



## Estimating Neotropical palaeotemperature and palaeoprecipitation using plant family climatic optima

Surangi W. Punyasena\*

Center for Tropical Paleocology and Archeology, Smithsonian Tropical Research Institute, Box 0843-03092, Balboa, Ancón, Panamá, Panamá

### ARTICLE INFO

#### Article history:

Received 2 November 2007

Received in revised form 18 April 2008

Accepted 23 April 2008

#### Keywords:

Palynology

Neotropics

Palaeotemperature

Palaeoprecipitation

Kernel density estimation

Maximum likelihood

A. H. Gentry

### ABSTRACT

The member species of individual plant families, particularly within the tropics, can be found in very different habitats. Despite this ecological and physiological diversity, the results of this study demonstrate that heterogeneity within families does not prevent climatic generalizations at higher taxonomic ranks. Modelled distributions that incorporate local abundance data can discriminate climatically among many plant families and can potentially be used to reconstruct palaeoclimate from sub-fossil and fossil pollen assemblages.

Presented are four probabilistic models of temperature and precipitation generated using published transect data of the spatial distributions of 154 Neotropical plant families and the Gaussian kernel probability densities of their occurrence (presence–absence), absolute abundance (density), proportional abundance (dominance), and basal area. Thirty-five published pollen surface samples, representing contemporary Central and South American lowland vegetation, were used to test the accuracy of the modelled climate estimates. The models are able to identify relative differences in temperature and precipitation, with absolute and proportional plant family abundances providing the most accurate estimates. There was little statistical difference among the different models and weighting methods used. The accuracy of individual estimates was more dependent on the number of pollen assemblage families shared with the modern vegetation dataset than any other investigated variable.

Family-based climatological proxies have several advantages over assemblage or biome-based reconstructions. Climatic estimates can potentially be derived even when novel combinations of taxa are observed in the fossil record, extending the utility of models based on plant family distributions to palaeobotanical data from deeper time. The approach is also dynamic. Plant communities do not need to be in equilibrium and individualistic dynamics can be incorporated into estimates. Finally, the models provide working hypotheses of how individual species may respond to climatic change based on the climatic specialization of their respective families.

© 2008 Elsevier B.V. All rights reserved.

### 1. Introduction

Plants are among the best organismal indicators of terrestrial climate (Holdridge, 1947; Walter, 1985; Woodward, 1987; Walter and Breckle, 2002). There are often clear physiological reasons for the biogeographic distributions of many plant species, including cold tolerance, heat sensitivity, and moisture requirements (Prentice et al., 1992, and references therein). This relationship between plant distribution and climate has also been observed through time. Dramatic shifts in species distributions have been documented globally following the Last Glacial Maximum and subsequent changes in global climate (e.g. in temperate biomes: Comes and Kadereit, 1998; Davis and Shaw, 2001; Jackson and Overpeck, 2000; Jackson et al., 2000; and in the Neotropics: Mayle et al., 2000; Bush et al., 2004). The responses

are individualistic, suggesting that ecological and physiological differences among taxa, and potentially a degree of stochasticity, are governing the observed changes in distribution (Huntley and Webb, 1989; Overpeck et al., 1992; Jackson and Overpeck, 2000; Jackson and Williams, 2004; Bush et al., 2004). The recognized relationship between the climate of a region and its resident plants makes botanical remains a useful tool of palaeontologists and Quaternary scientists in the reconstruction of palaeoclimate.

Comparative assessments of family-level biogeographic patterns indicate that even large families demonstrate climatic preferences (Punyasena et al., 2008). This study determines the extent of that specificity and demonstrates its applicability to palaeoclimate reconstructions. I present an alternative approach to determining Neotropical palaeotemperature and palaeoprecipitation from fossil pollen time series using the climatic affinities of 154 modern Neotropical plant families modelled from transect data collected by A. H. Gentry from 147 localities across Central and South America (Gentry, 1988;

\* Fax: +507 212 8154.

E-mail addresses: [PunyasenaS@si.edu](mailto:PunyasenaS@si.edu), [spunya1@illinois.edu](mailto:spunya1@illinois.edu).

Phillips and Miller, 2002). The accuracy of the model is evaluated by comparing modelled estimates for Neotropical surface pollen samples (representing modern vegetation and climate) to known precipitation and temperature data. Finally, I discuss the advantages and limitations associated with this method in the context of existing palaeoclimate proxy approaches.

### 1.1. Rationale for developing a plant family-based palaeoclimate model

The assumption implicit in the reconstruction of palaeoclimate from fossil plants is that climatic differences among taxa reside primarily at the species level (Bartlein et al., 1986; Prentice et al., 1992; Jackson and Overpeck, 2000; Williams et al., 2004). Most analyses of range changes in modern climate change studies (e.g. Parmesan and Yohe, 2003; Thomas et al., 2004) and descriptions of modern plant distributions (e.g. Thompson et al., 1999) have focused on species. However, given the potentially limited spatial range and lifespan of many species, species-level assessments of plant-climate relationships often cannot be generalized (Prentice et al., 1992). This becomes an even greater problem when studying tropical ecosystems, where little is known about modern species distributions, or the pre-Quaternary, when modern species are not present or exclusively identifiable. Explicitly incorporating higher taxonomic biogeographic and ecological patterns allows a more generalized approach to the reconstruction of climate.

In the pre-Quaternary, qualitative assessments of climatic conditions are often made using the presence of indicator taxa, which, because of the limitations of taxonomic identification, are often families. The most common example is the use of palms to identify minimum temperatures (Wing and Greenwood, 1993; Greenwood and Wing, 1995; Pearson et al., 2001; Wilf et al., 2003), because of their known low freezing tolerance (Sakai and Larcher, 1987). However, there are few comparable physiological studies for other plant families. The limited information available has meant that many reconstructions of deep-time environments have relied on morphological analyses which do not necessarily rely on taxonomic identifications, including stomata (McElwain and Chaloner, 1995; McElwain and Chaloner, 1996; McElwain et al., 1999), leaf margin (Wilf, 1997; Royer and Wilf, 2006), leaf size (Royer et al., 2005), or a combination of leaf morphological characteristics (Wolfe, 1995; Gregory-Wodzicki, 2000).

However, the majority of fossil plant data, in both the Quaternary and pre-Quaternary, are from pollen and spores. This means that most palaeobotanical climatic interpretations are by necessity taxon-based, relying on the taxonomic identification of pollen and spore morphotypes. Quaternary climate reconstructions from pollen employ the minimum recognizable taxonomic unit. If diversity is low, or taxa morphologically distinct, identifications can be made at the species level; for example, the use of *Curatella americana* as a dry forest indicator in lowland Bolivia (Mayle et al., 2000; Burbridge et al., 2004). However, most Quaternary pollen morphotypes are identified to genera, or for morphologically homogenous clades, such as Poaceae, to family (Traverse, 2007: 97–103).

Prior quantitative climatic reconstructions for the New World tropics have been specific to the individual study and its genera, and have used distributional ranges based on herbarium samples (Bush et al., 2004; Urrego et al., 2005). My goal was to develop a generalized palaeoclimate-vegetation model for the lowland Neotropics that was potentially applicable to both Quaternary and pre-Quaternary palynological records. Such a model would provide an additional means of understanding long-term trends in Neotropical climate using the growing Neotropical palynological dataset. I explore the limits of using plant family identifications, the taxonomic rank most readily available in the Quaternary and pre-Quaternary plant record, and examine key similarities and differences between the probabilistic palaeoclimate model described here and existing taxon-based methodologies. Additionally, by using different abundance measures (presence/absence, absolute

abundance, proportional abundance, and basal area), I am able compare these alternate descriptions of family distributions by the accuracy of the climatic information generated.

### 1.2. Approaches to palaeoclimate reconstruction

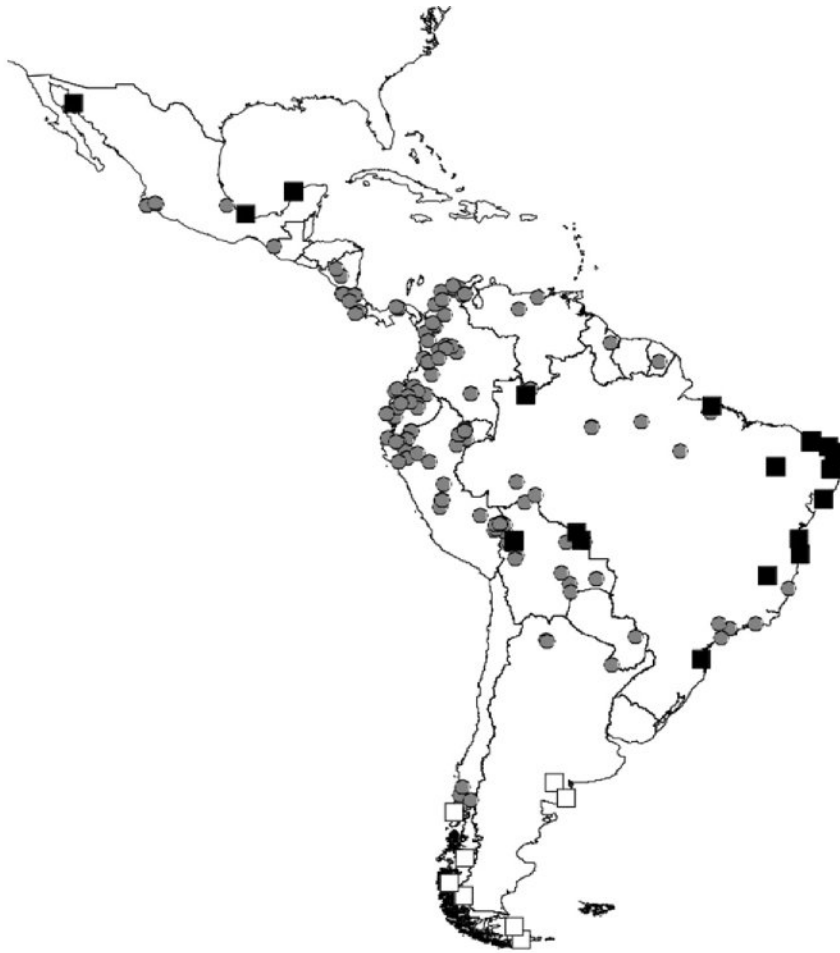
All taxon-based palaeoclimatic reconstruction methods employ a similar philosophy: the occurrence or abundance of modern plant clades is linked to a modern range of climates; the appearance of those taxa in the fossil record is then used to infer past climates. This relationship can be modelled in the forward, classical sense, where taxon abundances are a function of climate, and climate estimates are indirectly calculated, or the relationship can be modelled inversely, where climate is a function of taxon abundances, and estimates are calculated directly from fossil data (ter Braak, 1995).

Among the inverse models, regression-based approaches employ either multiple linear regression of pollen assemblage data (Davis and Anderson, 1985; Huntley and Prentice, 1988), which is most appropriate for modelling plant distributions along short climatic gradients (Birks, 1998; ter Braak and Prentice, 1988) or canonical techniques such as weighted averaging, which are designed for unimodal distributions (ter Braak, 1986, 1995; Birks, 1998). Alternatively, the modern analogue technique can be used to quantify similarity between the percent abundances of modern pollen assemblages and fossil assemblages based on a distance metric (Overpeck et al., 1985; Whitlock and Bartlein, 1993).

These pollen spectra approaches fail when “non-analogue communities,” plant assemblages unlike those seen in modern communities, appear in the fossil record (Jackson and Williams, 2004). Though often blamed on the low taxonomic resolution of pollen data and allochthonous pollen (Birks and Birks, 2000), non-analogue communities may also be a consequence of unique combinations of taxa resulting from climatic conditions that are outside the range of modern climates (Mosbrugger and Utescher, 1997; Jackson and Overpeck, 2000; Jackson and Williams, 2004). Models using plant functional types (PFTs) circumvent this problem to some extent by abstracting plant assemblages to biomes rather than maintaining their taxonomic identities (Prentice et al., 1992; Williams et al., 2004). However, even biome-level characterizations potentially face non-analogues as a result of extreme combinations of climatic influences, since component taxa assort independently along continuous gradients rather than strict biome categorizations (Williams et al., 2004).

The presence of non-analogue fossil communities means that community-aggregate analyses of pollen spectra may be less reliable under climatic scenarios which produce unique combinations of taxa. In contrast, the concept advocated by this paper is philosophically similar to the coexistence approach (Birks and Birks, 1980; Birks and Gordon, 1984; Kershaw and Nix, 1988; Mosbrugger and Utescher, 1997), which uses the nearest living relatives of fossil taxa to define climatic tolerances and overlapping plant occurrences to reconstruct palaeoclimate. Taxa are treated as individual units. The relationship between taxa and climate, and not communities and climate, are the basis of the palaeoclimate reconstruction.

The forward models of coexistence methods traditionally have used presence-absence data to define the upper and lower limits of a taxon's bioclimatic range. The procedure described in this study incorporates a probability distribution within these limits based on the abundance of individual families. In this respect, the methodology that I have adopted in this study is related to the likelihood surfaces, or “climatic response surfaces,” of Bartlein et al. (1986) and Webb et al. (1993), with three key differences. First, the focus of this study is on family-level generalizations; second, a kernel density function is used to calculate probability densities (Silverman, 1986); and third, maximum likelihood is used to calculate the most likely climatic values of a given pollen assemblage (McCullagh and Nelder, 1989; Bush et al., 2004). This forward modelling approach is not immune to



**Fig. 1.** Map of all data localities used in the study. Gentry forest transect sites are illustrated as grey circles. Temperate surface pollen data from the Latin American Pollen Database (LAPD) appear as white squares. Tropical surface pollen data from LAPD, Mayle et al. (2000), Urrego (2006), and Behling et al. (2000) appear as black squares.

problems posed by non-analogous combinations of taxa, but the uncertainty of the climatic estimates generated is explicit.

## 2. Methods

There are two parts to the study presented here:

- (1) The development of a climate-niche model of modern Neotropical families based on published forest transects. The model uses data on the occurrence (presence or absence), absolute abundance, relative abundance, and the proportional basal area of individual plant families at a given site to estimate the probability of finding a particular family along temperature or moisture gradients.
- (2) A test of the accuracy of modelled climate results against actual climate values using surface samples of pollen assemblages. Surface samples are from the sediment-water interface of a lake bottom and pollen within surface sediments represent contemporary vegetation.

### 2.1. Modern vegetation data

Information on plant family abundances within local Neotropical communities are from rapid transect data collected by A. H. Gentry (Gentry, 1988; Phillips and Miller, 2002). The dataset is one of the few tropical plant transect databases with extensive geographic coverage, a standardized collection methodology, and both taxonomic and abundance data and diameter measurements at breast height (dbh) (Phillips and Miller, 2002). As such, it is among the best ecological datasets available for developing a probabilistic model of Neotropical plant family

distributions, with specimen vouchers available for each unique species from each site. The data, however, are not of complete distribution ranges and are exclusively of forest communities. Thus, the final model will necessarily underestimate the total environmental range of Neotropical taxa and exclude non-forest taxa. Instead, these data provide specific information on the taxa which are known to be present at a single site. Data includes all stems > 2.5 cm dbh within ten 2 × 50 m transects. Small herbaceous individuals are excluded. Undescribed or unidentified species are catalogued as morphotypes (Gentry, 1982; Gentry, 1990).

I included all 147 transects in continental Central and South America, covering the area between 19.6°N–40.3°S and 40.1°W–105.1°W (Fig. 1). One hundred and forty-four of these sites fall within 20°N–30°S. The data include 154 families in total (Table 1). The taxonomic identification of individuals within the transect data were corrected and revised as described in Punyasena et al. (2008). In addition, six optional APG II synonymies were revised: Cochlospermaceae (Bixaceae); Krameriaceae (Zygophyllaceae); Quiinaceae (Ochnaceae); Rhizophoraceae (Erythroxylaceae); Trigoniaceae (Chrysobalanaceae); and Valerianaceae (Caprifoliaceae). All individuals, with the exception of 264 of 49,928 individuals, were uniquely assigned to a family.

### 2.2. Modern climate data

Temperature and precipitation data for each of the Gentry transect localities were taken from the University of East Anglia mean monthly climate dataset, available at a half-degree spatial resolution (New et al., 1999). Data are from published weather measurements from 1961–

**Table 1**

Plant families found in the Gentry transect dataset and used in developing the palaeoclimate models

Acanthaceae	Chrysobalanaceae	Lardizabalaceae	Podocarpaceae
Achatocarpaceae	Clethraceae	Lauraceae	Polygalaceae
Actinidiaceae	Clusiaceae	Lecythidaceae	Polygonaceae
Alzateaceae	Combretaceae	Liliaceae	Polypodiaceae
Amaranthaceae	Commelinaceae	Linaceae	Proteaceae
Anacardiaceae	Connaraceae	Loganiaceae	Pteridaceae
Anisophylleaceae	Convolvulaceae	Lomariopsidaceae	Ranunculaceae
Annonaceae	Coriariaceae	Loranthaceae	Rhamnaceae
Apocynaceae	Cornaceae	Lythraceae	Rosaceae
Aquifoliaceae	Cucurbitaceae	Magnoliaceae	Rubiaceae
Araceae	Cunoniaceae	Malpighiaceae	Rutaceae
Araliaceae	Cupressaceae	Malvaceae	Sabiaceae
Areaceae	Cyatheaceae	Marantaceae	Salicaceae
Aristolochiaceae	Cyclanthaceae	Marcgraviaceae	Santalaceae
Asparagaceae	Cyrillaceae	Melastomataceae	Sapindaceae
Aspleniaceae	Dichapetalaceae	Meliaceae	Sapotaceae
Asteraceae	Dicksoniaceae	Menispermaceae	Saxifragaceae
Basellaceae	Dilleniaceae	Monimiaceae	Scrophulariaceae
Begoniaceae	Dioscoreaceae	Moraceae	Simaroubaceae
Berberidaceae	Dryopteridaceae	Musaceae	Smilacaceae
Betulaceae	Ebenaceae	Myricaceae	Solanaceae
Bigoniaceae	Elaeocarpaceae	Myristicaceae	Staphyleaceae
Bixaceae	Ericaceae	Myrsinaceae	Strelitziaceae
Blechnaceae	Erythroxylaceae	Myrtaceae	Styracaceae
Boraginaceae	Euphorbiaceae	Nyctaginaceae	Symplocaceae
Brassicaceae	Fabaceae	Ochnaceae	Theaceae
Bromeliaceae	Fagaceae	Olacaceae	Theophrastaceae
Brunelliaceae	Gentianaceae	Oleaceae	Thymelaeaceae
Burmanniaceae	Gesneriaceae	Onagraceae	Ulmaceae
Burseraceae	Gnetaceae	Opiliaceae	Urticaceae
Buxaceae	Hamamelidaceae	Orchidaceae	Verbenaceae
Cactaceae	Heliconiaceae	Oxalidaceae	Violaceae
Campanulaceae	Hernandiaceae	Papaveraceae	Vitaceae
Caprifoliaceae	Humiriaceae	Passifloraceae	Vochysiaceae
Caricaceae	Hydrangeaceae	Phytolaccaceae	Winteraceae
Caryocaraceae	Icacinaceae	Pinaceae	Zingiberaceae
Caryophyllaceae	Juglandaceae	Piperaceae	Zygophyllaceae
Celastraceae	Lacistemataceae	Plantaginaceae	
Chloranthaceae	Lamiaceae	Poaceae	

1990, spanning the period when the Gentry transect data were collected (1971–1993). Annual values for minimum, mean, and maximum temperatures, precipitation, and total number of wet days are included in the analysis. Seasonal variables for temperature, vapour pressure, and precipitation were also investigated initially. However, seasonal measures strongly covary in their effect on tropical vegetation and the effect of vapour pressure was observed to covary as a function of temperature (Punyasena et al., 2008). For this reason, only the results of the annual analysis of temperature and precipitation are relevant and discussed here.

**2.3. Modern surface samples**

Sixteen of the pollen surface samples used to test the accuracy of the climate reconstructions are from the Latin American Pollen Database (LAPD) (<http://www.ncdc.noaa.gov/paleo/lapd.html>). The LAPD is a public repository of pollen counts taken predominately from lake sediment cores, but also include data from packrat middens and the Deep Sea Drilling Project (DSDP). Data are contributed by multiple authors. Ninety-five sites in Central and South America are available online. Of these, sixteen sites are found below 390 m in elevation and have recent surface sediments (<100 <sup>14</sup>C yr BP) (Godley and Moar, 1973; Markgraf, 1980; Markgraf, 1983; Mercer and Ager, 1983; Markgraf, 1989; Byrne and Horn, 1989; Ashworth et al., 1991; Markgraf, 1991; Graf, 1992; Markgraf, 1993; Schabitz, 1994; Anderson and Vandevender, 1995; Behling, 1995a; Behling, 1995b; Behling, 1996; Leyden et al., 1996). These were used to test the accuracy of the model. Eight of these sites are between 30°N–30°S, seven are from lowland tropical lakes (Byrne and Horn, 1989; Behling, 1995a; Behling, 1995b; Behling, 1996; Leyden et al., 1996) and one is from a packrat midden (Anderson and Vandevender,

1995). Four additional surface samples from lowland Bolivia were provided by F. E. Mayle (Mayle et al., 2000; Burbridge et al., 2004) and D. H. Urrego and M. A. Bush (Urrego, 2006). Fifteen additional sites from north-eastern Brazil were provided by H. Behling (Behling et al., 2000). The eight temperate samples lie outside the temperature range best represented by the Gentry data (Fig. 2). Angiosperm pollen taxonomy was revised to reflect APG II family classifications (APG II, 2003), with synonyms corrected.

**2.4. Modelling the climate space of modern families**

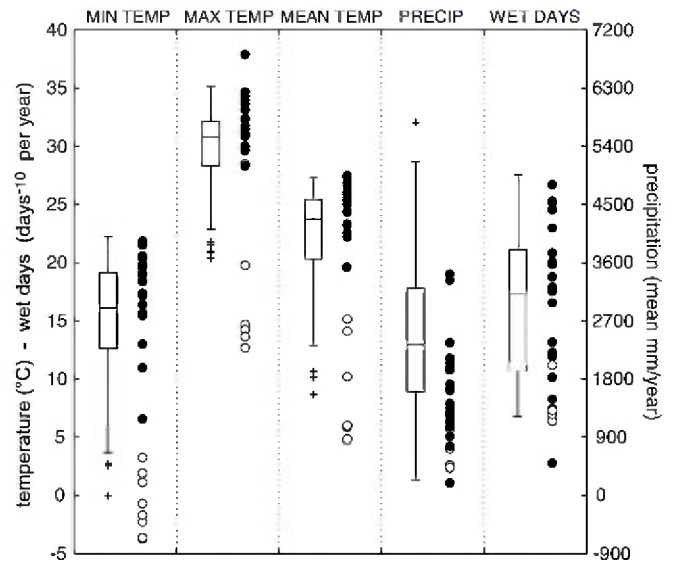
All 154 families within the Gentry dataset were individually modelled as kernel probability densities for the five annual measures of temperature and precipitation. The probability density reflects the likelihood of finding a given family at a given point along a climate gradient. Kernel density estimators are a nonparametric means of calculating a probability density function, in that they do not have a fixed structure and depend upon all data points to reach an estimate (Bowman and Azzalini, 1997). They are empirically derived from the vegetation data and avoid the unrealistic assumption of a fixed unimodal Gaussian distribution of climate tolerances for plant families. Kernel densities are analogous to histograms, smoothing the relative contribution of an individual datum over the local neighbourhood of that point. The density of a single data point *x* is calculated as

$$\hat{f}(x) = \frac{1}{n} \sum K\left(\frac{x-x(i)}{h}\right), i = 1, \dots, n \tag{1}$$

where *n* is the total number of observations, *K* is the kernel function, and *h* is the bandwidth used. The contribution of any given datum, *x*(*i*), to the density estimate at point *x* depends on the distance between *x*(*i*) and *x*, the kernel function used, and the bandwidth across which the kernel function is calculated. The density estimates can be fitted to a number of kernel shapes. The Gaussian kernel is used in this analysis, defined as

$$K(t) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-t^2/2}. \tag{2}$$

This kernel is used based on the theoretical assumption that the response surface of a plant family (*sensu* Bartlein et al., 1986) is the



**Fig. 2.** The range of the five climatic variables (annual minimum temperature, annual maximum temperature, annual mean temperature, annual precipitation, and annual number of wet days) encompassed by the 147 Gentry transect sites are shown by the white quartile plots. Outliers are marked by crosses. Adjacent are the respective temperature and precipitation values for the 35 surface pollen samples. The 27 tropical sites (located between 30° N–30° S) are represented by black circles, while those for the eight temperate sites are shown in white.

aggregation of the distribution of its respective species, which are presumed to be unimodal along climatic gradients (Prentice et al., 1992; Birks, 1998). Although the Gaussian kernel is used to weight the relative contribution of the abundance data, this does not mean that a normal probability distribution will be the result.

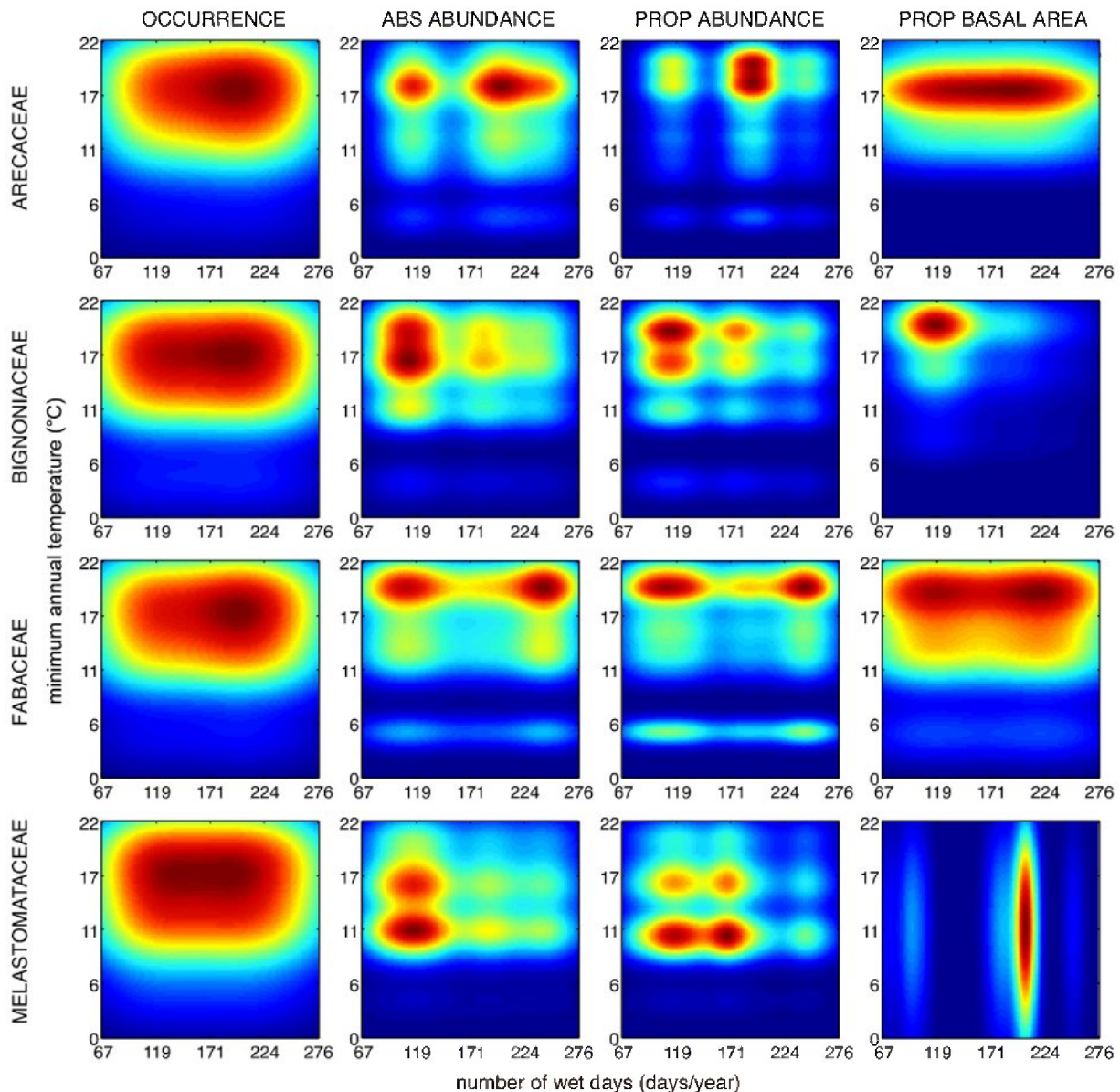
The shape of the kernel density function is affected more by the size of the bandwidth ( $h$ ) than by the kernel used. With a large enough  $h$  the density estimate will appear unimodal; with a small enough  $h$  the density estimate will appear bimodal or multimodal (Wang, 2003). The Silverman's Critical Bandwidth test is one means of testing for the degree of multimodality using a bootstrap approach (Silverman, 1981). However, there are reported limitations to Silverman (1981), given that the bootstrap distribution of a kernel density estimator is not a good approximation of the unconditional distribution of the estimator. As a result, the bandwidth used in this analysis was determined using the default method known as "Silverman's rule

of thumb" (Silverman, 1986), which is based on the standard deviation of the data ( $\sigma$ ) and number of observations ( $n$ ),

$$h = \sigma(4/3n)^{-1/5} \quad (3)$$

This approach is optimal for normally distributed data and will over smooth the probability distribution of families with few individuals and patchier distributions (Bowman and Azzalini, 1997). However, this option was more conservative than determining an *ad hoc* bandwidth for each family individually.

For this study, the kernel probability densities were calculated in four ways: based on (1) the occurrence (presence/absence) of a family at a site; (2) the absolute abundance of a family at a given site (density per 0.1 ha); (3) the proportional abundance of a family at a site (relative dominance) of the vegetation data; and (4) the proportional basal area of a family at a site. The data were not evenly



**Fig. 3.** Two-dimensional kernel density plots, with the annual number of wet days represented along the  $x$ -axis and annual minimum temperature represented by the  $y$ -axis. Probability densities were calculated using occurrence (presence/absence data), absolute abundance (number of individuals per 0.1 ha plot), proportional abundance (percentage of individuals per plot), and proportional basal area (percent area per plot). Four common families found across a range of precipitation and temperature values are illustrated. Arecaceae is most abundant and most dominant at warmer and wetter sites. Bignoniaceae, in contrast, is most abundant and most dominant at warmer and drier sites. Fabaceae is most abundant and most dominant at warmer sites; and Melastomataceae is most abundant and most dominant at cooler sites.

**Table 2**

Plant families found at a single Gentry transect site

Alzateaceae
Anisophylleaceae
Asparagaceae
Berberidaceae
Burmanniaceae
Caryophyllaceae
Commelinaceae
Coriariaceae
Cornaceae
Cupressaceae
Cyrillaceae
Hamamelidaceae
Lamiaceae
Lardizabalaceae
Liliaceae
Oxalidaceae
Plantaginaceae
Polypodiaceae
Pteridaceae
Scrophulariaceae

distributed across the range of temperature and precipitation levels observed, so were binned to prevent the calculated densities from solely reflecting the sampling distribution of the Gentry transects. Data within bins are averaged, so remove the sampling bias of multiple sites at a given temperature or precipitation range. Bins were equally spaced across the range of the Gentry data, with 100 bins per gradient. Temperature and precipitation are modelled separately in the analysis, which is not entirely appropriate from a physiological context. However, from Punyasena et al. (2008) we know that the variation in abundance for individual families along temperature and precipitation gradients is largely independent. Future refinement of the model will incorporate simultaneous modelling of precipitation and temperature variables.

### 2.5. Palaeoclimate estimates from surface pollen

Likelihood values for the climate estimates are the sum of the log likelihoods of the families found within a pollen sample (identical to the joint product of the kernel densities of these families). The

**Table 3**

Families that could not be modelled accurately, given the default parameters of the kernel density models

Achatocarpaceae
Begoniaceae
Buxaceae
Cactaceae (density)
Campanulaceae
Dicksoniaceae
Fagaceae
Gesneriaceae
Gnetaceae (dominance)
Hydrangeaceae
Juglandaceae
Lacistemaaceae
Linaceae
Loganiaceae
Myricaceae
Opiliaceae (density)
Smilacaceae
Theophrastaceae (density)
Winteraceae
Zygophyllaceae

**Table 4**

Families found at a limited number of Gentry transect sites

Aristolochiaceae (2)
Aspleniaceae (2)
Basellaceae (3)
Betulaceae (3)
Blechnaceae (4)
Clethraceae (5)
Dioscoreaceae (2)
Gentianaceae (4)
Lythraceae (4)
Musaceae (3)
Oleaceae (4)
Orchidaceae (4)
Papaveraceae (2)
Pinaceae (2)
Plantaginaceae (1)
Santalaceae (3)
Saxifragaceae (3)
Strelitziaceae (2)

Number of sites in parentheses.

probability of finding a given group of families ( $F_1 F_2 \dots F_n$ ) at a given temperature or precipitation level ( $\theta$ ) is then

$$L(\theta) = \Pr(F_1 F_2 \dots | \theta). \quad (4)$$

The most likely temperature and precipitation estimate for a fossil assemblage is the maximum sum of the log likelihoods of these densities or

$$\ln L(\theta) = \sum_{i=1}^n \ln \Pr(F_i | \theta) \quad (5)$$

(McCullagh and Nelder, 1989; Bush et al., 2004).

There were three methods of tabulating the pollen data: (1) equally weighted, using only the presence of a family in calculating the most likely climate value; (2) weighted as percentages of number of grains in the surface sample, normalizing the rows of the data matrix (the sites) which emphasizes the relative abundance of taxa; and (3) weighted as the standard deviation of the percentages, normalizing the columns of the data matrix (the families) which emphasizes relative changes within a family. The relative influence of rare taxa (those represented by 15 or fewer pollen grains) was also tested by iterating removing families with 15 or fewer grains from calculations of the climate estimates.

## 3. Results

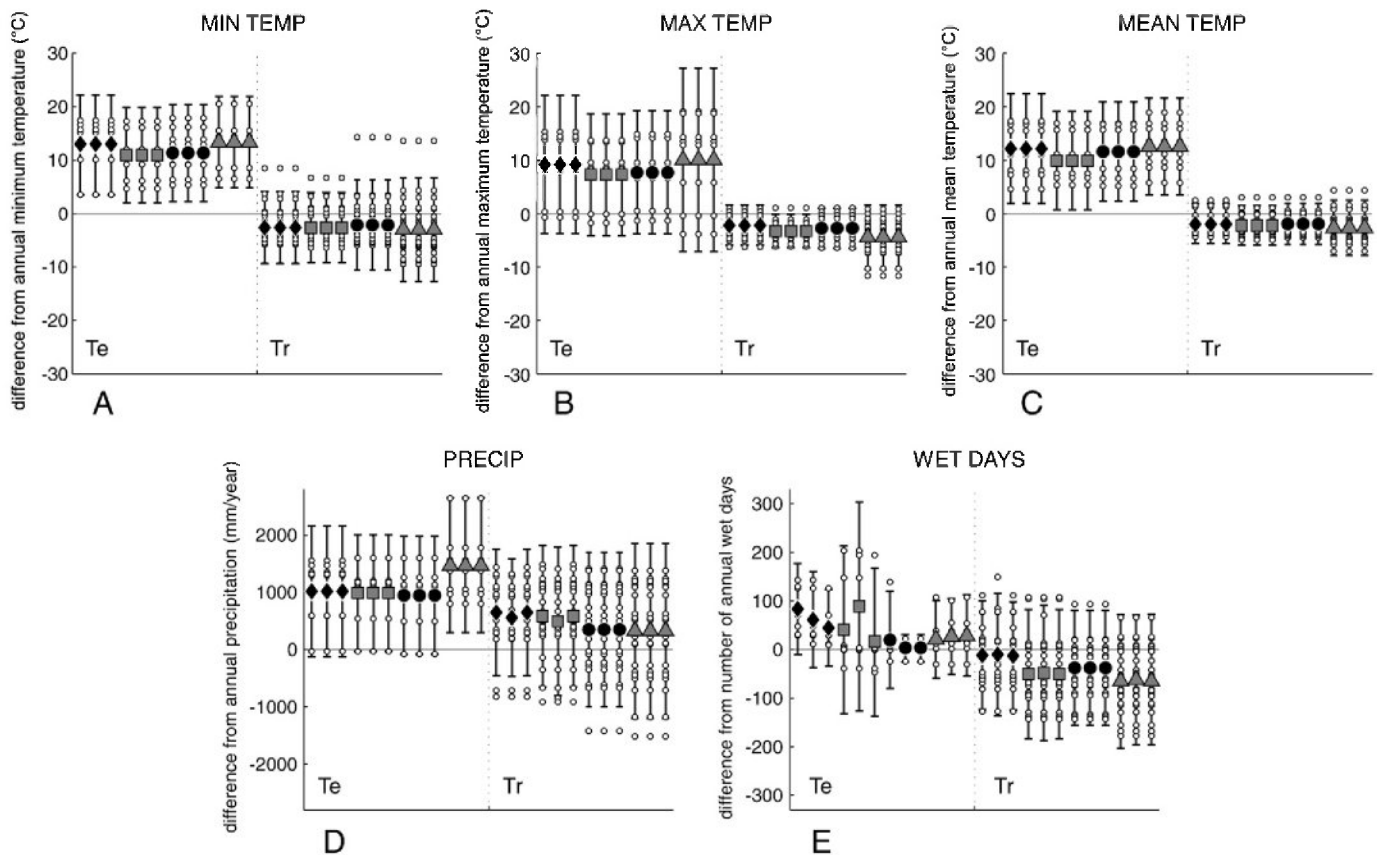
### 3.1. Climate space of modern families

The range of temperatures and precipitation levels in which a family was found was modelled using the kernel density function. Examples of the abstracted climate space of four large families (Arecaceae, Bignoniaceae, Fabaceae, and Melastomataceae) are illustrated in Fig. 3. Despite the widespread occurrence of these four families, there are clear climatic preferences evident in the density (absolute abundance), dominance (proportional abundance), and basal area occupied by each family. This is despite the fact that each family is composed of many species with presumably their own

**Table 5**

Rare families in the Gentry dataset which appear to serve as good indicators

Cactaceae (proportional abundance only)
Cupressaceae
Lardizabalaceae



**Fig. 4.** Accuracy of climatic estimates from modern surface pollen samples, with taxa represented by a single pollen grain excluded, for minimum temperature (A); maximum temperature (B); mean temperature (C); annual precipitation (D); and number of wet days (E). Difference between individual climate estimates and actual values are plotted as white circles, temperate sites (Te) are shown on the left and tropical sites (Tr) are shown on the right. Two standard deviations are illustrated by the black error bars. Mean values for estimates generated using the presence-absence (occurrence) model are marked as black diamonds. Mean values for estimates generated using the absolute abundance (density) model are marked as grey squares. Mean values for estimates generated using the proportional abundance (dominance) model are marked as black circles. Mean values for estimates generated using the proportional basal area model are marked as grey triangles. Pollen data for each model were weighted in three ways and are illustrated in order: (1) equally weighted, using only the presence of a family in calculating the most likely climate value; (2) weighted as percentages of number of grains in the surface sample; and (3) weighted as the standard deviation of percentage abundance.

individual climatic niches. The palms (*Arecaceae*) are most abundant and dominant at wet, warmer sites, while *Melastomataceae* is most abundant and dominant in cooler sites, while accounting for more proportional basal area in wetter areas. *Bignoniaceae* is most abundant and dominant in dry, warm sites, while the legumes (*Fabaceae*) are most abundant and dominant in warm sites, regardless of precipitation levels.

Many families are only found at a single site (Table 2) and some were too rare (low abundance or limited spatial distribution) to be modelled correctly (Table 3). Taxa which are predominately composed of small herbaceous species, such as *Poaceae*, may also be poorly represented by a model based only on the spatial distribution of individuals greater than 2.5 cm dbh. Several families were not modelled correctly by the default density bandwidth, with the calculated probability densities too flat (Table 4). Several taxa, including notably *Cupressaceae* and *Cactaceae*, appear to be good indicators of climate, given their climatic specificity, despite their limited inclusion in the Gentry dataset (Table 5).

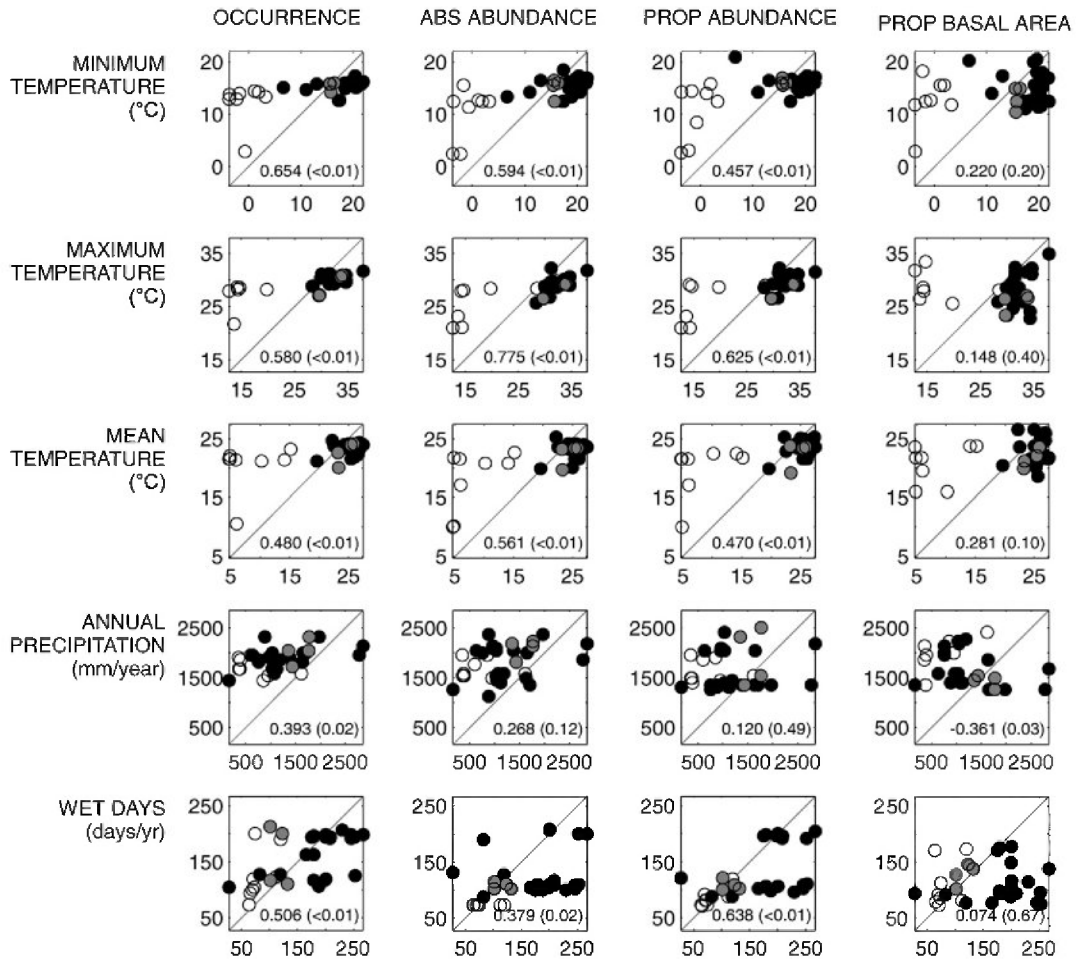
### 3.2. Evaluation using modern surface samples

The model evaluations are presented with caveats. Only 35 samples were used to test the analysis and the available sites were not evenly distributed geographically (Fig. 1). When the estimates generated for individual surface samples are compared to the actual climate values of each site, all measures are comparable (Fig. 4). The

model based on basal area performed the most poorly, perhaps reflecting the incompleteness of the Gentry data in providing total biomass. With the exception of the estimates for the annual number of wet days, the average results for all estimates were more accurate for tropical sites than for the temperate. Temperate temperatures were overestimated. This is not unexpected, given the spatial distribution of the Gentry sites relative to the surface samples. All 147 Gentry transect sites, with the exception of three Patagonian transects, fall within 20°N and 30°S (Fig. 1).

Mean differences for the three temperature estimates were marginally better using the dominance kernel density model, where the modelled probability densities of individual plant families were based on their proportional abundances at a given site. However, this cannot be distinguished statistically from the occurrence model (based on presence/absence) and density model (based on absolute abundances per 0.1 ha) (Fig. 4A–C). There was even less of an effect of weighting pollen data, with the results of the three weighting methods (equal weight, percent abundance in a sample, and standard deviation among samples) nearly identical (Fig. 4).

There was a large amount of error in the estimates of annual precipitation and the annual number of wet days, primarily resulting from the inability of any of the models to distinguish among high precipitation sites (Fig. 5). However, in estimating the severity of the dry season, reflected in the number of wet days, the estimates for temperate samples using proportional abundance kernel densities were remarkably accurate and precise (Fig. 4E). These sites are all within the drier



**Fig. 5.** Scatter plots of actual to estimated values of individual surface pollen samples for each of the climate variables for the four models: occurrence (presence–absence), absolute abundance (number of individuals per 0.1 ha plot), proportional abundance (percentage of individuals per plot) and proportional basal area (percent area per plot). Actual values for temperature and precipitation are plotted along the x-axis, estimated values are on the y-axis. All pollen data are weighted by the standard deviation of total family abundances. Temperate sites are shown in white; tropical sites in black. The four Bolivian sites are highlighted in grey. 1:1 equality is illustrated by a grey diagonal. Sample estimates falling above this line are overestimated. Those falling below this line are underestimated. In the lower right-hand corner is Spearman's rank correlation ( $\rho$ ), with  $p$ -values in parentheses. The density and dominance models best preserve the relative rank of the individual temperature variables. All three models had difficulty distinguishing among the high precipitation sites, although the occurrence model may best distinguish the annual number of wet days, suggesting that the presence of specific taxa may be most useful for distinguishing the severity of the dry season.

range of sites represented by the Gentry dataset (Figs. 2, 5), suggesting that extreme differences in the length of the dry season can potentially be identified using this methodology.

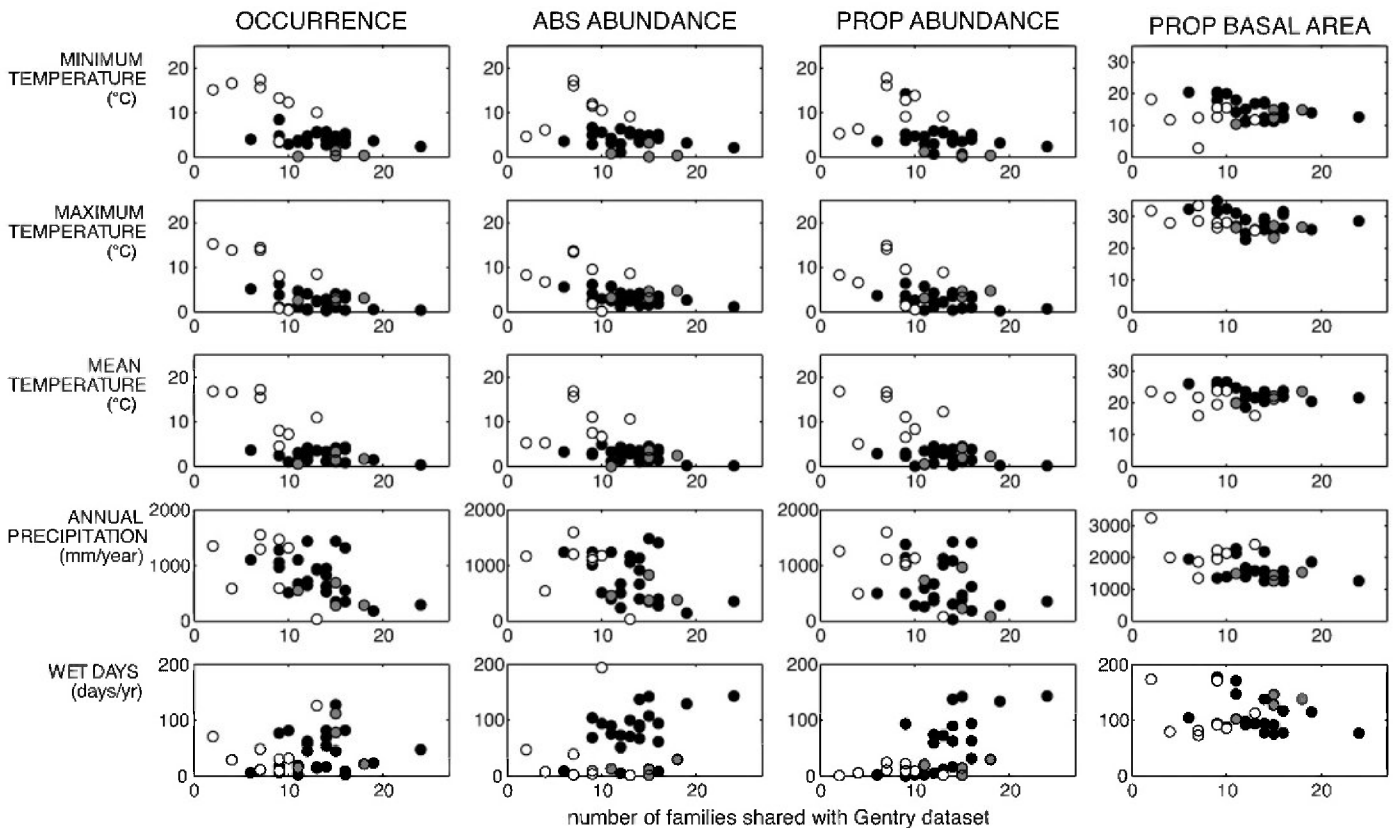
The influence of singletons and rare pollen taxa on the accuracy of these results was explored by iteratively removing taxa with fewer than two to fifteen grains per sample. However, there was little effect of the removal of rare taxa from estimates for either tropical or temperate sites, particularly when pollen families were weighted by the standard deviation of their relative abundances. Removing taxa with four or fewer grains (1.34% of a 300-count sample) improved the accuracy of precipitation estimates, which suggests that these estimates are sensitive to the presence of rare, potentially allochthonous taxa. However, increasing the threshold to five or more grains increased the range of error, suggesting that non-dominant, rarer taxa are also providing important climatic information, particularly for estimates of precipitation.

The absolute values of the temperature and precipitation estimates were not consistently accurate, with the most extreme temperature estimates off by as much as eighteen degrees and the most extreme estimates of annual wet days off by as many as 200 days. However, the relative ranking of individual sites in terms of temperature or precipitation can be reconstructed with varying degrees of fidelity to the original data (Fig. 5). Notably, the four Bolivian sites, whose geographic

locations fell within the greatest density of Gentry transects, were consistently among the most accurate estimates (Figs. 5, 6). All four models had difficulty distinguishing among the high precipitation sites, although the occurrence model may best distinguish the annual number of wet days. This suggests that the presence of specific indicator taxa, whatever their abundance, may be useful for distinguishing the severity of the dry season.

The accuracy of the climate estimates generated by the four probabilistic models is also affected by the number of taxa within a pollen sample that are shared with the 154 modelled families (Fig. 6). The tropical pollen samples have the most taxa shared with the Gentry dataset, which is expected given that only three temperate sites were available in modelling the bioclimatic distributions. Fewer taxa mean fewer data on which to calculate the climate estimate. This bias is most important for the three temperature variables (annual minimum, maximum, and mean), where tropical estimates of temperature were more accurate than temperate sites. However, the larger calculated error for tropical sites with few taxa also suggests that the compositional diversity of a site is highly indicative of temperature, with temperature estimates an average of the climatic preferences of many taxa. In the case of the estimates for precipitation, the lack of correlation between the number of taxa and accuracy of the estimates again suggests that there may be specific indicator taxa, particularly for estimates on the severity





**Fig. 6.** Accuracy of estimates based on the number of taxa shared between samples and the Gentry dataset for the proportional abundance model, with pollen samples weighted by standard deviation. Absolute difference from actual values are shown. Temperate sites are shown in white, tropical sites are shown in black. The four Bolivian sites are highlighted in grey. Temperate sites are the least accurate, and share the least number of families with the Gentry dataset. The number of taxa made the most difference for the temperature estimates, and were less important for estimates of annual precipitation and the annual number of wet days.

of the dry season (annual number of wet days), which may be more informative than the overall familial composition. However, among the twelve sites with the most accurate results (estimates within  $\pm 10$  days), there were no taxa that were shared among all the sites, although seven of these sites did contain Poaceae.

#### 4. Discussion

The coexistence approach draws from niche theory and follows the same basic principles as the predictive bioclimatic models used to study the impact of climatic change on species distributions (e.g. Box, 1981; Pearson et al., 2002; Berry et al., 2002; Pearson and Dawson, 2003; Elith et al., 2006; Guisan et al., 2006). Defining the climatic space occupied by taxa underpins all research on the ecology of organisms as they relate to the physical environment. Maguire (1973) first described this environmental niche as a “fitness response surface”. It is essentially the  $n$ -dimensional environmental space described by Hutchinson (1958), and an extension of the concept of the species niche of Hutchinson (1957). Among plant-climate models at the subfamily level, a growing number of researchers are developing probabilistic methods using both unimodal models (Vasko et al., 2000; Toivonen et al., 2001) and kernel density distributions (Bush et al., 2004). For family distributions, where the aggregate pattern is a mixed signal of species distributions, the latter approach is intuitively more appropriate.

There are genuine taphonomic uncertainties in translating the ecological and climatic information inherent in modern abundance distributions to fossil data. Pollen data in particular can be biased by the differential representation and preservation of taxa due to pollination mechanism (Bush and Rivera, 1998). Anecdotal evidence also suggests that preservational differences due to pollen morphology (e.g. thickness of the exine) can also play a role. These taphonomic

biases can be addressed in part if the probability densities of the modern climate space were calculated using pollen surface samples rather than transects of standing vegetation. Resorting to published modern transect data was necessitated by the limited amounts of surface sample data that exist for the Neotropics.

The results, however, show that despite the multiple taphonomic filters between a living forest and its pollen representation, and the reality that families are composed of many individual species, each with its own presumed climatic optima, generalized information on warming, cooling, wetting, and drying *can* be extracted from a coarse taxonomic analysis of the palynological record. In addition, differential weighting of pollen counts (presence–absence, percentage, or standard deviation) had less effect on climate estimates than taxonomic composition. This suggests that the differential representation of pollen taxa, as a result of taphonomic influences and differing reproductive strategies, may not be as influential as it may first appear.

Clearly, the distribution of individual plant families is not solely a function of these five measures of temperature and water availability. Biological interactions among plant species, pollinators, and herbivores are not considered, among a host of other variables affecting ecological assembly which are important in the construction of accurate bioclimatic models (Guisan et al., 2006). There is some evidence that biotic competition and stress may not be as critical in defining ecological niches for plants, particularly at regional distributions (Prinzing et al., 2001). This study also does not account for the interaction between temperature and precipitation, the seasonality of climate, and variation in carbon dioxide. All three factors are likely to change significantly over time and to affect plant physiology and distribution (Jackson and Williams, 2004). Nonetheless, I have attempted to delineate the range of uncertainty inherent in the methodology, a step which is not always explicitly documented (Guisan et al., 2006).

Ter Braak (1995) emphasized that communities are assumed to be in equilibrium when applying traditional biome-based models of climate reconstruction. Although the inherent time-averaging of many tropical pollen cores (from hundreds to thousands of years) (Burbridge et al., 2004) means that short-term variability and stochastic noise is unlikely to be detected within the pollen record, it is perhaps overly simplistic to characterize any vegetation assemblage as a static entity. Vegetation change on Quaternary timescales has been observed as highly individualistic and dynamic (Jackson and Williams, 2004) and when high-resolution time series are available for the pre-Quaternary, similar Gleasonian assembly of communities is observed among deep-time plant fossils (Wing et al., 2005). The methodology advocated here provides a dynamic means of characterizing changes and does not need to assume an artificial equilibrium. Moving away from biome-based reconstructions allows for a more flexible interpretation of the fossil record. The degree to which this method is successful in interpreting long time series of Quaternary climatic change is demonstrated in Punyasena et al. (in press).

The relative importance of less abundant taxa is also evident, corroborating Birks' (1998) observation that rare taxa provide a large amount of climatic information. Rather than adding statistical noise, rare taxa served as important climatic indicators. This was particularly true for the precipitation and wet day estimates, where removing rare taxa marginally degraded the accuracy of estimates. For both temperature and precipitation measures, removing rare taxa also tended to increase the confidence limits of the estimates.

The model may be improved with a denser sampling of transects and the addition of temperate and high-elevation Neotropical sites. Unlike taxa in comparable studies from North America and Europe, the abundance distributions of many Neotropical taxa are not well known, and as more transect data are incorporated, the accuracy of the estimates are likely to increase. Additionally, some of the error in the annual precipitation estimates may also be a result of the uncertainty inherent in the climatology used, with greater inaccuracy for precipitation extremes (New et al., 1999). Improvements in the climate data available, therefore, will also improve the model's accuracy.

The assumption of this and other taxon-based climate proxy methods is that the ecology of a plant clade is conserved through time. This is likely on Quaternary timescales, given the long molecular age of many tropical plant clades (e.g. Inga, Richardson et al., 2001), and evidence that North American taxa, at the very least, are found in climatically similar regions at the LGM as in the present (Martinez-Meyer and Peterson, 2006). Conservation of plant ecological niches may also be the case even deep in time, given evidence of the phylogenetic conservation of many ecologically important traits (Ackerly, 2003). This suggests that there may be utility in a family-level palaeoclimate estimator even beyond Quaternary timeframes, although the upper and lower limits of the model's predictive ability are set by the modern climates sampled.

Sub-family bioclimatic methods have been used successfully in deep time, based on plant occurrence distributions alone (Kershaw, 1997), although given the broad distribution of these families, my results show that incorporating measures of abundance into the model may provide more accurate estimates of climate, particularly of temperature. Since the climate space of an individual family was based on modern transect data that largely ignored herbaceous individuals, the model may be more applicable to the interpretation of fossil leaf data than of pollen. Most leaf assemblages are derived from trees from a single locale (Burnham et al., 1992), while pollen data contains a sizeable percentage of herbaceous taxa and taxa transported from the surrounding watershed (Bush and Rivera, 1998; Jackson and Williams, 2004). Applying a family-based climate model to leaf assemblages of the pre-Quaternary geologic record would be a possible next step.

Finally, the results support recent conclusions reached using an independent analysis of the differing climatic preferences of individual

families represented by the Gentry transect data (Punyasena et al., 2008). Even large and widespread families have climatic preferences, and there are areas of climate space in which they are more abundant or more diverse. Approximated bioclimatic distributions provide working hypotheses of the likely diversity and abundance distributions of Neotropical families, and can potentially be used to understand modern patterns of diversity and predict plant response to future environmental changes.

Whether we extrapolate forward or backward in time, the ecological diversity and commonalities within plant families shape tropical forest communities and are the foundation for understanding changing ecosystems. Explicitly reconstructing these bioclimatic distributions and their inherent uncertainties allows a flexible and nuanced view of these changing ecosystems and offers alternate possibilities for reconstructing climate with fossil plants.

## Acknowledgments

I would like to sincerely thank the Missouri Botanical Gardens and Jim Solomon for maintaining public online access to A. H. Gentry's forest transect data and the VAST (VAScular Tropicos) nomenclatural database. Herman Behling, Mark Bush, Dunia Urrego, and Francis Mayle kindly provided the raw data for 19 surface samples. Michael Foote, Jennifer McElwain, Susan Kidwell, Gidon Eshel, Francis Mayle, and Scott Lidgard provided feedback on early drafts. The comments of the editor and two anonymous reviewers expanded the analysis and discussion of the literature. Support for this research was provided by US EPA STAR Fellowship FP-91637701-0 and the Chicago ARCS Foundation, and was completed as a doctoral student at the University of Chicago in the Committee on Evolutionary Biology. Current support is from a Smithsonian Postdoctoral Fellowship.

## References

- Ackerly, D.D., 2003. Community assembly, niche conservatism, and adaptive evolution in changing environments. *International Journal of Plant Sciences* 164 (3), S165–S184.
- Anderson, R.S., Vandevender, T.R., 1995. Vegetation history and paleoclimates of the coastal lowlands of Sonora, Mexico: pollen records from pack-rat middens. *Journal of Arid Environments* 30 (3), 295–306.
- APG II, 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society* 141 (4), 399–436.
- Ashworth, A.C., Markgraf, V., Villagran, C., 1991. Late Quaternary climatic history of the Chilean channels based on fossil pollen and beetle analyses, with an analysis of the modern vegetation and pollen rain. *Journal of Biogeography* 13 (1), 35–57.
- Behling, H., 1995a. A high-resolution Holocene pollen record from Lago Do Pires, SE Brazil: vegetation, climate and fire history. *Journal of Paleolimnology* 14 (3), 253–268.
- Behling, H., 1995b. Investigations into the Late Pleistocene and Holocene history of vegetation and climate in Santa-Catarina (S. Brazil). *Vegetation History and Archaeobotany* 4 (3), 127–152.
- Behling, H., 1996. First report on new evidence for the occurrence of *Podocarpus* and possible human presence at the mouth of the Amazon during the Late glacial. *Vegetation History and Archaeobotany* 5 (3), 241–246.
- Behling, H., Arz, H.W., Patzold, J., Wefer, G., 2000. Late Quaternary vegetational and climate dynamics in northeastern Brazil, inferences from marine core Geob 3104-1. *Quaternary Science Reviews* 19 (10), 981–994.
- Berry, P.M., Dawson, T.P., Harrison, P.A., Pearson, R.G., 2002. Modelling potential impacts of climate change on the bioclimatic envelope of species in Britain and Ireland. *Global Ecology and Biogeography* 11 (6), 453–462.
- Birks, H.J.B., 1998. Numerical tools in palaeolimnology: progress, potentialities, and problems. *Journal of Paleolimnology* 20 (4), 307–332.
- Birks, H.H., Birks, H.J.B., 2000. Future uses of pollen analysis must include plant macrofossils. *Journal of Biogeography* 27 (1), 31–35.
- Birks, H.J.B., Birks, H.H., 1980. *Quaternary palaeoecology*. Edward Arnold, London. viii, 289 pp.
- Birks, H.J.B., Gordon, A.D., 1984. Numerical methods in Quaternary pollen analysis. Academic Press, London. Orlando, viii, 317 pp.
- Bowman, A.W. and Azzalini, A., 1997. Applied smoothing techniques for data analysis: the kernel approach with S-Plus illustrations. Oxford statistical science series; 18. Clarendon Press, Oxford University Press, Oxford, New York, xi, 193 pp.
- Box, E.O., 1981. Macroclimate and plant forms: an introduction to predictive modeling in phytogeography. Tasks for vegetation science, vol. 1. U.S. and Canada Kluwer, Boston. xiii, 258 pp.

- Burbridge, R.E., Mayle, F.E., Killeen, T.J., 2004. Fifty-thousand-year vegetation and climate history of Noel Kempff Mercado National Park, Bolivian Amazon. *Quaternary Research* 61 (2), 215–230.
- Burnham, R.J., Wing, S.L., Parker, G.G., 1992. The reflection of deciduous forest communities in leaf litter: implications for autochthonous litter assemblages from the fossil record. *Paleobiology* 18 (1), 30–49.
- Bush, M.B., Rivera, R., 1998. Pollen dispersal and representation in a Neotropical rain forest. *Global Ecology and Biogeography* 7 (5), 379–392.
- Bush, M.B., Silman, M.R., Urrego, D.H., 2004. 48,000 years of climate and forest change in a biodiversity hot spot. *Science* 303 (5659), 827–829.
- Byrne, R., Horn, S.P., 1989. Prehistoric agriculture and forest clearance in the Sierra de los Tuxtles, Veracruz, Mexico. *Palynology* 13, 181–193.
- Comes, H.P., Kadereit, J.W., 1998. The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in Plant Science* 3 (11), 432–438.
- Davis, R.B., Anderson, D.S., 1985. Methods of pH calibration of sedimentary diatom remains for reconstructing history of pH in lakes. *Hydrobiologia* 120 (1), 69–87.
- Davis, M.B., Shaw, R.G., 2001. Range shifts and adaptive responses to Quaternary climate change. *Science* 292 (5517), 673–679.
- Elith, J., et al., 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29 (2), 129–151.
- Gentry, A.H., 1982. Patterns of Neotropical plant species diversity. *Evolutionary Biology* 15, 1–85.
- Gentry, A.H., 1988. Changes in plant community diversity and floristic composition on environmental and geographical gradients. *Annals of the Missouri Botanical Garden* 75 (1), 1–34.
- Gentry, A.H., 1990. *Four Neotropical rainforests*. Yale University Press, New Haven. xiii, 627 pp.
- Godley, E.J., Moar, N.T., 1973. Vegetation and pollen analysis of two bogs in Chile. *New Zealand Journal of Botany* 11, 255–268.
- Graf, K., 1992. Pollendiagramme aus den Anden: Eine Synthese zur Klimageschichte und Vegetationsentwicklung seit der letzten Eiszeit. *Physische Geographie*, vol. 34. University of Zurich, Switzerland.
- Greenwood, D.R., Wing, S.L., 1995. Eocene continental climates and latitudinal temperature gradients. *Geology* 23 (11), 1044–1048.
- Gregory-Wodzicki, K.M., 2000. Relationships between leaf morphology and climate, Bolivia: implications for estimating paleoclimate from fossil floras. *Paleobiology* 26 (4), 668–688.
- Guisan, A., et al., 2006. Making better biogeographical predictions of species' distributions. *Journal of Applied Ecology* 43 (3), 386–392.
- Holdridge, L.R., 1947. Determination of world plant formations from simple climatic data. *Science* 105 (2727), 367–368.
- Huntley, B., Prentice, I.C., 1988. July temperatures in Europe from pollen data, 6000 years before present. *Science* 241 (4866), 687–690.
- Huntley, B., Webb, T., 1989. Migration: species response to climatic variations caused by changes in the Earth's orbit. *Journal of Biogeography* 16 (1), 5–19.
- Hutchinson, G.E., 1957. Population studies – animal ecology and demography: concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology* 22, 415–427.
- Hutchinson, G.E., 1958. Concluding remarks. *Cold Harbor Springs Symposium on Quantitative Biology* 22, 425–427.
- Jackson, S.T., Overpeck, J.T., 2000. Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology* 26 (4), 194–220.
- Jackson, S.T., Williams, J.W., 2004. Modern analogs in Quaternary paleoecology: here today, gone yesterday, gone tomorrow? *Annual Review of Earth and Planetary Sciences* 32, 495–537.
- Jackson, S.T., et al., 2000. Vegetation and environment in eastern North America during the Last Glacial Maximum. *Quaternary Science Reviews* 19 (6), 489–508.
- Kershaw, A.P., 1997. A bioclimatic analysis of early to Middle Miocene brown coal floras, Latrobe Valley, south-eastern Australia. *Australian Journal of Botany* 45 (3), 373–387.
- Kershaw, A.P., Nix, H.A., 1988. Quantitative paleoclimatic estimates from pollen data using bioclimatic profiles of extant taxa. *Journal of Biogeography* 15 (4), 589–602.
- Leyden, B.W., et al., 1996. A record of long and short-term climatic variation from Northwest Yucatan: Cenote San Jose Chulchca. In: Fedick, S.L. (Ed.), *The managed mosaic: ancient Maya agriculture and resource use*. University of Utah Press, Salt Lake City.
- Maguire, B., 1973. Niche response structure and analytical potentials of its relationship to habitat. *American Naturalist* 107 (954), 213–246.
- Markgraf, V., 1980. New data on the late and postglacial history of La Mision, Tierra del Fuego, Argentina. *Proceedings IV International Palynological Conference* 3, 68–74.
- Markgraf, V., 1983. Late and postglacial vegetational and paleoclimatic changes in subantarctic, temperate and arid environments in Argentina. *Palynology* 7, 43–70.
- Markgraf, V., 1989. Late Pleistocene/Holocene paleoclimates from subantarctic latitudes. *Antarctic Journal of the United States* 24 (5), 1–2.
- Markgraf, V., 1991. Younger Dryas in southern South America. *Boreas* 20 (1), 63–69.
- Markgraf, V., 1993. Paleoenvironments and paleoclimates in Tierra-del-fuego and southernmost Patagonia, South America. *Palaeogeography Palaeoclimatology Palaeoecology* 102 (1–2), 53–68.
- Martinez-Meyer, E., Peterson, A.T., 2006. Conservatism of ecological niche characteristics in North American plant species over the Pleistocene-to-Recent transition. *Journal of Biogeography* 33 (10), 1779–1789.
- Mayle, F.E., Burbridge, R., Killeen, T.J., 2000. Millennial-scale dynamics of southern Amazonian rain forests. *Science* 290 (5500), 2291–2294.
- McCullagh, P., Nelder, J.A., 1989. *Generalized Linear Models*. Chapman and Hall, London. 511 pp.
- McElwain, J.C., Chaloner, W.G., 1995. Stomatal density and index of fossil plants track atmospheric carbon dioxide in the Paleozoic. *Annals of Botany* 76 (4), 389–395.
- McElwain, J.C., Chaloner, W.G., 1996. The fossil cuticle as a skeletal record of environmental change. *Palaos* 11 (4), 376–388.
- McElwain, J.C., Beerling, D.J., Woodward, F.I., 1999. Fossil plants and global warming at the Triassic–Jurassic boundary. *Science* 285 (5432), 1386–1390.
- Mercer, J.H., Ager, T.A., 1983. Glacial and floral changes in southern Argentina since 14,000 years ago. *National Geographic Society Research Reports* 15, 457–477.
- Mosbrugger, V., Utescher, T., 1997. The coexistence approach: a method for quantitative reconstructions of Tertiary terrestrial palaeoclimate data using plant fossils. *Palaeogeography Palaeoclimatology Palaeoecology* 134 (1–4), 61–86.
- New, M., Hulme, M., Jones, P., 1999. Representing twentieth-century space-time climate variability. Part I: development of a 1961–90 mean monthly terrestrial climatology. *Journal of Climate* 12 (3), 829–856.
- Overpeck, J.T., Webb, R.S., Webb, T., 1992. Mapping eastern North American vegetation change of the past 18 Ka: no-analogs and the future. *Geology* 20 (12), 1071–1074.
- Overpeck, J.T., Webb, T., Prentice, I.C., 1985. Quantitative interpretation of fossil pollen spectra: dissimilarity coefficients and the method of modern analogs. *Quaternary Research* 23 (1), 87–108.
- Parmesan, C., Yohe, G., 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* 421 (6918), 37–42.
- Pearson, R.G., Dawson, T.P., 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography* 12 (5), 361–371.
- Pearson, P.N., et al., 2001. Warm tropical sea surface temperatures in the Late Cretaceous and Eocene epochs. *Nature* 413 (6855), 481–487.
- Pearson, R.G., Dawson, T.P., Berry, P.M., Harrison, P.A., 2002. SPECIES: a spatial evaluation of climate impact on the envelope of species. *Ecological Modelling* 154 (3), 289–300.
- Phillips, O., Miller, J.S., 2002. *Global patterns of diversity: Alwyn H. Gentry's forest transect data set*, vol. 89. Missouri Botanical Garden Press, St. Louis. 319 pp.
- Prentice, I.C., et al., 1992. A global biome model based on plant physiology and dominance, soil properties and climate. *Journal of Biogeography* 19 (2), 117–134.
- Prinzinger, A., Durka, W., Klotz, S., Brandl, R., 2001. The niche of higher plants: evidence for phylogenetic conservatism. *Proceedings of the Royal Society of London Series B-Biological Sciences* 268 (1483), 2383–2389.
- Punyasena, S.W., Eshel, G., McElwain, J.C., 2008. The influence of climate on the spatial patterning of Neotropical plant families. *Journal of Biogeography* 35 (1), 117–130.
- Punyasena, S.W., Mayle, F.E. and McElwain, J.C., in press. Quantitative estimates of glacial-Holocene temperature and precipitation change in lowland Amazonian Bolivia. *Geology*.
- Richardson, J.E., Pennington, R.T., Pennington, T.D., Hollingsworth, P.M., 2001. Rapid diversification of a species-rich genus of neotropical rain forest trees. *Science* 293 (5538), 2242–2245.
- Royer, D.L., Wilf, P., 2006. Why do toothed leaves correlate with cold climates? Gas exchange at leaf margins provides new insights into a classic paleotemperature proxy. *International Journal of Plant Sciences* 167 (1), 11–18.
- Royer, D.L., Wilf, P., Janesko, D.A., Kowalski, E.A., Dilcher, D.L., 2005. Correlations of climate and plant ecology to leaf size and shape: potential proxies for the fossil record. *American Journal of Botany* 92 (7), 1141–1151.
- Sakai, A., Larcher, W., 1987. Frost survival of plants: responses and adaptation to freezing stress. *Ecological studies*, vol. 62. Springer-Verlag, Berlin. New York, x, 321 pp.
- Schabitz, F., 1994. Holocene climatic variations in northern Patagonia, Argentina. *Palaeogeography Palaeoclimatology Palaeoecology* 109 (2–4), 287–294.
- Silverman, B.W., 1981. Using kernel density estimates to investigate multimodality. *Journal of the Royal Statistical Society Series B-Methodological* 43 (1), 97–99.
- Silverman, B.W., 1986. *Density estimation for statistics and data analysis*. Monographs on Statistics and Applied Probability. Chapman and Hall, London. New York, 175 pp.
- ter Braak, C.J.F., 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* 67 (5), 1167–1179.
- ter Braak, C.J.F., 1995. Nonlinear methods for multivariate statistical calibration and their use in paleoecology: a comparison of inverse (k-nearest neighbors, partial least-squares) and weighted averaging partial least-squares) and classical approaches. *Chemometrics and Intelligent Laboratory Systems* 28 (1), 165–180.
- ter Braak, C.J.F., Prentice, I.C., 1988. A theory of gradient analysis. *Advances in Ecological Research* 18, 271–317.
- Thomas, C.D., et al., 2004. Extinction risk from climate change. *Nature* 427 (6970), 145–148.
- Thompson, R.S., Anderson, K.H., Bartlein, P.J., 1999. Atlas of relations between climatic parameters and distributions of important trees and shrubs in North America. U.S. Geological Survey Professional Paper, vol. 1650. U.S. Geological Survey. A&B.
- Toivonen, H.T.T., Mannila, H., Korhola, A., Olander, H., 2001. Applying Bayesian statistics to organism-based environmental reconstruction. *Ecological Applications* 11 (2), 618–630.
- Traverse, A., 2007. *Paleopalynology*. Springer, New York. xviii, 813 pp.
- Urrego, D.H., 2006. Long-term vegetation and climate change in Western Amazonia. PhD Dissertation Thesis, Florida Institute of Technology, Melbourne, FL, 278 pp.
- Urrego, D.H., Silman, M.R., Bush, M.B., 2005. The Last Glacial Maximum: stability and change in a western Amazonian cloud forest. *Journal of Quaternary Science* 20 (7–8), 693–701.
- Vasko, K., Toivonen, H.T.T., Korhola, A., 2000. A Bayesian multinomial Gaussian response model for organism-based environmental reconstruction. *Journal of Paleolimnology* 24 (3), 243–250.
- Walter, H., 1985. *Vegetation of the earth and ecological systems of the geo-biosphere*. Springer-Verlag, Berlin. New York, xvi, 318 pp.
- Walter, H., Breckle, S.W., 2002. *Walter's Vegetation of the earth: the ecological systems of the geo-biosphere*. Springer, Berlin. New York, xx, 527 pp.
- Wang, S.C., 2003. On the continuity of background and mass extinction. *Paleobiology* 29 (4), 455–467.

- Webb, T., Bartlein, P.J., Harrison, S.P., Anderson, K.H., 1993. Vegetation, lake levels, and climate in eastern North America for the past 18000 years. In: Wright, H.E., et al. (Ed.), *Global climates since the last glacial maximum*. University of Minneapolis Press, Minneapolis, pp. 415–467.
- Whitlock, C., Bartlein, P.J., 1993. Spatial variations of Holocene climatic change in the Yellowstone region. *Quaternary Research* 39 (2), 231–238.
- Wilf, P., 1997. When are leaves good thermometers? A new case for leaf margin analysis. *Paleobiology* 23 (3), 373–390.
- Wilf, P., et al., 2003. High plant diversity in Eocene South America: evidence from Patagonia. *Science* 300 (5616), 122–125.
- Williams, J.W., Shuman, B.N., Webb, T., Bartlein, P.J., Leduc, P.L., 2004. Late-quaternary vegetation dynamics in north america: scaling from taxa to biomes. *Ecological Monographs* 74 (2), 309–334.
- Wing, S.L., Greenwood, D.R., 1993. Fossils and fossil climate: the case for equable continental interiors in the Eocene. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 341 (1297), 243–252.
- Wing, S.L., et al., 2005. Transient floral change and rapid global warming at the Paleocene-Eocene boundary. *Science* 310 (5750), 993–996.
- Wolfe, J.A., 1995. Paleoclimatic estimates from tertiary leaf assemblages. *Annual Review of Earth and Planetary Sciences* 23, 119–142.
- Woodward, F.I., 1987. *Climate and plant distribution*. Cambridge studies in ecology. Cambridge University Press, Cambridge; New York. xi, 174 pp.