Acta Botanica Brasilica, 2022, 36: e20220111
doi: 10.1590/1677-941X-ABB-2022-0111

Original article

Estimated degradation of the Caatinga based on modern pollen rain deposited in reservoirs

Marie-Pierre Ledru[1*], Raquel Franco Cassino[2], Vaneicia dos Santos Gomes[3, 4], Julia Caram Sfair[4] and Francisca Soares Araújo[4]

Received: May 01, 2022
Accepted: July 12, 2022

ABSTRACT

The vegetation cover is the result of many biological and abiotic interactions. To identify the different factors that cause changes is crucial when defining future sustainable development and protection of natural resources. In the Brazilian semiarid region, the vegetation cover has been subjected to drastic deforestation and land use at centennial-scale that has led to desertification. Pollen analysis is an efficient tool to reconstruct the different processes of degradation of the vegetation cover over time. We built a referential data set for the vegetation cover using 48 pollen surface samples collected in the reservoirs of the Ceará. We used satellite images for comparison with the pollen signatures and defined an alteration score to express the correlation between terrestrial pollen and anthropic cover. Our results showed our surface samples to be generally representative of the vegetation cover and of the general degradation of the landscape. Our study areas can be considered as degraded as the initial categories "preserved" or "intermediary" are not reflected in the pollen assemblages, in agreement with results of botanical surveys and the historical background. The on-going process of desertification is climate-independent and was initiated many decades or centuries ago by intensive land use for agriculture and grazing.

Keywords: palynology, calibration, tropical dry forest, human impact, Brazilian semiarid region, degradation.

Introduction

Modern vegetation cover is a result of many environmental and physical factors. Disentangling climatic from anthropogenic forcings on vegetation and biodiversity is a priority when planning reforestation and/or conservation policies. Long-term reconstruction of past agricultural activity is based on anthropogenic indicators, for instance *Zea mays* pollen grains or phytolith (Piperno & McMichael 2020), found in sediment cores. Quantitative reconstruction of human-induced landscapes from fossil pollen data in recent decades has enabled the development of a theoretical framework and of models of pollen-vegetation relationships

1 Institut des Sciences de l'Evolution Montpellier, Université de Montpellier CNRS IRD EPHE, Place Eugène Bataillon cc061, 34095 Montpellier, France.
2 Departamento de Geologia, Universidade Federal de Ouro Preto, 35400-000, Ouro Preto, MG, Brazil.
3 Faculdade de Educação, Ciências e Letras do Sertão Central, Universidade Estadual do Ceará, 63902-098, Quixadá, CE, Brazil.
4 Programa de Pós-Graduação em Ecologia e Recursos Naturais, Departamento de Biologia, Universidade Federal do Ceará, 60455-760, Fortaleza, CE, Brazil.

* Corresponding author: marie-pierre.ledru@ird.fr

at the Holocene time scale. However, these studies have primarily focused on temperate ecosystems in the Northern Hemisphere aided by the rich archeological background (*e.g.* Marquer *et al.* 2017; Harrison *et al.* 2020). In the Brazilian semiarid region (BSA), the vegetation cover has been subjected to drastic deforestation and land use at centennial scale and the lack of historical archives makes it impossible to fully evaluate the successive states that led to desertification in some areas (see map of the FUNCEME in CGEE 2016). Even though we cannot ignore the possibility that hunter-gatherer societies may also have impacted the landscape, anthropogenic and paleo studies showed that they were keen on preserving their resources (Bliege *et al.* 2008; Morrison *et al.* 2022). In the BSA region, over the last three centuries, extensive agriculture has been practised for commercial purposes with little concern for the sustainability of the resources. Analyzing historical changes in land use/land cover is essential to understand current human-environment dynamics and more specifically, the different steps that lead to desertification in a poor but densely populated region. There is a drastic lack of historical or natural archives to estimate changes in the landscape at millennial or decadal scales in tropical semiarid regions (Flantua *et al.* 2015). The absence of historical records is due to the fact that transmission was mostly oral, but also logical given that paleoecological reconstructions require permanent lakes or swamps for the preservation of the biological indicators. Changes in the landscape in a highly anthropized region of the BSA region, Vale do Acaraú, Ceará, were recently reconstructed from sediment cores collected in reservoirs (Ledru *et al.* 2020). They showed that at decadal scale, droughts do not affect the vegetation due to the resilience of the ecosystem whereas public policies have had a strong impact on the composition of the vegetation and biodiversity. To calibrate the responses of the vegetation to disturbances, pollen analysts usually infer modern pollen rain studies (Chevalier *et al.* 2020). However, the lack of such studies in the Caatinga biome (Flantua *et al.* 2015), the dry deciduous forest of the BSA, limited our interpretation (Ledru *et al.* 2020). In addition, when a relationship exists between vegetation and its associated pollen production/deposition, such relationships are usually analyzed in areas that have undergone little disturbance (*e.g.* Amaral *et al.* 2006; Gosling *et al.* 2009; Montade *et al.* 2016). Moreover, when dealing with mosaics, different vegetation (*e.g.* biomes) and land use (*e.g.* farming) conditions in space and over time means systematic approaches cannot be used (Deza-Araujo *et al.* 2020).

Soil samples or samples from pollen traps are usually used to build referential data sets. However, pollen traps were not used in the BSA region because they require regular collection over several months or years before a representative data set is obtained (*e.g.* Portes *et al.* 2020). Another problem with soil sampling in semi-arid regions is that high temperatures and low soil moisture prevent good preservation of the pollen grains in soil samples. However, in the BSA region, the large number of reservoirs built to retain the water during the long dry season created local permanent moisture conditions able to retain the pollen rain produced by the surrounding vegetation (Markgraf *et al.* 2002). The construction of reservoirs in Northeastern Brazil started in 1890 (Assunção & Livingstone 1993) and today there are more than 5,000 reservoirs from small to average capacity in the State of Ceará (COGERH). Before colonization, the tropical dry forest, Caatinga, covered 844,000 km^2, but was subsequently subjected to large-scale deforestation for wood extraction, farming and agriculture, and today a large area is under desertification (CGEE 2016; IPBES 2018). More than 40 % of the vegetation has been deforested (Overbeck *et al.* 2015), while protected areas account for only 9 % of the Caatinga region (Brasil 2000). Only 1.8 % of the Caatinga is fully protected (IUCN categories I to III), while 7.2 % is under sustainable use in which the direct use of natural resources is permitted according to Brazilian legislation (Brasil 2000). Given the current climate change and more specifically the predicted increase in temperature (IPCC 2018) it is urgent to characterize the degree of preservation required for the recovery and/or the protection of the Caatinga. In this context, we present a new data set built using reservoirs as potential pollen traps with the aim of defining the relation between vegetation degradation and pollen deposition and more particularly to characterize the degradation of the semi-arid vegetation using alteration scores.

Materials and methods

Modern vegetation and climate

The phytogeographical domain of the Caatinga is characterized by a hot semiarid climate (Bsh) according to Köppen's classification (Peel *et al.* 2007), mean annual temperature is 28 °C and mean annual precipitation ranges between 500 and 800 mm. The rainy season lasts from January to May, although with frequent inter- and intra-annual variability related to changes in the surface temperature of the tropical Atlantic and the El Niño Southern oscillation (Barbosa *et al.* 2006). For instance, 13 drought events have been recorded in the BSA region in the last 60 years (Marengo & Bernasconi 2015).

Two natural reserves in the crystalline Caatinga were created on ancient farming areas. Thirty-four botanical surveys were performed in the two reserves (Costa *et al.* 2007, Araujo *et al.* 2011, Moro *et al.* 2014). They showed the high frequency of the herbaceous layer among the plant communities. Among the 10 main herbaceous species were *Melochia tomentosa* (Malvaceae), *Tacinga palmadora, Tacinga inamoena* (Cactaceae), *Mollugo verticillata* (Molluginaceae), *Neoglaziovia variegate* (Bromeliaceae), *Waltheria americana* (Malvaceae), *Alternanthera brasiliana* (Amaranthaceae), *Delilia biflora* (Asteraceae), *Sida galheirensis* (Malvaceae), *Euphorbia hyssopifolia* (Euphorbiaceae). Among the main woody species are *Aspidosperma pyrifolium* (Apocynaceae), *Commiphora leptophloeos* (Burseraceae), *Jatropha mollissima* (Euphorbiaceae), *Cereus jamacaru* (Cactaceae), *Myracrodruon urundeuva* (Anacardiaceae), *Anadenanthera colubrine* (Fabaceae), *Bauhinia cheilantha* (Fabaceae), *Mimosa tenuiflora* (Fabaceae), *Poincianella pyramidalis* (Fabaceae), *Cynophalla flexuosa* (Capparaceae), *Spondias tuberosa* (Anacardiaceae), *Piptadenia stipulacea* (Fabaceae), *Schinopsis brasiliensis* (Anacardiaceae), *Libidibia ferrea* (Fabaceae), *Combretum leprosum* (Combretaceae), *Ziziphus joazeiro* (Rhamnaceae), *Pilosocereus gounellei* (Cactaceae), *Croton sonderianus* (Euphorbiaceae), *Lantana camara* (Verbenaceae), *Croton heliotropiifolius* (Euphorbiaceae), *Amburana cearenses* (Fabaceae) (Moro *et al.* 2014).

For this study, we selected 48 reservoirs located in the State of the Ceará, one of the driest semi-arid regions in Northeastern Brazil (Fig. 1, Tab. 1). The reservoirs are located within the crystalline basement where Caatinga vegetation grows in shallow, stony or fertile soils (Sampaio 1995). Our samples were collected in the top 2 cm of the soil-water interface in three different locations around the reservoir, at the end of the 2015 dry season. Studies performed in ponds showed that each sample represented between 5 and 20 years of pollen deposition (McLauchlan *et al.* 2007).

Pollen analyses

For each sample, we separated a volume of 2 cm^3 for pollen extraction following the method of Faegri & Iversen (1989). Samples were treated with fluorhydric acid (HF 70 %) for 12 hours to dissolve the silicate and silica, rinsed with 50 % HCl in a bain-marie at 80 °C to remove the silica colloids, then with 10 % KOH in a bain-marie at 80 °C to remove the humic acids and finally acetolysis with a 9:1 preparation of acetic anhydride and sulfuric acid. A spike (*Lycopodium clavatum*) was added to each sample to calculate the concentration (Stockmarr 1971). The pollen grains were identified and counted on microscope slides under an optical microscope at x60 magnification. A total of 300 grains of arboreal and non-arboreal pollen were counted in each sample, excluding aquatic and water-level related pollen and spores taxa. Eight samples were discarded due to their poor pollen content. Pollen grains were identified with the help of the pollen reference collections belonging to the Federal University of Ceará, Brazil, to the University of Montpellier, France, pollen atlases and catalogs covering different South American flora (*e.g.* Salgado-Labouriau 1973; Barth 1989; Silva 2007; Silva *et al.* 2016). Broken and deformed pollen grains were classified as indeterminate.

Reconstruction of the vegetation cover

The vegetation cover at each sampling site was characterized according to Mapbiomas (3rd collection) classification where the terms forest, savanna, grassland only relate to the vegetation cover and not to other biomes. The Caatinga domain is represented by both forested and non-forested formations; the term forest implies a continuous vegetation cover, while the term non-forested means the vegetation is dominated by herbaceous species and floodable areas. At each site, the vegetation cover was described in three areas (with a radius of respectively, 300 m, 500 m, and 1 km) around the collection point, and vegetation data were converted into percentages (surface areas of the water bodies were excluded when calculating the percentage).

Statistical analyses

Pollen counts were also converted into percentages; the total pollen sum used to calculate the percentage is represented by the sum of all pollen taxa except aquatic or water-level related taxa (*Agalinis*, *Cleome*, *Cuphea*, Cyperaceae, *Echinodorus*, *Ludwigia*, *Salvinia*, *Typha*, *Utricularia*) (Matias *et al.* 2021). Pollen frequencies are presented in diagrams drawn with C2 software (Juggins 2007).

We defined an alteration score to express the correlation between the percentage of terrestrial pollen and the percentage of anthropic cover around the sites according to the following method: first, pollen types were classified in two groups: anthropic vegetation taxa and natural vegetation taxa (Tab. 2). Pearson correlation coefficients were then calculated between the percentage of the individual pollen taxa and the percentage of anthropic cover. Anthropic vegetation taxa that were positively correlated and natural vegetation taxa that were negatively correlated with the percentage of anthropic cover were then used to calculate the alteration score.

The alteration score was used to calculate a linear regression model to reconstruct the area of anthropic cover from pollen percentages. For model calculations, the dataset was randomly split into a training dataset containing 80 % of the samples, and a test dataset containing the remaining 20 %. Linear regression coefficients were calculated first using the training dataset; the remaining samples (20 %) were then used to test the model and to calculate prediction errors. Accuracy was measured by calculating mean absolute errors (MAE). The model was then applied to the complete dataset to compare the predicted and real percentages of land cover.

To test the application of our results to fossil pollen data, we applied the linear regression model to a 60 cm sediment core collected in the reservoir of Araras near the city of Varjota (Ceará, Brazil) (4°11’59’’S, 40°27’59’’W) dated between 1959 and 2002 AD. (Ledru *et al.* 2020).

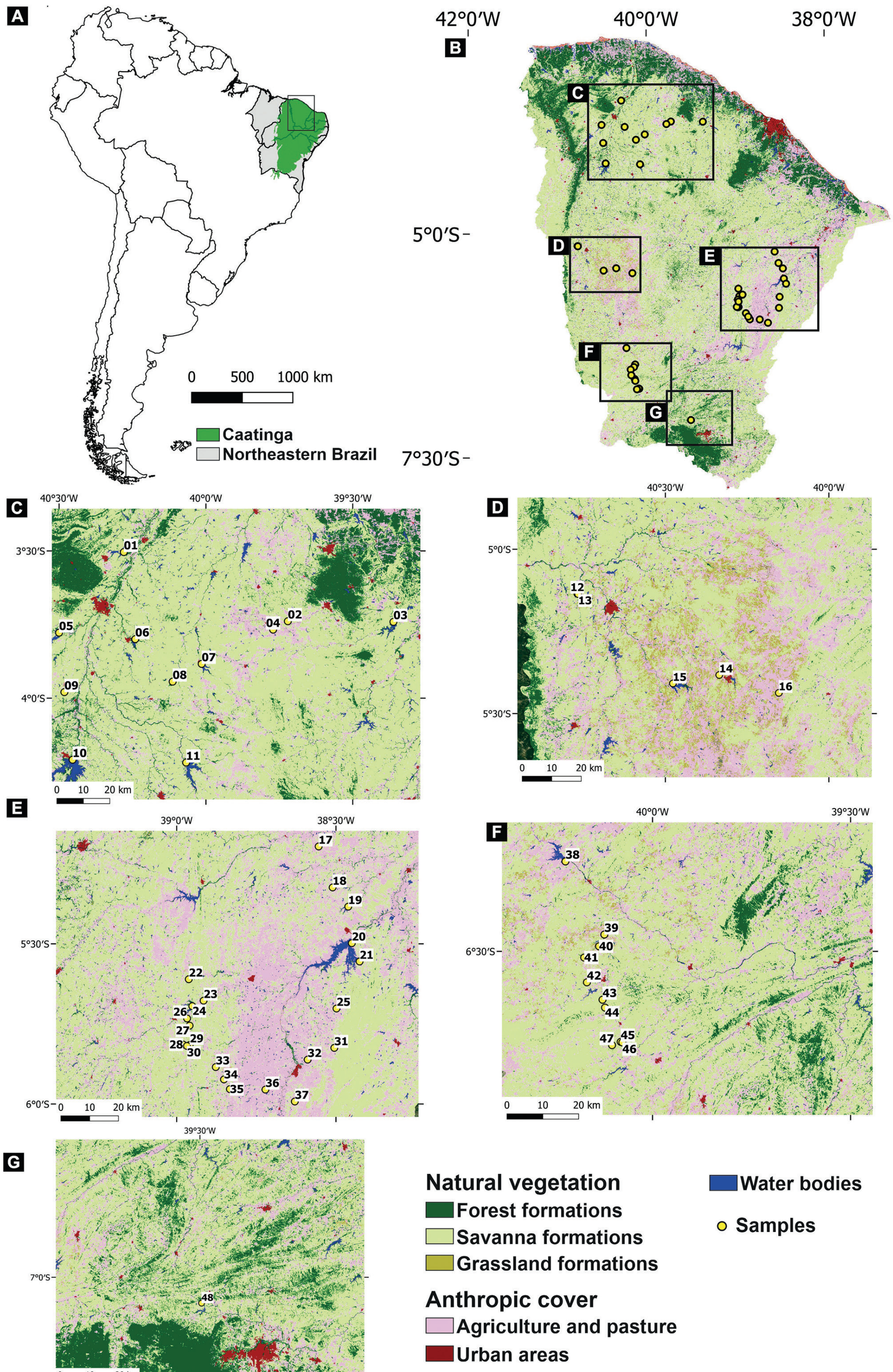

Figure 1. A. Map of South America showing the location of the State of Ceará. **B**. Map of the vegetation cover of the State of Ceará showing the location of the samples collected for this study. **C, D, E, F, G**. Zooms on the location of the samples and associated vegetation cover (Tab. 1).

Table 1. List of the 48 samples analyzed in the present study with their location and infor mation concerning the degree of preservation; D: degraded, P: preserved, IT: intermediate, and the nature of the surroundings.

Sample number	Location	Name	State	Vegetation
1	-3,503812 -40,279428	Acaraú-mirim	IT	*Neptunia* sp., *Cleome* sp., *Cryptostegia* sp., *Heliotropium* sp., *Tephrosia* sp., *Borreria* sp., Nymphaeaceae, *Angelonia* sp., *Attalea speciosa* Mart. ex Spreng on the hilltop
2	-3,738653 -39,720979	Saco Verde	IT	Open vegetation with abundant *Prosopis juliflora* (Sw.) DC. and *Cordia oncocalyx* Allemão
3	-3,740844 -39,361434	Caixitoré	IT	Cultivated plants: coconut, banana, *Azadirachta indica* A.Juss., *Cleome* sp., *Ipomeas* sp., *Chamaecrista* sp., cattle breeding and aquaculture
4	-3,76721 -39,771146	Mocó	D	Open vegetation with *Cryptostegia madagascariensis* Bojer ex Decne, *Jatropha* sp., *Prosopis juliflora* (Sw.) DC
5	-3,777572 -40,500035	Jaibaras	D	Caatinga with 2 to 3 years preservation. treelets with *Calotropsis procera* (Aiton) W.T.Aiton, *Heliotropium sp.*, Lamiaceae, *Borreria* on the margins
6	-3,799482 -40,240663	Forquilha	D	Open vegetation Grass for fodder and beans
7	-3,883624 -40,014361	Aracatiaçu	D	Open vegetation Cultivated area with corn, beans, grass, cattle-breeding
8	-3,943714 -40,113189	EMASA	D	Open vegetation with *Cleome* sp. on the margins of the reservoirs, *Senna alata* (L.) Roxb.
9	-3,979371 -40,481917	Jaibaras-Varjota	IT	Degraded Caatinga
10	-4,208846 -40,453924	Araras	D	Open vegetation
11	-4,218465 -40,067361	Edson Queiroz	IT	No habitation
12	-5,135178 -40,769448	Sobradinho	D	Open vegetation with Nympheaceae, *Trema* sp., the palm *Copernicia prunifera* (Mill.) H.E.Moore on the margins and *Licania, Anadenanthera colubrina* (Vell.) Brenan, *Cynophalla flexuosa* (L.) J.Presl, *Ziziphus joazeiro* Mart., *Calliandra spinosa* Ducke in the landscape
13	-5,136582 -40,766436	Crateus	D	Open vegetation
14	-5,382674 -40,335292	Independência	D	Open vegetation
15	-5,408405 -40,476955	Jaburu	D	Open vegetation. *Libidibia ferrea* (Mart. ex Tul.) L.P.Queiroz, *Cnidoscolus quercifolius* Pohl, *Commiphora leptophloeos* (Mart.) J.B.Gillett, *Aspidosperma pyrifolium* Mart. & Zucc., *Ziziphus joazeiro* Mat., Malvaceae, *Jatropha* sp., *Prosopis juliflora* (Sw.) DC., *Parkinsonia aculeata L.*, *Senna alata* (L.) Roxb., *Cleome* sp., *Piloscereus* sp.

Table 1. Cont.

Sample number	Location	Name	State	Vegetation
16	-5,437099 -40,15313	Independência - Tauá	D	Open vegetation, on the road margin
17	-5,292696 -38,921363	Banabuiú	D	Open vegetation
18	-5,32387 -38,510461	Poço Redondo	D	Inside a farm, deforested and burnt landscape
19	-5,383877 -38,461477	Poço de Barro	D	Reservoir since 1950
20	-5,497965 -38,450226	Castanhão. -areia	IT	Secondary vegetation
21	-5,555133 -38,425808	Castanhão - lama	IT	Secondary vegetation
22	-5,610476 -38,962084	Japão	P	Deciduous forest
23	-5,67738 -38,915552	Sítio Pasta	IT	Degraded area
24	-5,69445 -38,950995	Boqueirão 20	IT	Secondary vegetation. Reservoir built in 1915.
25	-5,701907 -38,498357	Castanhão-BR	IT	Secondary vegetation
26	-5,732898 -38,967316	Junco	D	On the road margins
27	-5,755312 -38,959446	São José	D	Aquaculture. *Mimosa tenuiflora*
28	-5,814402 -38,979705	Alfredo	IT	*Commiphora leptophloeos* (Mart.) J.B.Gillett, *Luetzelburgia auriculata* (Allemão) Ducke, *Mimosa* sp., *Cordia oncocalyx* Allemão
29	-5,81586 -38,959298	Caldeirões	IT	Secondary vegetation. Macrophytes and abundant *Mimosa tenuiflora*
30	-5,818772 -38,967736	Açude 805	D	
31	-5,824815 -38,505206	Seixo	D	Invasives and *Piptadenia* sp.
32	-5,861444 -38,588897	açude Estrada	D	Desert, exposed soil
33	-5,884668 -38,877293	Jaguaribe 1	D	

Table 1. Cont.

Sample number	Location	Name	State	Vegetation
34	-5,923215 -38,851595	Jaguaribe 2	D	
35	-5,953434 -38,833031	Feiticeiro	D	Open vegetation. *Ipomea* sp., *Cleome* sp.
36	-5,954901 -38,721432	Loca d'Agua	D	Open vegetation with exposed soil
37	-5,992088 -38,629156	Guariba	D	Near the road
38	-6,274404 -40,219854	Arneiroz	D	Open vegetation
39	-6,457772 -40,121623	Aiuaba	D	*Aspidosperma pyrifolium* Mart. & Zucc., *Jatropha* sp, *Ipomea* sp, *Mimosa caesalpiniifolia* Benth, *Mimosa tenuiflora* (Willd.) Poir., *Helicteres* sp., *Cnidoscolos quercifolius* Pohl.
40	-6,487218 -40,135008	Murzela	IT	Preserved vegetation with *Cnidoscolos quercifolius* Pohl, *Ziziphus*, *Anadenanthera colubrina* (Vell.) Brenan, cactaceae, *Licania rigida* Benth., *Aspidosperma pyrifolium* Mart. & Zucc., *Crateva tapia* L., *Handroanthus* sp., *Astronium urundeuva* (M.Allemão) Engl., *Croton* sp, *Cenostigma pyramidale* (Tul.) Gagnon & G.P.Lewis, *Cassia* sp., *Cynophalla flexuosa* (L.) J.Presl, *Clitoria* sp., *Pseudobombax marginatum* (A.St.-Hil., Juss. & Cambess.) A.Robyns
41	-6,515474 -40,172484	Andressa	IT	One side deforested, the other side secondary vegetation with *Jatropha* sp., *Anadenanthera colubrina* (Vell.) Brenan, abundant Cyperaceae
42	-6,577638 -40,165643	Autorama	D	*Heliotropium sp.*, *Anadenanthera colubrina* (Vell.) Brenan, *Aspidosperma pyrifolium* Mart. & Zucc., *Schinopsis brasiliensis* Engl., *Croton* sp.
43	-6,621212 -40,126234	Tonico	P	Deciduous forest with *Ipomea sp*, *Anadenanthera colubrina* (Vell.) Brenan, *Mimosa tenuiflora* (Willd.) Poir., *Croton* sp, *Ziziphus joazeiro* Mart., *Licania rigida* Benth., *Peltophorum dubium* (Spreng.) Taub, *Clitoria* sp., *Indigofera* sp., *Sida* sp., *Cleome* sp.
44	-6,640991 -40,119839	Boqueirão 45	P	Deciduous forest
45	-6,7268 -40,08028	São Raimundo	D	30-year old reservoir
46	-6,728287 -40,076116	Pista São Raimundo	IT	Nearby vegetation slashed and burned. On the other side of the road vegetation preserved for at least 25 years containing *Ipomea* sp., *Cleome* sp., *Bidens* sp.
47	-6,735128 -40,102142	Catolezeiro	D	Secondary vegetation
48	-7,077822 -39,495017	Cotia	D	

Table 2. List of the pollen taxa used for the calculation of the alteration score.

Anthropic indicators	Asteraceae Liguliflorae, Mollunginaceae, Urticaceae, Amaranthaceae, *Acalypha*, *Cuphea*, Anacardiaceae, *Gomphrena*, *Sebastiana*-t, Convolvulaceae, *Spermacoce*, Acanthaceae, *Ipomea*, *Evolvulus*, Boraginaceae, *Agalinis*, *Orbignya*, Solanaceae, Gymnosperma, Lamiaceae, *Zea mays*
Natural vegetation	*Chloroleucon, Myracrodruon urundeuva, Cassia, Poincianella pyramidalis, Mimosa acutistipula, Pithecelobium, Cordia, Amburana cearensis, Piptadenia, Salvinia, Diodella, Mimosa caesalpinifolia,* Alstroemeriaceae, *Ziziphus joazeiro,* Bignoniaceae, *Angelonia*-t, *Himatanthus, Copaifera,* Asteraceae

Results

Modern vegetation and degree of anthropic alteration (MapBiomas data)

In 33 of the 48 sites, the percentage of anthropic cover within the 300 m radius was higher than the percentage of natural vegetation (Fig. 2). In most sites, anthropic cover was characterized by pasture and agriculture. Among the natural vegetation, savanna (intermediary vegetation) was the most abundant type of vegetation in most sites.

Modern pollen assemblages

Among the 90 pollen taxa identified, 35 were arboreal pollen taxa (39 %); 36 non-arboreal pollen taxa (40 %); 10 indeterminate (11 %); 9 aquatic pollen taxa (10 %), all displaying good pollen preservation. The percentage of arboreal pollen taxa varied from 16 % to 67 %. The arboreal pollen taxa were dominated by Fabaceae, Anacardiaceae, Combretaceae, Apocynaceae and Euphorbiaceae. Fabaceae represented up to 64 % in sample 25, *Mimosa tenuiflora* was the dominant taxon in all the samples accounting for from 6 % to 37 %. The family Euphorbiaceae accounted for 41 % mainly by *Croton*. *Croton* was found in 19 samples at percentages of more than 1 % (between 2 % and 3 %) in samples 18, 30, 32, 33, 34, 47. Combretaceae was mainly represented by *Combretum* (5 % in sample 9), Anacardiaceae represented 10 % of pollen in sample 7 and 7 % in samples 14 and 16 (Fig. 3). Bignoniaceae (5 % in sample 47) and *Ziziphus joazeiro* in 7 samples were also observed but at lower percentages. Non-arboreal pollen taxa were present at high percentages (> 50 %) in all the samples (84.8 % in sample 36) except in samples 21 and 16 where they represented 35 %. The most abundant non-arboreal taxa were Poaceae and *Alternanthera*, which were both present in all samples at percentages ranging between 2 % (sample 4) and 30 % (sample 48), and 1 % (sample 37) and 41 % (sample 2), respectively. Other non-arboreal pollen taxa were dominated by Rubiaceae, Malvaceae and Asteraceae. Pollen grains of Rubiaceae (*Borreria* and *Mitracarpus*) were present in all the samples except four (15, 38, 47, 41) at percentages ranging between 1 % and 28 %. Asteraceae was represented by the group of the "*liguliflorae*" 3 % in sample 4, and "*tubiflorae*" 19 % in sample 43. Malvaceae represented up to 23 % (sample 45) of non-arboreal pollen. The highest percentages of non-arboreal pollen taxa were *Alternanthera* (41 %), *Mitracarpus* (18 % in sample 18) and *Borreria* (15 % in sample 3). Convolvulaceae was represented by three taxa *Evolvulus, Ipomea* (6 % in sample 39) while *Jacquemontia* was only found at low percentages in a few samples. Poaceae ranged between 2 % and 29 % with some pollen grains of *Zea mays* observed in 23 samples, at percentages ranging between 1 % and 3 % (the latter in sample 43). Low percentages (0.3-4 %) of allochthonous taxa of gymnosperms (*Podocarpus* and *Pinus*) were observed in 22 samples, the highest percentages being 4 % in sample 20 and 3 % in sample 4 (Fig. 3). Among the taxa observed only in one sample were Meliaceae (1 % in sample 2), *Ipomea* (6 % in sample 39), *Jatropha* (1 % in sample 37), *Amburana cearensis* (2 % in sample 6), Boraginaceae (1 % in sample 12) and Polygalaceae (1.65 % in sample 37).

Relationship between vegetation and pollen data

In the samples with high percentages of anthropogenic cover, pollen assemblages tended to present high frequencies of Amaranthaceae, *Acalypha*, *Gomphrena*, and Anacardiaceae (Fig. 4A-I). For Amaranthaceae, *Acalypha* and *Gomphrena*, we found a more significant correlation of pollen frequencies with the percentage of "pasture+agriculture" cover than with the total anthropic cover (*i.e.*, including urban areas) (Fig. 4A, C, E compared with 4B, D, F). The best correlation was found between the percentage of Amaranthaceae and the percentage of "pasture+agriculture", however, no pollen type alone was found to be a good predictor of the degree of anthropic cover. The total percentage of non-arboreal pollen, the total percentage of arboreal pollen, the AP/NAP ratio, and the Poaceae frequencies showed no correlation with the percentage of anthropic cover (Fig. 4I-L).

Figure 2. Land cover data (from Mapbiomas 2019) reconstructed at 1,000 m, 500 m, 300 m for each weir. Sites are organized according to latitude. In the legend, Forest, Savanna and Grassland refer to vegetation cover that is respectively, continuous, intermediary, and open.

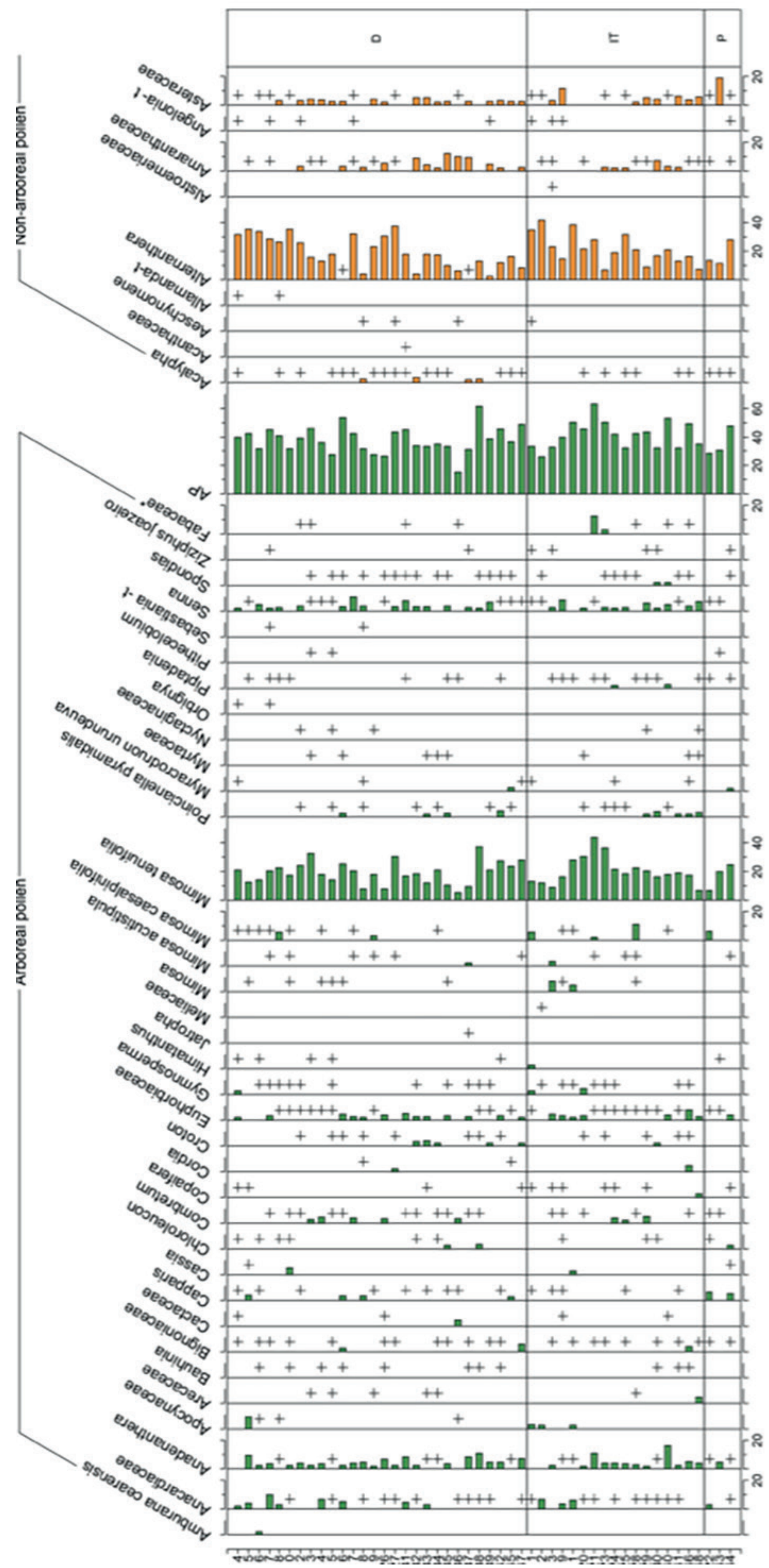

Figure 3. Diagram showing the percentages of the main taxa identified in the 48 surface samples. Green - arboreal pollen taxa; orange - non-arboreal pollen taxa; blue - aquatic or water-level related taxa. Sites are organized according to the degree of degradation delimited by horizontal black lines with from top to bottom: D: degraded, P: preserved, IT: intermediate.

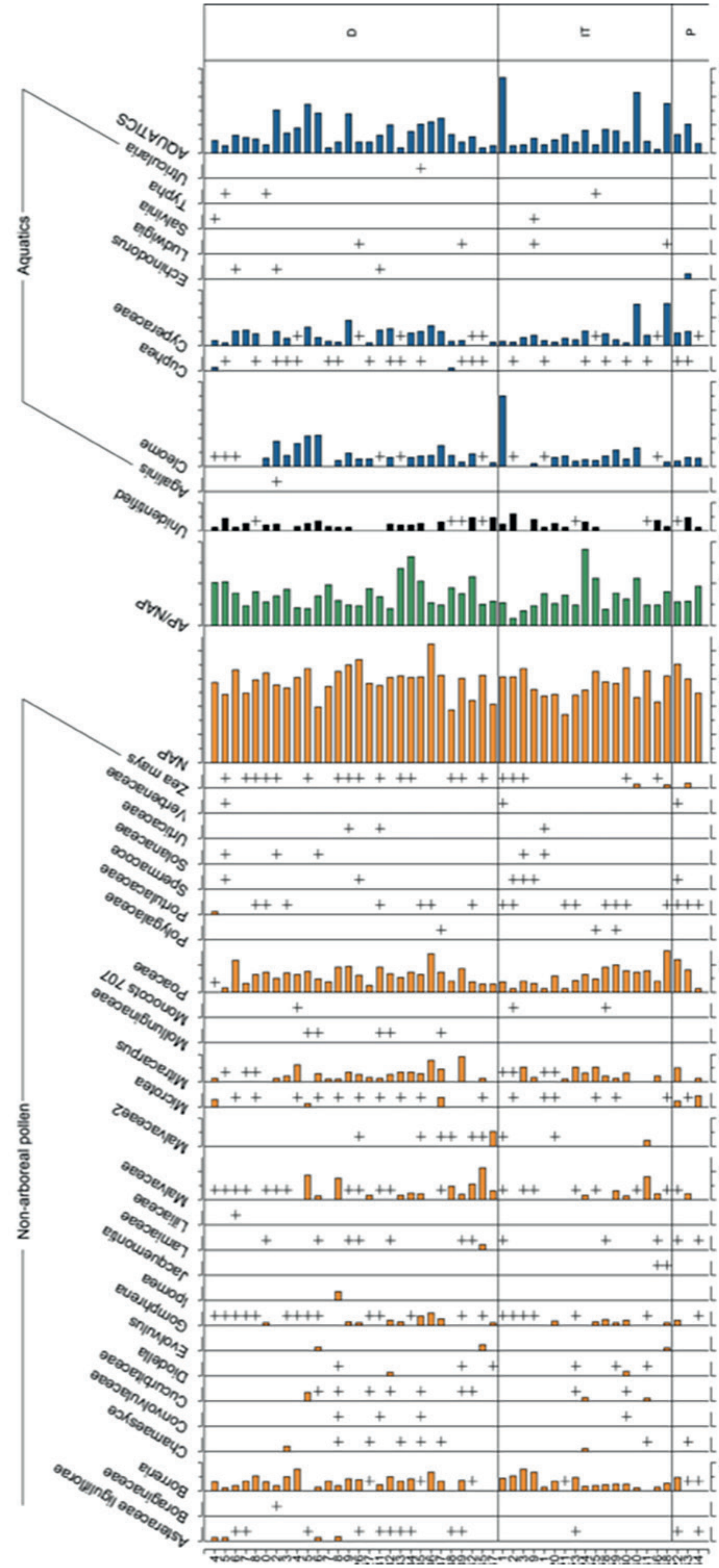

Figure 3. Cont.

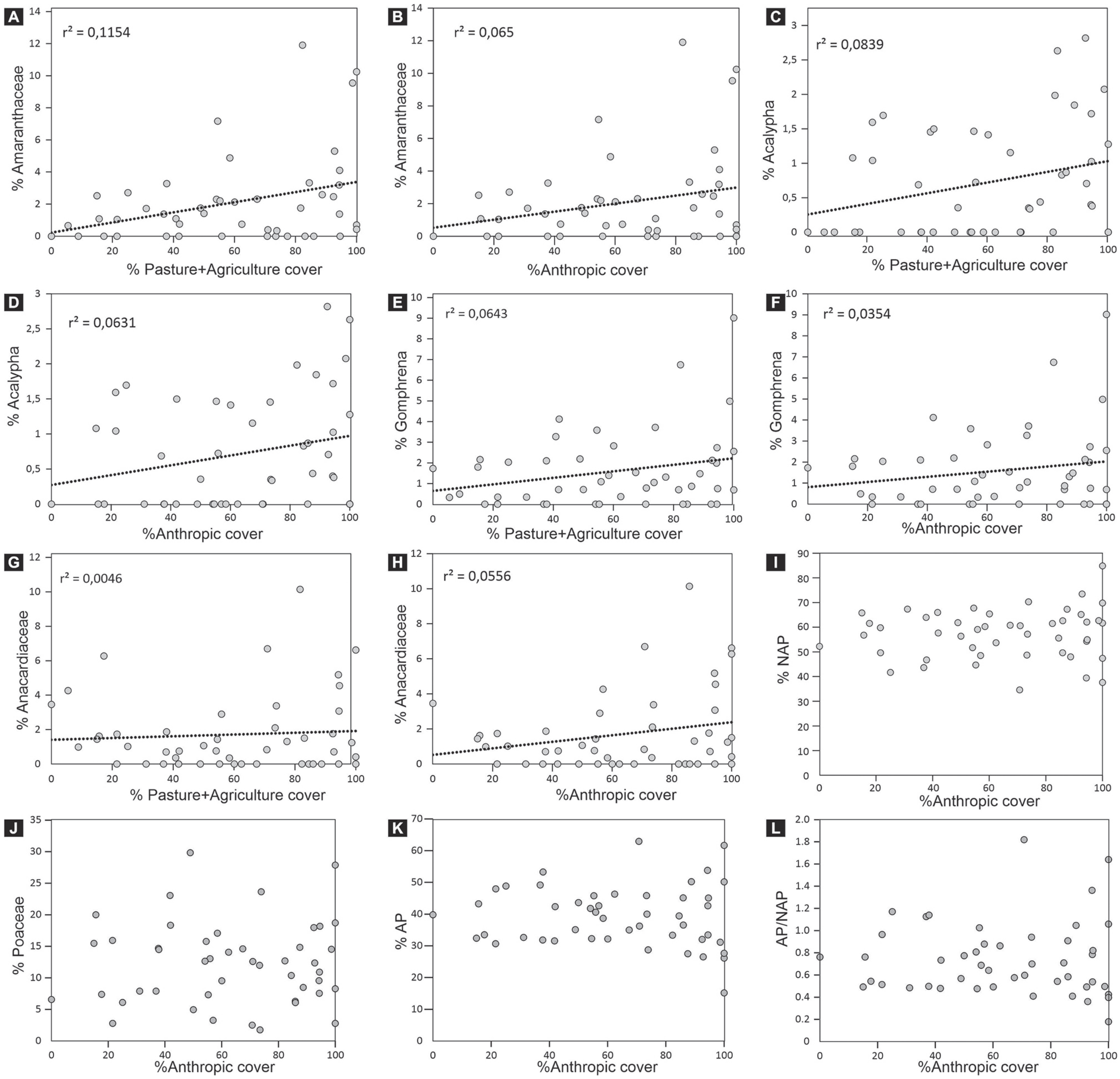

Figure 4. Relationship between pollen frequencies and the percentage of anthropic cover in the area with a radius of 300 m around the sites, considering the percentage of "pasture + agriculture" cover (A, C, F, H) or the total percentage of anthropic cover (*i.e.*, including pasture, agriculture, and urban areas) (B, D, E, G, I, J, K, L).

However, to predict the percentage of anthropic cover, it is possible to estimate the degree of anthropic disturbance from pollen assemblages when considering the frequencies of multiple pollen types. Hence, we listed and summarized the "anthropic indicators" and the "natural vegetation indicators" (Tab. 1).

The total anthropic indicators tended to be higher at sites with higher percentages of anthropic cover and the natural vegetation tended to be higher at sites with lower percentages of anthropic cover (Fig. 5).

Our definition of an "alteration score" is based on the following calculation: "sum (Anthropic indicators)" - "sum (Natural vegetation)". The "alteration score" was higher at sites with higher percentages of anthropic cover (Fig. 5) and can thus be used to predict the degree of anthropic disturbance around the sites.

The correlation between the alteration score and the percentage of anthropic cover around the sites was significant (Fig. 5) and was highest for the area with a radius of 300 m. (%Anthropic cover includes "pasture",

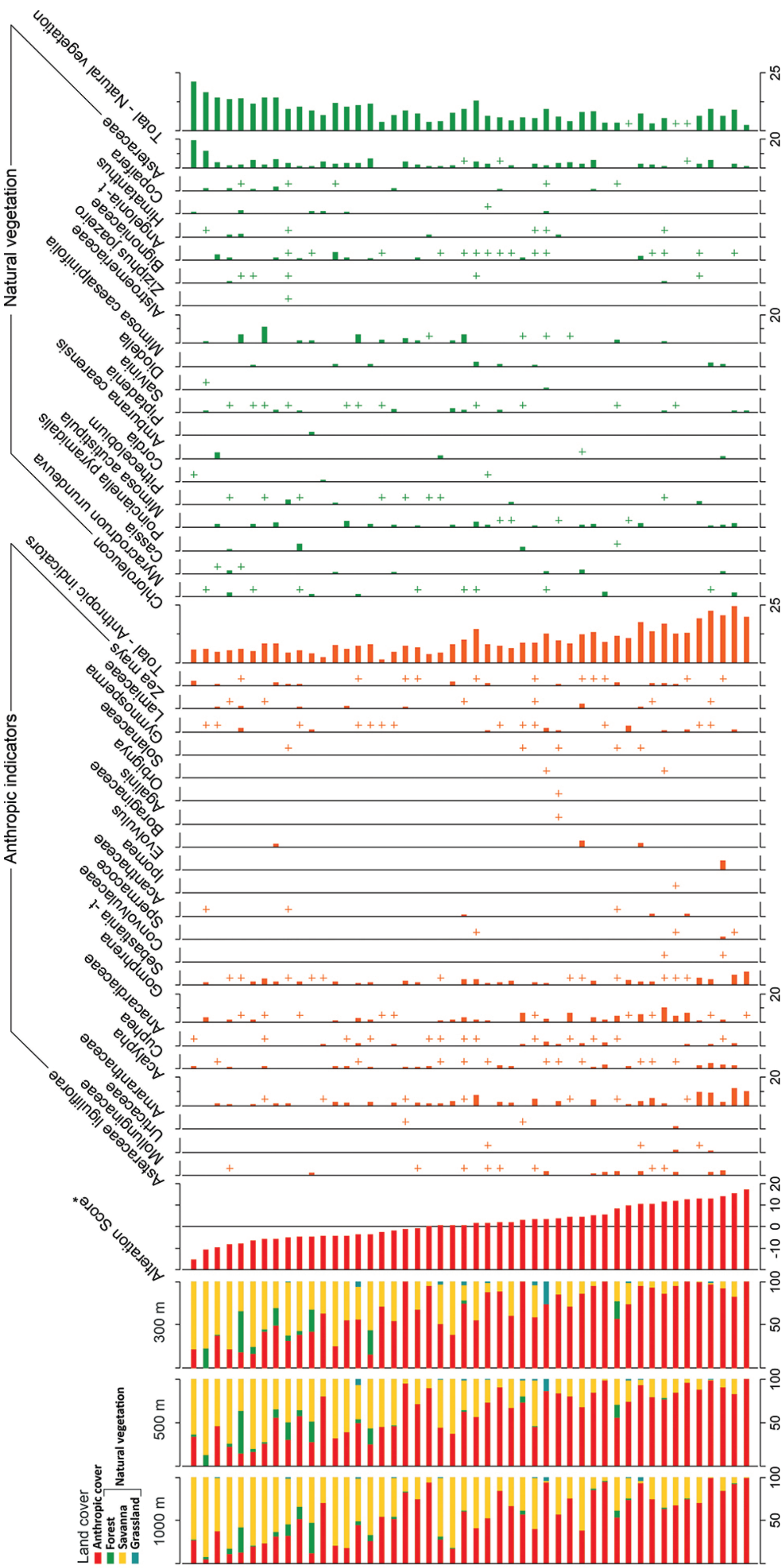

Figure 5. Synthetic diagram showing the results obtained from Mapbiomas (Fig. 2) and from the pollen analyses (Fig. 3) with anthropic indicator taxa (orange) and natural vegetation indicator taxa (green). Sites are organized according to the degree of degradation as defined by the alteration scores. Numbers refer to Table 1.

"agriculture", "pasture-agriculture", "urban areas", "other urban areas" and "rocky outcrops" from MapBiomas):

Reconstructing the percentage of anthropic cover using the pollen data - regression model

The alteration score was used to build a linear regression model and to reconstruct the percentage of anthropic cover in the area with a radius of 300 m from the point where the pollen was sampled. To calculate the model, the dataset was randomly split into training and test datasets with 80 % and 20 % of samples, respectively. Linear regression coefficients were calculated using the training dataset. The remaining samples (20 %) were used to test the model and to calculate prediction errors. Accuracy was measured by calculating mean absolute errors (MAE) (Fig. 6).

The regression model equation obtained was:

$$\text{AnthropicCov} = 3.17 \times \text{AlterationScore} + 60.04 \quad (1)$$

The model was applied to the complete dataset in order to compare the predicted and real percentages of anthropic land cover.

For 27 samples, the real percentage of anthropic cover fell within the margin of error of the predicted value (Tab. 3). For 10 samples, the real percentage of anthropic cover was less than 10 % higher or lower than the predicted value (including the error margin), and for the other 10 samples, the real percentage of anthropic cover was more than 10 % higher or lower than the predicted value (including the error margin) (Fig. 7). For at least some of the 10 samples whose percentage of anthropic cover was over- or underestimated by more than 10 % (10 % and 31 %), the difference between real and predicted values might be related to the presence of large water bodies within the area with a radius of 300 m, as water surfaces were not considered in the land cover.

Application to fossil pollen data

Application of the regression model to the ARA core (Ledru *et al.* 2020) indicated that in this area, the percentage of anthropic cover ranged from 62.6 % to 0 % (Fig. 8). Higher percentages of anthropic cover were found in the bottom part of the core that corresponded to the period when the dam was commissioned. A gradual decrease in anthropic cover was observed in the years following the building of the dam demonstrating recovery of the vegetation. Another increase in anthropic activities was clearly visible with the launching of the Araras Norte irrigation project in 1985. In the upper part of the core, natural vegetation predominated, evidence for recovery of the landscape. However, our results also revealed the limits of our surface sample dataset as today, the vegetation today is quite different from the original cover but is nevertheless characterized as "natural" by the alteration score. Our attempt to apply the score to the most recent decades also highlighted the need for additional surface samples from the native Caatinga to improve the calibration. However, we are confident that, once completed, our approach and the use of the alteration score will make it possible to characterize the state of the landscape in semi-arid regions.

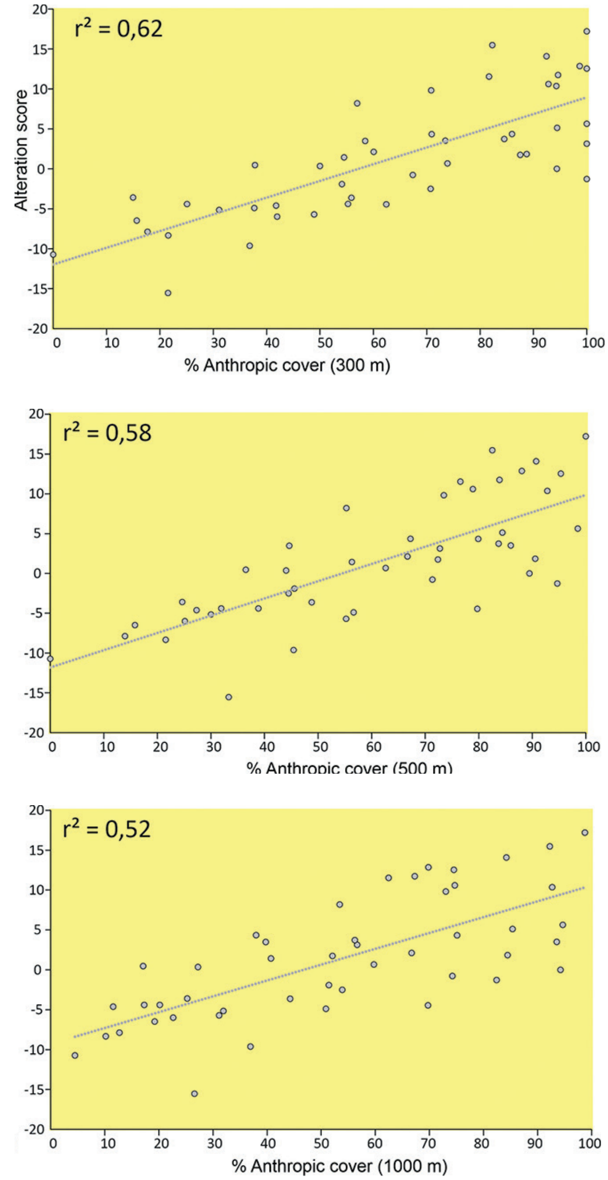

Figure 6. Regression curves between alteration scores and percentages of anthropic cover in areas with a radius of 300 m, 500 m and 1,000 m.

Table 3. Results of the regression model performed on our sample dataset.

Linear regression				
Coefficients		Std. Error	Pr(>|t|)	MAE
Intercept	60.04	3.01	2E-16	12.6
Regression coef.	3.17	0.39	1.85E-09	

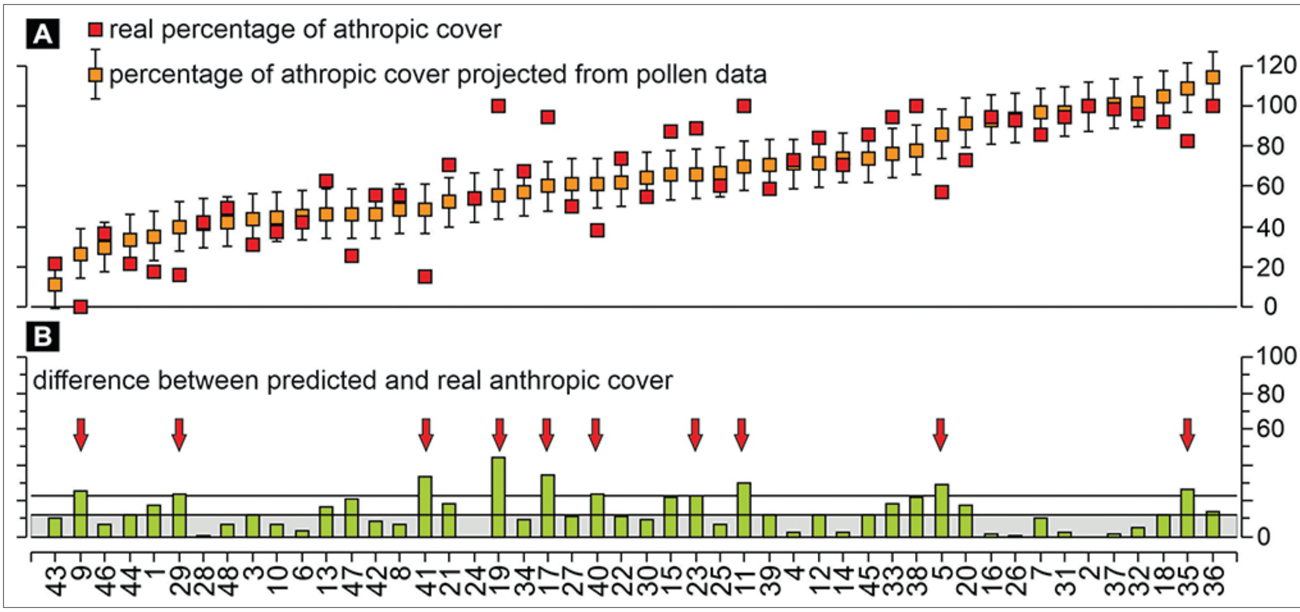

Figure 7. A) Calculated values of the percentage of anthropic cover (orange) and real values of anthropic cover (red). B) Difference between real and predicted (mean) values of anthropic cover. The grey area indicates the model error margin (12 %) and red arrows indicate samples for which the difference between real and predicted anthropic cover was more than 10 % higher than the margin of error.

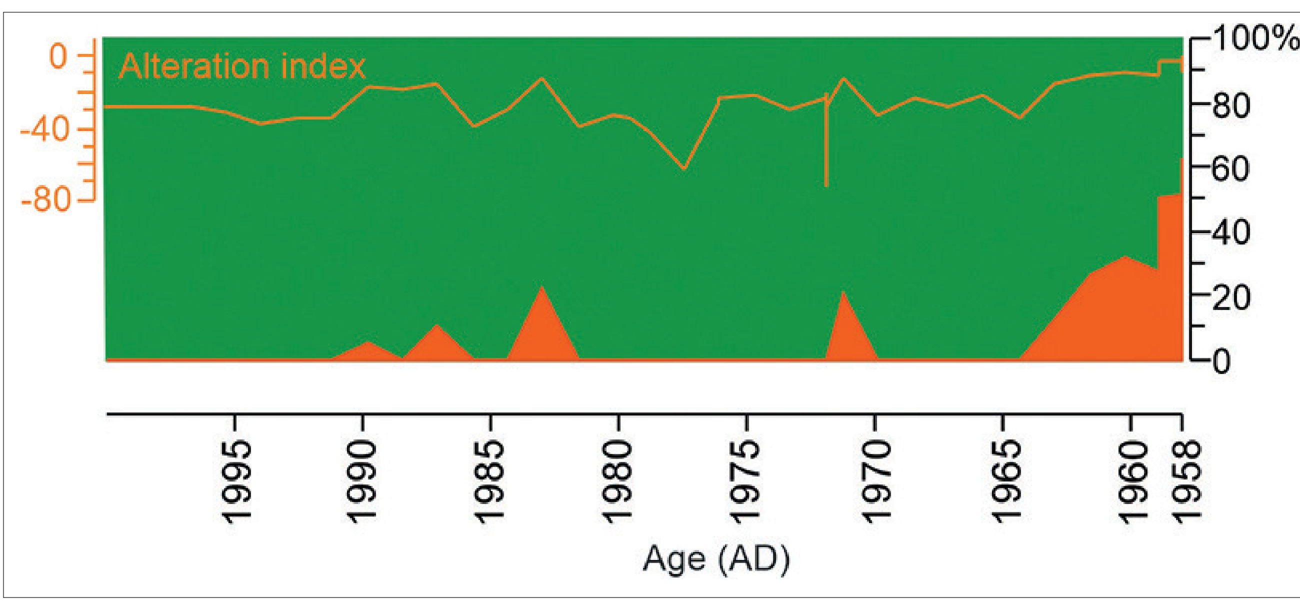

Figure 8. Graphical representation of the alteration score calculated for the sediment core collected in the reservoir of Araras (Varjota, Ceará) (see Ledru *et al.* 2020 for more information on the age model and pollen analyses).

Discussion

The Caatinga is one of the most degraded Brazilian biomes due on the one hand, to colonial agriculture organized around export-oriented production and on the other hand, the pressure on natural resources due to the recurrent threat of starvation in a highly populated region with more than 25 million people representing ~15 % of Brazil's population at the beginning of the 21st century (CGEE 2016). Our results show that even in areas where the vegetation cover appears to be relatively well preserved and hardly affected by anthropic activities, the alteration score remains relatively high (Tab. 1, Fig. 5). Consequently, we conclude that the BSA region is completely impacted by anthropogenic activities although at different levels depending on historical uses (Monteiro & Kurtz 2020). This is not surprising as the poor rural population had to count on natural resources for their survival, the most common activities being wood harvesting, hunting, slash and burn agriculture, and deforestation to create pasture for cattle and goats. Few phytosociological studies measured the herbaceous cover (Moro *et al.* 2016). However, when they did, they all showed a high percentage of herbaceous species, 60 % of the total species richness of the Caatinga, in line with our results, in which non-arboreal pollen accounted for more than 50 % of all pollen. Floristic composition in areas that were exploited for long periods of time (agriculture/planted pasture) and left to recuperate for a relatively short period of time (20 years) were strongly dominated by *Mimosa tenuiflora* (Pereira *et al.* 2002). The domination of *M. tenuiflora* in the pollen assemblages of all our samples thus tends to prove that all our study area was already cultivated and/or used for cattle. Indeed, whenever a piece of land was not cultivated, it was used for permanent grazing, thereby preventing natural regeneration and increasing desertification (CGEE 2016). Among the herbaceous species, the families with highest number of species were Euphorbiaceae, Fabaceae, Asteraceae, and Convolvulaceae, while 47 % of the families were only represented by a single species (Costa *et al.* 2007). This is in line with our results, if we consider that Euphorbiaceae is entomophilous and is a poor pollen producer (Barth 1989). In addition, among the 90 pollen taxa identified in our modern samples, the majority were represented by zero or only by single pollen grains only found once in the dataset. Other studies indicate that the tree species Fabaceae, Euphorbiaceae, and Cactaceae are among the most species rich families in the Caatinga (Moro *et al.* 2014). The pollen rain showed that Fabaceae is the dominant family while Euphorbiaceae and Cactaceae, two poor pollen producers, were scarce in our dataset. When we included herbaceous species, other families became floristically important, for instance, Asteraceae and Convolvulaceae (Araújo *et al.* 2005). However, these authors reported that ratios between herbaceous and tree species and the dominant species can vary considerably from one site to another (Moro *et al.* 2015; Monteiro & Kurtz 2020). Additionally, even areas that are considered as well-conserved show signs of past use (Guedes *et al.* 2012). We thus conclude that the historical background, climate, orography, and soils should be taken into account when defining a modern pollen rain sampling strategy in the BSA. Nevertheless, the surface samples we collected in the reservoirs are generally representative of the vegetation cover and of the general degradation of the landscape. In future studies, improvements such as increasing the number of samples in protected areas, diversifying the data set by sampling in all the different types of ecosystems found in the Caatinga, including the historical background in the metadata as well the age of the reservoirs and the successive agricultural activities, adding a botanical survey in each pollen sampling area. To summarize, our results show that our study areas can be considered as degraded and that the initial categories "preserved" or "intermediary" were not confirmed in the pollen assemblages, in line with results of botanical surveys and with the historical background of the Caatinga biozone. The on-going process of desertification is climate-independent and was initiated many decades or even centuries ago when the cultivated crops were replaced by grazing, thereby threatening the regional biodiversity.

Acknowledgements

This research was funded by a "sciences without boarder" (Portuguese acronym CSF) project "Evolution of biodiversity loss of areas in desertification processes" (CNPq 400890-2014-3), INCT IN-TREE: Interdisciplinary and transdisciplinary studies in ecology and evolution (CNPq CAPES FAPESBA) and covered a visiting foreign professor fellowship for MPL at the Federal University of Ceara (UFC) in the framework of the CSF project (2014–2019). This research is also part of the JEAI SANA (Le Semi-aride du Nordeste brésilien à l'Anthropocène) (IRD FUNCAP) and the Labex-CEBA.

Authors contribution

Marie-Pierre Ledru: Supervision, Visualization, Investigation, Conceptualization, Writing - Original Draft.
Raquel Franco Cassino: Formal analysis, Visualization, Investigation, Writing - Review & Editing.
Vaneicia Gomes: Formal analysis, Writing - Review & Editing.
Julia Sfair: Formal analysis, Writing - Review & Editing.
Francisca Soares Araújo: Supervision, Resources, Writing - Review & Editing, Project administration, Funding acquisition.

References

Amaral PGC, Ledru M-P, Ricardi Branco F, Giannini PC. 2006. Late Holocene development of a mangrove ecosystem in southeastern Brazil (Itanhaém, state of São Paulo). Palaeogeography, Palaeoclimatology, Palaeoecology 241: 608-620.

Araújo FS, Costa RC, Lima JR, *et al.* 2011. Floristics and life-forms along a topographic gradient, central western Ceará, Brazil. Rodriguésia 62: 341-66.

Araújo FS, Rodal MJN, Barbosa MRV, Martins FR. 2005. Repartição da flora lenhosa no domínio da caatinga. In: Araújo FS, Rodal MJN, Barbosa MRV (eds.). Análise das variações da biodiversidade do bioma Caatinga: suporte a estratégias regionais de conservação. Brasília, Ministério do Meio Ambiente. p. 15-33.

Assunção LM, Livingstone I. 1993.Desenvolvimento inadequado: construção de açudes e secas no sertão do Nordeste Revista Brasileira de Economia 47: 425-48.

Barbosa HA, Huete AR, Baethgen WE. 2006. A 20-year study of NDVI variability over the Northeast Region of Brazil. Journal of Arid Environment 67: 288-307.

Barth OM. 1989. O pólen no mel brasileiro. Rio de Janeiro, Instituto Oswaldo Cruz.

Bliege Bird R, Bird DW, Codding BF, Parker CH, Jones JH. 2008. The "fire stick farming" hypothesis: Australian Aboriginal foraging strategies, biodiversity, and anthropogenic fire mosaics. PNAS 105: 14796-14801.

Brasil. 2000. Lei N° 9.985, de 18 de Julho de 2000. https://www.planalto.gov.br/ccivil_03/leis/l9985.html.

CGEE – Centro de Gestão e Estudos Estratégicos. 2016. Desertificação, degradação da terra e secas no Brasil. Brasilia, Coronário Editora.

Chevalier M, Davis BAS, Heiri O, *et al.* 2020. Pollen-based climate reconstruction techniques for late Quaternary studies. Earth-Science Reviews 210: 103384.

Costa R, Araújo FS, Lima-Verde WL. 2007. Flora and life-form spectrum in an area of deciduous thorn woodland (caatinga) in northeastern. Brazil Journal of Arid Environments 68: 237-247.

Deza-Araujo M, Morales-Molino C, Tinner W, *et al.* 2020. A critical assessment of human-impact indices based on anthropogenic pollen indicators. Quaternary Science Reviews 236: 106291.

Faegri K, Iversen J. 1989. Textbook of pollen analysis. New Jersey, Blackburn Press.

Flantua SGA, Hooghiemstra H, Grimm EM, *et al.* 2015. Updated site compilation of the Latin American pollen database: challenging palynological research thriving. Review of Palaeobotany and Palynology 223: 104-115.

Gosling WD, Mayle FE, Tate NJ, Killeen TJ. 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.

Guedes RS, Zanella FCV, Costa Júnior JEV, Santana GM, Silva JA. 2012. Caracterização florístico-fitossociológica do componente lenhoso de um trecho de Caatinga no semiárido paraibano. Revista Caatinga 25: 99-108.

Harrison SP, Gaillard MJ, Stocker BD, *et al.* 2020. Development and testing scenarios for implementing land use and land cover changes during the Holocene in Earth system model experiments. Geoscientific Model Development 13: 805-824.

IPBES. 2018. The IPBES Assessment Report on Land Degradation and Restoration. Bonn, Secretariat of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services.

IPCC. 2018. Impacts of 1.5°C global warming on natural and human systems. In: Masson-Delmotte V, *et al.* (eds). Global warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. Geneva, World Meteorological Organization. pp. 175-311.

Juggins S. 2007. C2 Version 1.5 User Guide. Software for ecological and palaeoecological data analysis and visualisation. Newcastle upon Tyne, Newcastle University.

Ledru M-P, Jeske-Pieruschka V, Bremond L, *et al.* 2020. When archives are missing, deciphering the effects of public policies and climate variability on the Brazilian semi-arid region using sediment core studies. Science of the Total Environment 723: 137989.

MapBiomas. 2019. Project MapBiomas - Collection 2.3 of Brazilian Land Cover & Use Map Series. http://www.mapbiomas.org/

Marengo JA, Bernasconi M. 2015. Regional differences in aridity/drought conditions over Northeast Brazil: present state and future projections. Climate Change 129: 103-115.

Markgraf V, Webb RS, Anderson KH, Anderson L. 2002. Modern pollen/climate calibration for southern South America. Palaeogeography, Palaeoclimatology, Palaeoecology 181: 375-397.

Marquer L, Gaillard MJ, Sugita S, *et al.* 2017. Quantifying the effects of land use and climate on Holocene vegetation in Europe. Quaternary Science Reviews 171: 20-37.

Matias LQ, Guedes FM, Nascimento HP, Sfair JC. 2021. Breaking the misconception of a dry and lifeless semiarid region: the diversity and distribution of aquatic flora in wetlands of the Brazilian Northeast. Acta Botanica Brasilica 35: 46-61.

McLauchlan KK, Elmore AJ, Oswald WW, Sugita S. 2007. Detecting open vegetation in a forested landscape: pollen and remote sensing data from New England, USA. The Holocene 17: 1233-1243.

Montade V, Diogo IJS, Bremond L, *et al.* 2016. Pollen-based characterization of montane forest types in north-eastern Brazil. Review of Palaeobotany and Palynology 234: 147-158.

Monteiro DM, Kurtz BC. 2020. Phytosociology of Two Caatinga Phytophysiognomies with Different Histories of Anthropic Disturbance. Floresta e Ambiente 27: e20180045.

Moro MF, Araújo FS, Rodal MJN, Martins FR. 2015. Síntese dos estudos florísticos e fitossociológicos realizados no semiárido brasileiro. In: Eisenlohr PV, Felfili JM, Melo MMRF, Andrade LA, Meira Neto JAA. (eds.). Fitossociologia no Brasil: métodos e estudos de caso. Vol. II. Viçosa, Editora da Universidade Federal de Viçosa. pp. 412-51.

Moro MF, Lughadha EN, Araújo FS, Martins FR. 2016. A Phytogeographical Metaanalysis of the Semiarid Caatinga Domain in Brazil. The Botanical Review 82: 91-148.

Moro MF, NicLughadha E, File DL, Araújo FS, Martins FR. 2014. A catalogue of the vascular plants of the Caatinga Phytogeographical Domain: a synthesis of floristic and phytosociological surveys. Phytotaxa 160: 1-118.

Morrison KD, Hammer E, Popova L, *et al.* 2022. Global-scale comparisons of human land use: developing shared terminology for land-use practices for global change. Past Global Changes Magazine 26: 8-9

Overbeck GE, Vélez-Martin E, Scarano FR, *et al.* 2015. Conservation in Brazil needs to include non-forest ecosystems. Diversity and Distributions 21: 1455-1460.

Peel MC, Finlayson BL, McMahon TA. 2007. Updated world map of the Köppen-Geiger climate classification. Hydrology and Earth System Sciences 11: 1633-1644.

Pereira IM, Andrade LA, Barbosa MRV, Sampaio EVSB. 2002. Composição florística e análise fitossociológica do componente arbustivo-arbóreo de um remanescente florestal no Agreste Paraibano. Acta Botanica Brasilica 16: 241-369.

Piperno DR, McMichael C. 2020. Phytoliths in modern plants from amazonia and the neotropics at large: Implications for vegetation history reconstruction, Quaternary International 565: 54-74.

Portes MCG, Safford HD, Montade V, Behling H. 2020. Pollen rain–vegetation relationship along an elevational gradient in the Serra dos Orgãos National Park, southeastern Brazil. Review of Palaeobotany and Palynology 283: 104314.

Salgado-Labouriau ML. 1973. Contribuição à Palinologia dos Cerrados. Rio de Janeiro, Academia Brasileira de Ciências.

Sampaio VSBE. 1995. Overview of the Brazilian Caatinga. In: Bullock S, Mooney H, Medina E. (eds.). Seasonally dry tropical forests. Cambridge, Cambridge University Press. pp. 35-63.

Silva FHM, Santos FAR, Lima e Lima LC. 2016. Flora Polínica das Caatingas: Estação Biológica de Canudos (Bahia, Brasil). Feira de Santana, Micron Bahia.

Silva FHM. 2007. Contribuição à palinologia das Caatingas. PhD Thesis, Universidade Regional de Feira de Santana, Brazil.

Stockmarr J. 1971. Tablets with spores used in absolute pollen analysis. Pollen et Spores 13: 615-621.