Spatial and temporal abilities of proxies used to detect pre-Columbian Indigenous human activity in Amazonian ecosystems

C.N.H. McMichael a,*, C. Levis b, c, d, W.D. Gosling a, A.B. Junqueira e, D.R. Piperno f, E.G. Neves g, F. Mayle h, M. Peña-Claros b, F. Bongers b

a Department of Ecosystem and Landscape Dynamics, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, the Netherlands
b Forest Ecology and Forest Management Group, Wageningen University & Research, Wageningen, the Netherlands
c Forest Ecology and Forest Management Group, Wageningen University & Research, Wageningen, the Netherlands
d Post-graduate Program in Ecology, Universidade Federal de Santa Catarina, Florianópolis, Brazil
e Instituto de Ciencia 1 Tecnologia Ambienta, Universitat Autonoma de Barcelona, Barcelona, Spain
f Department of Anthropology, Smithsonian National Museum of Natural History, Washington, D.C., and Smithsonian Tropical Research Institute, Balboa, Panama
g Laboratory of Tropical Archaeology, Museum of Archaeology and Ethnology, University of Sao Paulo, Sao Paulo, Brazil
h Department of Geography and Environmental Science, University of Reading, Reading, United Kingdom

A R T I C L E   I N F O
Handling Editor: Dr Mira Matthews

Keywords:
Archaeology
Paleoecology
Ethnography
Human impacts
Plant genetics
Biological collections
Past human activity
Remote sensing

A B S T R A C T
Humans have been modifying ecosystems since before the Holocene began ca. 12,000 years ago, even in Neotropical regions. The Amazon was once thought to be ‘pristine’ and only lightly impacted by Indigenous people before European colonization in the Americas (e.g., pre-Columbian); however, multiple lines of evidence have shown that Indigenous human activities over the past millennia have left ecological legacies on modern ecosystems. We review the various lines of evidence used to reconstruct pre-Columbian Indigenous human activity in Amazonia, and assess the spatial and temporal resolution and limits of each one of them. We suggest that a multi-proxy approach is always preferred, and that lines of evidence that cover overlapping yet discrete spatial and temporal scales can provide a robust and comprehensive assessment of the nuances of pre-Columbian Indigenous human activities in Amazonia, and how they affect modern ecosystems.

1. Introduction
Humans are now considered the major driving force in many abiotic and biotic processes on Earth, and it has been suggested that this has created a distinct ecological era called the Anthropocene (e.g., Crutzen, 2002; Di Marco and Santini, 2015; Gallardo et al., 2015; Halpern et al., 2008; Lewis and Maslin, 2015; Pachauri et al., 2015; Vitousek et al., 1997). Humans contributed to the extinction of Pleistocene megafauna across the Neotropics (Barnosky and Lindsey, 2010; Rozas-Davila et al., 2016, 2021), and have been modifying landscapes in Neotropical ecosystems for at least the last 12,000 years (Ellis et al., 2021; Roosevelt, 2013). Forms of human activities, including plant cultivation and domestication, have also occurred in Neotropical regions for at least 10,000 years (Lombardo et al., 2020; Piperno, 2011; Roberts et al., 2017). Many of these activities, particularly those that have occurred over the last 2000 years, likely left persistent effects, or ecological legacies, on Neotropical ecosystems that are still visible today (Furquim et al., 2023; Levis et al., 2017; McMichael, 2021; McMichael et al., 2023; Ross, 2011).

Reconstructing human-environment interactions through time is particularly important in the ca. 6 million km² of Amazonian rainforests, which hold a large proportion of Earth’s biodiversity (Olson et al., 2001). The suggestion that pre-Columbian Indigenous human influence played a large role in shaping the biodiversity observed in today’s ecosystems has generated an important debate (e.g. Balée, 2010; Clement et al., 2015; Levis et al., 2017; Piperno et al., 2021; Piperno, 2019; Roosevelt, 2013). Most scholars from various disciplines agree that the pre-Columbian Indigenous Peoples of Amazonia influenced its ecosystems to some degree, but the intensity, cultural variability, spatial extensiveness, spatial variability, temporal duration and continuity remain debated (Barlow et al., 2012; Bush et al., 2015; Clement et al., 2015; Heckenberger et al., 2003, 2008; Levis et al., 2012, 2017).

* Corresponding author. Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, 904 Science Park, 1098 XH, Amsterdam, the Netherlands. E-mail address: c.n.h.mcMichael@uva.nl (C.N.H. McMichael).

https://doi.org/10.1016/j.quascirev.2023.108354
Received 8 May 2023; Received in revised form 10 October 2023; Accepted 10 October 2023
0277-3791/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
Much of the controversy results from the overall paucity of data in the region, differences in the interpretation of the same datasets, and on the ecological heterogeneity of this vast area. Even the largest ecological datasets within Amazonia reflect mostly relatively easily accessible areas (McMichael et al., 2017). Addressing aspects of this debate has important implications for how we conserve and manage modern ecosystems, as well as for Indigenous sovereignty.

Ecologists, paleoecologists, archaeologists, anthropologists, ethnographers, Indigenous peoples, and local communities can all provide valuable information about the activities of pre-Columbian Indigenous people and their influence on Amazonian vegetation and landscapes. These lines of information, however, come from a variety of sources, are measured at different spatial and temporal scales, and the capabilities of reconstructing pre-Columbian human activities vary between them. Integrating different data and knowledge sources, however, is vital for understanding multifaceted and complex problems, and to achieve a more reliable and robust assessment of the influence of pre-Columbian activities on Amazonian ecosystems (Mayle and Iriarte, 2014; McCleanchan et al., 2015; Tengo et al., 2014).

A large part of data interpretation is recognizing the advantages and limitations of the various proxies, methods, and settings used to reconstruct pre-Columbian Indigenous human activities and their effect on tropical systems. Here we review proxies (i.e., lines of evidence) and settings (i.e., depositional environments or archival material) commonly used to assess pre-Columbian human activities and influences on Amazonian ecosystems, with a specific focus on the spatial and temporal resolution (the potential sampling frequency of the archival material) and spatial and temporal limits (the total amount of space or time potentially captured within the record). We also discuss the detectability and variability of proxies used to assess pre-Columbian Indigenous human activities and their influences on Amazonian ecosystems (Fig. 1). Our goal is to provide a spatio-temporal framework for the various proxies used to assess long-term Indigenous human activity in Amazonian ecosystems so that future work can recognize the advantages and limitations of all of the proxies and integrate them more comprehensively.

2. Soils as local-scale archives of pre-Columbian Indigenous human activities and vegetation change

Soils contain a wealth of valuable information and proxies that can be used to reconstruct pre-Columbian human activities and the influence of Indigenous Peoples on the vegetation. Archaeological sites, artefacts, macrofossils, and microfossils that directly document pre-Columbian human activities are all found on, or in, tropical terrestrial soils (Figs. 1 and 2). Information from archaeological surveys is beginning to be compiled into online databases, including the AmazonArch (Amazonian Archaeological Sites Network), which contains the geographical location and basic archaeological information for over 10,000 sites (Clement et al., 2015; Riris and Arroyo-Kalin, 2019; WinklerPrins and Aldrich, 2010; https://sites.google.com/view/amazonarch). The variability and diversity of archaeological evidence includes sites containing artifacts such as lithics, ceramics, rock paintings, or petroglyphs, earthworks, and anthropogenic soils (called Amazonian Dark Earths, or ADE) (e.g., Neves et al., 2021) (examples of ceramics and

---

Fig. 1. The array of proxies that can be used to detect pre-Columbian Indigenous human activity in Amazonia, and the source, or archival material from which they are derived. The proxies are shown on a gradient from those which are used specifically to assess human activities (far right) to those which are used to assess both human and non-human processes (far left). ADE = Amazonian Dark Earths.

Fig. 2. Examples of directly documented evidence of human activity found in Amazonian soils. The top panel shows an archaeological excavation, and the middle panel shows ceramic artifacts that were uncovered during the excavation (Photos: Bernardo Oliveira/Instituto Mamiraua). The bottom panel shows a Google Earth image of earthworks that were uncovered in southwestern Amazonia following deforestation of the landscape.
earthworks shown in Fig. 2). At some of these sites, pre-Columbian people have modified soil texture, nutrient content, and stable isotopes, leaving persistent legacies of soil properties for thousands of years (Glaser and Birk, 2012; Glaser and Woods, 2004; Lehman et al., 2003).

Phytoliths are microscopic bodies of silica that are produced in the cells of many plant species and often preserved in high quantities in soil archives where pollen and macrofossils have largely decayed (Piperno, 2006) (example shown in Fig. 3). Phytolith analysis is commonly performed on samples obtained at archaeological sites, but also in soils that are not associated with archaeological settings such as lake sediments and terrestrial soils (Fig. 1). Phytoliths often reflect localized plant decay, particularly in terrestrial soils, as they do not have intrinsic dispersal mechanisms that would carry them considerable distances from the depositional environment as can be the case with pollen. Studies of phytoliths in surface soils from tropical forests in Panama showed that phytolith movement, via fire, or surface water transport, can be as little as 25–30 m from their source area (Piperno, 1998) or up to 100–120 m from their source plant (Piperno, pers. Comm.; Piperno and McMichael, manuscript in preparation). Further, lakes with in-flowing streams may contain phytoliths from considerable distances, and in fluvial forest soils from Brazil, phytolith movement from areas a substantial distance upstream is indicated (Watling et al., 2016). Seasonally flooded savannas might be expected to be similar.

Phytoliths directly document different types of vegetation, such as savanna, open forests with bamboo, evergreen, semi-evergreen, and deciduous forest, and early successional growth typical of human disturbance. Phytoliths also directly detect some major and now-minor crops and other economic plants, (e.g., maize [Zea mays L], Cucurbita spp. [squashes and gourds], manioc [Manihot esculenta Crantz], arrowroot [Maranta arundinacea L], and various palm species) (Carson et al., 2014; Dickau et al., 2013; Lombardo et al., 2020; McMichael et al., 2012a, 2012b; Piperno, 2006; Watling et al., 2016; Whitney et al., 2013, 2014). All palms, grasses, and sedges, and many arboreal basal angiosperms and eudicotyledons, produce high phytolith numbers diagnostic to at least the family, and not uncommonly, genus levels (Huisman et al., 2018; Morcote-Ríos et al., 2015, 2016; Piperno, 2006; Piperno and McMichael, 2023; Piperno et al., 2019; Witteveen et al., 2022). The high phytolith production of palms and grasses also means that their absence in the record actually represents the absence of these taxa in the vegetation (Piperno et al., 2015). Not all plants produce phytoliths, however, and some taxa will remain undetectable in reconstructions (Piperno, 2006). This is the case with some major economically important trees, e.g., Bertholletia excelsa (Brazil nut), Annona spp. (soursop, cherimoya), guava (Psidium guajava L).

Starch grains (example shown in Fig. 3) can identify certain cultivars, such as maize (Zea mays L.), squashes (Cucurbita spp.), manioc (Manihot esculenta Crantz), and other tuber crops (e.g., Pearsall et al., 2004; Piperno, 2006; Piperno, 2011). Some crops that do not produce phytoliths, such as peanuts (Arachis hypogaea L.) and chili peppers (Capsicum spp.), have diagnostic starch grains (e.g., Dickau et al., 2007; Ezell et al., 2006; Piperno, 2006). Beans, including Phaseolus species, and some palm trees that were used by pre-Columbian Indigenous Peoples also produce identifiable starch grains (Watling et al., 2018). The starch grains are usually isolated from ceramics or stone tools found at archaeological sites (Fig. 1) (Iriarte et al., 2004; Pearsall et al., 2004; Watling et al., 2018; Young et al., 2023). Macrobotanical remains, often carbonized, are typically recovered from site sediments and along with the starch grains and phytoliths, provide empirical evidence for the diets and lifestyles of pre-contact Indigenous Peoples (e.g., Furquim et al., 2021; Watling et al., 2018).

Macroscopic charcoal fragments (e.g., >500 μm) found in soils, including at archaeological sites, represent localized past fire events (e.g., Rhodes, 1998; Whitlock and Larsen, 2002) (example shown in Fig. 3). In the aseasonal forests with a limited dry season (for instance, those found in northwestern Amazonia), forest fire almost always starts with human intervention (Fig. 1) (Bush et al., 2008; Malhi et al., 2008). The
presence of charcoal in these aseasonal forests indicates human activity (Bush et al., 2008), but escaped wildfires in these areas can also occur during extreme droughts (Flores et al., 2017). In drier areas, i.e., the seasonal forests that border savannas in eastern Amazonia, fires are less dependent on human ignition, but are still more frequent when humans are present (Alencar et al., 2004; Maezumi et al., 2015, 2018b; Nepstad et al., 2004; Power et al., 2016; Ramos-Neto and Pivello, 2000). Paleoeocological and archaeological data show that fire frequency across Amazonia was more frequent in the late Holocene, when climate was wetter than the early-to mid-Holocene (Arroyo-Kalin and Riris, 2021; McMichael and Bush, 2019; Nascimento et al., 2022), highlighting that the primary source of ignition was human activity.

Charcoal retains diagnostic morphological features of the plant from which it is derived, and can be used to identify types of plants, e.g., woody versus non-woody taxa (Bodin et al., 2020; Di Pasquale et al., 2008; Orvis et al., 2005). Charcoal morphology can sometimes provide taxonomic identification to the family-level and sometimes genus or species level in tropical ecosystems, and has been used to characterize land use and successional forests at archaeological sites (Bachelet and Scheel-Ybert, 2017; Bodin et al., 2019; Cartwright, 2015; Fernandes, Caromao et al., 2013; Goulart et al., 2017; Iriarte et al., 2020; Scheel-Ybert et al., 2014). The chemical properties (i.e., FTIR spectroscopy) of charcoal fragments can also be used to infer burn temperature of the fire events, and can also distinguish plant types (e.g., woody vs grassy material) that were burned (Gosling et al., 2019; Maezumi et al., 2021).

With soil surveys to reconstruct past fire events, replicate soil cores are typically collected from a given site (100 m–200 m radius) (Hammond et al., 2006; McMichael et al., 2012a, 2012c) to account for the uneven deposition of charcoal that occurs on localized scales after burning of vegetation (e.g., Lynch et al., 2004). Thus, the repeated absence of charcoal from soil cores located in close proximity can be confidently interpreted that the sampled area truly lacked fire rather than the sampling was unable to detect the fire (McMichael et al., 2012a, 2015). Replicate soil cores can also indicate whether large tracts of vegetation were burned or repeatedly burned, especially if some of the particles are $^{14}C$ AMS dated (Feldpausch et al., 2022; Heijink et al., 2022; McMichael et al., 2012a; Sanford and Horn, 2006; Whitlock and Larsen, 2002). Replicate cores are often also analysed to look at how vegetation change has occurred over relatively small geographic scales or along environmental gradients (e.g., Heijink et al., 2022; McMichael et al., 2012a; McMichael et al., 2012b; Watling et al., 2017).

Stable carbon and nitrogen isotopes are commonly used to infer vegetation dynamics (Fig. 1) (de Freitas et al., 2001; Pesenda et al., 1998). Stable carbon isotopes from soils have also been used to infer landscape transformations by pre-Columbian Indigenous Peoples. These are particularly useful for documenting changes or shifts between C3 and C4 grass assemblages and vegetation changes on pre-Columbian raised fields (Iriarte et al., 2010; McKey et al., 2010; Watling et al., 2017). The analysis of stable carbon and nitrogen isotopes from bone collagen has also been used to reconstruct dietary changes in pre-Columbian Indigenous Peoples (e.g., Colonese et al., 2020; Roosevelt, 1989).

Soils have a very high spatial resolution as multiple samples can be collected within meters of each other (i.e., high potential sampling frequency) and a very high spatial limit because they are found almost everywhere in terrestrial systems (Fig. 4). The temporal limit of soils is also very high; soils capture evidence from modern times to thousands of years ago (Fig. 4). The uppermost 1 m of Amazonian soils typically represent the last several thousand years (Piperno, 2016; Piperno et al., 2021). Soils, however, have a low temporal resolution due to processes such as soil formation, erosion, and bioturbation, and establishing age-depth relationships is not always possible (e.g., Mayle and Iriarte, 2014; Sanford and Horn, 2000) (Fig. 4a). Radiocarbon ($^{14}C$ AMS) dates from archaeological sites or soil microfossils usually have a 2-sigma precision of ca. 100 years (Neves et al., 2004; Piperno, 2016; Schaan et al., 2012; Taylor and Bar-Yosef, 2016). When multiple dates are obtained from the same core, however, general trends in fire or vegetation of older to younger within soil cores can often be established (e.g., Hill et al., 2023; McMichael et al., 2012a; McMichael et al., 2012c; Piperno et al., 2021).

3. Lake sediments as local-to regional-scale archives of pre-Columbian Indigenous human activity and vegetation change

Like soils, lake sediments also contain microfossils that can be used to document pre-Columbian human activities and the resulting legacies on ecosystems. Lakes sufficiently old for palaeoecological studies are rare across much of Amazonia, limiting the spatial resolution of palaeo-vegetation reconstructions (Bush and Silman, 2007), although the temporal resolution can sometimes be high (Fig. 4). Unlike soils, lake sediments typically have continuous deposition and thus robust stratigraphic integrity, and age-depth relationships can be derived that place temporal frameworks on human activities and environmental change. Most lake sediment records from Amazonia contain samples analysed at

![Fig. 4.](image-url)
centennial scale temporal resolution (e.g., every century to several hundred years) (Nascimento et al., 2022). In rare deep lakes with anoxic conditions, however, the sediments may retain sub-decadal stratigraphy, allowing an almost continuous insight into the local dynamics of human activity and forest recovery (Fig. 4) (Åkesson et al., 2021; Bush et al., 2016, 2021a). The temporal limit of lake sediment records varies depending on lake type, local depositional environment and preservation conditions (Fig. 4). Few sites in Amazonia extend back to the Pleistocene (Mayle et al., 2000; Whitney et al., 2011), as most non-riverine lakes were formed (and sedimentation began) in the mid-to late-Holocene periods (ca. 8000-4000 years ago) (e.g. Bush and McMichael, 2016; Bush et al., 2007; Carson et al., 2014; Nascimento et al., 2022; Urrego et al., 2013).

Charcoal abundances and their changes within a sedimentary sequence are typically used to infer changes in the amount of biomass burning in the surrounding landscape (e.g. Marlon et al., 2013; Marlon et al., 2016); equated to fire severity (Keeley, 2009). Recent work has also estimated burn temperatures from charcoal found within lake sediments, and shown how vegetation composition and traits change depending on the intensity of the fire (Nascimento et al., 2023). Charcoal particles are deposited into lake sediments from airborne or terrestrial sources, and assessing size classes of charcoal particles is commonly used to distinguish local from regional input (Clark and Royall, 1996; Sanford and Horn, 2000). The relationships between charcoal abundance and biomass burned, and the source area of charcoal particles for Amazonian lakes remain poorly documented and need further exploration.

Organic macrofossils and microfossils (e.g., pollen and spores) that decay in soils typically preserve in lake sediments. Fungal spores can be associated with fire and thus human activity (Fig. 1) (Brugger et al., 2016; Loughlin et al., 2018). Tree and shrub pollen can generally be identified to a more specific taxonomic level than phytoliths, although the inverse is true for herbaceous taxa (especially grasses and sedges) and palm taxa (Piperno, 2006). Over 1000 pollen morphotypes from Amazonia have been identified and catalogued (Bush and Weng, 2007), with up to over 100 pollen types being identified within a single pollen sample (Åkesson et al., 2021). Pollen from domesticated maize (Zea mays) (see example in Fig. 5) can be reliably identified because wild Zea does not occur in South America, but major crop plants such as manioc (Manihot esculenta) and sweet potato (Ipomoea batatas) cannot be differentiated from their wild varieties using pollen because the latter are native to South America, and taxonomic distinctions cannot reliably be made (Mayle and Iriarte, 2014; Whitney et al., 2012).

The spatial representation of pollen and phytolith data depends on site-specific characteristics of the lake. Phytoliths from lake sediment cores have varying source areas that depend on lake size and the presence of in-flowing streams; the assemblages are often mixtures of these source areas (Carson et al., 2014; Mayle and Iriarte, 2014; Piperno, 2006; Plumpton et al., 2020; Whitney et al., 2013, 2014). Pollen assemblages also reflect a combination of local and regional inputs, and this varies, not only depending upon lake size and the presence or absence of inflowing streams, but also on the relative proportion of wind-pollinated taxa in the parent vegetation (Bush et al., 2021b; Jacobson and Bradshaw, 1981). In the forest-savanna ecotone regions of Bolivia, where the dominant forest taxa (e.g. Moraceae) have wind-dispersed pollen, large lakes may have a pollen source area of up to 40 km (Whitney et al., 2019). However, in many areas of Amazonia, closed-canopy forests are dominated by insect-pollinated taxa, where very small lakes register much smaller pollen source areas of potentially only 1–2 km² (Blaus et al., 2023). Differentiating between local- and regional-scale human land-use and deforestation is possible, however, with pollen analyses from tight clusters of small lakes (Bush et al., 2007) or pairs of small and large lakes (Carson et al., 2014). It is also important, where possible, to pair lake sediments from small lakes with nearby archaeological sites to provide matching spatial resolution and a continuous temporal framework of past land use (Carson et al., 2014; Mayle and Iriarte, 2014; Whitney et al., 2014).

Diatoms are siliceous microalgae that are found in water bodies that provide information about environmental or hydrological conditions (e.g., Battarbee, 1986; Benito et al., 2018) (Fig. 1, for example see Fig. 5). Diatoms are commonly assessed in lake sediment reconstructions, and can indicate changes in hydrology (e.g., lake level recorded by changing proportions of deep versus shallow water taxa) or water quality that are related to climate dynamics (Castro et al., 2013; Nascimento et al., 2021) (Fig. 1). Diatoms, however, can also provide information about pre-Columbian human activity in Amazonia (Fig. 1). They have been used to document changes in wetland management (Duncan et al., 2021), and nutrient status and productivity (Bush et al., 2016). Diatom assemblages can also be used to parse apart climatic and human-induced vegetation changes in lake sediment records (e.g., Bush et al., 2000) (Fig. 1).

Lake sediments also contain stable isotopes and chemical elements that can shed light on pre-Columbian human activity (Fig. 1) ( Hodell et al., 1995, 2005). For example, Ca++, and K+ concentrations were used to provide information on lake level changes related to climatic fluctuations alongside human activity (Bush et al., 2000; Sahoo et al., 2019). More recently, micro- X-ray fluorescence (XRF) has become a standard tool in paleolimnology, including multivariate analysis of XRF data (Parsons et al., 2018) or ratios of cation concentrations, such as Ca/Ti (proxy for drought), Fe/Mn (proxy for lake depth) or Rb/Sr (proxy for grain size) (Davies et al., 2015). XRF data have been used to reconstruct human-induced soil runoff and erosion (Åkesson et al., 2019), and to place human activities in a context of environmental change.
change (Aniceto et al., 2014; Maezumi et al., 2018b; Rodríguez-Zorro et al., 2015). Sediment color, which reflects abundances of organic material and clays, can also provide information on environmental rhythms or human-induced change (Bush et al., 2000, 2017; Rodbell et al., 1999).

4. Ethnographic, ethnohistorical and ethnoecological data provide insights into pre-Columbian Indigenous human activity and vegetation change

Ethnographic, ethnohistorical and ethnoecological studies can be used to assess, document, interpret and obtain insights from Indigenous and local knowledge systems (for example see Fig. 6). Indigenous and local knowledge systems provide valuable information that can be used to interpret pre-Columbian human activities (Fig. 1) (Cassino et al., 2019), such as: (i) resource use and management, including past and modern distributions of plant resources (Cassino et al., 2019; Levis et al., 2018, 2020); (ii) how lifestyles and resource management systems influence, and are influenced by landscapes (Balée, 2006); (iii) the technological and labour constraints on resource use (Junqueira et al., 2016); and (iv) the population densities that can be sustained in different ecosystems and by different production systems (Heckenberger et al., 2008).

Ethnography refers to the in-depth description of everyday life and practice of a given culture or society (Oxford, 2016; see also Hammerle and Atkinson (2007) for a broader definition), and ethnohistory combines ethnography with the scrutiny of historical records and other sources of information (Axtell, 1979). Ethnoecology is the study of people’s interactions with their environment, often with special attention to current knowledge and practices concerning resource use and management, including the subdisciplines of ethnobotany and ethnozoology (Martin, 2004). The subdiscipline of ethnoarchaeology involves ethnographic studies performed by archaeologists with an explicit focus on material culture (Politis, 2014). Much of the culture and resource management practices of the current inhabitants of Amazonia have been inherited from pre-Columbian populations, but transformed to different extents because of the heterogeneous and discontinuous history of human occupation in the region (Cleary, 2001; Denevan, 2001). For example, European arrival caused a massive die-off of Indigenous populations (Denevan, 1992), and many of the remaining groups were fragmented, displaced, or enslaved during European colonization (Dobyns, 1966). Thus, the projection of ethnographic data to past lifestyles and production systems requires caution (McClennen et al., 2015). Ethnographic data cannot always be tied to human activity at specific times in the past, but local Indigenous knowledge is crucial to locating ancestral forests, anthropogenic soils and other signs of human activity, especially in remote regions (Franco-Moraes et al., 2019; Kopenawa and Albert, 2023).

Indigenous Peoples are the descendants of native ethnic groups that retain historical and cultural connections with pre-Columbian Indigenous societies, though major social disruptions and upheavals occurred with European colonizations (Cook, 1998; Livi-Bacci, 2016). Most ethnographic and ethnoecological studies were done by non-Indigenous scientists. Recent collaborative studies with contemporary Indigenous people have posited that Amazonian ecosystems have been transformed by an extensive and long-term network of social relations and interconnections between humans (e.g., Heckenberger et al., 2008; Ribeiro et al., 2023). Future research should promote participatory and collaborative approaches, as Indigenous people and their knowledge systems can inspire new ideas to enhance understanding of human activities and vegetation changes (for example see Fig. 6).

Archaeological evidence combined with observation of Indigenous production systems suggests, for example, that pre-Columbian production systems in some regions seem to have been much more based on agroforestry and on the management of forest and aquatic resources compared with colonial systems (Maezumi et al., 2018a; Moraes, 2015; Neves, 2013; Shepard et al., 2020). Ethnographic work on current soil and waste management systems has also been essential to our understanding of the processes that led to the formation of Amazonian Dark Earth (ADE) (Hecht, 2003; Schmidt et al., 2014; Winklerprins, 2009).

The variance and scarcity of ethnographic data limits extrapolations of locally derived information to other regions within Amazonia. For example, while major crops like maize and manioc were grown in most regions of the basin, some minor crops, or useful native species, that are culturally or economically important in a certain region may not be so in others. Instead of a direct projection of the present into the past, ethnographical research provides insights to further understand ancient Indigenous livelihoods, resource management strategies and their potential impacts in past and current landscapes (McClennen et al., 2015). Ethnographic studies also provide valuable information on Indigenous resource use and societal practice since European colonization, and facilitate disentangling the effects of pre- and post-Colonial human activities in current landscapes (Forline, 2008). Ethnographic data can be paired with linguistics, as it is known that groups within the same language families are more likely to share similar resource management systems, e.g., the Arawak (Eriksen and Danielsen, 2014). Historical distributions of Indigenous languages (e.g., Eriksen, 2011), may thus facilitate ethnographic projections across space since European colonization.

Historical documents originating during the early colonization of
Amazonia may also provide insight into pre-Columbian Indigenous land use systems. Francisco Orellana led the first expedition down the Amazon River in 1540 CE, which was recorded by Gaspar de Carvajal (Medina, 1934). Early colonists moved in and established Jesuit missions later in the 1600s CE in several regions (Reeve, 1994), and the Amazonian Rubber Boom occurred from ca. 1850–1920 CE (Hecht and Cockburn, 2010; Weinstein, 1983). All these events have corresponding historical documents that recorded aspects of Indigenous Peoples and their interactions with the colonists.

The spatial resolution of ethnographic records and historical data has the potential to be very high (Fig. 4). There is also a high potential temporal resolution of both historical and ethnographic records, though available data through time are relatively sparse. The temporal limit of historical documents encompasses only the last several hundred years since European arrival, and the temporal limit of ethnographic records is bound by the memories of local residents, although it can extend much further back in time through knowledge transmission across generations (Fig. 4b).

Because of the fragmentation and upheaval of Indigenous populations after European colonization, it remains contentious as to how far back in time ethnography can be extrapolated back. In some cases there is a strong cultural and historical continuity between ancient and contemporary Indigenous Peoples, but in other cases contemporary Indigenous Peoples have inherited their current lands for only several decades (e.g., the Kichwa People of Andean descent that now inhabit areas of lowland Ecuador).

5. Modern datasets provide insight to pre-Columbian Indigenous human activities and vegetation change: biological collections and remote sensing data

Biological collection records include plant and animal surveys (inventories), biological or ecological monitoring networks (e.g., Anderson-Teixeira et al., 2015; Malhi et al., 2002; ter Steege et al., 2013), and herbarium or museum specimens (e.g., Feeley and Silman, 2011) (for example see Fig. 7). These biological collection records are used to assess the presence, absence, and abundances of plants and animals in modern ecosystems (or those during the historic period). Although biological records provide valuable information about past (pre- and post-colonial) activities, they do not directly measure pre-Columbian Indigenous influences on ecosystems (Fig. 1). Biological records need linkages with archaeological, paleoecological, genetic, or ethnographic data on the degree and form of past human activities to infer cause and effect relationships (e.g., Heijink et al., 2022; Heijink et al., 2020; Levis et al., 2017; Piperno et al., 2021).

Past and current human activities are increasingly recognized as factors influencing species’ distributions (Boivin et al., 2016; Di Marco and Santini, 2015; Gallardo et al., 2015; Guisan and Thuiller, 2005; Halpern et al., 2008). Humans have modified the distribution range and abundance of several plant species, expanding the distribution of useful and domesticated plants more often than plants without a documented use to humans (Balée, 1989; Levits et al., 2017, Coelho et al., 2021). For instance, manioc (Manihot esculenta) was cultivated by Indigenous people throughout the Holocene (Piperno, 2011). Currently, domesticated manioc populations are cultivated throughout the tropics, while the direct ancestor of domesticated manioc (M. Rabellifolia) is limited to South America (Olsen and Scharf, 1999). Differences between the natural distribution and the human-modified distribution of cultivated species, and especially those with domesticated populations can indicate past human activities. Several palms and trees that are used for food are abundant in plant assemblages of archaeological sites (Balée, 1989; Junqueira et al., 2010). Modern plant inventories that are closer to archaeological sites also tend to have a higher abundance and diversity of useful and domesticated plants (Levis et al., 2012, 2017; Thomas et al., 2015).

Plant genetic material (DNA) is typically derived from plants but can also be found in soils and sediments. Genetic material from plant remains can be linked with past events in human history, such as plant domestication and migrations (Fig. 1) (e.g. Clement, 1988a; Clement, 1988b; Gutaker and Burbano, 2017; Moreira et al., 2017; Roullier et al., 2013). Genetic studies can also provide information about species with populations that were domesticated by humans, because selection and cultivation of desirable phenotypes results in changes in morphology, physiology, and genotype of descendent populations (Emshwiller, 2006; Harlan, 1992; Olsen and Schaal, 2001). The whole set of selected phenotypic changes in a species is termed its domestication syndrome, which can be studied with morphological, chemical, archaeobotanical, and molecular genetic methods (e.g. Emshwiller, 2006; Meyer et al., 2012; Smith, 2006).

Species with populations with some degree of domestication show patterns of morphological variation and genetic diversity and structure across geographical space that result from domestication events, dispersal and subsequent diversification (Meyer and Purugganan, 2013). Economically important domesticates are more likely to show dramatic morphological changes, such as a 2000% increase in fruit size from wild source populations of peach palm (Bactris gasipaes) compared with some domesticated populations (Clement, 1988b). Dispersal events are often accompanied by other natural and human selection pressures, resulting in diversification of uses, variation in morphology, chemical composition and physiology (Meyer and Purugganan, 2013), and adaptation to domesticated landscapes (Clement, 1999).

Until very recently, plant geneticists worked exclusively with living plants or samples collected over the last 200–300 years (e.g., Roullier et al., 2013). Over the last decade, new molecular genetic methods have allowed the extraction and analysis of DNA from archaeobotanical remains (Wales et al., 2014). Ancient DNA (aDNA) is increasing the ability to document the genetic history of plants, can differentiate crops from their wild ancestors, or estimate genetic change and migration of domesticated plants or cultivars over time (Freitas et al., 2003; McLachlan and Clark, 2005; Piperno, 2011). It is now even possible to extract and date aDNA recovered from pollen grains found in lake sediments and historical samples (Gutaker and Burbano, 2017; Parducci et al., 2017). aDNA is denatured very rapidly in tropical lake sediments (due to the hot, wet conditions), though extraction has been successful in the African tropics (Bremond et al., 2017). Geographic representation of aDNA samples is patchy, although many crop plants have been databased.

Biological collection records can be collected from anywhere on Earth, so their potential spatial limit is endless (Fig. 4). To date, however, sampling covers only a small portion of the 6 million km² of

Fig. 7. Researcher Carolina Levis measures and identifies a piqia tree (Car- yocar villosus) with a local resident of the Tapajós National Forest of Brazil (Photo: Bernardo Flores).
Amazonia (Carvalho et al., 2023; ter Steege et al., 2013). Ground-based surveys of plant and animal distributions and abundances can also have high spatial resolution, with hierarchies of transects or 1-ha plots commonly clustered within a region (e.g., ter Steege et al., 2013). Biological inventories can also be measured at yearly frequencies, and have high temporal resolution, but most repeated censuses only span a few decades (Malhi et al., 2002; Phillips et al., 1994) (Fig. 4).

Ground-based biological collections and satellite imagery are often paired in modern ecological and global change studies. Satellite imagery has also been used to infer soil and vegetation legacies of pre-Columbian land use (Iriarte, 2016; Palace et al., 2017; Thayn et al., 2011). Satellite data from Landsat, Sentinel, and MODIS, typically capture landscape features at 30-m to 1-km spatial resolution (Fig. 4). Landsat has been used to detect legacies of pre-Columbian land use (Heckenberger et al., 2003; Söderstrom et al., 2016), and MODIS has been used to detect or predict Amazonian Dark Earth (ADE) (Palace et al., 2017; Thayn et al., 2011). This detection is possible because pre-Columbian Indigenous Peoples permanently changed the soil characteristics, which affects the types of vegetation that can grow on those soils (Junqueira et al., 2011). The differences in modern biomass between ADE and forested non-ADE sites within the same region is also detectable using MODIS satellite imagery (Palace et al., 2017).

Remotely sensed data has low temporal limits (Fig. 4). Satellite images are only available for the last few decades, although aerial images may extend further back in time. However, they have higher spatial limits than biological collections, and usually have Amazonian-wide spatial coverage (Fig. 4).

6. Dating data sources and proxies used in assessing pre-Columbian Indigenous human activity and vegetation change

Different scientific disciplines assessing pre-Columbian Indigenous human activities place differing emphases on dating or age control of their data sources or proxies. In soils, archaeologists and paleoecologists typically use $^{14}$C AMS dating (radiocarbon dating) to obtain ages on specific material(s) of interest. In archaeological surveys, multiple ages are usually derived from specific horizons of interest where artifacts are recovered (e.g., Roosevelt et al., 1996). Sometimes, however, ages of specific horizons are inferred from a known type of pottery or artifact that has been recovered and dated from another location (e.g., McEwan, 2001). Paleoecologists will typically obtain dates from individual charcoal fragments or conglomerations of phytoliths recovered from soil cores or profiles (e.g., Heijink et al., 2022; McMichael et al., 2012a; Piperno et al., 2021; Watling et al., 2017). In both archaeological and paleoecological surveys in soils, repeated dating across sites can help determine the synchronicity of events across space. Because of soil bioturbation, age-depth models are not applied to soil profiles. General stratigraphic trends, however, are often intact (Piperno et al., 2021; Watling et al., 2017).

Lake sediments typically retain stratigraphic integrity, and age-depth models (e.g., Blaauw and Christen, 2011) are often used to reconstruct a temporally continuous sedimentary sequence. The age-depth models typically require fewer $^{14}$C dates than may be obtained with soil surveys, but the more dates obtained on a sequence, the more confidence can be placed in the model. With lake sediments, specific layers of change or markers of human activity can be directly dated to provide additional confidence for the timing of those events. Additional temporal control can also be placed on the younger sections of lake sediment cores using $^{210}$Pb dating, which can be used on sediments less than 150 years old (e.g., Sanchez-Cabezudo and Ruiz-Fernandez, 2012). Age-depth models can incorporate mixtures of $^{210}$Pb and $^{14}$C dates (Aquino-Lopez et al., 2018), which can be particularly useful for increasing confidence in the ages of sediments around the period of European Contact.

The other data sources included in this review are not based on laboratory dating of materials or sediments. Historical records have specific ages corresponding to the dates they were produced. It is not possible to place specific dates or ages on Indigenous knowledge, or its interpretation through the studies of ethnography, ethnohistory, or ethnoecology. It is possible, however, to correlate some historical events or personal events, such as the arrival of missionaries in an Indigenous community or abandonment of a village with changes in resource use and management. This information is crucial to understanding how pre-Columbian people lived, even though specific ages are difficult to obtain.

Modern datasets, including living plants and museum specimens, may have specific dates recorded. Inferring an age when pre-Columbian Indigenous people shaped plant abundances or plant growth patterns (for instance) is possible with the direct dating of living materials using a combination of dendrochronology, radiocarbon analysis, stable isotope analysis and DNA analysis (Caetano-Andrade et al., 2020). Recent advances in genetic methods have ages inferred also from molecular clocks and DNA-based dating method for ancient genomes (Kistler et al., 2020).

7. Discussion: moving forward

Understanding the interactions of pre-Columbian people and their environments in Amazonia is important for sustainability science, conservation biology and cultural anthropology (Levis et al., 2017; Mayle and Iriarte, 2014; McMichael et al., 2017; Roberts et al., 2017; Szabó and Hedl, 2011; Watling et al., 2017; WinklerPrins and Levis, 2020). Here we have provided a review of the most commonly used sources and proxies for reconstructing pre-Columbian human activity in Amazonia (Fig. 1) and have described the associated advantages and limitations of each by assessing their potential spatial and temporal resolution and limits (Fig. 4). We acknowledge that a comprehensive assessment of all proxies of Indigenous human activity is beyond the scope of this manuscript. For example, proxies such as lipid biomarkers have been used to assess the components of anthropogenic soils (Glaser, 2007). Fecal biomarkers are a newly emerging proxy that is being used to detect past human activity (Argiriadis et al., 2018; Zocatelli et al., 2017). These proxies have not yet been used in Amazonian systems, though show great promise in other geographic regions. The preservation of biomarkers and sterols in the humid tropics, however, is likely poorer than in other areas.

Assessments of the long-term Indigenous history in Amazonia would be stronger if ‘absence data’ from all lines of evidence were reported or archived in data repositories. Ecological datasets, including plant inventory records, include the presence, absence, and abundance of species within a given study area (e.g. Hubbell, 1979; ter Steege et al., 2013), allowing for more robust statistical analyses than presence-only analysis often applied to archaeological datasets (e.g., McMichael et al., 2014a; McMichael et al., 2017). To fully understand the impact of pre-Columbian people in Amazonian landscapes, the reporting of ‘absence data’ (i.e., when there is no evidence of past human activity) is crucial. For example, when identifying earthworks using remotely sensed data, the total area surveyed and total number of earthworks found should be reported so that site densities can be calculated and compared across regions. The entire area sampled and information on the absence of ADEs using field-based surveys is also rarely reported, and the varying densities of ADEs across the landscape cannot yet be calculated. The same approach should apply for archaeological surveys, in which generally a wide area is initially surveyed before determining excavation locations. We suggest that efforts to compile and build repositories of archaeological information should develop guidelines and protocols for reporting and documenting absence data.

The ‘absence’ of evidence of human activities from paleoecological proxies, including charcoal, pollen, and phytoliths is reported. However, because these lines of evidence come from a limited amount of sample material, it is possible that they are present but not detected. Thus, the absence of evidence is not necessarily evidence of absence. This issue is partially overcome with repeated sampling in the case of soil cores (i.e., multiple cores collected and analysed per locality) (e.g., Heijink et al.,...
Several predictive models have been developed for various types of archaeological features in Amazonia and for the overall likely distribution of pre-Columbian Indigenous Peoples in the region, which provide targets for future archaeological surveys in the ca. 6 million km² of Amazonian forests (McMichael et al., 2014a, 2014b, 2017; Souza et al., 2018; Walker et al., 2023). Due to the lack of ‘absence data’, the only approaches available for these predictions are models that require presence-only data (i.e., that do not require absence data). Absence data of all types will help to validate and refine these models on both regional and continental-wide scales, and open doors to an array of additional modelling approaches that could be employed (McMichael et al., 2017). All types of models and other macro-paleoecological and macro-archaeological syntheses would also further benefit by including more precision in the geographic coordinates of localities being studied (i.e., to ca. ± 200 m spatial resolution), and more precision in the time bins of analysis.

Perhaps the best way to strengthen assessments of past human activity is via an interdisciplinary approach, pairing multiple proxies and multiple types of data sources (Carson et al., 2014; Mayle and Iriarte, 2014; Watling et al., 2017; Whitney et al., 2014). Without this pairing, it can be difficult to determine whether the observed pattern is related to human activity or other forcing mechanisms. Ideally, though not necessarily, those pairings should be spatially overlapped. For example, Levis et al. (2012) paired plant distribution information with charcoal recovered from soils beneath them, and found a higher percentage of useful species in plots that contained higher amounts of charcoal in the vicinity of archaeological sites. It is also well documented how phytolith and pollen data complement one another in paleoecological and archaeological reconstructions (Åkesson et al., 2021; Mayle and Iriarte, 2014; Piperno, 2006). Phytoliths tend to be more sensitive to detecting cultivation, forest openings, and some basal angiosperm and eudicotyledon tree taxa, particularly in wet closed canopy forests (as opposed to the savanna ecotone regions), whereas pollen can detect changes in tree taxon abundances that remain undetectable in phytolith analyses (Åkesson et al., 2021; Piperno and McMichael, 2023). Phytoliths identify basal angiosperm and eudicotyledon taxa that pollen does not, and the two are highly complementary when they can be studied together (Piperno and McMichael, 2023). Phytolith analysis has also been paired with starch grain and stable carbon isotope analysis in archaeological settings for a more comprehensive view of the diet and lifestyle of pre-Columbian Indigenous people in Amazonia (Iriarte et al., 2010; McKey et al., 2010). The pairing of proxies, or using multi-proxy approaches, can also aid in detectability of past human influences whereas single proxies may lack detectability of specific lines of evidence.

To move forward, the integration of data needs to occur across proxies, data sources, and consider both spatial and temporal scales (Fig. 4). Either within or between archives, a multi-proxy approach is more sensitive to detecting past human activity than a single-proxy approach, providing more confidence in conclusions on the presence or absence of pre-Columbian Indigenous human activity and the impact that they caused on the landscape (Fig. 1) (Clement et al., 2015; Iriarte, 2016; Mayle and Iriarte, 2014; Piperno, 2006). Recent research has already begun to integrate: (i) lake sediment data with soil survey data from areas within the watershed (McMichael et al., 2012b), (ii) terrestrial soil archives with archaeological sites containing earthworks (Watling et al., 2017), (iii) lake sediment records, even with lakes of different sizes reflecting different source areas, with archaeological sites (Carson et al., 2014; Maezumi et al., 2018b; Whitney et al., 2013), (iv) geospatial patterns of plant distributions with archaeological sites across Amazonia (Levis et al., 2017). As trees in Amazonia can also be directly dated (Brienen and Zuidema, 2006; Chambers et al., 1998; Schöngart et al., 2015), pairing the age of the modern forest with archaeological, paleoecological and historical data could also prove invaluable. In a recent example, Caetano Andrade et al. (2019) integrated dendrochronology and historical data to evaluate Indigenous and traditional management of a Brazil nut (Bertholletia excelsa) stand near an archaeological site south of Manaus.

If the advantages and limitations of data used to infer pre-Columbian Indigenous human activity are recognized and acknowledged, particularly regarding spatial and temporal scale (Fig. 4), then disagreement among existing and future datasets may be minimized. We also highlight the potential and importance of bridging ethnography and ethnoecology with historical records, archaeological data, and paleoecological data. Importantly, but often not considered, the voices and knowledge of Indigenous Peoples should also be integrated into scientific research designs (Trisos et al., 2021). Together, these recommendations can advance the understanding of the complexity and variation of pre-Columbian Indigenous human influences in tropical ecosystems, such as Amazonia.

Author contributions

All authors conceived the ideas. CNHM led the writing, with input from all coauthors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data were used for the research described in the article.

Acknowledgements

CNHM is funded by a European Research Council Starting Grant (ERC-StG 853394) and the Netherlands Organisation for Scientific Research (NWO) award ALWOP.322. CI thanks the Brazilian National Research Council (CNPq) for a doctoral research scholarships (proc. No. 141652/2014–4 and 201573/2014–8) and CNpq and Capes for a post-doctoral fellowships (proc. No. 159440/2018–1, 88887.474568/2020 and 400369/2021–4). ABJ received funding from Terra Preta Program funded by the Interdisciplinary Research and Education Fund (INREF) of Wageningen University, from the Netherlands Organisation for INTERNATIONAL COOPERATION in Higher Education (NUFFIC), and from the European Research Council under an ERC Consolidator Grant (FP7–771056-LICCI). DRP thanks the Smithsonian Institution. We would all like to thank Bernardo Oliveira, Bernardo Flores, Majoi de Novaes Nascimento, and Mark Bush for photo contributions.

References

C.N.H. McMichael et al.
Quaternary Science Reviews 321 (2023) 108354