Modern pollen spectra of the Cerrado vegetation in two national parks of Central Brazil, and implications for interpreting fossil pollen records

Raquel Franco Cassino^{a,*}, Caroline Thaís Martinho^b, Silane A.F. da Silva Caminha^c

^a Instituto de Geociências, Universidade de Brasília, Campus Universitário Darcy Ribeiro, CEP 70910-900 Brasília, DF, Brazil

^b Instituto do Petróleo e dos Recursos Naturais, Pontifícia Universidade Católica do Rio Grande do Sul, Av. Ipiranga, 6681, prédio 96J, CEP 90619-900 Porto Alegre, RS, Brazil

^c Programa de Pós Graduação em Geociências, Instituto de Ciências Exatas e da Terra, Universidade Federal de Mato Grosso, CEP 78060-900 Cuiabá, MT, Brazil

ARTICLE INFO

Article history:

Received 17 April 2014

Received in revised form 8 September 2015

Accepted 20 September 2015

Available online 26 September 2015

Keywords:

Cerrado

Modern pollen spectra

Palm swamp

Central Brazil

ABSTRACT

Twenty four surface samples were collected for palynological analysis from different environments in the *Cerrado* biome, which extends over a vast area in Central Brazil. These samples were used to investigate the relationship between vegetation and pollen spectra in the *Cerrado* and, thereby, contribute to a more precise interpretation of fossil pollen records from this region, particularly those collected on palm swamps. Pollen grains from the phytophysognomy that surrounds the palm swamp generally occur in low percentages in surface samples, whereas local plants are very abundant in the pollen spectra. Although local taxa tend to predominate, two aspects of the landscape may be reconstructed from the pollen spectra using numerical methods: the local environment of the deposition site and the surrounding regional vegetation. Using PCA and dissimilarity coefficient analysis, a group of seventeen taxa and a group of fourteen arboreal taxa are proposed as those that provide improved results when investigating the local environment and regional vegetation, respectively. Because herbaceous taxa are very abundant in the palm swamps, an analysis of the content of arboreal taxa in the pollen spectra is more appropriate for reconstructing the openness of the regional vegetation.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The *Cerrado* is the second largest biome in Brazil, and it covers over 2,000,000 km² of the central part of the country (Fig. 1A). The *Cerrado* vegetation is broadly classified as moist savanna, and it encloses a great diversity of flora (Furley, 1999; Sano et al., 2008). Indeed, this Brazilian savanna is a mosaic of several phytophysognomies characterized by peculiar vegetation structures and floras; based primarily on the degree of arboreal cover, eleven principal phytophysognomies have been recognized ranging from forests to grasslands (Ribeiro and Walter, 2008). The most typical are: the *Cerrado* stricto sensu, the *Cerradão*, the gallery forests, the *Campo Sujo* and the *Campo Limpo*. The *Cerrado* stricto sensu is a typical savanna physiognomy with small and gnarled trees, large shrubs and a prominent herbaceous stratum. The *Cerradão* has a dense cover of sclerophyllous trees that form a continuous canopy over the herbaceous-shrub stratum. The *Campos* are grasslands characterized by few trees and shrubs (*Campo Sujo*) or by the absence of trees (*Campo Limpo*). The gallery forests thrive alongside the small rivers of the Brazilian plateau and contain taller trees with an important arboreal cover.

These different phytophysognomies coexist in the same geographic space, and their distribution is due to a complex interaction of several factors, such as climate, soil characteristics, water-table depth, fire incidence and local history (Henriques, 2005).

Although the *Cerrado* and its vegetation were largely affected by the climate changes of the late Quaternary (c.f. Salgado-Labouriau, 1997; Ledru, 2002; Meyer et al., 2014), there are few records and little knowledge of the history of these changes. The majority of the available information comes from palynological studies that have been performed in lakes and in the palm swamps of the *Cerrado* core area (Ledru, 1993; Behling, 1995, 2002, 2003; Ferraz-Vicentini and Salgado-Labouriau, 1996; Parizzi et al., 1998; Barberi et al., 2000; Horák, 2009; Lorente et al., 2010; Raczka et al., 2012; Cassino and Meyer, 2013).

The palm swamps, known locally as *Veredas*, are a distinctive feature of the *Cerrado* landscape: they develop in marshy areas containing stream headwaters, and they are characterized by a typical flora dominated by the *buriti* palm (*Mauritia flexuosa*). The palm swamps, with the accumulation of peat in their central marshy zone, constitute a particularly favorable environment for the preservation of pollen grains and other palynomorphs. They are common features on all *Cerrado*'s plateaus and plains and thus contain valuable records of the *Cerrado* Quaternary history that remain underexplored. There are, however, several difficulties in interpreting the fossil pollen records collected in the palm swamps. Firstly, as in other biomes, the production and dispersion of pollen are variable between species, so the percentage of pollen

* Corresponding author at: Departamento de Geologia, Universidade Federal de Ouro Preto, Campus Morro do Cruzeiro, CEP 35400-000 Ouro Preto, MG, Brazil. Tel.: +55 3135591881.

E-mail address: raquelcassino@degeo.ufop.br (R. Franco Cassino).

Fig. 1. A) Distribution of the Cerrado biome in Brazil. B) Localization of the parks where samples were collected (arrows). C) Satellite image (Google Earth) of the Parque Nacional de Brasília with the location of the sampling sites. D) Satellite image (Google Earth) of the Parque Nacional Grande Sertão Veredas with the location of the sampling sites.

deposited is not directly related to the abundance of the species in the vegetation. Species that are wind pollinated tend to produce and disperse more pollen and so are generally over-represented in pollen spectra. In the Cerrado vegetation, herbaceous species of Poaceae and Cyperaceae family are commonly anemophilous, while other herbaceous and arboreal species are usually entomophilous or zoophilous, the most common exception being *Podocarpus*, the only gymnosperm genera occurring in the Cerrado biome.

Other difficulties in the interpretation of palm swamp pollen records arise from the complexity of the Cerrado vegetation and from the particular nature and vegetation of these sedimentary sites. The main difficulty is to interpret from the pollen record which phytophysiology was present in the paleovegetation. As mentioned earlier, the different phytophysionomies have different degrees of arboreal cover, but in the pollen spectra, the percentage and concentration of arboreal and herbaceous taxa are also influenced by the local vegetation of the palm swamp. In summary, with respect to past climate reconstruction, the correct interpretation of the palm swamp pollen records relies on the ability to accurately evaluate how local and regional vegetations

are represented in the pollen spectra, considering the differential production and dispersion of pollen of the different taxa.

One mechanism to evaluate this is through the analysis of modern pollen spectra extracted from surface samples. The study of surface samples for comparison with fossil pollen spectra has been used, in other regions, by several authors (i.e. Stutz and Pietro, 2003; Weng et al., 2004; Gosling et al., 2009; Burn et al., 2010; Ortuño et al., 2011) to generate modern analogs for the fossil record and to understand the relationship between vegetation characteristics and pollen spectra. Regarding the Cerrado vegetation, the first attempt to study the pollen rain was conducted by Salgado-Labouriau (1973) who analyzed the pollen rain collected on three pollen traps placed amidst a typical Cerrado in Goiás State for three years. A large predominance of Poaceae, that represented over 70% of the total pollen spectra, was observed in the samples; the second most representative family was Fabaceae, that accounted only for 3.2%. This work was focused on the variability of pollen precipitation throughout the year and indicated May as the month with more abundant pollen precipitation whereas September was the one with the more diverse pollen spectra. Ferraz-Vicentini (1993),

which presented the palynological study of a core collected in a palm swamp in Cromínia (southern Goiás State), studied ten surface samples, four of them collected on the margin of a lake situated upstream of the palm swamp and four collected in a gallery forest downstream of the palm swamp. The area around the palm swamp was occupied by a farm and the original *Cerrado* vegetation had been replaced by plantations. The other two samples analyzed by Ferraz-Vicentini (1993) were from a palm swamp of a park situated near the federal capital, Brasília (*Águas Emendadas Biological Reserve*); Barberi et al. (2000) further presented another two surface samples collected in this same palm swamp. In the pollen spectra extracted from these samples, Poaceae was the most abundant taxon in all environments, with, however, higher percentage in the palm swamp and lake samples (around 60%) than in the gallery forest samples (40%). Forest trees were abundant and diversified in the gallery forest samples but rare in the other ones, whereas typical *Cerrado* trees were rare or absent in all environments; Cyperaceae, Melastomataceae and *Mauritia flexuosa* were important components of the pollen spectra of all samples. Ferraz-Vicentini (1999) presented data from other four surface samples collected in the margins and in the water/sediment interface of a lake in *Chapada dos Viadinhos* Park, northern Goiás State. Poaceae and Cyperaceae together represented 70 to 80% of the pollen spectra in all samples; other important taxa were Melastomataceae, Myrtaceae, *Byrsonima* and *Eriocaulaceae*. Finally, Ledru (2002) presented data from modern pollen spectra of the *Cerrado* vegetation obtained from surface samples collected in different types of *Cerrado* in Distrito Federal, Mato Grosso and São Paulo States. The *Cerrado* pollen spectra were compared with those from Araucaria and semi-deciduous forest, allowing the designation of *Cerrado* indicators pollen types: *Byrsonima* and *Didymopanax* (equal *Schefflera*), over-represented in the pollen spectra, and *Caryocar* and *Qualea*, that, although exclusive of the *Cerrado* samples, were under-represented considering its abundance in the vegetation. Poaceae pollen was abundant in the samples, with significant variations between phytogeographies: around 20% in the *Cerradão* samples and over 50% in the more open *Cerrado* (*Campo Cerrado*).

Pollen spectra from surface samples collected among the *Cerrado* vegetation or from pollen traps, such as those presented by Ledru (2002) or Salgado-Labouriau (1973), are very useful to learn about the general characteristics of the pollen rain, but they are not directly comparable with fossil pollen spectra. Fossil pollen records are necessarily collected in specific depositional sites that harbor specific kinds of vegetation that have great influence in the composition of the pollen spectra. For cores collected on palm swamps, local taxa such as Cyperaceae and *Mauritia flexuosa*, among others, are present and abundant, influencing the percentage of all the other taxa, as showed by Ferraz-Vicentini (1993) and Barberi et al. (2000), resulting in pollen spectra different from those found in samples that were collected directly on the *Cerrado* soils. Therefore, when the study of pollen spectra aims to support the interpretation of fossil samples, it is more appropriate to use surface samples collected in depositional environments that present the same conditions in which the fossil samples were deposited. Despite the mentioned works of Ferraz-Vicentini (1993, 1999) and Barberi et al. (2000), which present only a few surface samples collected on undisturbed areas, to date, no systematic studies of surface samples that could be used as suitable parameters for the interpretation of fossil records collected on palm swamps have been conducted. Therefore, the purpose of the research described in this report is to present a palynological analysis of surface samples from depositional environments of the *Cerrado* biome that may contribute to generate reliable parameters for interpreting fossil pollen records from the palm swamps. To accomplish this, twenty four surface samples were collected from different botanical contexts within the *Cerrado* biome. Statistical analyses of the pollen content of these samples were used to compare pollen spectra from different local environments and different phytogeographies in order to highlight the influence of local and regional vegetations in the spectra. This comparison enabled better understanding of how

local and regional vegetations are reflected in these spectra and also permitted the determination of pollen taxa that are important for distinguishing between different *Cerrado* phytogeographies and local environments. This research also provides a set of data that can be directly used for the comparison with fossil samples and the determination of modern analogs and establishes the parameters that must be used in this comparison in order to associate correctly each fossil sample to a local environment and to a regional vegetation.

2. Materials and methods

2.1. Sampling, chemical processing and analyses

Despite the large territorial range of the *Cerrado* biome, pastures and crops have replaced much of its original vegetation. Therefore, the sample collection was conducted preferably in National Parks (Conservation Areas) within the *Cerrado* where the natural landscape and vegetation remain preserved.

From twenty four samples, twenty were collected in two National Parks with slightly different climatic and environmental conditions: the *Parque Nacional de Brasília* (PNB) and the *Parque Nacional Grande Sertão Veredas* (PGSV; Fig. 1). The first one, located near the federal capital Brasília, extends over a plateau with an altitude of approximately 1000 m. The mean annual temperature is 21.7 °C, and the mean annual precipitation is 1600 mm. The *Parque Nacional Grande Sertão Veredas* is located on the border between the Minas Gerais and Bahia States in a plateau at a lower altitude of approximately 700 m. The climate there is slightly hotter and drier, with an annual mean temperature of 23 °C and an annual mean precipitation of 1400 mm. With respect to the four other samples, one was from a federal reserve located near the *Parque Nacional de Brasília*, the *Reserva Biológica da Contagem* (RBC), and the three others were collected in sites outside conservation areas. The most northern came from a *Cerrado* enclave in the *Caatinga* region on the *Chapada Diamantina* plateau (CD) and was collected in a valley, with an altitude of 880 m, surrounded by plateaus that reach over 1000 m. Another was collected in a plateau at 700 m altitude between the Paracatu and São Francisco rivers in the Minas Gerais State (MG); and the final sample was collected in the same state, but further south, on the Serra do Espinhaço, which is a mountain range with an altitude of approximately 1400 m (SE; Fig. 1). In all these regions, the length of the dry season is of four to five months, characterizing a semi-humid climate. The sampling was not conducted with the intention of representing all the different climatic regions of the *Cerrado* biome since the main goal of the study was to analyze how the different phytogeographies that coexist within the parks are represented in the pollen spectra.

A central question in this study was to evaluate how the sedimentary site vegetation (local vegetation) and the surrounding phytogeography (regional vegetation) are represented in the pollen spectra; hence, the sampling sites were classified based on two parameters: local environment and regional vegetation (Table 1, Fig. 2). Three types of local environments were sampled: palm swamps, swamps and gallery forests. The purpose of this study being also to provide a set of modern pollen spectra that can be directly compared with fossil samples, the surface sampling was performed in order to better reflect the conditions in which the fossil samples are deposited. To do so, the local environments of the selected sampling sites all represented possible stages of a palm swamp's evolution.

The palm swamp samples were separated between those collected in the herbaceous border of the palm swamp and those collected in its center, where the *Mauritia flexuosa* palms concentrate. The palm swamps were also classified as dense arboreal palm swamps, typical palm swamps or sparse palm swamps (Table 1). These environments represent different stages of the evolution of a palm swamp: the sparse palm swamp represents an initial stage during which herbaceous vegetation predominates with few spaced *M. flexuosa* palms; the typical

Table 1

Local environments and regional vegetation where the twenty four samples were collected. Regional vegetation: G – grasslands and shrubby grasslands; C – arboreal *Cerrado*; and F – forests.

Sample code	Local environment	Phytophysiognomies	Regional vegetation
PNB01	Swamp edge zone	<i>Campo Limpo</i> and swamp	G
PNB02	Swamp	<i>Campo Limpo</i> and forests	G
PNB03	Gallery forest	<i>Cerrado</i> stricto sensu	C
PNB05	Swamp	<i>Campo Sujo</i>	G
PNB06	Arboreal palm swamp edge zone	<i>Campo Sujo</i>	G
PNB08	Arboreal palm swamp	<i>Campo Sujo</i>	G
PNB09	Gallery forest	<i>Campo Sujo</i>	G
PNB10	Gallery forest	<i>Campo Limpo</i>	G
PNB12	Typical palm swamp	<i>Campo Sujo</i>	G
PNB13	Typical palm swamp edge zone	<i>Campo Sujo</i>	G
PNB14	Typical palm swamp edge zone	<i>Campo Sujo</i>	G
PGSV01	Typical palm swamp	<i>Cerrado</i> stricto sensu	C
PGSV02	Typical palm swamp	<i>Cerrado</i> stricto sensu	C
PGSV03	Sparse palm swamp	<i>Cerrado</i> stricto sensu	C
PGSV04	Typical palm swamp edge zone	<i>Cerrado</i> stricto sensu	C
PGSV05	Typical palm swamp	<i>Cerrado</i> stricto sensu	C
PGSV07	Swamp	Forests and <i>Cerrado</i> stricto sensu	F
PGSV09	Swamp	Forests and swamp	F
PGSV10	Swamp	Forests and <i>Cerrado</i> stricto sensu	F
PGSV11	Typical palm swamp	<i>Campo Sujo</i> and <i>Cerrado</i> stricto sensu	G
RBC01	Gallery forest	<i>Cerrado</i> stricto sensu	C
CD01	Typical palm swamp edge zone	<i>Campo Rupestre</i>	G
SE01	Swamp	<i>Campo Rupestre</i>	G
MG01	Typical palm swamp	<i>Cerrado</i> stricto sensu	C

palm swamp shows an inner zone with numerous aligned *M. flexuosa* palms and shrubby vegetation; the arboreal palm swamp is the final stage of development and resembles a gallery forest with many arboreal taxa coexisting with the *M. flexuosa* palm. All sampled local environments, including swamps and gallery forests, represent possible local environments of the fossil samples of palm swamps' cores.

Each of these local environments presents particular edaphic conditions. The central zones of swamps and palm swamps are characterized by peat soil, permanently waterlogged, while their edge zones have wet soils, subject to seasonal drying. Gallery forests may grow on various types of soils that are usually well-drained.

The sampling sites were also classified based on their regional vegetation, i.e. the predominant phytophysiognomy present in the area adjacent to the palm swamp, swamp or gallery forest. The following phytophysiognomies were sampled: the previously described *Campo Sujo*, *Campo Limpo* and *Cerrado* stricto sensu and additionally, *Campo Rupestre*, a shrubby and herbaceous vegetation adapted to rocky soils, and forests, term used to refer to any of the forested phytophysiognomies of the *Cerrado* (Table 1). For the numerical analyses, these phytophysiognomies were grouped into three types of regional vegetation: “forests”, “arboreal *Cerrado*” and “grasslands and shrubby grasslands” (Table 1).

In each sampling site, three small volumes of soil were collected a few centimeters away from each other, which were then mixed together to avoid overrepresentation of plants situated just above the collection point. Following the method described in Salgado-Labouriau (2007), the samples were extracted using small 2-cm high cylindrical plastic collectors such that the soil samples were up to 2 cm deep. Using this procedure, samples should reflect the pollen deposition of a few tens of years, avoiding the influence of the variation of pollen production throughout the different seasons.

The samples were processed for palynological analysis using methods similar to those used traditionally for fossil samples, as

described in Bennett and Willis (2001), using HCl-30%, HF-40%, KOH-10% and acetolysis treatments. Prior to treatment, 1 mg of *Kochia scoparia* pollen was added to each sample to estimate pollen concentration (c.f. Salgado-Labouriau and Rull, 1986). The pollen counting was continued until a saturation point was reached; more detail on the use of saturation curves and on the pollen sums is provided in Cassino (2014). The pollen sums, that include all angiosperm and gymnosperm taxa, varied between 338 and 1448 pollen grains.

2.2. Numerical methods

Percentage and concentration pollen diagrams were generated using C2 software (Juggins, 2007). Pollen types were separated into seven groups based on the ecological preferences of the taxa: *Cerrado* arboreal taxa; forest arboreal taxa; *Cerrado* herbs and shrubs; grasslands taxa; swamp taxa; taxa of various environments; and exotic taxa (Fig. 3). The association of each taxon to one of these groups was based on the data from Mendonça et al. (2008) that presents a checklist of *Cerrado* species and indicates the phytophysiognomies and habitats in which they are found. The association of taxa to ecological groups was a complex task for two reasons: firstly, some taxa were identified only at the family or genera level and so can include plants from various habitats. Secondly, most taxa that could be identified at the species level are not exclusive of one habitat and cannot be associated exclusively to one type of environment or phytophysiognomy. Poaceae was associated to the group “grassland taxa” because they are one of the predominant components of grasslands but they are also abundant in palm swamps, swamps and in other *Cerrado* phytophysiognomies. The group “swamp taxa” includes all taxa associated to humid environments, some of them being more specifically associated to the palm swamps. This is the case of *Mauritia flexuosa* that, moreover, although being the most characteristic taxon of the palm swamps, can also be commonly found in gallery forests.

To test whether the modern pollen spectra could accurately reflect the local environment and/or the regional vegetation of the samples, two multivariate statistical analyses were used: Principal Component Analysis and dissimilarity coefficients.

Species-centered Principal Component Analyses (PCAs) of pollen percentages were generated using C2 software and displayed as biplot diagrams that, according to Ter Braak (1983), reflect sample diversity and the contribution of each species to the overall diversity. In the PCA, pollen spectra are compared with each other and the more similar spectra are grouped together in the biplot. PCA was used in three successive phases: initially, it was used to test if the more similar pollen spectra, i.e. those plotted closer in the biplot, were from samples of the same local environment or regional vegetation. In a second phase, various tests were conducted in order to define which taxa should be used to obtain a PCA biplot in which the samples are grouped according to their local environment. In the third phase, other tests were performed to define which taxa should be used to obtain a PCA biplot in which the samples are grouped according to their regional vegetation.

Dissimilarity coefficients, as introduced by Overpeck et al. (1985), are useful when quantifying the dissimilarity between modern pollen spectra and to determine modern analogs for fossil samples. Overpeck et al. (1985) proposed, after testing eight dissimilarity coefficient formulas, that the square-chord distance was the one that enabled the best differentiation between vegetation types of North America. That same coefficient formula was used here to test whether a good differentiation between local environments and regional vegetation of the *Cerrado* could be achieved. The square-chord distance formula is $d_{ij} = \sum_k (p_{ik}^{1/2} - p_{jk}^{1/2})^2$, where d_{ij} is the dissimilarity coefficient between samples i and j , and p_{ik} is the percentage of pollen taxon k in sample i .

Fig. 2. Environments and vegetation of selected surface samples. On the left, satellite images (Google Earth) of the sites and on the right, schematic representations of the local and regional vegetations. A and B) Samples PGSV04 and PGSV05, collected respectively on the edge zone and in the central zone of a typical palm swamp (local vegetation), surrounded by *Cerrado stricto sensu* (regional vegetation) in *Parque Nacional Grande Sertão Veredas*. The extension of the palm swamp central zone on the collection site is of approximately 50 m, and the extension of the edge zone, where herbaceous vegetation predominates, is of approximately 20 m in each side of the central zone. C and D) Sample PNB09, collected in the gallery forest of a small stream, surrounded by *Campo Sujo*, in *Parque Nacional de Brasília*. The gallery forest has an extension of approximately 50 m and the sample was collected in its center. E and F) Sample PNB05, collected in a swamp surrounded by *Campo Sujo*, in *Parque Nacional de Brasília*.

3. Results

3.1. Percentage and total concentration diagrams

Palynological analysis of the twenty four surface samples resulted in the identification of 236 types of pollen grain. The pollen types included different taxonomic levels and ranged from species to family; descriptions of the pollen morphology and illustrations of these pollen types were published elsewhere (Cassino, 2014). Many pollen taxa only occurred in one or in a few samples, and the diversity (number of taxa) is highly variable among the samples: in the least-diverse pollen spectra, only nineteen pollen types were observed, and in the most diverse, a total of sixty one pollen types were identified (Fig. 3). The samples with higher pollen diversity came from the gallery forests and the samples with the lowest taxa diversity were from the drier soils present in the border zone of swamps or palm swamps; therefore, it is noteworthy that there is a relationship between the pollen diversity and local environment. The majority of the pollen types were 'rare taxa': they occurred in very low percentages and were represented by one or two

grains in the pollen spectra. One or two very abundant taxa, typically Poaceae, Melastomataceae/Combretaceae or *Mauritia flexuosa* (except for samples PNB03 and PGSV01, that were dominated by *Alchornea* and *Cecropia*, respectively), dominated the pollen spectra (Fig. 3).

The total concentration of pollen grains in the sediment also varied greatly. The concentration does not directly reflect the amount of local pollen deposited because it is also influenced by sedimentation rates, which differed from site to site. However, of note, the samples collected in swamps or in the border zone of palm swamps typically showed lower total pollen concentrations that ranged from 11,000 to 36,000 pollen grains/cm³. The other samples collected in palm swamps or gallery forests showed variable total pollen concentrations that ranged from 35,000 to 135,000 pollen grains/cm³. The sample PGSV01, from a palm swamp, was very different from the other samples and showed a pollen concentration of more than 400,000 pollen grains/cm³. This anomalous concentration was due to the abundance of *Cecropia* pollen in the sample.

The pollen diagrams of the twenty four samples (Fig. 3) suggest that the pollen spectra are typically dominated by local taxa. Of course,

depending on the type of local environment, different taxa should be considered “local taxa”; thus, for swamps, Poaceae and Cyperaceae are local taxa, while for the gallery forests, forest trees should be considered as local. For example, in the pollen spectra of the palm swamps, generally the most abundant taxa are *Mauritia flexuosa*, Cyperaceae, Melastomataceae/Combretaceae and Poaceae, all taxa that are abundant in this environment. Samples collected in the edge zone of the palm swamp tend to have lower *M. flexuosa* and higher Poaceae, reflecting the fact that the palm *M. flexuosa* is concentrated in the central zone of the palm swamp and its edge zone is predominantly herbaceous. Pollen spectra of swamp samples are dominated by Poaceae; and gallery forest samples have low percentages of Poaceae and more significant percentages of forest trees such as *Tapirira*, *Alchornea*, *Hedyosmum brasiliense*

and Myrtaceae (Fig. 3). Therefore, for all the considered environments, the local taxa are predominant, and consequently, samples collected in the same type of local environment are more alike than samples that come from sites with the same regional vegetation (Fig. 3) and there is no direct relationship between pollen spectra and regional vegetation. For example, there were no higher percentages of *Cerrado* arboreal taxa in the samples collected on sites surrounded by a dense arboreal *Cerrado* and, likewise, there were no higher percentages of herbaceous taxa in samples collected on sites characterized by grassland regional vegetation (Fig. 3). It is of note, however, that despite the local pollen predominance and the similarities between pollen spectra of samples from the same type of local environment, there are, between samples of the same local environment, significant differences in the percentages

Fig. 3. Percentage and total concentration pollen diagrams of the twenty-four surface samples. The taxa included in the statistical analysis are assigned with symbols: * (a) *Tocoyena*, *Alibertia*, *Bowdichia*, *Cordia*, *Ericaceae*, *Pterogyne nitens*, *Qualea*, *Rhamnium*, and *Rourea*. (b) *Eichornia*, *Eriocaulaceae* type 1. (c) *Siphocampylus*, *Crumenaria*, and *Dalechampia*. (d) *Zornia*, *Crotalaria*, *Dorstenia*, *Eriosema*, *Helicteres*, *Ipomoea*, *Melochia*, *Poeppigia procera*, *Richardia*, *Smilax*, type *Sida*, and *Toulicia*. (e) *Virola*, type *Chorisia*, *Pyrostegia*, *Rhamnus*, *Piptadenia*, *Picrasma*, *Neea*, *Odontadenia*, *Passifloraceae*, *Muntingia calabura*, *Licania*, *Inga*, *Centrosema*, *Cedrela*, *Anadenanthera*, *Acnistus*, and *Caraipa*. (f) *Sterculiaceae* type 1, *Thunbergia*, *Ricinus communis*, *Solanaceae* type 1, *Solanaceae* type 2, *Papilionoidea* type 1, *Lythraceae* type 2, *Malvaceae*, *Chenopodium*, *Croton*, *Bromeliaceae*, and *Acanthaceae* type 1. (g) *Eucalyptus* and *Pinus*.

Fig. 3 (continued).

of certain taxa (for example Melastomataceae/Combretaceae; Fig. 3) and in the presence and diversity of rare taxa.

3.2. Principal Component Analysis numerical methods

Considering the diversity of pollen types and the similarities and differences between pollen spectra observed in the pollen diagrams of the twenty four surface samples, and given the use of pollen spectra as parameters for paleo-vegetational reconstruction, it is important to test if it is in fact possible to determine the vegetation of the sites where surface samples were collected from their pollen spectra. For this to be

considered possible, when comparing the pollen spectra from surface samples, pollen spectra that come from sites with the same kind of vegetation must be more alike. In a Principal Component Analysis (PCA), this means that pollen spectra from sites with the same kind of vegetation must group together in the biplot. If pollen spectra deposited in the same vegetational context are sufficiently homogeneous and different from pollen spectra deposited in other contexts, they will plot together in the PCA biplot, indicating that it is possible to reconstruct the paleovegetation of a fossil sample from its pollen spectra.

In the present case of study, the vegetational context of each site is characterized by two independent entities: the vegetation of the

depositional environment (local vegetation) and the phyto-physiognomy that surrounds the site (regional vegetation). As already discussed, the pollen diagrams suggest that the pollen spectra have a minor influence of the regional vegetation and thus samples that come from the same kind of regional vegetation, but from different local environments, can have very different pollen spectra. Thus, it is expected that, comparing the samples in a PCA, they will not be grouped according to their regional vegetation in the biplot. On the other hand, samples collected in the same type of local environment tend to have similar pollen spectra, although there are differences in the abundance of some taxa and overlaps between pollen spectra from different local environments. Thus, initially, in the first phase of the numerical analysis, PCA was used to test if the more similar pollen spectra, i.e. those plotted closer in the biplot, were in fact from samples of the same local environment and not from samples of the same regional vegetation, thus confirming the bigger influence of local pollen in the samples. That being the case, the PCA was also used to test if the pollen spectra are sufficiently different to allow precise and correct assignment of the samples to their local environment.

In this first phase, the forty five most-common taxa were included in the PCA, all taxa that occurred in five or more samples were considered (Table 2, Fig. 3) with the exception of *Cecropia*, which was excluded because it was overrepresented in sample PGSV01; the percentages were recalculated to sum 100%.

In the obtained biplot (PCA-45; Fig. 4), the first principal component accounted for 36% of the variance in the pollen spectra, and the second principal component accounted for 27%. Poaceae, Melastomataceae/Combretaceae and *Mauritia flexuosa* showed the highest species scores, and the separation between samples was primarily due to the percentages of these taxa (Fig. 4). Other taxa that have achieved a high score, i.e., whose percentages were important to differentiate between samples, were Cyperaceae, usually abundant in the same samples that Poaceae; *Erythroxylum* which was very abundant only in sample PGSV11; and *Cuphea micrantha*, which occurred in a few samples of different environments (Fig. 4). The biplot PCA-45 confirms that pollen spectra are more influenced by the local vegetation as the distribution of the samples in the biplot is more related to their local environment: the majority of the typical palm swamp samples are in the bottom left of the diagram due to their high content of *M. flexuosa*; samples from swamps and from the edge zone of palm swamps are on the right side, reflecting their high content of Poaceae and Cyperaceae; samples from gallery forests concentrated near the origin because they showed lower percentages of these three taxa. However, two samples with high Melastomataceae/Combretaceae and low Poaceae percentages placed in the top left of the diagram although they were not from the same environment and many samples (e.g., PGSV02, PGSV04, PNB02 and PNB06) were not placed within their groups (i.e., within the group of samples from the same local environment). Therefore, the comparison of pollen spectra based on these forty-five taxa did not allow a correct assignment of samples to their local environment. This method is thus not adequate to determine the local environment of fossil samples from the comparison with modern spectra.

The reason for this result is likely the fact that the group of forty-five taxa used in the PCA included taxa, such as Cyperaceae and Melastomataceae/Combretaceae, that are abundant in various local environments and, also, the fact that the regional vegetation also has an influence, although limited, on the percentages of these forty-five taxa.

Therefore, modifying the basis of comparison of samples, that is, restricting the comparison to the percentages of some selected taxa, it is possible that a biplot in which samples would be grouped according to their local environment could be obtained. PCA was thus used, in the second phase of the numerical analysis, to determine which group of taxa provides, when used to compare samples, the best grouping according to the local environment of the samples. The PCA was thus repeated multiple times, removing taxa progressively, until a group of taxa that best separated the local environments in the biplot

Table 2

Lists of the pollen taxa used in the different numerical analyses. The letters in the left column refer to the ecological group of the taxon: C – Cerrado arboreal taxa; F – forest taxa; G – grassland taxa; H – Cerrado herbs and shrubs; S – swamp taxa; and V – various environments.

Most common taxa (PCA-45; CD-45)		Seventeen selected taxa (PCA-17; CD-17)	
C	Arecaceae	C	Arecaceae
C	<i>Erythroxylum</i> (Erythroxylaceae)	F	<i>Allophylus</i> (Sapindaceae)
C	<i>Tabebuia</i> (Bignoniaceae)	F	<i>Euplassa</i> (Proteaceae)
C	Type <i>Eriotheca</i> (Bombacaceae)	F	<i>Hedyosmum brasiliense</i> (Chloranthaceae)
C	Vochysiaceae	F	<i>Ilex</i> (Aquifoliaceae)
F	<i>Alchornea</i> (Euphorbiaceae)	F	<i>Protium</i> (Burseraceae)
F	<i>Cestrum</i> (Solanaceae)	F	<i>Symplocos</i> (Symplocaceae)
F	<i>Hedyosmum brasiliense</i> (Chloranthaceae)	F	<i>Symplocos uniflora</i> (Symplocaceae)
F	<i>Hyeronima</i> (Euphorbiaceae)	F	<i>Tapirira</i> (Anacardiaceae)
F	<i>Ilex</i> (Aquifoliaceae)	G	<i>Alternanthera</i> (Amaranthaceae)
F	Moraceae/Urticaceae	G	Poaceae
F	<i>Myracrodruon urundeuva</i> (Anacardiaceae)	S	<i>Drosera</i> (Droseraceae)
F	<i>Myrsine</i> (Myrsinaceae)	S	<i>Eryngium</i> (Apiaceae)
F	<i>Piper</i> (Piperaceae)	S	<i>Mauritia flexuosa</i> (Arecaceae)
F	<i>Podocarpus</i> (Podocarpaceae)	S	<i>Xyris</i> (Xyridaceae)
F	<i>Protium</i> (Burseraceae)	V	<i>Byrsonima</i> (Malpighiaceae)
F	<i>Schefflera</i> (Araliaceae)	V	Myrtaceae
F	<i>Serjania</i> (Sapindaceae)	Fourteen arboreal taxa (PCA-14; CD-14)	
F	<i>Sloanea</i> (Elaeocarpaceae)	C	<i>Caryocar brasiliense</i> (Caryocaraceae)
F	<i>Tapirira</i> (Anacardiaceae)	C	<i>Curatella americana</i> (Dilleniaceae)
F	<i>Trema micrantha</i> (Ulmaceae)	C	<i>Plenckia populnea</i> (Celastraceae)
G	<i>Alternanthera</i> (Amaranthaceae)	C	Type <i>Eriotheca</i> (Bombacaceae)
G	<i>Cuphea micrantha</i> (Lythraceae)	F	<i>Chrysophyllum</i> (Sapotaceae)
G	<i>Cuphea</i> (Lythraceae)	F	<i>Euplassa</i> (Proteaceae)
G	<i>Gomphrena</i> (Amaranthaceae)	F	<i>Ilex</i> (Aquifoliaceae)
G	Poaceae	F	<i>Matayba</i> (Sapindaceae)
G	<i>Polygala</i> (Polygalaceae)	F	<i>Maytenus</i> (Celastraceae)
H	<i>Hyptis</i> (Lamiaceae)	F	<i>Myrsine</i> (Myrsinaceae)
H	<i>Mimosa</i> (Fabaceae)	F	<i>Podocarpus</i> (Podocarpaceae)
H	<i>Sebastiania corniculata</i> (Euphorbiaceae)	F	<i>Roupala</i> (Proteaceae)
H	<i>Senna</i> (Fabaceae)	F	<i>Schefflera</i> (Araliaceae)
H	Type <i>Maprounea</i> (Euphorbiaceae)	F	<i>Trema micrantha</i> (Ulmaceae)
H	Type <i>Borreria</i> (Rubiaceae)		
H	Type <i>Peixotoa</i> (Malpighiaceae)		
H	Type <i>Spermacoce</i> (Rubiaceae)		
S	Cyperaceae		
S	<i>Drosera</i> (Droseraceae)		
S	<i>Ludwigia</i> (Onagraceae)		
S	<i>Mauritia flexuosa</i> (Arecaceae)		
S	<i>Xyris</i> (Xyridaceae)		
V	Asteraceae		
V	<i>Byrsonima</i> (Malpighiaceae)		
V	Melastomataceae/Combretaceae		
V	Myrtaceae		
V	<i>Solanum</i> (Solanaceae)		

was obtained. This was achieved using a group of seventeen taxa (Table 2, Fig. 3) that included some abundant taxa, such as *Mauritia flexuosa* and Poaceae, and excluded others, such as Cyperaceae and Melastomataceae/Combretaceae. Besides those two taxa, the group comprised other taxa abundant in the palm swamps and swamps (*Alternanthera*, *Drosera*, *Eryngium*, *Xyris* and *Byrsonima*) and some arboreal taxa that occur in gallery forests (Arecaceae, *Allophylus*, *Euplassa*, *Hedyosmum brasiliense*, *Ilex*, *Protium*, *Symplocos*, *Symplocos uniflora*, *Tapirira*). In all analyses, the percentages were recalculated to sum 100%.

The analysis that included this group of seventeen taxa (PCA-17) resulted in a PCA biplot that accurately separated the samples by their local environment (Fig. 5); the first component accounted for 74% of the variance among the spectra, and the second component accounted for 20% of the variance. The samples from swamps grouped in the left bottom part of the diagram, and the samples from palm swamps grouped in the right bottom (Fig. 5). Samples collected in the border zone of palm swamps or in sparse palm swamps grouped to the left

Fig. 4. Principal component biplot calculated using the forty five most-common taxa in the pollen spectra: A) sample scores for the first and second components (the circles enclose samples from similar environments) and B) species scores (only taxa with higher scores are shown (arrows): 1. Poaceae; 2. Melastomataceae/Combretaceae; 4. Cyperaceae; 8. *Mauritia flexuosa*; 17. *Erythroxylum*; 25. *Cuphea micrantha*).

side, nearer to the origin. The samples from gallery forests formed a distinct group in the top of the diagram (Fig. 5). Sample PGSV08, collected in the center of an arboreal palm swamp, which is in the process of becoming a gallery forest, placed close to the gallery forest group. Sample PGSV10, collected on a swamp, placed within the samples from palm swamps borders, which may be explained by the presence of a palm swamp nearby. Therefore, when comparing the pollen spectra considering only the percentages of these seventeen taxa, the position of samples along the axes of the PCA reflects their local environment and the correct assignment that was not achieved in the previous analysis was reached in this biplot.

A direct relationship between the pollen spectra and the surrounding phytophysognomy (regional vegetation) was not observed in the pollen diagrams and the PCA-45 confirmed that, when the percentages of the most common taxa are considered, the pollen spectra are not grouped in the biplot according to the regional vegetation of the samples. Hence, a third phase of the numerical analysis was conducted, in which PCA was used to test if removing the local taxa and restricting the analyses to arboreal taxa, it would be possible to recognize properly and accurately the regional vegetation of each sample from its pollen spectra. Again, various PCAs with different groups of taxa were run until a biplot in which the samples were grouped according to their

Fig. 5. Principal component biplot calculated using the group of seventeen selected taxa: A) sample scores for the first and second components (the circles enclose samples from similar environments) and B) species scores (only taxa with higher scores are shown (arrows): 4. *Mauritia flexuosa*; 7. Poaceae; 10. *Hedyosmum brasiliense*; 12. *Protium*; 15. *Tapiria*; 16. *Byrsonima*; 17. *Myrtaceae*).

regional vegetation was obtained. The best result was reached with a PCA biplot that included a group of fourteen arboreal taxa (PCA-14; Table 2, Fig. 6), although the variance accounted for by the first and second principal components (respectively 29 and 21%) was reduced compared with the previous analysis.

The first feature of note in this diagram (Fig. 6) is that the samples of the two parks that concentrate the majority of the collection sites are separated along the first component axis, with samples from *Parque Grande Sertão Veredas* (PGSV) concentrated in the positive side of the axis and the *Parque Nacional de Brasília* (PNB) in the negative side. An exception is made for samples PGSV07 and PNB03 that plotted to the 'wrong' side. This separation between parks was not obtained in the previous analyses and reflects a separation between two different climatic and vegetation zones. Moreover, the samples are, in most cases, grouped according to the predominant phytophysiology around the sites: samples surrounded by *Cerrado* stricto sensu plotted to the right bottom of the diagram (reflecting high percentages of *Caryocar* or *Ilex*, *Eriotheca* and *Plenckia*). Samples surrounded by more open vegetation (*Campo Limpo* and *Campo Sujo*) grouped to the bottom left due to the absence of *Cerrado* arboreal taxa and the presence of *Podocarpus*. The presence of *Podocarpus* pollen in those samples, from open vegetation sites, is interesting as this taxon is usually found in forests and gallery forests in this region; however, being an anemophilous tree, its pollen is more easily dispersed than those from most typical *Cerrado* species, and can consequently be present in depositional environments surrounded by grassland vegetation and distant from its occurrence area. Two samples, collected near Woodlands, plotted to the right top, with high percentages of *Trema* and *Matayba*. The samples from gallery forests are the ones that least reflected the regional vegetation and tended to plot near to the origin, which is explained by the fact that the denser vegetation may prevent the accumulation of regional pollen.

3.3. Dissimilarity coefficients

In addition to Principal Component Analysis, dissimilarity coefficients were calculated to assess the dissimilarity between the pollen

spectrum of a given sample and all the other samples in the dataset. The results were plotted in diagrams that show the dissimilarity coefficients in ascending order, which enables grouping of the less dissimilar samples. In the diagrams, presented for eight selected samples (PGSV02, PGSV04, PGSV05, PGSV09, PNB03, PNB05, PNB13 and RBC01), the sample that is being compared has a coefficient equal to zero and is plotted in the left extremity of the diagram (Figs. 7, 8, 9 and 10). For example, in Fig. 6A, sample PGSV02 is being compared with all the other samples of the dataset; the diagram shows that the most similar sample to PGSV02 is PGSV03 that have a dissimilarity coefficient around 30, and the most dissimilar sample is PNB03, which presents a dissimilarity coefficient of 160.

Dissimilarity coefficients were used with the same objectives of the Principal Component Analysis, in order to see if different results would be obtained when using a different numerical method. As done for PCAs, initially, dissimilarity coefficients were used to test if, when the most common taxa are considered in the analyses (DC-45), pollen spectra from the same type of local environment have lower dissimilarity coefficients. Additionally, dissimilarity coefficients considering all the taxa were also calculated (DC-all). Then, dissimilarity coefficients were calculated considering only the group of seventeen taxa (DC-17) determined from the PCAs, to test if samples would be grouped according to their local environment in the diagrams. Finally, dissimilarity coefficients were also calculated considering only the group of fourteen arboreal taxa (DC-14) determined from the PCAs, to test if samples would be grouped according to their regional vegetation.

When the calculation of the dissimilarity coefficient is based on all taxa (DC-all), in the majority of cases the nearest samples (with lower dissimilarity coefficients) are not of the same local environment or the same regional vegetation. There are some exceptions, for example, for samples PGSV09, PNB05 and PNB03 (Figs. 8 and 9) pollen spectra from similar local environments tended to concentrate on the left of the diagram. However, in general, if all the taxa are considered, the dissimilarity coefficients do not provide a diagram with samples grouped according to their vegetational context. When the forty-five

Fig. 6. Principal component biplot calculated from the group of fourteen arboreal taxa: A) sample scores for the first and second components (the circles enclose samples from the same regional vegetation) and B) species scores (arrows): 1. *Podocarpus*; 2. *Ilex*; 3. *Trema micrantha*; 4. *Myrsine*; 5. *Schefflera*; 6. *Eriotheca*; 7. *Caryocar brasiliense*; 8. *Curatella americana*; 9. *Roupala*; 10. *Matayba*; 11. *Maytenus*; 12. *Euplassa*; 13. *Plenckia populnea*; and 14. *Chrysophyllum*.

A) PGSV02 - all taxa included

B) PGSV02 - most-common taxa

C) PGSV02 - seventeen selected taxa

D) PGSV02 - fourteen arboreal taxa

E) PGSV04 - all taxa included

F) PGSV04 - most-common taxa

G) PGSV04 - seventeen selected taxa

H) PGSV04 - fourteen arboreal taxa

Fig. 7. Diagrams of dissimilarity coefficients calculated from all the taxa, the forty five most-common taxa, the seventeen selected taxa and the fourteen arboreal taxa for samples PGSV02 and PGSV04. The vertical line in the seventeen-taxa diagrams separates samples with coefficients lower than the critical value (30).

most common taxa are used in the coefficient calculations (DC-45), different results are obtained (Figs. 7, 8, 9 and 10); in these diagrams, samples from the same type of local environment tend to have lower

dissimilarity coefficients. As observed for the PCAs, this confirms the hypothesis that pollen spectra are primarily influenced by the local vegetation of the sites.

A) PGSV05 - all taxa included

B) PGSV05 - most-common taxa

C) PGSV05 - seventeen selected taxa

D) PGSV05 - fourteen arboreal taxa

E) PGSV09 - all taxa included

F) PGSV09 - most-common taxa

G) PGSV09 - seventeen selected taxa

H) PGSV09 - fourteen arboreal taxa

Fig. 8. Diagrams of dissimilarity coefficients calculated from all the taxa, the forty five most-common taxa, the seventeen selected taxa and the fourteen arboreal taxa for samples PGSV05 and PGSV09. The vertical line in the seventeen-taxa diagrams separates the samples with coefficients lower than the critical value (30).

When the seventeen selected taxa (DC-17) are used, the dissimilarity coefficients display a better grouping of samples with respect to local environment. Indeed, dissimilarity coefficients <40 were

obtained for all samples from similar environments (Figs. 7, 8, 9 and 10). For example, the coefficients calculated for sample PNB05 (Fig. 9G) are <40 for all samples from swamps and from the edge

Fig. 9. Diagrams of dissimilarity coefficients calculated from all taxa, the forty five most-common taxa, the seventeen selected taxa and the fourteen arboreal taxa for samples PNB03 and PNB05. The vertical line in the seventeen-taxa diagrams separates the samples with coefficients lower than the critical value (30).

zone of palm swamps, which are very similar local environments. Conversely, samples collected on palm swamps and gallery forests, when compared to sample PNB05, show dissimilarity coefficients >60. The same pattern was observed in other samples (PGSV05, PGSV09, PNB03, RBC01 and PNB13). For PGSV02 and PGSV04, the

dissimilarity coefficient that separates the samples of the same environment from the other samples is lower (approximately 30).

The attempt to separate the samples based on their regional vegetation was less conclusive when based on the dissimilarity coefficients than in the PCA. In the DC-14 analysis of samples PNB05, PNB13 and

Fig. 10. Diagrams of dissimilarity coefficients calculated from all taxa, the forty five most-common taxa, the seventeen selected taxa and the fourteen arboreal taxa for samples RBC01 and PNB13. The vertical line in the seventeen-taxon diagrams separates the samples with coefficients lower than the critical value (30).

RBC01, which had more open regional vegetation, lower coefficient values (<100) were obtained for other samples from grasslands and shrubby grasslands (Figs. 9H, 10H and D) although some samples

were 'misplaced'. In the analysis of samples PGSV02 and PGSV04, which are surrounded by an arboreal *Cerrado*, samples that showed the same regional vegetation also tended to present dissimilarity

coefficients of <100 (Fig. 7D and H). However, for the other samples, this critical value could not be used to determine samples of the same regional vegetation type.

4. Discussion

The pollen diagrams of the twenty-four surface samples presented here showed that the pollen spectra have more influence of the local vegetation, i.e. the vegetation of the depositional site, so that samples collected in the same local environment, for example the central zone of a palm swamp, are more alike, regardless of the phytophysiology present around the site. In the previously mentioned data presented by Ledru (2002), a clear difference in the percentage of Poaceae was found in surface samples from different phytophysionomies, however, our study shows that in surface samples collected on depositional sites similar to those where cores are collected, this difference is no longer visible, as it is masked by local pollen. Previously published data from surface samples collected on palm swamps or similar environments (Ferraz-Vicentini, 1993, 1999; Barberi et al., 2000) had already demonstrated the abundance in the pollen spectra of local taxa such as *Mauritia flexuosa* and Cyperaceae. However, as there was no comparison of samples from different phytophysionomies (and moreover, most of the samples were from disturbed areas), it was not clear if the influence of the regional vegetation could be visible in the percentage of Poaceae or of the arboreal taxa.

As local pollen is more abundant in the pollen spectra of the surface samples and samples from the same local environment are more alike, the vegetation of the depositional site can more easily be reconstructed from the pollen spectra than the regional vegetation. Even so, the PCA that compares the samples based on the percentages of the forty-five more common taxa (PCA-45, Fig. 4) did not allow the correct grouping of all samples according to their local environments. This indicates that, firstly, that are heterogeneities among samples from the same local environments and overlapping between samples from different environments, and, secondly, that other factors, such as the surrounding phytophysionomy, also have an influence in the pollen spectra, which is, of course, expected. If the composition of the pollen spectra had no other influence than that of the local vegetation, the samples would be grouped according to their local environment in the PCA and in the dissimilarity coefficient diagrams that considered the most common taxa (PCA-45; DC-all; DC-45), which was not the case in this study.

However, parameters of comparison can be chosen in order to obtain diagrams in which samples are correctly grouped according to their local environment. To achieve this, instead of comparing samples in relation to the percentages of all taxa, only the taxa whose percentages are more directly related to the vegetation of the depositional environments were selected and used in the statistical analyses. Our results (PCA-17, Fig. 5) showed that, to separate between the most common depositional environments of the *Cerrado* (palm swamps, swamps, and gallery forests), the percentages of the seventeen following taxa should be used in the statistical analyses: Arecaceae; *Allophylus*; *Euplassa*; *Hedyosmum brasiliense*; *Ilex*; *Protium*; *Symplocos*; *Symplocos uniflora*; *Tapirira*; *Alternanthera*; Poaceae; *Drosera*; *Eryngium*; *Mauritia flexuosa*; *Xyris*; *Byrsonima*; and Myrtaceae. The percentages of these taxa are useful for differentiating between the different types of depositional environments of the *Cerrado*, which means that this group of taxa can be used to compare a fossil sample, whose depositional environment is unknown, with the surface samples from this database. From the fossil sample position in the PCA diagram (PCA-17), it will be possible to determine exactly in which type of environment the fossil sample was deposited. It is important to note that in the PCA-17, the determination of the local environment is quite precise. For example, samples PGSV04 and PGSV05 were collected a few meters away from each other, the former in the herbaceous edge zone of a palm swamp and the other in its inner shrubby-arboreal zone (Fig. 2); despite the

geographical proximity of the samples, they plotted to different groups in the biplot. This indicates that the analysis has sufficient accuracy to separate between closely located environments that harbor different kinds of vegetation and shows that PCA-17 will be useful to visualize expansions and retractions of the central zone of palm swamps in fossil cores, since this analysis allows the separation between samples deposited in the central zone of the palm swamp from those deposited in its edges.

In addition to PCA-17, our results show that the determination of the local environment of fossil samples can also be obtained with the dissimilarity coefficient analysis based on the same group of seventeen taxa (DC-17). In this analysis, a critical value of 30 can be defined from our study, as the value below which the pollen spectra are from the same local environment. This value may be used as a parameter to compare fossil samples with the modern pollen spectra and to determine modern analogs for past deposition environments.

Although it is interesting to reconstruct the expansions and retractions of the palm swamps and the variations of the vegetation of the depositional site, since these are influenced, among other factors, by changes in climate humidity, the principal goal of palaeoenvironmental studies is usually to reconstruct the changes in the regional vegetation and their climatic implications. It is therefore valid to attempt the reconstruction of the regional vegetation from the samples collected in palm swamps and similar depositional environments, although, as noticed, the predominance of local pollen complicates this task. The case of samples PGSV04 and PGSV05, discussed above, reflects this difficulty, as these samples, which come from the exact same regional vegetation, have quite different pollen spectra, as shown in PCA-45 (Fig. 4). However, the influence of regional vegetation in the pollen spectra exists and can be retrieved. Presumably, the openness of the regional vegetation could be reflected in the pollen spectra as the percentage of herbaceous taxa and in particular, Poaceae. However, these elements are also very abundant in the depositional sites where the samples were collected, i.e., their occurrence and abundance in the pollen spectra primarily reflect the local vegetation and have little influence of the degree of openness of the regional vegetation. Thus, to assess the openness of the regional vegetation, and so to determine the phytophysionomy, the most important parameter to consider is the percentage of certain arboreal taxa that are common in the savannah or in the forest phytophysionomies. Indeed, depending on the degree of openness of the vegetation, these arboreal taxa are more or less abundant or even absent in the pollen spectra, and our results show that they are the most suitable parameter to reconstruct the regional paleovegetation. This group of fourteen arboreal taxa includes some typical *Cerrado* stricto sensu taxa (*Caryocar brasiliense*, *Curatella Americana*, *Plenckia populnea* and Type *Eriotheca*) and some arboreal taxa typical of the forests phytophysionomies (*Chrysophyllum*, *Euplassa*, *Ilex*, *Matayba*, *Maytenus*, *Myrsine*, *Podocarpus*, *Roupala*, *Schefflera* and *Trema micrantha*). The percentages of these taxa are useful to determine the regional vegetation of the samples, as the first ones are present in the samples from savannah phytophysionomies, the second ones in the samples from forest phytophysionomies and only *Podocarpus*, because of its anemophilous type of pollination, is present in the pollen spectra from grasslands phytophysionomies.

5. Conclusions

The study of pollen spectra from twenty four surface samples collected in different sites within the *Cerrado* biome has important implications for the interpretation of fossil pollen records and the reconstruction of past vegetation in Central Brazil. A great diversity of pollen types is present in the modern samples; however, in general, the spectra are dominated by one or two local taxa. The regional vegetation surrounding the deposition site is poorly represented in the spectra, even when a dense *Cerrado* is present. This provides an extra importance to rare pollen taxa and emphasizes the importance of using

saturation curves for pollen counting, as done in this study, in sediment samples from the *Cerrado*. Counting with saturation curves usually exceeds the pollen sum of 300 grains traditionally used for other vegetation types, providing a more complete representation of rare taxa.

Despite this predominance of local taxa, the use of numerical methods has demonstrated that two aspects of the landscape may be reconstructed based on the pollen spectra: the local environment of the deposition site and the regional vegetation that surrounds the site. The reconstruction of these two aspects is important in paleoenvironmental research because not only the regional vegetation but also the local environment is influenced by climatic and other environmental changes. The central question when using numerical methods for comparing pollen spectra is the taxa that are considered for the calculations. The PCA and dissimilarity coefficient calculations presented here showed that little information is retrieved from comparisons when a great number of taxa are used. The pollen spectra were very diverse; thus, there was too much noise in the analyses, and this noise prevented good separation between the sites.

Different analyses that consider different groups of taxa should be used to group similar samples based on the two different aspects of the landscape that were investigated. A set of seventeen taxa is proposed as being optimal for statistical analyses when investigating the local environment. When using this group of taxa for calculating dissimilarity coefficients, a critical value of 30 was established, above which two samples may be considered as originating from the same local environment.

A set of fourteen arboreal taxa is proposed as being optimal for the investigation of regional vegetation. Although in palynological analysis the abundance of herbaceous taxa, and in particular Poaceae, is frequently used as an indicator of the openness of regional vegetation, the data presented here showed that in the *Cerrado* pollen spectra, the abundance of Poaceae pollen is influenced primarily by local vegetation and does not reflect the openness of the landscape. Therefore, in numerical analyses, the use of herbaceous taxa is not appropriate for assessing regional vegetation openness.

The data presented here also suggest that arboreal taxa are more appropriate when assessing climatic regions from pollen spectra; however, a more complete data set is required to determine a group of taxa that can be used effectively in statistical analyses to separate the different climatic regions of the *Cerrado*.

Acknowledgments

This research was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), process number 475929/2011-0 and grant number 151478/2011-2. The authors thank the Instituto de Pesquisa Chico Mendes de Biodiversidade (ICMBio) for authorizing the sampling on the national parks. We also thank the two anonymous reviewers for their constructive comments and suggestions that helped to improve the early version of this manuscript.

References

- Barberi, M., Salgado-Labouriau, M.L., Suguio, K., 2000. Paleovegetation and paleoclimate of "Vereda de Águas Emendadas", Central Brazil. *J. S. Am. Earth Sci.* 13, 241–254.
- Behling, H., 1995. A high resolution Holocene pollen record from Lago do Pires, SE Brazil: vegetation, climate and fire history. *J. Paleolimnol.* 14 (3), 253–268.
- Behling, H., 2002. Late Quaternary vegetation and climate dynamics in Southeastern Amazonia inferred from Lagoa da Confusão in Tocantins State, Northern Brazil. *Amazoniana* 17 (1/2), 27–39.
- Behling, H., 2003. Late Glacial and Holocene vegetation, climate and fire history inferred from Lagoa Nova in the Southeastern Brazilian lowland. *Veg. Hist. Archaeobot.* 12, 263–270.
- Bennett, K.D., Willis, K.J., 2001. Pollen. In: Smol, J.P., Birks, H.J.B., Last, W.M. (Eds.), *Tracking Environmental Change Using Lake Sediments: Terrestrial, Algal, and Siliceous Indicators* vol. 3. Kluwer Academic Publishers, Dordrecht, pp. 5–32.
- Burn, M.J., Mayle, F.E., Killeen, T.J., 2010. Pollen-based differentiation of Amazonian rainforest communities and implications for lowland palaeoecology in tropical South America. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 295, 1–18.
- Cassino, R.F., 2014. Reconstituição paleoambiental da região dos cerrados do norte de Minas Gerais baseada na análise palinológica de sedimentos de veredas e na comparação com conjuntos polínicos atuais Ph.D thesis Instituto de Geociências, Universidade de Brasília.
- Cassino, R.F., Meyer, K.E.B., 2013. Reconstituição Paleoambiental do Chapadão dos Gerais durante o Quaternário Tardio, a Partir da Análise Palinológica da Vereda Laçador. *Rev. Bras. Paleontol.* 16 (1), 127–146.
- Ferraz-Vicentini, K.R., 1993. Análise palinológica de uma vereda em Cromínia - GO Masters thesis, Departamento de Ecologia, Universidade de Brasília.
- Ferraz-Vicentini, K.R., 1999. História do fogo no Cerrado: uma análise palinológica Ph.D thesis Departamento de Ecologia, Universidade de Brasília.
- Ferraz-Vicentini, K.R., Salgado-Labouriau, M.L., 1996. Palynological analysis of a palm swamp in central Brazil. *J. S. Am. Earth Sci.* 9 (3/4), 207–219.
- Furley, P.A., 1999. The nature and diversity of neotropical savanna vegetation with particular reference to the Brazilian cerrados. *Glob. Ecol. Biogeogr.* 8, 223–241.
- Gosling, W.D., Mayle, F.E., Tate, N.J., Killeen, T.J., 2009. Differentiation between Neotropical rainforest, dry forest, and savannah ecosystems by their modern pollen spectra and implications for the fossil pollen record. *Rev. Palaeobot. Palynol.* 153, 70–85.
- Henriques, R.P.B., 2005. Influência da história, solo e fogo na distribuição e dinâmica das fitofisionomias no bioma do Cerrado. In: Scariot, A., Sousa-Silva, J.C., Felfili, J.M. (Eds.), *Cerrado: Ecologia, Biodiversidade e Conservação*. Editora do Ministério do Meio Ambiente, Brasília, pp. 75–92.
- Horák, I., 2009. Relações pedológicas, isotópicas e palinológicas na reconstrução paleoambiental da turfeira da Área de Proteção Especial (APE) Pau de Fruta, Serra do Espinhaço Meridional Masters thesis, Escola Superior de Agricultura 'Luiz de Queiroz', Universidade de São Paulo.
- Juggins, S., 2007. C2 Version 1.5 User Guide. Software for Ecological and Palaeoecological Data Analysis and Visualisation. Newcastle University, Newcastle upon Tyne, UK (73 pp.).
- Ledru, M.P., 1993. Late Quaternary environmental and climatic changes in Central Brazil. *Quat. Res.* 39 (1), 90–98.
- Ledru, M.P., 2002. Late Quaternary history and evolution of the cerrados as revealed by palynological records. In: Oliveira, P.S., Marquis, R.S. (Eds.), *Ecology and Natural History of a Neotropical Savanna: The Cerrados of Brazil*. The University of Columbia Press, pp. 33–50.
- Lorente, F.L., Meyer, K.E.B., Horn, A.H., 2010. Análise Palinológica da Vereda da Fazenda Urbano, Município de Buritizinho, Minas Gerais, Brasil. *Geonomos* 18 (2), 57–72.
- Mendonça, R.C., Felfili, J.M., Walter, B.M.T., Silva Júnior, M.C., Rezende, A.V., Filgueiras, T.S., Nogueira, P.E., Fagg, C.W., 2008. Flora Vascular do Cerrado. Checklist com 12.356 espécies. In: Sano, S.M., Almeida, S.P., Ribeiro, J.F. (Eds.), *Cerrado: Ecologia e Flora* vol. 2. Embrapa Informação Tecnológica, Brasília, pp. 423–1279.
- Meyer, K.E.B., Cassino, R.F., Lorente, F.L., Raczka, M., Parizzi, M.G., 2014. Paleoclima e paleoambiente do Cerrado durante o Quaternário com base em análises palinológicas. In: Carvalho, I.S., Garcia, M.J., Lana, C.C., Strohshoen, O. (Eds.), *Paleontologia: Cenários da Vida vol. 5. Interiência, Rio de Janeiro*, pp. 403–420.
- Ortuño, T., Ledru, M.P., Cheddadi, R., Kuentz, A., Favier, C., Beck, S., 2011. Modern pollen rain, vegetation and climate in Bolivian ecoregions. *Rev. Palaeobot. Palynol.* 165, 61–74.
- Overpeck, J.T., Webb, T., Prentice, I.C., 1985. Quantitative interpretation of fossil pollen spectra: dissimilarity coefficients and the method of modern analogs. *Quat. Res.* 23, 87–108.
- Parizzi, M.G., Salgado-Labouriau, M.L., Kohler, C.H., 1998. Genesis and environmental history of Lagoa Santa, SE Brazil. *The Holocene* 8 (3), 311–321.
- Raczka, M.F., De Oliveira, P.E., Bush, M., McMichael, C.H., 2012. Two paleoecological histories spanning the period of human settlement in southeastern Brazil. *J. Quat. Sci.* 28, 144–151.
- Ribeiro, J.F., Walter, B.M.T., 2008. As principais fitofisionomias do Cerrado. In: Sano, S.M., Almeida, S.P., Ribeiro, J.F. (Eds.), *Cerrado: Ecologia e Flora* vol. 1. Embrapa Informação Tecnológica, Brasília, pp. 151–199.
- Salgado-Labouriau, M.L., 1973. Contribuição à Palinologia dos Cerrados. *Academia Brasileira de Ciências, Rio de Janeiro*.
- Salgado-Labouriau, M.L., 1997. Late Quaternary palaeoclimate in the savannas of South America. *J. Quat. Sci.* 12 (5), 371–379.
- Salgado-Labouriau, M.L., 2007. Critérios e Técnicas para o Quaternário. Editora Edgard Blucher, São Paulo.
- Salgado-Labouriau, M.L., Rull, V., 1986. A method of introducing exotic pollen for palaeoecological analysis of sediments. *Rev. Palaeobot. Palynol.* 47, 97–103.
- Sano, S.M., Almeida, S.P., De Ribeiro, J.F., 2008. Cerrado: Ecologia e Flora vol. 1. Embrapa Informação Tecnológica, Brasília.
- Stutz, S., Prieto, A.R., 2003. Modern pollen and vegetation relationships in Mar Chiquita coastal lagoon area, southeastern Pampa grasslands, Argentina. *Rev. Palaeobot. Palynol.* 126, 183–195.
- Ter Braak, C.J.F., 1983. Principal components biplots and alpha and beta diversity. *Ecology* 64 (3), 454–462.
- Weng, C., Bush, M.B., Silman, M.R., 2004. An analysis of modern pollen rain on an elevational gradient in Southern Peru. *J. Trop. Ecol.* 20, 113–124.