Updated site compilation of the Latin American Pollen Database


⁎ Corresponding authors.
E-mail addresses: S.G.A.Flantua@uva.nl (S.G.A. Flantua), H.Hooghiemstra@uva.nl (H. Hooghiemstra).

1. Introduction

1.1. Global pollen databases

Although analyses of individual sites remain crucially important and are the basis of all integrative studies, pollen analysis is moving from an emphasis on single-site analyses to multisite integrative analyses and the generation of spatial summaries of ecosystem dynamics (Bradshaw, 2013). Integration of multi-site and multiproxy data facilitates investigation of a range of key research questions, such as response lags (Seppä and Bennett, 2003), the direction and rates of species migration and population expansions and contractions (Blosi et al., 2011), acclimation and eventual microevolutionary adaptation (Feeley et al., 2012), ecological resilience (Willis et al., 2010), asynchronous ecosystem responses to (abrupt) climate change (Kohfeld and Harrison, 2000), microrefugia (Rull, 2009; Mosblech et al., 2011; Collins et al., 2013), and no-analog ecosystem formation (Williams and Jackson, 2007; Correa-Metrio et al., 2012). Furthermore, a continent-wide overview of sites highlights
available data for both generating and answering research questions (Seddon et al., 2014). Therefore, global databases of environmental data from the past and present will become increasingly important for enabling new research directions, to facilitate large accessible archives and collaborative thinking between research groups.

1.2. Regional pollen database of Latin America

The lack of a comprehensive inventory of pollen sites from "Latin America" (Caribbean region, México, and Central and South America) has inhibited synoptic studies of the region and has caused an unawareness of the spatial and temporal coverage potentially available for global-scale assessments (Kohfeld and Harrison, 2000; Gajewski, 2008; Bartlein et al., 2011). After the Latin American Pollen Database (LAPD) was initiated in 1994, it passed through several institutions, to finally reappear again without new updates in 2007, when it was incorporated into the Neotoma Paleoecology Database (www.neotomadb.org; Grimm et al., 2013). Beginning in 2014, new data contributed by individual investigators was added to LAPD/Neotoma; however most of the LAPD data in Neotoma are from a 2002 version of LAPD. It is important to present the full extent of the growing palynological efforts in Latin America so that the potential for continental-scale synoptic studies can be assessed and researchers are aware of the available data for inclusion in their research. A comprehensive and updated database would be also useful to better address the origin and temporal evolution of "Latitudinal Diversity Gradients," as well as to analyze the relationships between pollen diversity and vegetation diversity in time and space. Finally, it is important to have identified those areas where sampling effort should be focused in the near future so as to complete the picture of different biotic and climate dynamic changes.

1.3. Aims of paper

The aim of this paper is to present an updated inventory and synthesis of palynological sites in Latin America, identifying spatial patterns, research tendencies and potential research opportunities. Unique and contrasting findings are highlighted and metadata are summarized. The objectives of this paper are to stimulate collaborative palynological opportunities, foster intra-, and inter-database comparisons, and apprise the best available data for multidisciplinary research.

2. Methods

2.1. Database compilation

2.1.1. History of LAPD

In 1994, Vera Markgraf at the University of Colorado (USA) started an open database of Latin American pollen data. During the period between 1998 and 2003, management of the LAPD was based at the University of Amsterdam, where Robert Marchant served as coordinator. Among numerous papers by Marchant and collaborators two are key for this paper. One was a synthesis of the distribution and ecology of taxa in the LAPD (Marchant et al., 2002), and the other was pollen-based biome reconstructions for Latin America (Marchant et al., 2009). The version of the LAPD in Amsterdam ultimately incorporated about 500 pollen sites, including data from cores, archeological, cave and alluvial sites and surface samples and sites from Central America, the Caribbean, and South America. However, this database was never publicly available. Through the NOAA Paleoclimatology website and several workshops, the idea of sharing data was promoted, especially before data become lost or abandoned over time (Michener et al., 1997). Active management of LAPD at Amsterdam ended in 2003, and no further updates were made to the website and related database. Meanwhile, publication of new data from Latin America has accelerated, and the current LAPD is outdated, a common problem of shared databases (Costello et al., 2014).

2.1.2. Description of the data compilation

As a first step to updating LAPD and for developing a plan for acquisition of new data, we have developed an inventory of existing data for potential incorporation into LAPD. We carried out an exhaustive search for published pollen datasets and compiled the following metadata: site latitude-longitude coordinates and altitude, proxies analyzed, chronological control, and estimated time range covered. The inventory includes both paleodata and modern pollen samples and data from terrestrial and marine environments. Although historically LAPD has focused on the Quaternary period, the new inventory also includes pollen records from the Pliocene and late Miocene, especially long records that also span the Quaternary. Direct access to Colombian universities facilitated the addition of "gray" literature (including Master and Ph.D. theses and unindexed national journals). The inventory of such resources from other countries is limited to that revealed by digital resources, and is less exhaustive than that available from Colombia. Thus, opportunity exists in other countries to inventory more datasets from these kinds of publications. Therefore we emphasize the opportunity to extend the compilation provided in this paper based on unpublished literature.

2.2. Spatial and temporal assessment of the LAPD

2.2.1. Paleomapping

Across Latin America, investigators have employed various techniques to elucidate emergent properties from spatial arrays of pollen data, including i) isopollen maps of modern pollen samples (Paez et al., 2001; Hooghiemstra et al., 2006; Tonello and Prieto, 2008; Urrego et al., 2009), ii) translation of pollen data into biomes (Marchant and Hooghiemstra, 2004; (Marchant et al., 2002, 2009), iii) quantitative climate reconstruction combining modern pollen assemblages and modern climatological data in a pollen-climate training set (Punyasena et al., 2008; Tonello et al., 2009; Tonello and Prieto, 2010; Schäbitz et al., 2013) and iv) CO₂ concentrations in paleo-atmospheres (Boom et al., 2002; Beerling and Mayle, 2006). More advanced geographical information system (GIS) modeling tools or dynamic vegetation models were used for species and biome distribution mapping for the Guyana highlands (Rull and Nogué, 2007), Amazonia (Beerling and Mayle, 2006), the Yucatan peninsula (Carrillo-Bastos et al., 2012), and the Colombian lowlands (Flantua et al., 2007) and highlands (Flantua et al., 2014).

2.2.2. Spatial metadata

The geographical extent of the inventory was confined between latitudes 32°N to 63°S and longitudes 30°W to 116°W. For many sites, especially those published before 1980, accurate locational information is not available, but insofar as possible, the published geographic coordinates were recorded. Significant discrepancies in published descriptive locations and latitude-longitude coordinates were checked, for example sites with coordinates in the wrong country or terrestrial sites with coordinates in the sea. In such cases, other references and Google Earth were consulted where possible to obtain estimated coordinates. Most modern samples have excellent metadata for their location with the advent of Geographical Positioning System (GPS) and online mapping tools such as Google Earth. Published altitudinal data have not been adjusted even if coordinates were incorrect (a typo in a longitude or latitude coordinate does not necessarily imply an error in altitude). Here again, the pre-GPS altitude assessments were often based on altimeters that could drift in accuracy by >100 m during a day.

The temporal coverage of datasets was recorded as published without further evaluation of radiocarbon accuracy. Sites were categorized into broad time bins, including the last glacial cycle (300–50 thousand calibrated years before the Present, here indicated as 300–50 ka), the transition of the Last Glacial Maximum (LGM; 22–20 ka), the Holocene (12–10 ka), and the period of most recent human impact (2–0 ka). The latter time bin is currently of specific interest at a global scale (PAGES-
2.2.3. Geochronologic database of the LAPD

Mapped syntheses are critically limited by the accuracy and precision of $^{14}C$ ages (Grimm and Jacobson, 2004; Giesecke et al., 2012; Flantua et al., 2015a), especially for those aimed at assessing the synchrony of change across a region. In this first phase of temporal assessment, we evaluated only the “age control density” for each dataset. This score indicates the density of the geochronologic ages ($^{14}C$, $^{210}Pb$, U-series, etc.) along the total core length and is quantified by dividing the length of time represented by the sedimentary sequence by the number of geochronologic ages. The density of ages was calculated using the original author-reported age controls. Sites ($n = 97$) with only one geochronologic date were excluded.

2.2.4. Additional metadata

If the publication named the biome in which the site occurred, that biome was recorded. If no specific biome was named, a biome was assigned as per Table 1 from Marchant et al. (2009). Datasets were categorized as monoproxy if they included only pollen data or multiproxy if other biological, physical, or chemical data are available. Figures throughout this paper were created using ArcGIS® software by Esri (2012).

3. Results

3.1. Spatial assessment of the LAPD

3.1.1. Spatial coverage

The density and the spatial distribution of pollen sites over a geographical region are important variables for palaeomapping and spatial interpolation. The LAPD inventory comprises datasets from 30 countries, including 4817 modern pollen surface samples (Fig. 1A) and 1379 paleo-datasets (Fig. 1B). Surface-sample coverage in terms of density of samples and representation of biomes is especially good from Argentina (21%), Colombia (18%), and Chile (9%) (Table 1). A limitation, however, is that although published surface-sample datasets often include a large number of samples, precise geographical coordinates are not always provided, in which case, a number of points may be assigned identical latitude-longitude coordinates and are mapped as single point in Fig. 1, e.g. in the Cordillera Oriental of Colombia (Grabandt, 1985), Venezuelan Andes (Rull, 2006), and Tierra del Fuego (Heusser, 1989). Differences with previously announced LAPD versions (Flantua et al., 2013) are owing to continuous updates and revisions of literature. The long history of palaeoecological research in Latin America contributes to the impressive number of paleo-sites investigated, particularly from Chile (19%), Colombia (18%), Argentina (14%), Brazil (13%), and Mexico (10%). Palynological research at fossil animal middens ($n = 88$) contributes to the large amount of paleo-sites in Chile. New countries to the database are mainly from the Caribbean (e.g. Cuba and Haiti) and Central America (mainland Honduras and El Salvador). The density of sites is especially high in central and southern Mexico, the Colombian Andes, the region of Lake Titicaca in Peru and Bolivia, the southern pampas of Argentina, and the southern Patagonia of Argentina and Chile. Paraguay and Jamaica are not represented in the inventory. The sites are widely distributed in altitude (Fig. 2), in terms of proportionate areal coverage, high altitudes are better documented than low altitudes. The number of sites per country is assigned.

The inventory includes 38 marine records, the earliest of which are distributed from the northern and southern Atlantic Ocean (Venezuela: Muller, 1959; Argentina: Groot et al., 1965) to the Pacific Ocean (Costa Rica: Horn, 1985; Ecuador: Heusser and Shackleton, 1994). Recent marine cores from the Cariaco basin in the Caribbean (68 ka: González et al., 2008), western Patagonia (22 ka: Montade et al., 2012, 2013), and the Panama basin in the Pacific (40 ka: González et al., 2006) provide insights on terrestrial environmental changes during the last glacial interglacial cycle.

3.1.2. Spatial distance assessment

The nearest neighbor index of the LAPD is $< 1$ for all time bins, indicating clustering of sites at all times. At 0 BP, mean nearest neighbor distance is 67 km; whereas at 50 ka, the mean distance is 218 km, with only 145 sites available older than 25 ka (Fig. 3A). Thus, with increasing age, the density of sites decreases, while the degree of

Table 1

<table>
<thead>
<tr>
<th>Country</th>
<th>LAPD 2014</th>
<th>LAPD 1999</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Modern – Total</td>
<td>Modern – known coordinates</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Argentina</td>
<td>1032</td>
<td>102</td>
</tr>
<tr>
<td>Bahamas</td>
<td>130</td>
<td>3</td>
</tr>
<tr>
<td>Barbados</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Barbuda</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>Brazil</td>
<td>253</td>
<td>163</td>
</tr>
<tr>
<td>Chile</td>
<td>432</td>
<td>146</td>
</tr>
<tr>
<td>Colombia</td>
<td>873</td>
<td>273</td>
</tr>
<tr>
<td>Costa Rica</td>
<td>118</td>
<td>57</td>
</tr>
<tr>
<td>Cuba</td>
<td>100</td>
<td>3</td>
</tr>
<tr>
<td>Dominican Republic</td>
<td>49</td>
<td>48</td>
</tr>
<tr>
<td>Ecuador</td>
<td>265</td>
<td>117</td>
</tr>
<tr>
<td>El Salvador</td>
<td>210</td>
<td>117</td>
</tr>
<tr>
<td>French Guiana</td>
<td>107</td>
<td>18</td>
</tr>
<tr>
<td>Guatemala</td>
<td>118</td>
<td>77</td>
</tr>
<tr>
<td>Guyana</td>
<td>100</td>
<td>3</td>
</tr>
<tr>
<td>Haiti</td>
<td>49</td>
<td>48</td>
</tr>
<tr>
<td>Honduras</td>
<td>365</td>
<td>174</td>
</tr>
<tr>
<td>Mexico</td>
<td>210</td>
<td>18</td>
</tr>
<tr>
<td>Peru</td>
<td>182</td>
<td>77</td>
</tr>
<tr>
<td>Surinam</td>
<td>33</td>
<td>6</td>
</tr>
<tr>
<td>Trinidad y Tobago</td>
<td>49</td>
<td>48</td>
</tr>
<tr>
<td>Falklands Islands/Islas Malvinas</td>
<td>1452</td>
<td>1379</td>
</tr>
</tbody>
</table>
| Argentina: Muller, 1959; Argentina: Groot et al., 1965) to the Pacific Ocean (Costa Rica: Horn, 1985; Ecuador: Heusser and Shackleton, 1994). Recent marine cores from the Cariaco basin in the Carribean (68 ka: González et al., 2008), western Patagonia (22 ka: Montade et al., 2012, 2013), and the Panama basin in the Pacific (40 ka: González et al., 2006) provide insights on terrestrial environmental changes during the last glacial interglacial cycle.
uncertainty for paleomapping increases. At the Pleistocene–Holocene transition (11.6 ka), mean distance is higher in the central regions of Latin America and the Amazon basin compared to the Colombian Andes and Southern Chile and Argentina. Most sites (n > 800) include the last two millennia, while the LGM (n = 120) and Pleistocene–Holocene transition (n = 310) is represented in smaller numbers of datasets (Fig. 3B). Time control for some datasets is uncertain, however, because of few 14C or other geochronological control. In some regions, scarcity of suitable depositional environments for pollen accumulation contributes to low density of sites. In particular, the Amazon lowlands are scarce in LGM records (Ledru et al., 1998). The studies by Colinvaux et al. (1996), Burbridge et al. (2004), Mayle et al. (2007), Whitney et al. (2011) and Hermanowski et al. (2012) stand as the few long records available from the lowlands. Additionally, marine cores by Haberle and Maslin

Fig. 1. Sites (cores and sections) in the 2014 Latin American Pollen Database. (A) Modern pollen rain data from core tops, traps and surface sediments shown in red dots. (B) Drilled cores and exposed sections are shown in blue triangles.

Fig. 2. Distribution pattern of all sites, including modern pollen rain sites and cores and sections lodged in the LAPD organized by latitude and elevation. There are no known coordinates for Surinam and marine cores not shown. Negative values indicate southern latitudes.
(1999) and Hoorn (1997) were explored filling the void on our understanding on long term climate change dynamics of the Amazon basin. The sites with the longest continuous records are located at higher altitudes in Colombia (284 ka in Lake Fúquene: core Fq-9C; 2225 ka in paleolake Bogotá: core Funza09) and Bolivia (151 ka in Lake Titicaca: core LT01). From lower elevation, El Valle in Panama (Bush and Colinvaux, 1990) stands out with a record extending to ~150 ka, as well as Colónia CO3-1 in Brazil (Ledru et al., 2005) and Lake Petén-Itza in Guatemala (Correa-Metrio et al, 2013), 100 ka and 85 ka respectively. The Cariaco record (68 ka, González et al., 2008) is an example of an exceptional regional reference for the palaeoecclimatology of the area, and also as a chronological landmark for its annually laminated sediments (Hughen et al., 2004) that have been used to calibrate the worldwide 14C database (Reimer et al., 2013).

3.2. Temporal assessment of the LAPD

The new LAPD inventory contains 5003 chronological controls from 1062 sites (some sites contain multiple records) including tephras ($n = 30$) and uranium series dates ($n = 6$). Relative chronologic controls, such as biostratigraphic dates, were not included. During recent years, the number of ages used for stratigraphic age models has trended upwards; since 2010, the mean and median number of dates per published age model is five and three, respectively (Fig. 4); however, the number varies widely, between 1 and 50. Wide variation in dating resolution has important implications for paleomapping (Fig. 5).

Publications report geochronologic metadata inconsistently. Important details are often missing, including material dated, laboratory number, calibration curve and program, and age modeling algorithm. Some publications report only the calibrated ages and not the original 14C ages reported by the laboratory. Other publications only present the dates or age models graphically. Missing geochronologic data and metadata complicate efforts to evaluate dating accuracy and precision, which is necessary for synoptic multi-site analyses (Flantua et al., 2015a). For example, if a study is trying to assess climate change during the Younger Dryas interval, it is important to know how secure the dating for this interval is, as errors of a few hundred years could be consequential. Another important issue is that published age models are
based on different versions of the radiocarbon calibration curve, which has been updated every few years. To ensure comparability for synoptic studies, age models may need to be updated to the current calibration curve, for which sufficient data must be available for updating age models and chronologies. This exercise has recently been presented by Flantua et al. (2015a) for the northern and central Andes within a framework of a newly composed LAPD geochronologic database. Suggestions for the data needed for reconstructing chronologies can be found in Blaauw (2010), Buck and Millard (2004) and Grimm et al. (2014). The need for protocols on procedures for constructing chronologies is imperative and has become the focus of attention of different international workgroups, e.g. Neotoma Paleoecology Database (Grimm et al., 2014) and the project Integration of Ice-core, Marine and Terrestrial records (INTIMATE; Blockley et al., 2012).

3.3. Publications, proxies, and biomes

3.3.1. Publications

The LAPD inventory includes over 1200 publications published prior to April 2014 or known to be in press (Figs. 6; 7). Of these, 58% were published since 2000 with the peak years being 2010 and 2012, with 66 and 67 publications respectively. We expect to find more studies published in 2012–2014. Receiving early attention in the 1960s and 1970s were the Patagonian lowlands, the Amazon basin, and the high Andes, which are still foci of investigation. Early palynologists worked on pre-Quaternary and Quaternary sediments in the course of oil exploration (e.g. De Boer et al., 1965; Germeraad et al., 1968); early compilation by Archangelsky (1968). Islands and their unique biogeography have attracted attention, including the Easter Islands (Flenley and King, 1984; Cañellas-Boltà et al., 2013), the Galapagos (Colinvaux, 1972; Restrepo et al., 2012), and the Falkland Islands (Barrows, 1977). In the past decade, the Brazilian coastal region and the Caribbean islands have received increased investigation. Few studies exist from the arid zones of Chile and Argentina mainly because of the absence of lakes or strong reservoir effect on sediments from highlands such as Altiplano lakes. However, pollen from rodent middens (e.g. Markgraf et al., 1997; Maldonado et al., 2005) provide a basis for reconstructing the climate and environment of the driest zone of the Atacama desert (Chile) for the past 55 ka. Midden studies have been carried out primarily in northern Chile and Argentina but have potential in other arid zones such as eastern Brazil.

3.3.2. Multiproxy studies

A developing trend has been the shift from monopropoxy to multiproxy studies. Since 2000, a half of all papers report multi-proxy analyses compared to about a fifth of all papers in the 1990s (Fig. 9). In addition to pollen, charcoal and physical sedimentology have been the most common additional proxies; however, studies of 3 to 4 proxies are not uncommon (e.g. Conroy et al., 2009; Enters et al., 2010; Vriend et al., 2012; Ledru et al., 2013). Palynological findings have been supported by biomarkers to define upper forest-line migration (Jansen et al., 2010), by charcoal to indicate human influences (e.g. Montoya and Rull, 2011) or seasonality (e.g. Iglesias et al., 2012), and a variety of different proxies to reveal the cultural history of America (e.g. Goman et al., 2010; Sedov et al., 2010; Friarte et al., 2012; Whitney et al., 2013).

3.3.3. Representativeness of biomes and ecosystems

The wide diversity of biomes across Latin America is unequally represented in the inventory (Fig. 10). The páramos (cool grasslands), Amazonian rainforest (tropical rainforest), and warm temperate evergreen broadleaf forest are relatively well studied. Whereas, the tropical dry forest and warm temperate mixed forest are geographically restricted and less well represented. The latter is characteristic of the mid to high altitudes in Mexico, while dry forest is extensive in central Brazil. Pollen data from the deserts and vegetation communities from the arid regions of Chile and Mexico are equally sparse, although palynological research on fossil animal middens is quickly expanding. Investigators have often targeted transition zones or ecotones, such as the savanna-forest boundary in the Colombian Llanos (e.g. Behling and Hooghiemstra, 2000) and in the Venezuelan Gran Sabana (e.g. Leal Rodríguez, 2010; Montoya et al., 2011); the upper Andean forest and páramos in the Ecuadorian Andes (e.g. Brunschöhn and Behling, 2010; González-Carranza et al., 2012); the transition from pampa grasslands to xerophytic woodlands (e.g. Schabitz and Liebricht, 1998; Tonello and Prieto, 2010); and the Patagonian forest-steppe ecotone (e.g. Markgraf et al., 2007; Iglesias et al., 2012). Scaling down, some investigations have focused on ecological dynamics within biomes, which can encompass significant variability. For example, the tropical rainforest biome includes Amazonian rainforest and coastal mangrove ecosystems, while the tropical dry forest biome encompasses “Cerrado,” Caatinga, and “Campo rupestre” (Huber and Rüsha, 1997, 2003). Similarly the Pampa grasslands, “Espinal,” “Monte” and Patagonian Steppe are grouped into temperate grasslands, savannas and shrublands biome (Olson et al., 2001). Furthermore, at longer time-scales poor-analog or no-analog systems may be “forced” into a modern biome structure rather than being treated as a novel assemblage. Thus, the biome approach is interesting as a generalized categorization but due the large climatic amplitudes of the Plant Functional Types are not apt for detailed ecosystem level assessments.

4. Discussion

4.1. The variety of data and questions in the LAPD

The heterogeneity of the LAPD data has implications for multisite comparisons. The sites differ greatly in spatial and temporal coverage, sample and geochronologic resolution, and research questions. The use of palynology for novel conservation problems such as invasive species (Van Leeuwen et al., 2008), human–landscape interactions (e.g. Haug, 2003; Friarte and Behling, 2007; Lozano-García et al., 2007; Vegas-Vilarrúbia et al., 2011; Gosling and Williams, 2013; Carson et al., 2014), impact of sea level rise (Cohen et al., 2009; Behling,
and conservation and management of ecosystems (Behling et al., 2007; van Leeuwen et al., 2008; Vegas-Vilarrubia et al., 2011), are just some of the few exciting new research topics in Latin America. No-analog ecosystems (Correa-Metrio et al., 2012; Velásquez-R. and Hooghiemstra, 2013) and microrefugia (Rull, 2009; González-Carranza et al., 2012; Whitney et al., 2014) deal with more regional questions but have important global implications. Quantitative paleoclimatic reconstructions based on pollen from Latin America are scarce and focused on Southern Patagonia (Tonello et al., 2009; Schäbitz et al., 2013).

Long sediment sequences (> 20 ka) are crucial for identifying and understanding long-term ecological and evolutionary processes (Torres et al., 2013; Seddon et al., 2014). Such sequences from Colombia and Mexico were essential in relating high-mountain vegetation change to global ice volume and concentrations of greenhouse gasses (Groot et al., 2011), and identifying temperature change velocity (Correa-Metrio et al., 2013). A deep core from the south Argentinean Potrok Aike Maar is presently under study by the international Potrok Aike Maar Lake Sediment Archive Drilling Project (PASADO) team (Recasens et al., 2012). Regardless of the length of the time span, records may detect certain vegetation dynamics scarcely observed in other records. An example are the short pollen records in Chocó, Colombia (Behling et al., 1998; Berrío et al., 2000), which register minimal biome change although there are indications of environmental change related to tectonically driven changes in the coastal areas. Pollen assemblages from fossil rodent middens record altitudinal changes in vegetation belts and infer therefore the interaction between summer and winter precipitation (Maldonado et al., 2005) and hydrologic changes associated to human responses (Maldonado and Uribe, 2012; Mujica et al., 2015). Additionally, sites can be present in an area of low representativity in the palynological coverage. Besides newly added pollen data from rodent middens in arid regions, pollen sites from the Venezuelan “tepuis” (Rull, 2010) and
lake Timbio (Wille et al., 2001) are among the few sites available to study the Guyana highlands and lower montane (subandean) forest respectively. These sites represent ongoing “silent” regions and indicate the importance of expanding the database and research efforts.

4.2. Re-visit or re-evaluation of sites

When disagreements on regional climate dynamics persist, such as the magnitude of temperature or precipitation change, a revisit or re-evaluation of sites may be in order. Examples of such controversial areas are the northwest corner of Brazil (e.g. Bush et al., 2004; D’Apolito et al., 2013), southeast Amazonia (Hermanowski et al., 2012), Easter Island (Rull et al., 2013) and southern Patagonia (Kilian and Lamy, 2012). Other important sites that have not been substantiated by later studies may provide insights into key to palaeoecological understanding. They might, however, not meet modern standards of inquiry in terms of taxonomy or adequate dating control, particularly if they were studied before the advent of accelerator mass spectrometry (AMS) radiocarbon dating. Examples are the Lusitania record in Colombia (Monsalve, 1985) and Crominia in Brazil (Ferraz-Vicentini and Salgado-Labouriau, 1996). These sites prove the value of conclusions derived from pollen site compilations, and that revisiting lakes with specific research questions can be more rewarding than the search for new lakes to core.

4.3. Potential limitations and weaknesses

With respect to current global climate change and controversies over anthropogenic forcing vs. natural variation, documentation and understanding of decadal to centennial scale climate variability has become increasingly urgent. However, few currently studied sites have the requisite temporal resolution, such as Cambará do Sul (Behling et al., 2007), La Cocha (González-Carranza et al., 2012), Papallacta (Ledru et al., 2013) and El Junco (Restrepo et al., 2012). High resolution pollen records from Southern Patagonia have been developed in the last two decades (Heusser et al., 1999) but late Holocene high resolution records appear only the last few years (e.g. Moreno et al., 2009). Nevertheless, the publication of new radiocarbon calibration curves in 2013 (Hogg et al., 2013; Reimer et al., 2013) will necessitate development of new chronologies and age models. An inventory of associated geochronological data is necessary to assess the temporal resolution of questions that can be addressed. An assessment of chronological data derived from the LAPD is presented by Flantua et al. (2015a).

Physiographic variability and vegetation mosaics present challenges for paleomapping, which interpolates between sites that represent local vegetation. Such variability is evidenced in studies by D’Apolito et al. (2013) for lowland Amazonia and by Marchant et al. (2001) for Andean sites, and Venezuelan savannas (Montoya et al., 2011). Even with...
relatively short distances between sample sites in the last 10 ka, the topographical diversity of the continent adds an increased element of uncertainty for paleomapping. Nevertheless, advanced geospatial analysis of pollen is hardly explored in Latin America, although some research hypothesis can only be solved with a spatial focus.

Finally, we identified: i) ecosystem responses to millennial-scale and centennial climate changes, ii) rates of change, iii) changes in connectivity and its impact on gene flow and isolation, iv) changes in altitudinal distribution of individual taxa as well as biomes, v) changes in biomes and its impact on ancient human populations, vi) impact of pCO2 and frequency of nightfrost on the altitudinal distribution of taxa, vii) warm interglacial refugia (páramo at high elevation and C4 vegetation at low elevations) and cold glacial refugia as still poorly document-ed forces of vegetation change.

4.4. Grey literature and missing research

Data from grey literature are often the most difficult to identify and acquire, especially from countries in which national journals are published in the local language. Based on our increased access to grey literature in Colombia, we suspect that substantial literature of similar nature, including theses, exists in other Latin American countries as well as in North America and Europe, where many researchers who have worked in Latin America are based. The access to unpublished studies and grey literature may be advantageous but it can also have a handicap. On the one hand, the information is frequently more complete but, on the other hand, often has not undergone rigorous peer review, thus potentially compromising scientific quality. Additionally, archeological literature may be a source of pollen data not adequately surveyed for this inventory (e.g., Enters et al., 2010). Latin American researchers often face obstacles to publishing in English-language journals (Vasconcelos et al., 2008; Clavero, 2010; Arbeláez-Cortés, 2013). The LAPD inventory will especially attract attention to papers not written in English or in international journals and will create increased awareness of this important literature, even if the raw data are not yet available in a public database. The updated LAPD provides the ideal literature platform to make the wide range of studies known and accessible for researchers.

4.5. Integrative comparison of multiple pollen data

This paper provides an initial analysis of the updated inventory to identify locations, ecosystems, and research topics that are well studied, under discussion, or under-represented in existing data networks. This first level inventory is supported by a literature database and corresponding metadata of the sites. To address macro-scale questions that require integrative comparison of multiple pollen data, sites should be comparable, reproducible and ideally in a shared database environment. The Neotoma Paleoecology Database (www.neotomadb.org) is global paleodatabase initiative that offers a readily available platform to reduce overall IT overhead for community-wide database maintenance while enhancing capabilites for data ingest, display, analysis, and distribution. Not only does the Neotoma Paleoecology Database provide a data curation system to counter data loss through time (Michener et al., 1997; Jahnke et al., 2012), it addresses commonly identified problems of open access databases (Porter and Callahan, 1994; Costello et al., 2014). However, the Neotoma Paleoecology Database only facilitates a set of issues for an open-access palynological database to be success-ful. For the LAPD to be a sustainable database for researchers to rely upon, contribute to, and participate in, the scientific community is responsible for both the quality and comprehensiveness of the data-base and its continuation. Database cyberinfrastructure can support urgent inquiry, long-term needs, and future visions of the community, but the level of data sharing should be a guided by a regional science organizations or institutions with recognized mandates. Neotoma provides an online platform for palynological exchange, but the answers to macro-scale research depend on the collaborative will of the regional data generators. Already existing groups in collaboration with the newly updated LAPD are the Long-Term Climate Reconstruction and Dynamics of South America (LOTRED-SA) initiative (Flantua et al., 2015b; Villalba et al., 2009; www.pages-igbp.org/workinggroups/lotred-sa) and the Global Charcoal Database (http://gpwg.org/). Doubts on the ethical re-use of research data is a common problem of open-access databases (Jahnke et al., 2012), but possible incorrect interpretations of either a publications or data should not interfere with the opportunity to contribute of a collaborative long-term effort. Therefore a proactive communication with contributors is essential during any synthesis of data, and will ultimately enrich the interpretation and understanding of the compilation of information.

Shared palynological databases aim to facilitate integration, synthesis, understanding of the palerecord, and to promote information sharing and collaboration. To achieve the best projections and predictions of future vegetation responses to ongoing warming and other global changes, we need to make use of the best available data (Wolkovich et al., 2012) and exploit the full potential of palynological data. For the continuation of an updated LAPD, an interactive participation and interchange of the palynological community is essential and will provide an enrichment of the Latin American pollen data use and analysis. It is important that the palynological community becomes aware of regional and global initiatives to enable joint analysis of multiproxy datasets to address paleoenvironmental questions that transcend those possible with single-proxy databases.

5. Conclusions

The updated inventory of Latin American palynological research facilitates exploration of the temporal and spatial coverage of pollen data. The publically available LAPD has not been updated for over a decade. The current inventory identifies a large quantity of data that potentially could be incorporated into LAPD; the number of bibliographic records has expanded from ~190 publications to over 1200. The new LAPD inventory shows an impressive increase in the spatial distribution of sites, currently 1379 paleo-sites and over 4800 modern pollen samples. An increasing number of studies in Latin America are addressing broad-scale questions with multisite comparison, multi-proxy approaches, and multidisciplinary interests; and research practices and methodologies are commensurate with other continents (Seddon et al., 2014). We anticipate that this inventory data will prompt new research questions and opportunities.

All scientific research is built on the existing body of scientific data. LAPD strives to make these data more accessible and thereby foster and support research. In addition, by providing spatial and temporal visualizations of the existing data, LAPD aims to identify gaps in current knowledge and to provide a mechanism for generating new hypotheses by displaying emergent properties in arrays of data. New research might involve revisiting strategic sites for reanalysis with newer technologies, especially acquisition of improved geochronologic data (e.g., Sedov et al., 2010; Bogotá-A et al., 2011), or have the potential for a deeper cores, such as Lake Chalo (Brown et al., 2012). Finally, synoptic studies comparing pollen data with other proxies contained either in the Neotoma Paleoecology Database or residing in other databases, such as the Global Charcoal Database (Power et al., 2008) will provide new insights into more robust reconstructions of paleoenvironmental change. The list of sites will become available through Neotoma while a new LAPD literature website aims to facilitate an overview of expanding research.