 8.9 ± 4.3, N=9     | 7.6 ± 4.6, N=11     |
| RS Phos % canopy                     | 0.120 ± 0.004      |                    | 0.117 ± 0.004       |
| RS LMA g m <sup>-2</sup>             | 135 ± 2            |                    | 131 ± 7             |
| RS wood density g cm <sup>-3</sup>   | 0.577 ± 0.006      |                    | 0.577 ± 0.008       |
| RS biomass Mg ha                     | 137 ± 88           |                    | 148 ± 89            |

**Table 2** –Coefficients that minimize AIC after accounting for spatial autocorrelation with a neighborhood distance of 40km to predict ADE (N=85 TP, C=1020) and geoglyphs (G=52, C = 583). We test total cation exchange capacity (CEC) for the top 0-5 cm soil layer (mmol(c) kg<sup>-1</sup>), Vapor Pressure Deficit (VPD in kPa), Mean Monthly Precipitation (MMP) (mm/month), maximum temperature (Tmax °C), remotely sensed canopy traits such as phosphorus (%P), leaf mass area (LMA g m<sup>-2</sup>), wood density (WD), and structure (biomass, forest height (m), plant area volume density (PAVD) from 0-5m.

|                              | Geoglyphs<br>r <sup>2</sup> =0.19, AIC-<br>91.9 | Amazonian Dark<br>Earth r <sup>2</sup> =0.10,<br>AIC-313 |
|------------------------------|-------------------------------------------------|----------------------------------------------------------|
| RS LMA g m <sup>-2</sup>     |                                                 | 51                                                       |
| RS P % canopy                |                                                 | 3.3                                                      |
| RS biomass Mg ha             | 0.00014                                         |                                                          |
| RS WD g cm <sup>-3</sup>     | 4.92                                            | 1.3                                                      |
| Forest height (m)            | -0.0067                                         |                                                          |
| PAVD 05                      | -1.28                                           | 0.96                                                     |
| MMP mm/month                 |                                                 | 0.0026                                                   |
| CEC mmol(c) kg <sup>-1</sup> | 0.002                                           | -0.003                                                   |
| VPD kPa                      | -.33                                            |                                                          |

**Table 3** –Coefficients that minimize AIC after accounting for spatial autocorrelation with a neighborhood distance of 40km to predict biomass with archeology as a possible variable for ADE (TP=85, C=1020) and geoglyphs (G=52, C = 583). We test archeology likelihood from figure 7a, total cation exchange capacity (CEC) for the top 0-5 cm soil layer (mmol(c) kg<sup>-1</sup>), Mean Monthly Precipitation (MMP) (mm/month), maximum temperature (Tmax °C), remotely sensed canopy traits such as phosphorus (% P), leaf mass area (LMA g m<sup>-2</sup>), wood density (WD), and structure (biomass, forest height (m), plant area volume density (PAVD) from 0-5m.

|                              | Geoglyphs<br>r <sup>2</sup> =0.20, AIC=6972 | Amazonian Dark<br>Earth, r <sup>2</sup> =0.45,<br>AIC=11306 |
|------------------------------|---------------------------------------------|-------------------------------------------------------------|
| archeology                   | 33                                          |                                                             |
| RS LMA g m <sup>-2</sup>     |                                             |                                                             |
| RS P % canopy                |                                             |                                                             |
| RS biomass Mg ha             |                                             |                                                             |
| RS WD g cm <sup>-3</sup>     |                                             | -129                                                        |
| Forest height (m)            | 5.9                                         | 9.5                                                         |
| PAVD 05                      | -527                                        | -206                                                        |
| MMP mm/month                 |                                             |                                                             |
| Tmax (°C)                    | -2.8                                        |                                                             |
| CEC mmol(c) kg <sup>-1</sup> |                                             |                                                             |



615  
616 **Figure 1 – Map of sites.** High resolution lidar collected near Xmakabatun, in Guatemala for  
617 structures (A), agriculture (B), and control (C) (Canuto *et al.*, 2018). Red dots indicate 30m  
618 transects. We compare the impact of civilizations on forest carbon structure and leaf traits for  
619 four regions. (D) Polygons from Canuto et al 2018 with estimates of population densities of the  
620 classic lowland Maya regions for low (blue) and high (red). (E) Larger region shown. Locations  
621 of geoglyphs (F) and ADE sites known in Brazil(G).

622

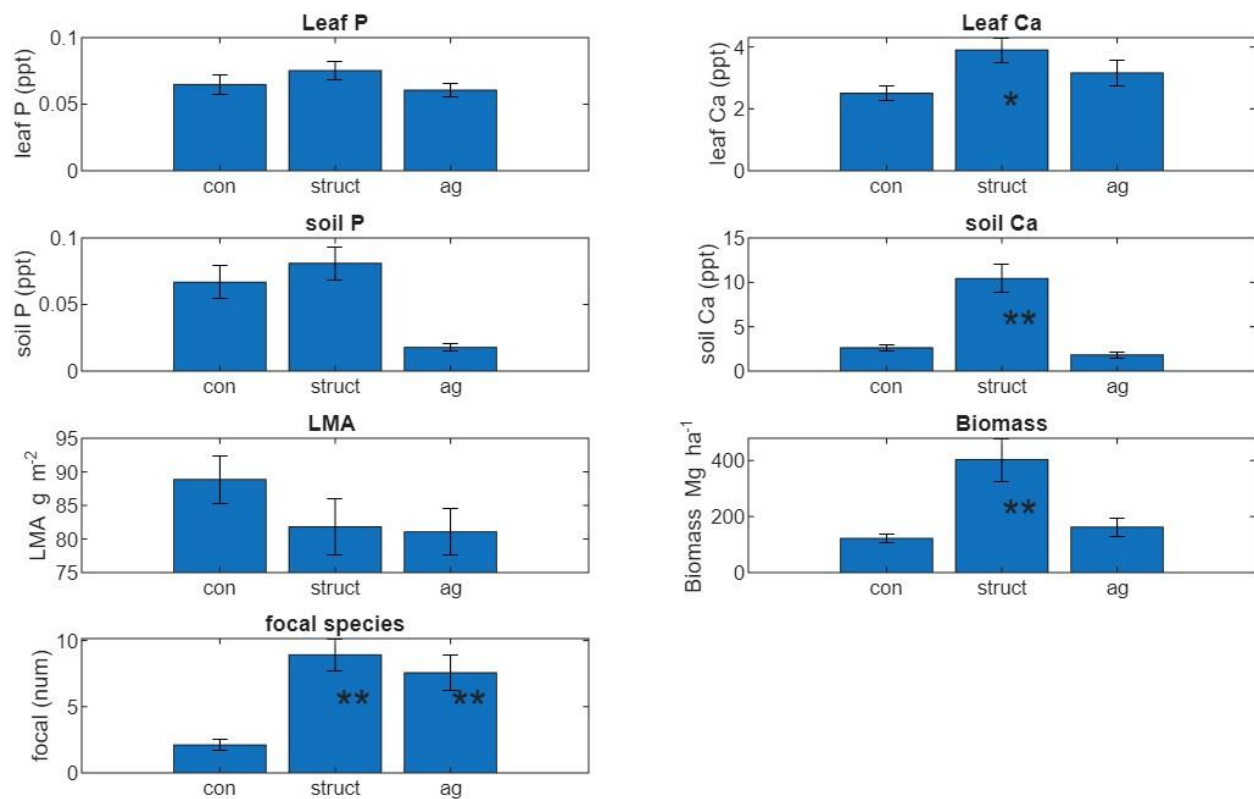


**Figure 2** – Field results from Guatemala from 29, 30m transects where we measured leaf and soil phosphorus and calcium with an XRF, LMA, plot biomass and number of individuals of the species of interest (Copál, Corozo, Ramón, Chico zapote, Guano/Botán). Stars indicate (\*= $0.05$ , \*\*= $0.01$ ) significant differences from the control using an anova and a multiple comparison with Tukey's honestly significant difference procedure.

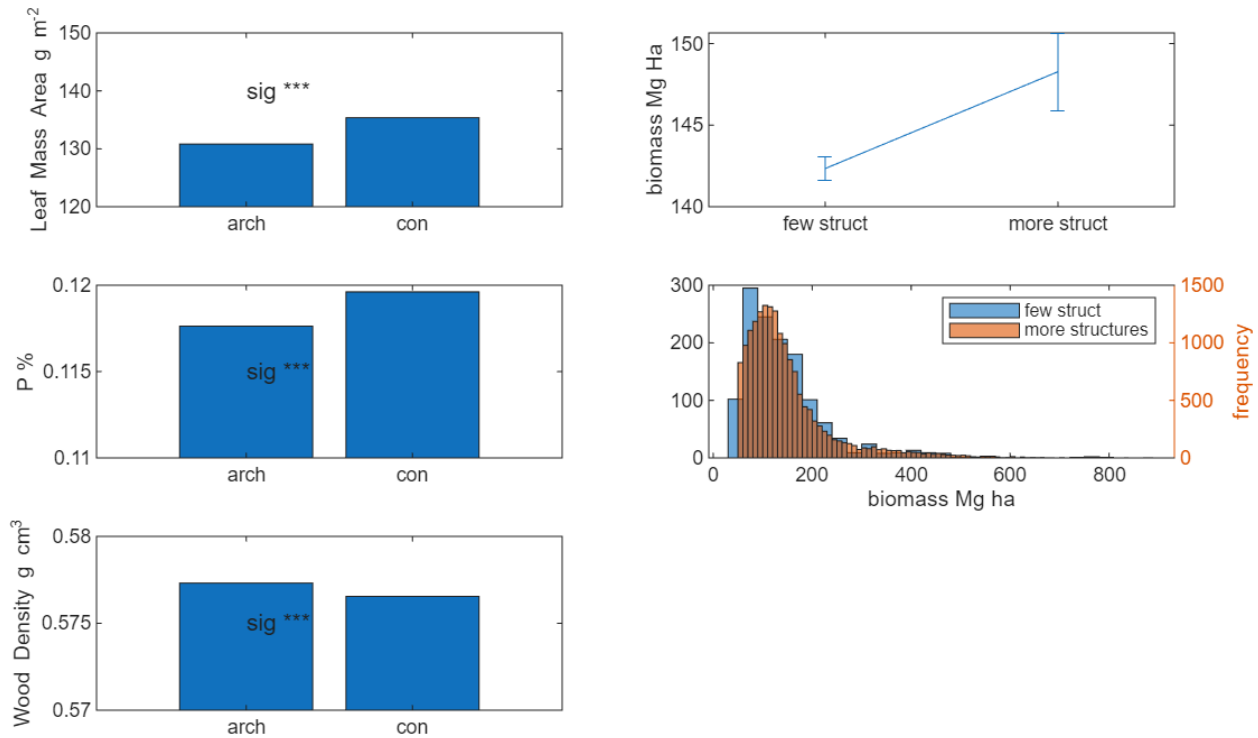
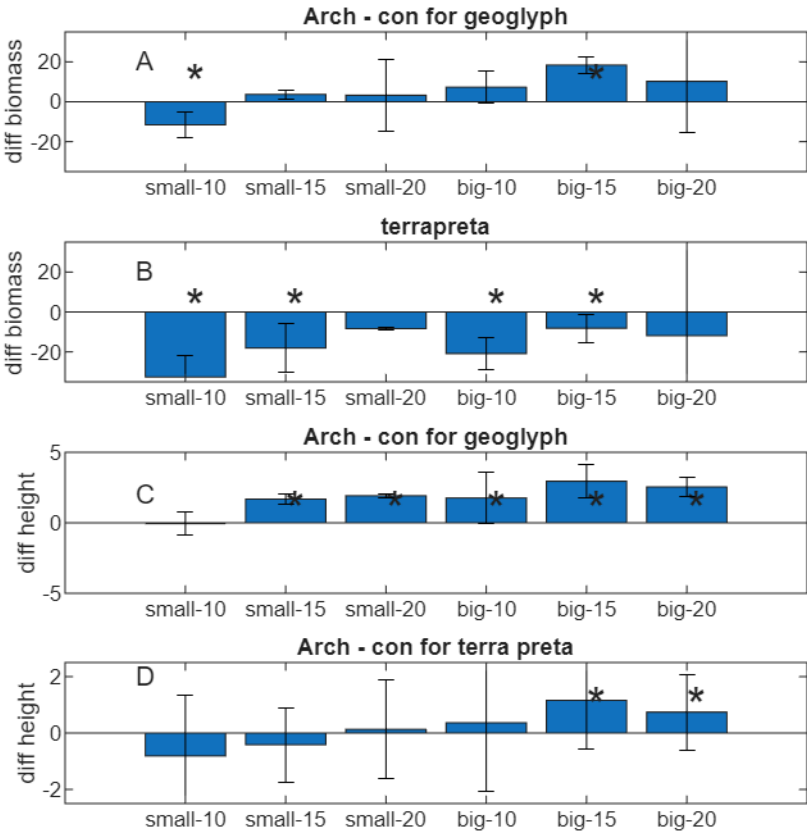


**Figure 3**– (left) Differences in average remote sensed trait values for regions in Peten, Guatemala near archeology (arch – the blocks with red numbers in Fig 1D) or the control areas (con - the blocks with blue numbers in Fig 1D) for LMA g m<sup>-2</sup> (top), Phosphorus % (middle) and wood density g cm<sup>-3</sup> (bottom). (right) Differences in average biomass values for regions near archeology (arch) or the control areas (con) averages (top) and histogram (middle).

637



638

639

640

641

642

643

644

645

**Figure 4–** (a) Archeology in the Amazon biomass minus control for a small area (small) and big (big) area with tree height thresholds of 10,15 and 20m for geoglyph biomass (a), terra preta biomass (c), geoglyph forest height and (d) terra preta forest height. A “small” region of 3x3 pixels around the centroid, resulting in 900x900m<sup>2</sup> pixels, and a “large” region of 21x21 pixels to obtain 6300x 6300m<sup>2</sup>. Error bars are standard deviations. Stars indicate significance.

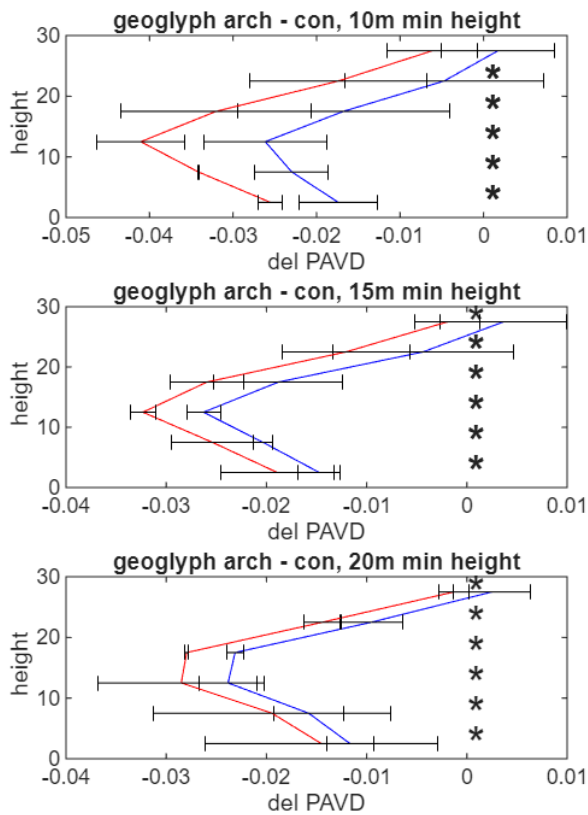
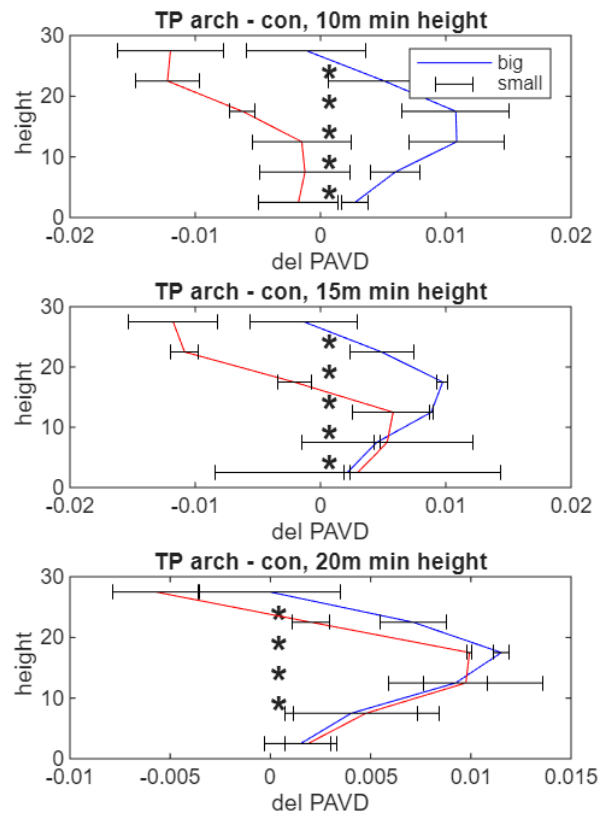
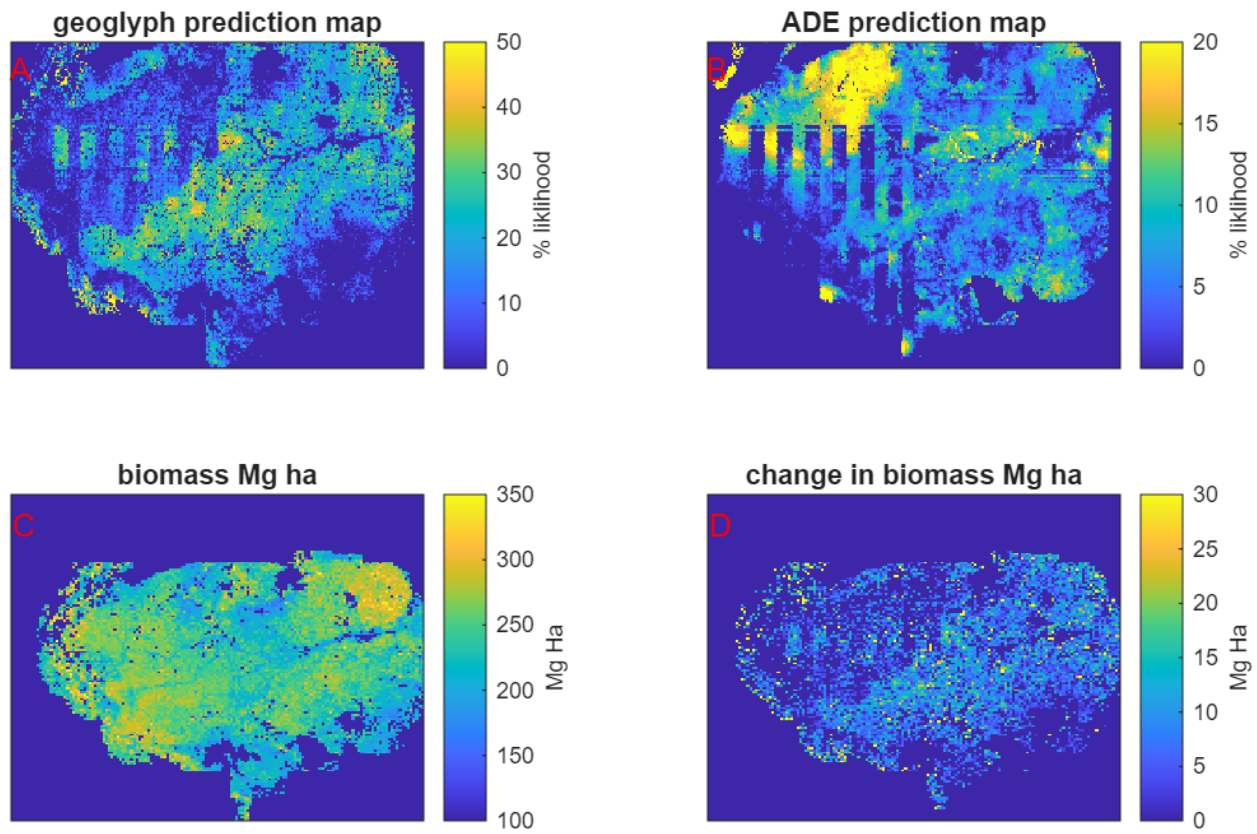


**Figure 5** – Change in forest structure in the Amazon for a small (red) and large (blue) area surrounding the coordinates of geoglyphs (left) and terra preta sites (right) with forest height thresholds of 10m (top), 15m (middle) and 20m (bottom). Error bars are standard deviations. Stars indicate significant different from 0 ( $P < 0.05$ ) for the blue lines.

657



658

659 **Figure 6** –A prediction map for Geoglyphs sites (A) and ADE sites (B) with warmer colors  
 660 indicating a higher likelihood of finding archeological sites. Predicted biomass (C) with  
 661 archology as a predictor and change in biomass (D) (a biomass map with no archeology  
 662 predicted subtracted from map C). Note – the western Amazonian stripes are due to missing  
 663 CEC data and Fig S6 shows the maps A and B with fixed CEC and no stripes.

664

## References

- Abatzoglou, J.T. *et al.* (2018) “TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015,” *Scientific Data*, 5(1), p. 170191. Available at: <https://doi.org/10.1038/sdata.2017.191>.
- Aguirre-Gutiérrez, J. *et al.* (2025) “Canopy functional trait variation across Earth’s tropical forests,” *Nature*, 641(8061), pp. 129–136. Available at: <https://doi.org/10.1038/s41586-025-08663-2>.
- Auld-Thomas, L. *et al.* (2024) “Running out of empty space: environmental lidar and the crowded ancient landscape of Campeche, Mexico,” *Antiquity*. 2024/10/29, 98(401), pp. 1340–1358. Available at: <https://doi.org/DOI: 10.15184/aqy.2024.148>.
- Balée, W., (2013), *Cultural Forests of the Amazon: A Historical Ecology of People and their Landscape*. The University of Alabama Press, Tuscaloosa.
- Bartoń, K. (2009) “MuMIn: Multi-model inference.” *R package version 0.12.2*. Retrieved from <http://r-forge.r-project.org/projects/mumin/>
- Basso, L.S. *et al.* (2023) “Atmospheric CO<sub>2</sub> inversion reveals the Amazon as a minor carbon source caused by fire emissions, with forest uptake offsetting about half of these emissions,” *Atmospheric Chemistry and Physics*, 23(17), pp. 9685–9723. Available at: <https://doi.org/10.5194/acp-23-9685-2023>.
- Batjes, N.H., Ribeiro, E. and van Oostrum, A. (2020) “Standardised soil profile data to support global mapping and modelling (WoSIS snapshot 2019),” *Earth System Science Data*, 12(1), pp. 299–320. Available at: <https://doi.org/10.5194/essd-12-299-2020>.
- Burns, P., Hakkenberg, C.R. and Goetz, S.J. (2024) “Multi-resolution gridded maps of vegetation structure from GEDI,” *Scientific Data*, 11(1), p. 881. Available at: <https://doi.org/10.1038/s41597-024-03668-4>.
- Campbell, D.G. *et al.* (2006) “The feral forests of the eastern Petén,” *Time and Complexity in Historical Ecology: Studies in the Neotropical Lowlands*, pp. 21–55.
- Canuto, M.A. *et al.* (2018) “Ancient lowland Maya complexity as revealed by airborne laser scanning of northern Guatemala,” *Science*, 361(6409), p. eaau0137. Available at: <https://doi.org/10.1126/science.aau0137>.
- Chave, J. *et al.* (2005) “Tree allometry and improved estimation of carbon stocks and balance in tropical forests,” *Oecologia*, 145(1), pp. 87–99. Available at: <https://doi.org/10.1007/s00442-005-0100-x>.
- Choi, J., Wright, D.K. and Lima, H.P. (2020) “A new local scale prediction model of Amazonian landscape domestication sites,” *Journal of Archaeological Science*, 123, p. 105240. Available at: <https://doi.org/https://doi.org/10.1016/j.jas.2020.105240>.

700 Cobb, A.R. *et al.* (2023) “Improved terrain estimation from spaceborne lidar in tropical peatlands  
 701 using spatial filtering,” *Science of Remote Sensing*, 7, p. 100074. Available at:  
 702 <https://doi.org/https://doi.org/10.1016/j.srs.2022.100074>.

703 Denevan, W.M. (2014) “Estimating Amazonian Indian Numbers in 1492,” *Journal of Latin*  
 704 *American Geography*, 13(2), pp. 207–221. Available at: <http://www.jstor.org/stable/24395921>.

705 Doughty, C.E. *et al.* (2014) “The production, allocation and cycling of carbon in a forest on  
 706 fertile terra preta soil in eastern Amazonia compared with a forest on adjacent infertile soil,”  
 707 *Plant Ecology & Diversity*, 7(1–2), pp. 41–53. Available at:  
 708 <https://doi.org/10.1080/17550874.2013.798367>.

709 Doughty, C.E. *et al.* (2016) “Megafauna extinction, tree species range reduction, and carbon  
 710 storage in Amazonian forests,” *Ecography*, 39(2), pp. 194–203. Available at:  
 711 <https://doi.org/https://doi.org/10.1111/ecog.01587>.

712 Doughty, C.E. *et al.* (2023) “Tropical forests are mainly unstratified especially in Amazonia and  
 713 regions with lower fertility or higher temperatures,” *Environmental Research: Ecology*, 2(3), p.  
 714 35002. Available at: <https://doi.org/10.1088/2752-664X/ace723>.

715 Doughty, C.E. *et al.* (2024) “Satellite Derived Trait Data Slightly Improves Tropical Forest  
 716 Biomass, NPP and GPP Estimates,” *Journal of Geophysical Research: Biogeosciences*, 129(7),  
 717 p. e2024JG008108. Available at: <https://doi.org/10.1029/2024JG008108>.

718 Doughty, C. E. (2026). Ancient human occupation alters modern forest traits, structure and  
 719 biomass in neotropical forests [Dataset]. *Figshare*. <https://doi.org/10.6084/m9.figshare.32013042>

720 Dubayah, R. *et al.* (2020) “The Global Ecosystem Dynamics Investigation: High-resolution laser  
 721 ranging of the Earth’s forests and topography,” *Science of Remote Sensing*. Available at:  
 722 <https://doi.org/10.1016/j.srs.2020.100002>.

723 Dubayah, R.O. *et al.* (2023) “GEDI L4B Gridded Aboveground Biomass Density, Version 2.1.”  
 724 ORNL Distributed Active Archive Center. Available at:  
 725 <https://doi.org/10.3334/ORNLDAAC/2299>.

726 Duncanson, L. *et al.* (2022) “Aboveground biomass density models for NASA’s Global  
 727 Ecosystem Dynamics Investigation (GEDI) lidar mission,” *Remote Sensing of Environment*, 270,  
 728 p. 112845. Available at: <https://doi.org/https://doi.org/10.1016/j.rse.2021.112845>.

729 Estrada-Belli, F. *et al.* (2025) “New regional-scale Classic Maya population density estimates  
 730 and settlement distribution models through airborne lidar scanning,” *Journal of Archaeological*  
 731 *Science: Reports*, 66, p. 105288. Available at:  
 732 <https://doi.org/https://doi.org/10.1016/j.jasrep.2025.105288>.

733 F. Dormann, C. *et al.* (2007) “Methods to account for spatial autocorrelation in the analysis of  
 734 species distributional data: a review,” *Ecography*, 30(5), pp. 609–628. Available at:  
 735 <https://doi.org/https://doi.org/10.1111/j.2007.0906-7590.05171.x>.

736 Fendrich, A., Feng, Y., Wigneron, JP. *et al.* (2025) Human influence on Amazon’s aboveground  
737 carbon dynamics intensified over the last decade. *Nat Commun* **16**, 6681 . Available at:  
738 <https://doi.org/10.1038/s41467-025-61856-1>

739 Fox J and S, W. (2019) “An R Companion to Applied Regression.”

740 Fyllas, N.M., Quesada, C.A. and Lloyd, J. (2012) “Deriving Plant Functional Types for  
741 Amazonian forests for use in vegetation dynamics models,” *Perspectives in Plant Ecology,*  
742 *Evolution and Systematics*. Available at: <https://doi.org/10.1016/j.ppees.2011.11.001>.

743 Heckenberger, M., *et al.* (2008) “Pre-Columbian Urbanism, Anthropogenic Landscapes, and the  
744 Future of the Amazon”, *Science* 321: 1214-1217. Available at:  
745 <https://www.science.org/doi/10.1126/science.1159769>

746 Hirst J., *et al.* (2025) “Constraining the population size estimates of the pre-Columbian Casarabe  
747 Culture of Amazonian Bolivia”. *PLoS One* 20(5): e0325104. Available at:  
748 <https://doi.org/10.1371/journal.pone.0325104>

749 Howey, M.C.L. *et al.* (2020) “Remotely Sensed Big Data and Iterative Approaches to Cultural  
750 Feature Detection and Past Landscape Process Analysis,” *Journal of Field Archaeology*,  
751 45(sup1), pp. S27–S38. Available at: <https://doi.org/10.1080/00934690.2020.1713435>.

752 Hubau, W. *et al.* (2020) “Asynchronous carbon sink saturation in African and Amazonian  
753 tropical forests,” *Nature*, 579(7797), pp. 80–87. Available at: [https://doi.org/10.1038/s41586-](https://doi.org/10.1038/s41586-020-2035-0)  
754 [020-2035-0](https://doi.org/10.1038/s41586-020-2035-0).

755 Iriarte, J, *et al.* (2020) “Geometry by Design: Contribution of Lidar to the Understanding of  
756 Settlement Patterns of the Mound Villages in SW Amazonia”, *Journal of Computer Applications*  
757 *in Archaeology*, 3(1), pp. 151–169. Available at: <https://doi.org/10.5334/jcaa.45>.

758 Iriarte, José. *The Archaeology of Amazonia: A Human History*. London: Bloomsbury Academic,  
759 2024.

760 Junqueira, A.B., Shepard, G.H. and Clement, C.R. (2010) “Secondary forests on anthropogenic  
761 soils in Brazilian Amazonia conserve agrobiodiversity,” *Biodiversity and Conservation*, 19(7),  
762 pp. 1933–1961. Available at: <https://doi.org/10.1007/s10531-010-9813-1>.

763 Koch, A., Hubau, W. and Lewis, S.L. (2021) “Earth System Models Are Not Capturing Present-  
764 Day Tropical Forest Carbon Dynamics,” *Earth’s Future*, 9(5), p. e2020EF001874. Available at:  
765 <https://doi.org/https://doi.org/10.1029/2020EF001874>.

766 Liu, A., Cheng, X. and Chen, Z. (2021) “Performance evaluation of GEDI and ICESat-2 laser  
767 altimeter data for terrain and canopy height retrievals,” *Remote Sensing of Environment*, 264, p.  
768 112571. Available at: <https://doi.org/https://doi.org/10.1016/j.rse.2021.112571>.

769 Lombardo, U., *et al.* (2022) Evidence confirms an anthropic origin of Amazonian Dark Earths.  
770 *Nature Communications* 13 : 3444. Available at : <https://doi.org/10.1038/s41467-022-31064-2>

771 Marselis, S.M. *et al.* (2018) “Distinguishing vegetation types with airborne waveform lidar data  
 772 in a tropical forest-savanna mosaic: A case study in Lopé National Park, Gabon,” *Remote*  
 773 *Sensing of Environment*, 216, pp. 626–634. Available at:  
 774 <https://doi.org/https://doi.org/10.1016/j.rse.2018.07.023>.

775 Marselis, S.M. *et al.* (2020) “Evaluating the potential of full-waveform lidar for mapping pan-  
 776 tropical tree species richness,” *Global Ecology and Biogeography*, 29(10), pp. 1799–1816.  
 777 Available at: <https://doi.org/https://doi.org/10.1111/geb.13158>.

778 McMichael, C.H. *et al.* (2014) “Predicting pre-Columbian anthropogenic soils in Amazonia,”  
 779 *Proceedings of the Royal Society B: Biological Sciences*, 281(1777), p. 20132475. Available at:  
 780 <https://doi.org/10.1098/rspb.2013.2475>.

781 McMichael, C.H. *et al.* (2015) “Phytolith Assemblages Along a Gradient of Ancient Human  
 782 Disturbance in Western Amazonia,” *Frontiers in Ecology and Evolution*, Volume 3-2015.  
 783 Available at: <https://doi.org/10.3389/fevo.2015.00141>.

784 McMichael, C.N.H. *et al.* (2017) “Ancient human disturbances may be skewing our  
 785 understanding of Amazonian forests,” *Proceedings of the National Academy of Sciences*, 114(3),  
 786 pp. 522–527. Available at: <https://doi.org/10.1073/pnas.1614577114>.

787 McMichael, Crystal N H *et al.* (2023) “Past human-induced ecological legacies as a driver of  
 788 modern Amazonian resilience,” *People and Nature*, 5(5), pp. 1415–1429. Available at:  
 789 <https://doi.org/https://doi.org/10.1002/pan3.10510>.

790 McMichael, C N H *et al.* (2023) “Spatial and temporal abilities of proxies used to detect pre-  
 791 Columbian Indigenous human activity in Amazonian ecosystems,” *Quaternary Science Reviews*,  
 792 321, p. 108354. Available at: <https://doi.org/https://doi.org/10.1016/j.quascirev.2023.108354>.

793 Montoya, E. *et al.* (2020) “Human Contribution to Amazonian Plant Diversity: Legacy of Pre-  
 794 Columbian Land Use in Modern Plant Communities,” in V. Rull and A.C. Carnaval (eds.)  
 795 *Neotropical Diversification: Patterns and Processes*. Cham: Springer International Publishing,  
 796 pp. 495–520. Available at: [https://doi.org/10.1007/978-3-030-31167-4\\_19](https://doi.org/10.1007/978-3-030-31167-4_19).

797 Norby, R.J. and Zak, D.R. (2011) “Ecological Lessons from Free-Air CO<sub>2</sub> Enrichment (FACE)  
 798 Experiments,” *Annual Review of Ecology, Evolution, and Systematics*, 42(Volume 42, 2011), pp.  
 799 181–203. Available at: <https://doi.org/https://doi.org/10.1146/annurev-ecolsys-102209-144647>.

800 Palace, M.W. *et al.* (2017) “Ancient Amazonian populations left lasting impacts on forest  
 801 structure,” *Ecosphere*, 8(12), p. e02035. Available at:  
 802 <https://doi.org/https://doi.org/10.1002/ecs2.2035>.

803 Pärsinnen, M., et al. (2026) “Over 20,000 precolonial earthworks in the Southwest Amazonia,”  
 804 *Nature*, 657, pp. 420–425. Available at: <https://doi.org/10.1038/s41586-026-10835-7>.

805 Peripato, V. *et al.* (2023) “More than 10,000 pre-Columbian earthworks are still hidden  
 806 throughout Amazonia,” *Science*, 382(6666), pp. 103–109. Available at:  
 807 <https://doi.org/10.1126/science.ade2541>.

- Quesada, C.A. *et al.* (2010) “Variations in chemical and physical properties of Amazon forest soils in relation to their genesis,” *Biogeosciences*, 7(5), pp. 1515–1541. Available at: <https://doi.org/10.5194/bg-7-1515-2010>.
- Saatchi, S.S. *et al.* (2011) “Benchmark map of forest carbon stocks in tropical regions across three continents,” *Proceedings of the National Academy of Sciences*, 108(24), pp. 9899–9904. Available at: <https://doi.org/10.1073/pnas.1019576108>.
- Schimel, D.S. and Poulter, B. (2022) “The Earth in Living Color - NASA’s Surface Biology and Geology Designated Observable,” in *2022 IEEE Aerospace Conference (AERO)*, pp. 1–6. Available at: <https://doi.org/10.1109/AERO53065.2022.9843640>.
- Schmidt, M.J. *et al.* (2025) “Intentional creation of carbon-rich dark earth soils in the Amazon,” *Science Advances*, 9(38), p. eadh8499. Available at: <https://doi.org/10.1126/sciadv.adh8499>.
- Schroder, W. *et al.* (2020) “The lowland Maya settlement landscape: Environmental LiDAR and ecology,” *Journal of Archaeological Science: Reports*, 33, p. 102543. Available at: <https://doi.org/10.1016/j.jasrep.2020.102543>.
- Silva, L.C.R. *et al.* (2021) “A new hypothesis for the origin of Amazonian Dark Earths,” *Nature Communications*, 12(1), p. 127. Available at: <https://doi.org/10.1038/s41467-020-20184-2>.
- Söderström, M. *et al.* (2016) “Sensor mapping of Amazonian Dark Earths in deforested croplands,” *Geoderma*, 281, pp. 58–68. Available at: <https://doi.org/10.1016/j.geoderma.2016.06.024>.
- ter Steege, H. *et al.* (2006) “Continental-scale patterns of canopy tree composition and function across Amazonia,” *Nature*, 443(7110), pp. 444–447. Available at: <https://doi.org/10.1038/nature05134>.
- Walker, R.S. *et al.* (2023) “Predicting the geographic distribution of ancient Amazonian archaeological sites with machine learning,” *PeerJ*. Edited by M. Sponheimer, 11, p. e15137. Available at: <https://doi.org/10.7717/peerj.15137>.
- Watling, J., *et al.* (2017) “Impact of pre-Columbian “geoglyph” builders on Amazonian forests”, *Proceedings of the National Academy of Sciences of the U.S.A.* 114 (8) 1868-1873, Available at: <https://doi.org/10.1073/pnas.1614359114> (2017).
- WinklerPrins, A.M. and Aldrich, S.P. (2010) “Locating Amazonian Dark Earths: Creating an interactive GIS of known locations,” *Journal of Latin American Geography*, 9(3), pp. 33–50.