

# *Amazonian pollen assemblages reflect biogeographic gradients and forest cover*

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Blaus, A., McMichael, C. N. H. ORCID logoORCID: <https://orcid.org/0000-0002-1064-1499>, Raczka, Marco ORCID logoORCID: <https://orcid.org/0000-0002-6602-8087>, Herrick, C., Palace, M., Witteveen, N. H., Nascimento, M. N. and Bush, M. B. ORCID logoORCID: <https://orcid.org/0000-0001-6894-8613> (2023) Amazonian pollen assemblages reflect biogeographic gradients and forest cover. *Journal of Biogeography*. ISSN 1365-2699 doi: <https://doi.org/10.1111/jbi.14701> Available at <https://centaur.reading.ac.uk/112930/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1111/jbi.14701>

Publisher: Wiley

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online



# Amazonian pollen assemblages reflect biogeographic gradients and forest cover

A. Blaus<sup>1</sup> | C. N. H. McMichael<sup>2</sup>  | M. F. Raczka<sup>2,3</sup> | C. Herrick<sup>4</sup> | M. Palace<sup>4,5</sup> | N. H. Witteveen<sup>2</sup> | M. N. Nascimento<sup>2</sup> | M. B. Bush<sup>1</sup> 

<sup>1</sup>Institute for Global Ecology, Florida Institute of Technology, Melbourne, Florida, USA

<sup>2</sup>Department of Ecosystem and Landscape Dynamics, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands

<sup>3</sup>School of Archaeology, Geography and Environmental Science, University of Reading, Reading, UK

<sup>4</sup>Earth System Research Center, Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, New Hampshire, USA

<sup>5</sup>Department of Earth Sciences, College of Engineering and Physical Sciences, University of New Hampshire, Durham, New Hampshire, USA

## Correspondence

C. N. H. McMichael, Department of Ecosystem and Landscape Dynamics, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands.  
Email: [c.n.h.mcmichael@uva.nl](mailto:c.n.h.mcmichael@uva.nl)

## Funding information

European Research Council Starting, Grant/Award Number: StG 853394; National Science Foundation, Grant/Award Number: HEGS 2148984; NASA, Grant/Award Number: NASA-IDS 80NSSC20K1266; National Geographic Society, Grant/Award Number: #W393-15; The Netherlands Organisation for Scientific Research, Grant/Award Number: ALWOP.322

Handling Editor: Brian Huntley

## Abstract

**Aim:** Pollen assemblages are commonly used to reconstruct past climates yet have not yet been used to reconstruct past human activities, including deforestation. We aim to assess (i) how pollen assemblages vary across biogeographic and environmental gradients, (ii) the source area of pollen assemblages from lake sediment samples and (iii) which pollen taxa can best be used to quantify deforested landscapes.

**Location:** Amazonia.

**Taxon:** Plantae.

**Methods:** Pollen assemblages ( $N = 65$ ) from mud-water interface samples (representing modern conditions) of lake sediment cores were compared with modern gradients of temperature, precipitation and elevation. Pollen assemblages were also compared with local-scale estimates of forest cover at 1, 2, 5, 10, 20 and 40 km buffers around each lake.

**Results:** Over 250 pollen types were identified in the samples, and pollen assemblages were able to accurately differentiate biogeographic regions across the basin, corresponding with gradients in temperature and precipitation. Poaceae percentages were the best predictor of deforestation, and had a significant negative relationship with forest cover estimates. These relationships were strongest for the 1 km buffer area, weakening as buffer sizes increased.

**Main conclusions:** The diverse Amazonian pollen assemblages strongly reflect environmental gradients, and percentages of Poaceae best reflect local-scale variability in forest cover. Our results of modern pollen-landscape relationships can be used to provide a foundation for quantitative reconstructions of climate and deforestation in Amazonia.

## KEYWORDS

Amazon, forest cover, modern pollen, paleoecology, tropical forests

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Journal of Biogeography* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

The vastness of lowland Amazonia embraces a diversity of naturally-occurring habitats ranging from savannas to dense rainforests. Gradients of temperature, elevation, soil properties, precipitation and seasonality contribute to species distribution patterns, forest structural differences and propensity for fire (e.g. Gentry, 1988; Nepstad et al., 2004; ter Steege et al., 2006). This diversity of habitats has played a role in determining the distribution and density of past and present human occupation of the Amazon Basin, and also in the manner of land use (e.g. Bush et al., 2015; Neves & Petersen, 2006; Watling et al., 2015). For example, on the most nutrient-poor soils, people enriched the native clays with ash, faeces and silage to create Amazonian Dark Earths, but this practice was not followed on the more fertile soils emanating from the Andes (e.g. Glaser & Woods, 2004; Lehman et al., 2003). A commonality across Amazonia is that pre-Columbian settlers appeared to prefer close proximity to rivers or lakes as they offered additional food resources and, in the case of rivers, a means of transportation (e.g. Denevan, 1996).

In most settings, the characteristic signature of human occupation was forest clearance, commonly associated with fire or signals of cultivation. The scale of such disturbances in pre-Columbian Amazonia is actively debated, as is the detectability of these disturbances in paleoecological records (e.g., Clement et al., 2015; McMichael et al., 2012; Piperno et al., 2015). While identifying periods of human occupation in Amazonian forests can be relatively straightforward using archaeological or paleoecological approaches, quantifying past landscape modifications has proved challenging. Lake sediments can be particularly useful for quantifying aspects of past human activity as they offer the best possibility of obtaining an uninterrupted archive spanning millennia. Investigations of pre-Columbian land use and occupation patterns in Amazonia using lake sediments are routinely based on fossil pollen assemblages (Nascimento et al., 2022 and references within). The pollen assemblages are commonly analysed alongside charcoal fragments, which can be used to estimate the severity, frequency and intensity of fire events in Amazonian landscapes (Gosling et al., 2021; Power et al., 2008, 2013). Cultivation practices can also be identified (though not quantified) through the analysis of pollen grains or phytoliths found in lake sediments.

One parameter that has been lacking, however, is quantifying the extent of deforestation that occurred alongside occupation and cultivation. Using pollen assemblages to quantify past land cover has been successful in temperate ecosystems (Dawson et al., 2019; Gaillard, Sugita, Bunting, Dearing, et al., 2008; Gaillard, Sugita, Bunting, Middleton, et al., 2008; Liu et al., 2022; Prentice, 1988; Sugita, 1993, 2007; Trondman et al., 2015). The only attempt to quantify land cover in Amazonia was performed in the savanna-forest landscape of Bolivia, which revealed that deforestation was easier to detect in the closed-canopy rainforests than in the savannas (Whitney et al., 2019).

One of the problems encountered in quantitatively reconstructing aspects of climate or land cover is that for a given taxon (which

can be at the species, genus, or family level), pollen percentages do not reflect a 1:1 relationship with plant abundances in the landscape (Bush, 1991; Bush & Rivera, 1998, 2001; Gosling, 2004). In tropical settings such as Amazonia, many forest taxa produce very little pollen, or produce many pollen grains that cannot be identified with certainty (Bush et al., 2021; Bush & Weng, 2007). Most Amazonian tree species have hermaphroditic flowers (both stamens and anthers on the same flower), and these species tend to be severely under-represented in pollen assemblages though they may comprise the majority of aboveground biomass of the system (Bush, 1995). The pollen abundances of monoecious tree species (having both male and female flowers on the same individual plant) may be close to parity with their abundances in the landscape (Bush, 1995). Dioecious species have either male or female flowers on an individual plant, and as the pollen must travel some distances for pollination to occur, these species are likely to be over-represented in pollen assemblages. Many pioneer or early successional species in Amazonia are dioecious, and because of their over-representation, pollen assemblages may suggest a more disturbed landscape than is the reality (Bush & Rivera, 1998).

A further complication in translating pollen assemblages into total amounts of forest cover or relative amounts of vegetation types is that many pollen grains can only be identified at the family level, and some families include congeners that are shrubs, trees or vines, for example Combretaceae. Thus, natural biogeographic variation in species composition, with different balances of dioecy, monoecy, or hermaphroditism between habitats across Amazonian pollen assemblages, may indicate varying amounts of forest cover even when no differences exist. An alternative approach is to look at the relative abundances of pollen that are derived from non-tree taxa (e.g. grasses, the Poaceae family). Poaceae percentages within a pollen assemblage are often interpreted as indicators of disturbances or open landscapes within the closed-canopy forests of Amazonia (Bush, 2002). Poaceae pollen within closed-canopy forests, however, can also be derived from floodplains, marshes and floating aquatic grass mats and may not be a direct reflection of deforestation due to human activity (Bush, 2002; Colinvaux et al., 1985, 1988). Further, in the savannas located along the periphery of Amazonia, the grass is naturally abundant, and Poaceae pollen may not be able to detect forest openings caused by human activities or other processes (Whitney et al., 2019).

Quantifying the amount of deforested areas using pollen assemblages in Amazonia is also dependent on the source area of pollen (Liu et al., 2022). Studies in the forest-savanna ecotone areas of Bolivia suggested that anemophilous (wind-dispersed) pollen found in lake sediments may be derived from distances of over 40km (Whitney et al., 2019). In the closed-canopy forests that comprise most of the Amazon, pollen dispersal distances are believed to be smaller, limited by the high density of trees, low sub-canopy wind velocities, and predominance of entomophilous taxa (Bush et al., 2021; Gosling et al., 2009). There has been no assessment of the variability of the source area of pollen deposited into lake sediments across various regions of Amazonia.

Here we assess how pollen assemblages across Amazonia vary in aspects of biogeography, source area, forest cover and amounts of cleared landscapes. We use modern pollen assemblages from 65 Amazonian lakes to assess the variability in composition of Amazonian pollen assemblages and how they may be used to reconstruct past landscape deforestation. We specifically address the following questions: How do pollen assemblages derived from lake sediments vary across biogeographic regions in the Amazon? What spatial scale are pollen assemblages representing across the biogeographic regions of Amazonia? Which pollen taxa can be used to best quantify deforested landscapes?

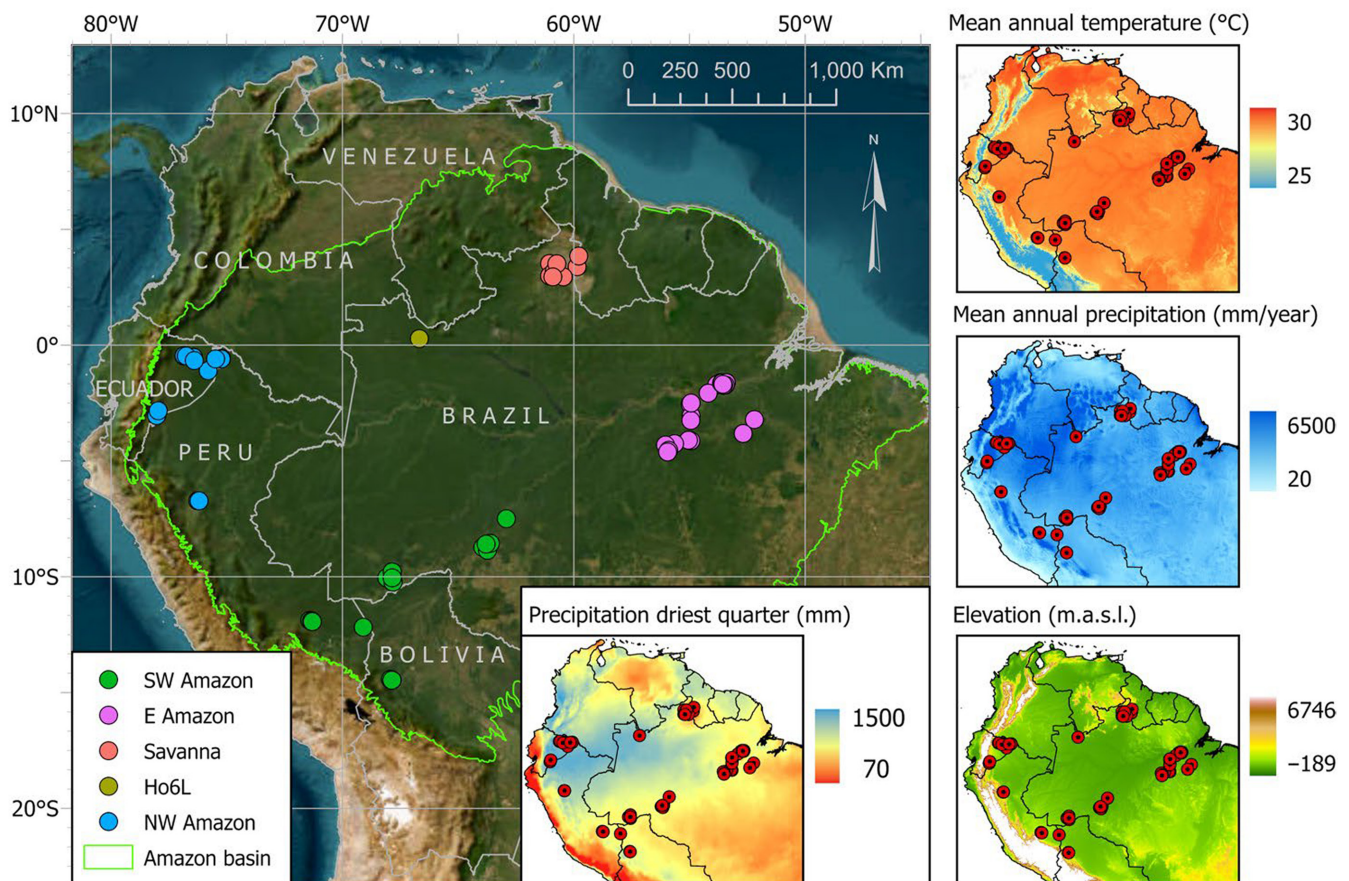
## 2 | MATERIALS AND METHODS

### 2.1 | Data collection and preparation

A total of 65 sites located within Amazonia sensu stricto (Eva et al., 2005) were used in our analysis (Figure 1, Table S1). Samples from new sites ( $n=44$ ) were collected with a Universal corer during 2018 and 2019 (Table S1). Samples from previously published sites

( $n=21$ ) were collected with a Colinvaux-Vohnout piston corer or a Universal corer (Table S1) (Bush et al., 2021). Our sites were located within evergreen tropical rainforests ( $n=56$ ) or savannas close to the ecotone with forests ( $n=9$ ; Figure 1a). Forty-four of the forested sites were situated within terra firme soils, and 21 on floodplains (Table S1). The sites included permanent lakes, oxbow lakes and flooded marshes in the Brazilian savannas (Table S1).

We divided our sites into three geographic regions (northwestern, southwestern and eastern) based on previous classifications on the basis of the biogeographical distinctness of vegetation (Gentry, 1988; ter Steege et al., 2000, 2013, 2016; Figure 1). We included additional classifications of 'savanna' and 'Hill of Six Lakes' (hereafter Ho6L) sites, which are also biogeographically distinct from the surrounding landscape (Figure 1). The 'savanna' sites of Roraima (Brazil) are located in non-forested savanna areas, and the Ho6L sites are located atop an inselberg in northwestern Amazonia (Figure 1). We characterized the climatic and topographic conditions of each site (Table S1). The climatic variables of mean annual temperature, mean annual precipitation and dry quarter precipitation were derived from CHELSA, a free high resolution (30 arc sec, ~1 km) global climate dataset (Karger et al., 2017). Averages were derived



**FIGURE 1** Map of 65 lakes within Amazonia sensu stricto (Eva et al., 2005, green outline) with modern pollen assemblages used in the analysis. Sites are colour-coded based on assigned biogeographic regions. Inset maps show environmental gradients of mean annual temperature, mean annual precipitation, dry quarter precipitation and elevation within Amazonia (see Table S1 for values for individual sites). All maps were generated using ArcGIS Pro 3.0 and the environmental data were obtained from CHELSA V 2.1 (Karger et al., 2017) or Shuttle Radar Topography Mission digital elevation models (Jarvis et al., 2008).

from climatological data obtained from 1981 to 2010. Elevation data were obtained from Shuttle Radar Topography Mission S hole-filled ~1 km resolution digital elevation models (Jarvis et al., 2008). Climate and topographic data for a 40 km radius around each site were downloaded, and the average values from the grid cells falling within each of the buffers were calculated (see below for description of buffers).

Elevation ranged from 10 to 700 m a.s.l. across all sites, with the western and Ho6L regions bordering the Andes primarily containing sites with higher elevations and the eastern region containing sites at the lower elevation range. Mean annual temperatures were highest in the eastern and savanna sites (27°C) and decreased toward the northwestern sites (22–25°C; Figure 1; Karger et al., 2017). Annual precipitation ranged from 1600 to 3300 mm, with the northwestern sites receiving the highest amounts of precipitation and the savanna sites receiving the lowest (Figure 1). Seasonal precipitation (precipitation of the driest quarter) ranged from 138 to 923 (mm) and was highest in the northwestern region but lower in the eastern and savanna regions (Figure 1).

From each of the 65 sites, we collected the upper 1 cm of sediment (mud-water interface) for pollen analysis. Pollen samples were prepared in the Palaeoecological Laboratory at the Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Netherlands or in the Neotropical Paleoecology Laboratory at Florida Institute of Technology, USA, and followed standard procedures as documented in Faegri and Iversen (1989). For each sample, 300 to 500 pollen grains were counted to generate a pollen assemblage, and tablets of *Lycopodium clavatum* spores (Batch no: 483216, 18,583 spores per tablet) were added prior to sample preparation to allow for calculations of pollen concentration (Stockmarr, 1972). Pollen identification of the new samples was at a finer taxonomic resolution than the previously published samples, and data were harmonized (e.g. merged to a coarser taxonomic resolution) prior to analysis (e.g. Bush et al., 2021).

To assess the spatial scale of pollen representation, we generated a series of buffers around the perimeter of each of the 65 lakes at distances of 1, 2, 5, 10, 20 and 40 km using Google Earth Engine (Gorelick et al., 2017; Figure S1). We used the Global Forest Change v1.8 database, which is publicly available at 30 m spatial resolution and generated from Landsat satellite imagery, to calculate forest cover for each buffer around each site (Hansen et al., 2013). Within the database, forest cover is expressed as percentages per pixel, where 0 values represent no forest cover within the pixel, and 100 represents a completely forested pixel. We considered pixels containing >50% of forest cover as forested. The number of forested pixels (number of pixels where forest canopy height exceeds 5 m) for each buffer was calculated by taking the number of forested pixels present in the year AD 2000 and subtracting the number of forest pixels lost each year until the time of sample collection. For example, if the lake sediment sample was collected in AD 2018, the number of forested pixels was the number of forested pixels present in the AD 2000 minus the number of forested pixels lost each year until AD 2018 (Figure S1). JRC Global Surface Water Mapping Layers, v1.3

(Pekel et al., 2016) was used to delineate water body boundaries and to calculate the number of water-based pixels within each buffer (Figure S1). The final metric of relative forest cover was a proportion of the total number of forested pixels out of the total number of non-water pixels for each buffer (Figure S1).

## 2.2 | Data analysis

All pollen counts were converted to percentage data, and all analyses were run on pollen percentages. We performed Detrended Correspondence Analysis (DCA) to assess (dis)similarity between samples within and between the biogeographic regions. We also fitted smooth surfaces of elevation, mean annual temperature, annual precipitation and precipitation seasonality on DCA site score ordinations (Oksanen et al., 2013) to assess (dis)similarity among sites along the environmental gradients. All ordinations were performed using pollen taxa that exceeded 5% in at least one sample. We used this threshold to reduce the possibility that rare taxa (which are also potentially misidentified) would drive the sample differences (Correa-Metrio et al., 2014). To reduce the number of taxa shown on the plot (overcrowding), only taxa with a significant correlation ( $p < 0.05$ ) with either of first two axes were shown on DCA outputs.

For each pollen type identified, we categorized the growth form and pollination strategy based on expert knowledge and published literature (Bush, 1995; Ollerton et al., 2011; Renner & Feil, 1993) (Table S2) and used these to form pollen taxa groups (i.e., grasses, herbaceous, woody, wind pollinated, animal pollinated, etc.). We assessed how the percentages of specific pollen taxa (Poaceae, *Cecropia* and Moraceae/Urticaceae) or percentages of the pollen taxa groups (Table S2) could be used to predict relative amounts of forest cover. Poaceae (the grasses) are typically indicative of open areas, *Cecropia* is a light-loving disturbance tree that is indicative of early successional regrowth, and the Moraceae/Urticaceae families, which are not distinguishable using pollen, is a group composed primarily of Amazonian trees. The pollen taxa groups were either related to forest structure (e.g. herbaceous, woody and trees) or the pollination syndrome (i.e., anemophilous and entomophilous), as both could be linked to amounts of forest cover in the surrounding landscape. We used a beta regression model (Cribari-Neto & Zeileis, 2010), which was developed for response variables that are proportions or rates, to determine how well the pollen taxa and taxa groups could predict the proportion of forest cover pixels in the landscape. Lake size can have a strong influence on pollen source area (Jacobson & Bradshaw, 1981; Prentice, 1988); therefore, we also ran models with lake area (km<sup>2</sup>) as a secondary predictor variable to determine if size affected the apparent source area of pollen. We ran models for each of the buffers around the lakes where forest cover was measured (1, 2, 5, 10, 20 and 40 km). Model performance was assessed using pseudo- $R^2$  values and  $p$ -values of the beta regression coefficients for the predictor variables (Cribari-Neto & Zeileis, 2010). Pseudo- $R^2$  values for the beta regression model are squared correlation values of the linear predictor and link-transformed response (in this case, logit). To show

proof of concept, we applied our beta regression model to previously published pollen data from Lake Kumpak<sup>a</sup> in Ecuador (Åkesson et al., 2021). The superscripted “a” in Kumpak<sup>a</sup> reflects pronunciation of the lake name in the Shuar language (Colinvaux et al., 1985).

Data preparation and handling were conducted using Google Earth Engine (Gorelick et al., 2017) and ESRI ArcGIS Pro Version 3.1.1 software. Statistical analyses were performed in R (R Development Core Team, 2017) using the ‘rioja’ (Juggins, 2020), ‘betareg’ (Cribari-Neto & Zeileis, 2010) and ‘vegan’ (Oksanen et al., 2013) packages, and visualizations were made using the ‘ggplot2’ (Wickham, 2011) package.

### 3 | RESULTS

Pollen assemblages, based on 252 pollen types found across all sites, showed high taxonomic diversity and variation both within and among the biogeographic regions (Figure 2). The northwestern, southwestern and Ho6L sites generally contained higher abundances of Moraceae/Urticaceae pollen (10%–60%) compared with the eastern and savanna sites (Figure 2). The dry forest shrub *Byrsonima* was present in higher abundances (5%–35%) in the savanna and eastern sites compared with the other regions where it rarely exceeded 5%. *Alchornea*, a tree, was present in all regions, but highest in the Ho6L sites (Figure 2). Percentages of the hardwood palm *Iriartea deltoidea* were higher in the northwestern and southwestern sites (5%–20%) compared with other regions, matching its known geographic distribution. *Euterpe*, a genus of palm that includes *E. precatoria*, the most common tree in the Amazon, had higher pollen abundances (5%–25%) in the eastern sites compared with other regions (Figure 2). Percentages of the weedy herbaceous taxon *Ambrosia* and the sedges (Cyperaceae family) were generally higher in the eastern and savanna sites compared with the other regions. All regions contained grasses (Poaceae), though abundances were highest in the savanna region (25%–68%) and ranged from 0%–40% in the other regions (Figure 2). Some taxa were present in all regions but had highly variable percentages within the region. For example, *Cecropia*, a genus of fast-growing, disturbance-oriented trees, was present in all regions, but within a region ranged from 0% to >30%. The palms *Mauritia flexuosa* and *Mauritiella* sp., which only grow in swampy conditions, were also present in all regions but were absent from many individual sites (Figure 2).

The DCA analysis provided some separation according to the biogeographic affiliation of the 65 sites (Figure 3a). The northwestern and southwestern sites had vegetation that was more similar to each other and was located on the positive end of Axis 1, whereas the eastern Amazonian and savanna sites were overlapping and both were located on the negative end of Axis 1 (Figure 3a). In total, the 28 taxa used in this analysis were able to significantly separate the western sites from the eastern sites on DCA Axis 1.

Forest cover values ranged from <1% to 99% within each of the buffer sizes (1, 2, 5, 10, 20 and 40 km) around the 65 lakes (Figure S2). Generally, sites within the NW Amazon and Ho6L regions had the highest percentages of surrounding forest cover (75%–99%) and the E Amazon sites typically had lower percentages (30%–78%) (Figure S2).

The savanna sites had the lowest forest cover values, ranging from <1% to 26%. The SW Amazon sites exhibited the most variability, where forest cover percentages ranged from 30% to 90%, depending on the size of the buffer. The percentages of forest cover around each lake, and thus site rankings, varied depending on the size of the buffer (Figure S2). Forest cover values around each lake were generally positively correlated, though some spatial heterogeneity existed in NW Amazonia, E Amazonia and the savanna sites (Figure S3).

The beta regressions indicated that Poaceae percentage was the best predictor of forest cover, based on the pseudo- $R^2$  values, model coefficients and  $p$ -values of model predictors (Table 1, Table S3). Poaceae percentages were able to predict forest cover at all buffer sizes (1, 2, 5, 10, 20 and 40 km), but model performance decreased as buffer size increased, and performance was highest for the 1 km buffer (Figure 4, Table 1). *Cecropia* was also able to predict forest cover, though with lower pseudo- $R^2$  values and only for the 1, 2, 5 and 10 km buffers (Table 1). Poaceae plus *Cecropia* values (Poaceae + *Cecropia*) were able to predict forest cover at all buffer sizes better than *Cecropia* alone but not as well as with Poaceae alone (Table 1). Pseudo- $R^2$  values for all significant models (Poaceae, *Cecropia*, Poaceae + *Cecropia*) increased when lake size was included as a secondary predictor, though lake size itself was not a significant predictor in any models (Table 1). All other pollen groupings based on pollination syndrome or ecological characteristics (Table S2) performed worse than Poaceae, *Cecropia*, or Poaceae + *Cecropia* percentages (Table 1, Table S3).

### 4 | DISCUSSION

Our analysis of 65 mud-water interface samples collected across a broad range of Amazonian habitats provides novel insights into the biogeographic and local-scale variability of pollen percentages across this heterogeneous landscape. To date, most quantitative studies of modern pollen representation in Amazonia have been on relatively local scales comparing sites within 10s of km (Bush, 2002; Gosling et al., 2009; Whitney et al., 2019). The greater spatial scale of this study, across >2000 km, embraces long environmental gradients (Figure 1, Figures S2 and S3). Whereas the beta diversity of habitats in local analyses for example the difference between igapo and terra firme forest in pollen traps (Bush, 2002; Gosling et al., 2009) was readily identifiable, the larger catchment area afforded by lake sediments provides a more blended signal. In this analysis, we use the commonest pollen types to attempt to determine gamma diversity and relate those changes to long environmental gradients.

#### 4.1 | Broad biogeographic patterns

Floristic studies show that environmental gradients across Amazonia are reflected in altered forest composition resulting in biogeographic sub-regions (Gentry, 1988; ter Steege et al., 2006). These evaluations are based on species-level turnover and changes

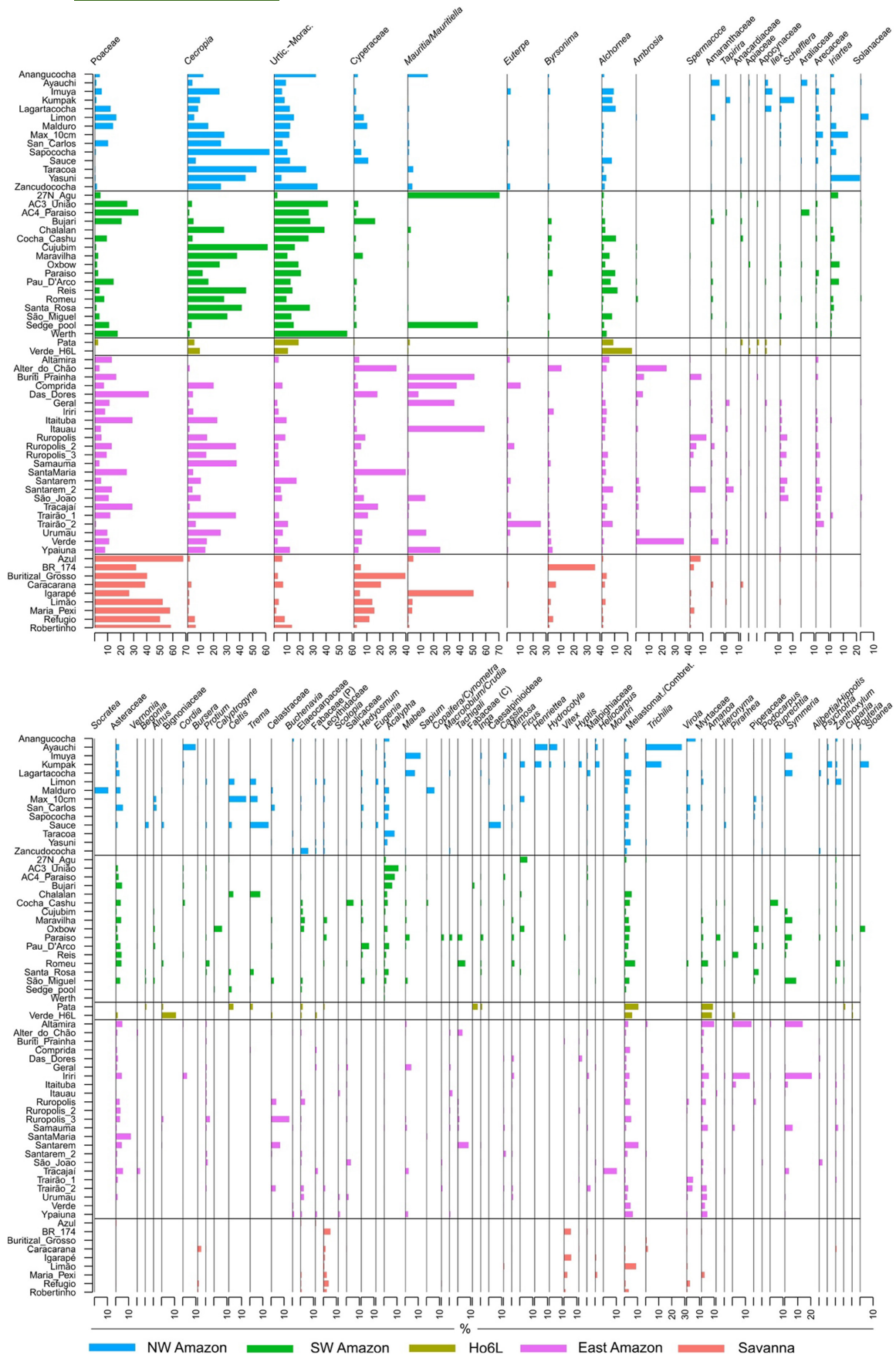


FIGURE 2 Pollen percentages of taxa reaching at least 5% abundance from 65 lakes in Amazonia. These taxa were included in the Detrended Correspondence Analysis (see Figure 3). Sites are colour-coded based on assigned biogeographical region.



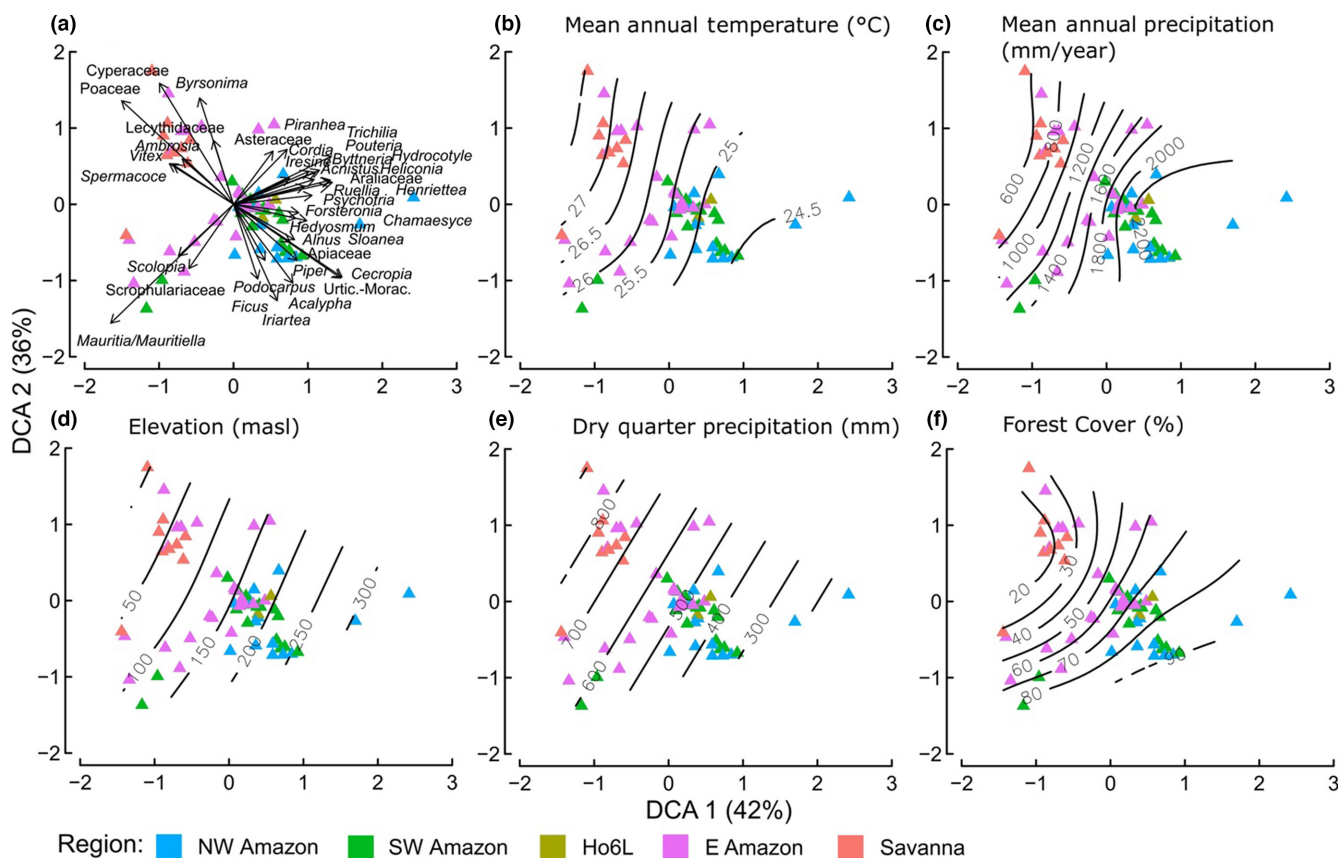
in stem density within forest plots (Gentry, 1988; Ruokolainen et al., 2019; ter Steege et al., 2006; Tuomisto et al., 2016). While we are usually unable to identify pollen to the species or genus level, we find that when using the 30 commonest pollen types, the biogeographic sub-regions can still be differentiated (Figures 2 and 3).

In many DCA analyses the Axis 1 versus Axis 2 site scores form a diamond shape (Gauch, 1982), but in practice often form a triangle. Somewhat unusually, in our plot of the DCA Axis 1 versus Axis 2 results, the species scores form a diagonal with forest taxa. Taxa such as Urticaceae/Moraceae, *Cordia*, *Ficus*, *Acalypha* and *Iriartea*, plot at the lower right, while indicators of openness, for example Poaceae, Cyperaceae and *Byrsonima*, plot at the upper left (Figure 3a). We explain this diagonal trend by observing that the large environmental gradients are somewhat, but not completely, autocorrelated. Due to this partial autocorrelation there was enough variance for the first two axes of the DCA to cause this diagonal pattern (Figure 3b–f). Consequently, we find that the site scores plot from the hottest, lowest, driest, most seasonal savannas at the top left, that is low scores on Axis 1 and high scores on Axis 2, to the coolest, highest, wettest and least seasonal forests plot at the bottom right with high scores on Axis 1 and low scores on Axis 2 (Figure 3b–f).

Our data reflect known distributional patterns. Poaceae abundances track the gradient of increasing precipitation and seasonality across the savanna, cerrado, dry forest, rainforest continuum (Figures 2 and 4; Flenley, 1979; Gosling et al., 2009; Rodgers & Horn, 1996). Poaceae pollen percentages are highest in the savanna region, where forest cover is lowest (Figure 2, Figure S2). Similarly, the western Amazonian forests have higher stem densities of *Cordia*, *Ficus*, *Acalypha* and *Iriartea*, than eastern forests (ter Steege et al., 2013, 2016), and this is reflected in the pollen analysis (Figure 2). The pollen of taxa derived from the eastern flank of the Andes mountains, such as *Podocarpus*, *Alnus* and *Hedyosmum*, is also more abundant in the western Amazonian sites. Collectively, these taxa could be used to differentiate western Amazonian sites from those in the Ho6L, eastern Amazon, or savanna sites (Figure 3a).

## 4.2 | Local gradients of forest cover

Detecting disturbance in paleoecological records is important for assessing the impacts of climate change or human impacts on forests (e.g. Davis, 1981; Fyfe et al., 2015). A key metric of disturbance is the amount of openness in a forest or its reciprocal, percent forest



**FIGURE 3** Detrended Correspondence Analysis showing site-taxa relationships and site-environment relationships for pollen assemblages analysed from 65 lakes across Amazonia. Sites are colour-coded by assigned biogeographic regions. (a) Site-taxa relationships where vectors indicate taxa that were significant ( $p < 0.05$ ) in separating sites from each other; (b–f) Site (dis)similarity across environmental gradients of: mean annual temperature (b), annual precipitation (c), elevation (d), dry quarter precipitation (e) and forest cover (f) for the 2 km buffer.

TABLE 1 Summary of beta regression results using groups of pollen taxa to predict forest cover values at 1, 2, 5, 10, 20 and 40 km buffers around Amazonian lakes (N = 65). For specifics on pollen groupings, see Table S2. For the full set of results showing all pollen groups, see Table S3.

Pollen types or groups (predictor)	Buffer size (km)	Model using only pollen			Model using pollen + lake size			p-value (pollen)	p-value (lake size)
		Coefficient (pollen)	Pseudo-R <sup>2</sup>	p-value	Coefficient (pollen)	Coefficient (lake size)	Pseudo-R <sup>2</sup>		
Poaceae	1	-0.069	0.610	<0.001	-0.071	-0.055	0.620	<0.001	0.173
Poaceae	2	-0.068	0.579	<0.001	-0.070	-0.050	0.589	<0.001	0.213
Poaceae	5	-0.068	0.543	<0.001	-0.070	-0.046	0.552	<0.001	0.231
Poaceae	10	-0.070	0.554	<0.001	-0.072	-0.053	0.565	<0.001	0.169
Poaceae	20	-0.068	0.538	<0.001	-0.070	-0.047	0.549	<0.001	0.229
Poaceae	40	-0.062	0.505	<0.001	-0.064	-0.049	0.517	<0.001	0.198
<i>Cecropia</i>	1	0.027	0.122	<b>0.002</b>	0.029	-0.040	0.131	<b>0.002</b>	0.467
<i>Cecropia</i>	2	0.024	0.099	<b>0.007</b>	0.026	-0.034	0.105	<b>0.006</b>	0.531
<i>Cecropia</i>	5	0.020	0.075	<b>0.020</b>	0.022	-0.030	0.080	<b>0.016</b>	0.570
<i>Cecropia</i>	10	0.019	0.067	<b>0.031</b>	0.021	-0.032	0.073	<b>0.025</b>	0.552
<i>Cecropia</i>	20	0.016	0.053	0.060	0.018	-0.026	0.056	0.051	0.627
<i>Cecropia</i>	40	0.015	0.042	0.084	0.016	-0.028	0.047	0.068	0.588
Poaceae + <i>Cecropia</i>	1	-0.031	0.180	<0.001	-0.032	0.042	0.188	<0.001	0.417
Poaceae + <i>Cecropia</i>	2	-0.032	0.193	<0.001	-0.033	0.042	0.201	<0.001	0.404
Poaceae + <i>Cecropia</i>	5	-0.032	0.206	<0.001	-0.033	0.043	0.214	<0.001	0.380
Poaceae + <i>Cecropia</i>	10	-0.033	0.226	<0.001	-0.034	0.041	0.233	<0.001	0.399
Poaceae + <i>Cecropia</i>	20	-0.034	0.244	<0.001	-0.036	0.047	0.251	<0.001	0.342
Poaceae + <i>Cecropia</i>	40	-0.034	0.243	<0.001	-0.035	0.041	0.249	<0.001	0.375

Bold p-values indicate significance of predictors used in beta regressions.

cover (Broström et al., 1998; Fyfe et al., 2015). It might be thought that simply adding all 'forest pollen' taxa together would provide the best proxy for forest cover. There is considerable uncertainty, however, as to the correct classification of pollen types that could be large trees or vines (e.g., Combretaceae), or herbs, shrubs or vines (e.g., Asteraceae). Consequently, the 'forest pollen' group is not significant in predicting forest cover. Furthermore, many forest taxa are silent in the pollen record because of their reproductive mechanisms (Bush, 1995), whereas others, such as *Acalypha* and *Cecropia* are strongly over-represented (Bush & Rivera, 1998; Bush & Rivera, 2001). Weighting for the area and the perimeter of the lake in case pollen taxa were growing on the shoreline does not improve correlations. Of the many combinations of taxa, Poaceae percentages are the best predictor of forest cover in our samples, with an inverse relationship, regardless of the scale of examination (Table 1, Figure 4). Our data highlight that floating grass mats found on many Amazonian lakes do not contribute significantly to increased Poaceae percentages even when deforestation is minimal (e.g. Colinvaux et al., 1985; Colinvaux et al., 1988). As grasses occasionally grow around the edges of the lake, our results also show that this is not a determining input of Poaceae pollen into our sites.

We found that as Poaceae percentages increased within the dense closed-canopy forests of the non-savanna sites, forest cover in the surrounding landscape decreased significantly (Figure 4). These results suggested that the relationships between forest cover

and Poaceae are stronger in the closed-canopy forest than in the savanna. The uppermost 1–2 cm of sediment contains pollen deposited over multiple years, and the forest cover estimates covered approximately 10-years (Hansen et al., 2013). Most of our samples were collected in 2018 and 2019, and forest opening was calculated collectively since the onset of forest cover database in 2010 until the time of sample collection. It is thus possible that our analysis was affected by the timing of deforestation at the sites. For example, a site that has been continuously deforested for the last 10 years may have a stronger signal of forest opening than a site which has only had some opening in one out of the 10 years.

Whitney et al. (2019) found that in a forest-savanna ecotone, small openings in dense forests were detectable by small increases in Poaceae, but in seasonally flooded savanna landscapes that were more open the Poaceae representation was not negatively proportional to forest cover. Our findings are consistent with those results. The savanna sites in our study exhibited a wide range of Poaceae percentages (5%–68%) alongside little variation in forest cover (0%–22%; Figure 4). Unlike the Whitney et al. (2019) study, where Poaceae representation saturated in landscapes with >5% forest openings, we found a more persistent relationship. One explanation may be that our sites were small, closed-basin lakes (especially in the savanna settings), whereas the sample sites exhibiting this lack of sensitivity used by Whitney et al. had more seasonal flooding or were much larger waterbodies.

The source area from which the Poaceae were most indicative of forest openings is also of great interest as there is the perennial question in palynology, do signals reflect small scale changes near the sampling point or large-scale changes at distance (e.g. Bunting & Middleton, 2009; Hellman et al., 2009)? Whitney et al. (2019) found that samples collected within forests were sensitive to small scale clearance events close to the sampling point, but for sample points that lay outside forest in Bolivian savannas, the source area of pollen could be derived from up to 40km. Our data were not inconsistent with these findings, though we can refine this understanding. We found that the Poaceae percentages from sites collected in a forested setting best reflected forest cover within a 1–2km radius of the sampled sites (Table 1, S3). The best fits between pollen percentages that had significant model coefficients (Poaceae, *Cecropia* and Poaceae+*Cecropia*) and forest cover were all strongest within the 1–2km buffer (Table 1, Table S3). Our findings likely differed from Whitney et al. (2019) because most sites were in predominantly closed-canopy rainforests that were not near forest-savanna boundaries, apart from the nine savanna sites (Figure 1). These results provide increased confidence that the Poaceae percentages can reflect small canopy openings at distances of 1–2km from the lake edge, increasing the detectability of small-scale land use in paleoecological contexts.

Our findings correspond with the fact that in closed-canopy forests, up to 98% of taxa have pollen that is dispersed by animals (i.e.

insects, birds or bats; Bush & Rivera, 2001). The potential for long distance dispersal of pollen released within 1–2m of the ground, that is grasses and herbs, is greater in savanna systems, where wind speeds are higher, there are fewer tree trunks to intercept pollen and convective heating is more likely to reach the herb layer and generate more turbulence, thereby helping pollen to disperse longer distances (Hoffmann et al., 2012; Tauber, 1977). Our dataset did not contain enough sites to model savanna separately from forested settings, but we consider it likely that the wind-dispersed pollen types would have shown a stronger relationship with the larger-sized buffers in our savanna sites, and that entomophilous taxa would have had a stronger relationship with the smaller buffers.

### 4.3 | An applied case study

Even within regions, for any given amount of forest cover there is considerable variability in Poaceae pollen representation. Local factors, that is basin shape, the presence of input streams, or fringing marsh extent, provide enough variability that we do not recommend taking an observed value of Poaceae and translating that into forest cover via the regression line without taking note of the confidence intervals. For example, Lake Kumpak<sup>a</sup> had two overflight photographs taken during coring operations in 1983 and 2014 and the imagery revealed ca. 85% forest cover within 1 km of the lake during both flights

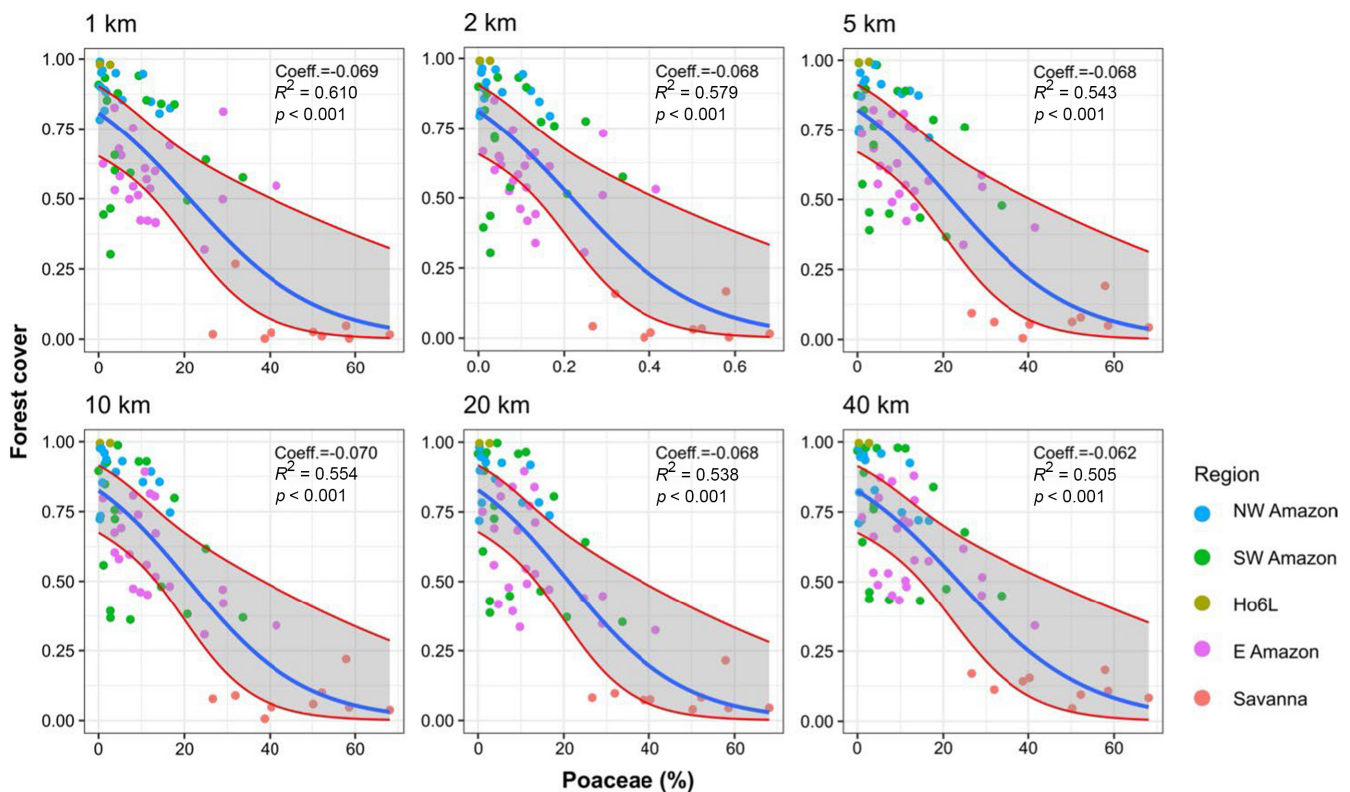
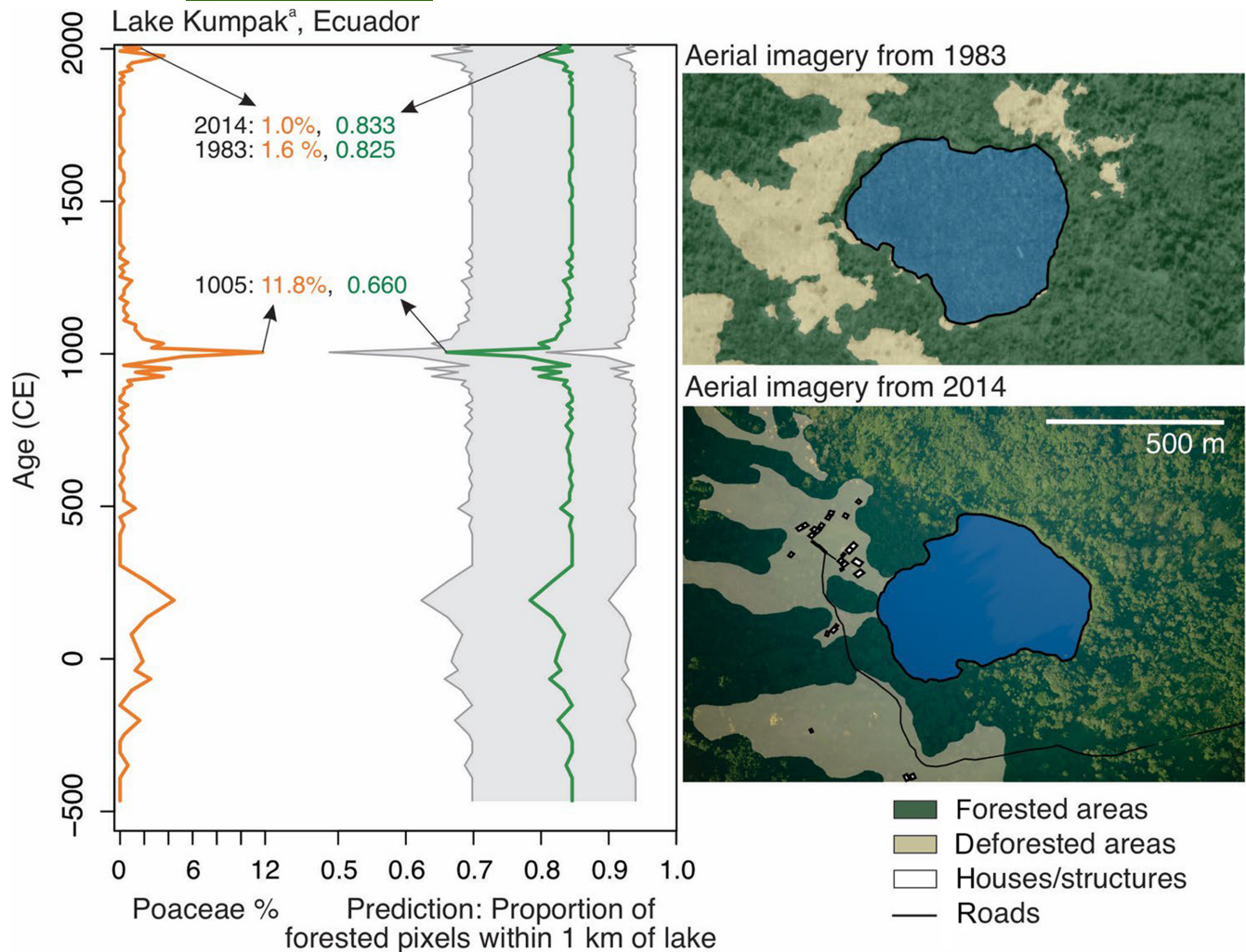


FIGURE 4 Beta regression models using Poaceae percentages to predict forest cover values (the proportion of forest pixels) calculated at 1, 2, 5, 10, 20 and 40km buffers around the 65 Amazonian lakes. Sites are colour coded by assigned biogeographic regions. Pseudo-R<sup>2</sup> values, model coefficients, and p-values are shown for each regression.



**FIGURE 5** The fossil Poaceae record from Lake Kumpak<sup>a</sup>, Ecuador, over the last 2200 years (orange line) alongside the aerial imagery collected in 1983 and 2014 (data from Åkesson et al., 2021). Predictions from the beta regression generated in this study show the likely amount of forest cover during the last 2200 years at the lake (green line with confidence intervals). Arrows on the Poaceae pollen curve (orange) and predicted forest cover (green) indicate the peak of human disturbance at c. 1005 CE and the years when overflight images are available (1983 and 2014).

(Figure 5; Åkesson et al., 2021). Pollen analysis revealed the Poaceae pollen percentages were functionally the same in 1983 and 2014 (i.e. 1.6% and 1.0%, respectively; Figure 5; Åkesson et al., 2021). These data give us confidence in our beta regression modelling results—that Poaceae representation is correlated with forest cover. Our regression predicts that 1.6% and 1% of Poaceae pollen represents 82.5%–83% forest cover within 1 km of the lake (Figure 5). Based on our regression, the least forest cover of the last 2000 years, occurred c. 1005 CE, corresponding to c. 66% forest cover within a kilometre of the lake (potential range 48%–80%; Figure 5).

Regardless of these uncertainties, our results show promising relationships between pollen percentages, particularly Poaceae, and local forest cover. We provide a dataset that can be used to move forward with quantitative reconstructions of forest cover in Amazonian paleoecological records using Poaceae pollen, which readily preserves in lake sediments over thousands of years and is easily identifiable even with novice pollen analysts. We

hypothesize that more nuanced estimates of forest cover and forest type can be estimated if classifications of forest type (using different remote sensing products or areal imagery) are matched to pollen assemblages across our sites (and other sites). Quantitative estimates of forest cover and forest type in a paleoecological context are important for addressing research questions regarding the extent, duration and frequency of pre- and post-Colonial deforestation in Amazonia.

#### 4.4 | The path forward

We can identify limitations within our study and offer some thoughts on profitable directions for future research. One limitation of our dataset was that the distribution of forest cover values for our sites was skewed toward high values of forest cover (Figure S2). The total amount of forest cover surrounding our sites was locally variable

at various buffer sizes, although it was generally high everywhere but in the savanna sites (Figures S2 and S3). Accordingly, Poaceae pollen comprised <10% of the total pollen sum at more than half of the sites used in this study (Figure 2). Moderate values of forest cover and Poaceae percentages were sparse in our dataset, and filling this gap in the experimental design would have likely increased the strength of the relationships that we found between forest cover and Poaceae.

With our current capabilities, grouping pollen into categories such as 'woody', 'herbaceous', 'wind-pollinated', etc. (Table S2) did not provide better predictions of forest cover in the landscapes surrounding our sites (Table 1). As pollen taxonomy improves, whether through emerging artificial intelligence technologies or improved reference collections, there will be increased confidence in our ability to group pollen types into such categories (Table S2). Another possibly interacting factor is the coarseness of the forest cover dataset used in the analysis (Hansen et al., 2013), which is at 30 m spatial resolution but, perhaps more critically, is ecologically simplistic. The forest cover database classifies all forest >5 m in height simply as 'forested' and does not account for the variability in forest structure and forest type. The integration of paleoecological data with: (i) multispectral remote sensing products, for example Sentinel or Landsat, (ii) light detection and ranging imagery, or (iii) unpiloted aerial systems (or drone) imagery, offers exciting possibilities for calibrating forest type and structure with pollen and other microfossil representation (Khan et al., 2017; Witteveen et al., 2023).

## ACKNOWLEDGEMENTS

This work was supported by The Netherlands Organisation for Scientific Research (NWO) award ALWOP.322, European Research Council Starting Grant StG 853394, and a National Geographic Expedition Award to C.N.H.M. This work was also funded by National Science Foundation (HEGS 2148984) to M.B.B., NASA (NASA-IDS 80NSSC20K1266) to M.P. and M.B.B., and National Geographic (NatGeo Grant #W393-15) to M.N.N. We thank Rudney Almeida dos Santos, Luiz C. Lobato, Yoshi Maezumi, Luciane Fontana and everyone else directly and indirectly involved in the fieldwork, Kees Nooren for assistance with pollen analysis, and Christine Åkesson for providing aerial imagery of Lake Kumpak<sup>a</sup>. Ecuadorian samples were collected under permits 08-2017-IC-FLO-DNB/MA and 18-2018-IC-FLO-DNB/MA. Peruvian and Bolivian samples were previously published. Brazilian citizens (M.F.R. and M.N.N.) collected samples in Brazil for their postdoctoral research so no permits were required.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

## DATA AVAILABILITY STATEMENT

All pollen data are uploaded to Data Dryad ([datadryad.org](https://datadryad.org), <https://doi.org/10.5061/dryad.tb2rbp05n>) and will also be uploaded to the Neotoma pollen database ([www.neotomadb.org](http://www.neotomadb.org)).

## ORCID

C. N. H. McMichael  <https://orcid.org/0000-0002-1064-1499>

M. B. Bush  <https://orcid.org/0000-0001-6894-8613>

## REFERENCES

- Åkesson, C. M., McMichael, C. N., Raczka, M. F., Huisman, S. N., Palmeira, M., Vogel, J., Neill, D., Veizaj, J., & Bush, M. B. (2021). Long-term ecological legacies in western Amazonia. *Journal of Ecology*, 109, 432–446.
- Broström, A., Gaillard, M.-J., Ihse, M., & Odgaard, B. (1998). Pollen-landscape relationships in modern analogues of ancient cultural landscapes in southern Sweden—A first step towards quantification of vegetation openness in the past. *Vegetation History and Archaeobotany*, 7, 189–201.
- Bunting, M. J., & Middleton, R. (2009). Equifinality and uncertainty in the interpretation of pollen data: The multiple scenario approach to reconstruction of past vegetation mosaics. *The Holocene*, 19, 799–803.
- Bush, M. B. (1991). Modern pollen-rain data from south and Central America: A test of the feasibility of fine resolution lowland tropical palynology. *The Holocene*, 1, 162–167.
- Bush, M. B. (1995). Neotropical plant reproductive strategies and fossil pollen representation. *The American Naturalist*, 145, 594–609.
- Bush, M. B. (2002). On the interpretation of fossil Poaceae pollen in the humid lowland neotropics. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 177, 5–17.
- Bush, M. B., Correa-Metrio, A., van Woesik, R., Collins, A., Hanselman, J., Martinez, P., & McMichael, C. N. (2021). Modern pollen assemblages of the Neotropics. *Journal of Biogeography*, 48, 231–241.
- Bush, M. B., McMichael, C. H., Piperno, D. R., Silman, M. R., Barlow, J. B., Peres, C. A., Power, M. J., & Palace, M. W. (2015). Anthropogenic influence on Amazonian forests in prehistory: An ecological perspective. *Journal of Biogeography*, 42, 2277–2288.
- Bush, M. B., & Rivera, R. (1998). Pollen dispersal and representation in a neotropical rain forest. *Global Ecology and Biogeography Letters*, 7, 379–392.
- Bush, M. B., & Rivera, R. (2001). Reproductive ecology and pollen representation among neotropical trees. *Global Ecology and Biogeography*, 10, 359–368.
- Bush, M. B., & Weng, C. (2007). Introducing a new (freeware) tool for palynology. *Journal of Biogeography*, 34, 377–380.
- Clement, C. R., Denevan, W. M., Heckenberger, M. J., Junqueira, A. B., Neves, E. G., Teixeira, W. G., & Woods, W. I. (2015). The domestication of Amazonia before European conquest. *Proceedings of the Royal Society B: Biological Sciences*, 282, 20150813.
- Colinvaux, P. A., Frost, M., Frost, I., Liu, K.-b., & Steinitz-Kannan, M. (1988). Three pollen diagrams of forest disturbance in the western Amazon basin. *Review of Palaeobotany and Palynology*, 55, 73–81.
- Colinvaux, P. A., Miller, M. C., Liu, K.-B., Steinitz-Kannan, M., & Frost, I. (1985). Discovery of permanent Amazon lakes and hydraulic disturbance in the upper Amazon basin. *Nature*, 313, 42–45.
- Correa-Metrio, A., Dechnik, Y., Lozano-García, S., & Caballero, M. (2014). Detrended correspondence analysis: A useful tool to quantify ecological changes from fossil data sets. *Boletín de la Sociedad Geológica Mexicana*, 66, 135–143.
- Cribari-Neto, F., & Zeileis, A. (2010). Beta regression in R. *Journal of Statistical Software*, 34, 1–24.
- Davis, M. B. (1981). Quaternary history and the stability of forest communities. In D. C. West, H. H. Shugart, & D. B. Botkin (Eds.), *Forest succession: Concepts and application* (pp. 132–154). Springer Verlag.
- Dawson, A., Paciorek, C. J., Goring, S. J., Jackson, S. T., McLachlan, J. S., & Williams, J. W. (2019). Quantifying trends and uncertainty in

- prehistoric forest composition in the upper Midwestern United States. *Ecology*, 100, e02856.
- Denevan, W. M. (1996). A bluff model of riverine settlement in prehistoric Amazonia. *Annals of the Association of American Geographers*, 86, 654–681.
- Eva, H. D., Colchester, M., Duivenvoorden, J., Hoogmoed, M., Junk, W., Kabat, P., Kruijt, B., Malhi, Y., Müller, J., & Pereira, J. (2005). *A proposal for defining the geographical boundaries of Amazonia*. Office for Official Publications of the European Communities.
- Faegri, K., & Iversen, J. (1989). *Textbook of pollen analysis* (4th ed.). Wiley.
- Flenley, J. R. (1979). *A geological history of tropical rainforest*. Butterworths.
- Fyfe, R. M., Woodbridge, J., & Roberts, N. (2015). From forest to farmland: Pollen-inferred land cover change across Europe using the pseudo-biomization approach. *Global Change Biology*, 21, 1197–1212.
- Gaillard, M.-J., Sugita, S., Bunting, J., Dearing, J., & Bittmann, F. (2008). *Human impact on terrestrial ecosystems, pollen calibration and quantitative reconstruction of past land-cover* (pp. 415–418). Springer.
- Gaillard, M.-J., Sugita, S., Bunting, M. J., Middleton, R., Broström, A., Caseldine, C., Giesecke, T., Hellman, S. E., Hicks, S., & Hjelle, K. (2008). The use of modelling and simulation approach in reconstructing past landscapes from fossil pollen data: A review and results from the POLLANDCAL network. *Vegetation History and Archaeobotany*, 17, 419–443.
- Gauch, J. H. G. (1982). *Multivariate analysis in community ecology*. Cambridge University Press.
- Gentry, A. H. (1988). Changes in plant community diversity and floristic composition on environmental and geographical gradients. *Annals of the Missouri Botanical Garden*, 75, 1–34.
- Glaser, B., & Woods, W. I. (2004). *Amazonian dark earths: Explorations in space and time*. Springer-Verlag.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18–27.
- Gosling, W. D. (2004). *Characterisation of Amazonian forest and savannah ecosystems by their modern pollen spectra*. Unpublished PhD Thesis. University of Leicester.
- Gosling, W. D., Maezumi, S. Y., Heijink, B. M., Nascimento, M. N., Raczka, M. F., van der Sande, M. T., Bush, M. B., & McMichael, C. N. (2021). Scarce fire activity in north and north-western Amazonian forests during the last 10,000 years. *Plant Ecology & Diversity*, 14, 143–156.
- Gosling, W. D., Mayle, F. E., Tate, N. J., & Killeen, T. J. (2009). Differentiation between Neotropical rainforest, dry forest, and savannah ecosystems by their modern pollen spectra and implications for the fossil pollen record. *Review of Palaeobotany and Palynology*, 153, 70–85.
- Hansen, M., Potapov, P., Moore, R., Hancher, M., Turubanova, S., Tyukavina, A., Thau, D., Stehman, S., Goetz, S., & Loveland, T. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, 15, 850–853.
- Hellman, S., Bunting, M., & Gaillard, M.-J. (2009). Relevant source area of pollen in patchy cultural landscapes and signals of anthropogenic landscape disturbance in the pollen record: A simulation approach. *Review of Palaeobotany and Palynology*, 153, 245–258.
- Hoffmann, W. A., Jaconis, S. Y., McKinley, K. L., Geiger, E. L., Gotsch, S. G., & Franco, A. C. (2012). Fuels or microclimate? Understanding the drivers of fire feedbacks at savanna–forest boundaries. *Austral Ecology*, 37, 634–643.
- Jacobson, G. L., & Bradshaw, R. H. W. (1981). The selection of sites for paleovegetational studies. *Quaternary Research*, 16, 80–96.
- Jarvis, A., Reuter, H. I., Nelson, A., & Guevara, E. (2008). Hole-filled SRTM for the globe version 4. In CGIAR-CSI SRTM 90m database. CGIAR-CSI (Consortium for Spatial Information).
- Juggins, S. (2020). Rioja: Analysis of quaternary science data, CRAN [code], R package version (0.9–15.1).
- Karger, D. N., Conrad, O., Böhrner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., & Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, 4, 1–20.
- Khan, S., Aragão, L., & Iriarte, J. (2017). A UAV–lidar system to map Amazonian rainforest and its ancient landscape transformations. *International Journal of Remote Sensing*, 38, 2313–2330.
- Lehman, J., Kern, D. C., Glaser, B., & Woods, W. I. (2003). *Amazonian dark earths: Origin, properties, management*. Kluwer Academic Publisher.
- Liu, Y., Ogle, K., Lichstein, J. W., & Jackson, S. T. (2022). Estimation of pollen productivity and dispersal: How pollen assemblages in small lakes represent vegetation. *Ecological Monographs*, 92, e1513.
- McMichael, C., Piperno, D. R., Bush, M. B., Silman, M. R., Zimmerman, A. R., Raczka, M. F., & Lobato, L. C. (2012). Sparse pre-Columbian human habitation in western Amazonia. *Science*, 336, 1429–1431.
- Nascimento, M. N., Heijink, B. M., Bush, M. B., Gosling, W. D., & McMichael, C. N. H. (2022). Early to mid-Holocene human activity exerted gradual influences on Amazonian forest vegetation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377, 20200498.
- Nepstad, D., Lefebvre, P., Silva, U. L. D., Tomasella, J., Schlesinger, P., Solorzano, L., Moutinho, P., Ray, D., & Benito, J. G. (2004). Amazon drought and its implications for forest flammability and tree growth: A basin-wide analysis. *Global Change Biology*, 10, 704–717.
- Neves, E., & Petersen, J. (2006). Political economy and pre-columbian landscape transformation in Central Amazonia. In W. Balee & C. L. Erickson (Eds.), *Time and complexity in historical ecology: Studies in the neotropical lowlands* (pp. 279–310). Columbia University Press.
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R., Simpson, G. L., Solymos, P., Stevens, M., & Wagner, H. (2013). Package 'vegan'.
- Ollerton, J., Winfree, R., & Tarrant, S. (2011). How many flowering plants are pollinated by animals? *Oikos*, 120, 321–326.
- Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A. S. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*, 540, 418–422.
- Piperno, D. R., McMichael, C., & Bush, M. B. (2015). Amazonia and the Anthropocene: What was the spatial extent and intensity of human landscape modification in the Amazon Basin at the end of prehistory? *The Holocene*, 25, 1588–1597.
- Power, M., Marlon, J., Ortiz, N., Bartlein, P., Harrison, S., Mayle, F., Ballouche, A., Bradshaw, R. H. W., Carcaillet, C., & Cordova, C. (2008). Changes in fire regimes since the last glacial maximum: An assessment based on a global synthesis and analysis of charcoal data. *Climate Dynamics*, 30, 887–907.
- Power, M., Mayle, F., Bartlein, P., Marlon, J., Anderson, R., Behling, H., Brown, K., Carcaillet, C., Colombaroli, D., & Gavin, D. (2013). Climatic control of the biomass-burning decline in the Americas after AD 1500. *The Holocene*, 23, 3–13.
- Prentice, C. (1988). Records of vegetation in time and space: The principles of pollen analysis. In *Vegetation history* (pp. 17–42). Springer.
- R Development Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Renner, S. S., & Feil, J. P. (1993). Pollinators of tropical dioecious angiosperms. *American Journal of Botany*, 80, 1100–1107.
- Rodgers, J. C., & Horn, S. P. (1996). Modern pollen spectra from Costa Rica. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 124, 53–71.
- Ruokolainen, K., Moulatlet, G. M., Zuquim, G., Hoorn, C., & Tuomisto, H. (2019). Geologically recent rearrangements in central Amazonian river network and their importance for the riverine barrier hypothesis. *Frontiers of Biogeography*, 11, 1–11. <https://doi.org/10.21425/F5FBG45046>
- Stockmarr, J. (1972). Tablets with spores used in absolute pollen analysis. *Pollen et Spore*, 13, 615–621.
- Sugita, S. (1993). A model of pollen source area for an entire lake surface. *Quaternary Research*, 39, 239–244.

- Sugita, S. (2007). Theory of quantitative reconstruction of vegetation I: Pollen from large sites REVEALS regional vegetation composition. *The Holocene*, 17, 229–241.
- Tauber, H. (1977). Investigations of aerial pollen transport in a forested area. *Dansk Botanisk Arkiv*, 32, 1–121.
- ter Steege, H., Pitman, N. C. A., Phillips, O. L., Chave, J., Sabatier, D., Duque, A., Molino, J.-F., Prevost, M.-F., Spichiger, R., Castellanos, H., von Hildebrand, P., & Vasquez, R. (2006). Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature*, 443, 444–447.
- ter Steege, H., Pitman, N. C. A., Sabatier, D., Baraloto, C., Salomão, R. P., Guevara, J. E., Phillips, O. L., Castilho, C. V., Magnusson, W. E., Molino, J.-F., Monteagudo, A., Núñez Vargas, P., Montero, J. C., Feldpausch, T. R., Coronado, E. N. H., Killeen, T. J., Mostacedo, B., Vasquez, R., Assis, R. L., ... Silman, M. R. (2013). Hyperdominance in the Amazonian tree flora. *Science*, 342, 1243092.
- ter Steege, H., Sabatier, D., Castellanos, H., van Andel, T., Duivenvoorden, J., de Oliveira, A. A., Ek, R., Lilwah, R., Maas, P., & Mori, S. (2000). An analysis of the floristic composition and diversity of Amazonian forests including those of the Guiana Shield. *Journal of Tropical Ecology*, 16, 801–828.
- ter Steege, H., Vaessen, R. W., Cárdenas-López, D., Sabatier, D., Antonelli, A., de Oliveira, S. M., Pitman, N. C. A., Jørgensen, P. M., & Salomão, R. P. (2016). The discovery of the Amazonian tree flora with an updated checklist of all known tree taxa. *Scientific Reports*, 6, 29549.
- Trondman, A. K., Gaillard, M. J., Mazier, F., Sugita, S., Fyfe, R., Nielsen, A. B., Twiddle, C., Barratt, P., Birks, H. J. B., & Bjune, A. E. (2015). Pollen-based quantitative reconstructions of Holocene regional vegetation cover (plant-functional types and land-cover types) in Europe suitable for climate modelling. *Global Change Biology*, 21, 676–697.
- Tuomisto, H., Moulatlet, G. M., Balslev, H., Emilio, T., Figueiredo, F. O., Pedersen, D., & Ruokolainen, K. (2016). A compositional turnover zone of biogeographical magnitude within lowland Amazonia. *Journal of Biogeography*, 43, 2400–2411.
- Watling, J., Saunaluoma, S., Pärssinen, M., & Schaan, D. (2015). Subsistence practices among earthwork builders: Phytolith evidence from archaeological sites in the southwest Amazonian interfluvies. *Journal of Archaeological Science: Reports*, 4, 541–551.
- Whitney, B. S., Smallman, T. L., Mitchard, E. T., Carson, J. F., Mayle, F. E., & Bunting, M. J. (2019). Constraining pollen-based estimates of forest cover in the Amazon: A simulation approach. *The Holocene*, 29, 262–270.
- Wickham, H. (2011). ggplot2. *Wiley Interdisciplinary Reviews: Computational Statistics*, 3, 180–185.
- Witteveen, N. H., White, C., Sanchez Martinez, B. A., Booij, R., Philip, A., Gosling, W. D., Bush, M. B., & McMichael, C. N. (2023). Phytolith assemblages reflect variability in human land use and the modern environment. *Vegetation History and Archaeobotany*, 1–16. <https://doi.org/10.1007/s00334-023-00932-2>

#### BIOSKETCH

**A. Blaus** is a postdoctoral researcher exploring patterns of climatic and anthropogenic change in the Neotropical Paleoeology Research Group led by Mark Bush at Florida Institute of Technology. He is also working on a project aiming to quantify tropical pollen assemblages, led by Crystal McMichael at the University of Amsterdam.

**Author Contributions:** C. N. H. McMichael conceived the ideas. A. Blaus, M. F. Raczka, M. N. Nascimento, C. N. H. McMichael and M. B. Bush collected the data. A. Blaus, M. F. Raczka, M. N. Nascimento, C. Herrick, M. Palace, M. B. Bush and C. N. H. McMichael analysed the data. A. Blaus and C. N. H. McMichael led the writing.

#### SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Blaus, A., McMichael, C. N. H., Raczka, M. F., Herrick, C., Palace, M., Witteveen, N., Nascimento, M. N., & Bush, M. B. (2023). Amazonian pollen assemblages reflect biogeographic gradients and forest cover. *Journal of Biogeography*, 00, 1–13. <https://doi.org/10.1111/jbi.14701>