

A few prospective ideas on climate reconstruction: from a statistical single proxy approach towards a multi-proxy and dynamical approach

J. Guiot^{1,2}, H. B. Wu³, V. Garreta¹, C. Hatté⁴, and M. Magny⁵

¹CEREGE, CNRS/Aix-Marseille Université UMR 6635, BP 80, 13545 Aix-en-Provence cedex, France

²ECCOREV, CNRS/Aix-Marseille Université FR 3098, BP 80, 13545 Aix-en-Provence cedex, France

³Institut des Sciences de l'Environnement, UQAM, Montréal PQ, H3C 3P8, Canada

⁴LSCE, CNRS/CEA UMR 1572, Domaine du CNRS, 91198 Gif-sur-Yvette, France

⁵CNRS, UMR 6249, Laboratoire Chrono-Environnement, UFR des Sciences et Techniques, 16 Route de Gray, 25030 Besançon, France

Received: 29 October 2008 – Published in *Clim. Past Discuss.*: 8 January 2008

Revised: 21 August 2009 – Accepted: 3 September 2009 – Published: 8 October 2009

Abstract. Important progresses have been made in palaeoclimatological studies by using statistical methods. But they are in somewhere limited as they take the present as an absolute reference. This is particularly true for the modern analogue technique. The availability of mechanistic models to simulate the proxies measured in the sediment cores gives now the possibility to relax this constraint. In particular, vegetation models provide outputs comparable to pollen data (assuming that there is a relationship between plant productivity and pollen counts). The input of such models is, among others, climate. The idea behind palaeoclimatological reconstructions is then to obtain inputs, given outputs. This procedure, called model inversion, can be achieved with appropriate algorithms in the frame of the Bayesian statistical theory. But we have chosen to present it in an intuitive way, avoiding the mathematics behind it. Starting from a relative simple application, based on an equilibrium BIOME3 model with a single proxy (pollen), the approach has evolved into two directions: (1) by using several proxies measured on the same core (e.g. lake-level status and $\delta^{13}\text{C}$) when they are related to a component of the vegetation, and (2) by using a more complex vegetation model, the dynamic vegetation model LPJ-GUESS. Examples presented (most of them being already published) concern Last Glacial Maximum in Europe and Africa, Holocene in a site of the Swiss Jura, an

Eemian site in France. The main results are that: (1) pollen alone is not able to provide exhaustive information on precipitation, (2) assuming past CO_2 equivalent to modern one may induce biases in climate reconstruction, (3) vegetation models seem to be too much constrained by temperature relative to precipitation in temperate regions. This paper attempts to organise some recent ideas in the palaeoclimatological reconstruction domain and to propose perspectives in that effervescent domain.

1 Introduction

For a long time, Quaternary palaeoecologists, and in particular palynologists, have used intuitive methods to reconstruct palaeoclimates or paleoenvironments from biological data. We focus here on pollen bioindicator as it has profited, during the last two decades, of progresses of outstanding vegetation models. The most common approach was to compare the present-day distribution of selected species with the corresponding distribution of climate variables thought to be determinant for them, according to the niche's theory. The species are analyzed separately and related to one climatic variable. But the species respond to a combination of climatic variables and their distributions are controlled by different climatic factors in different parts of their ranges. Moreover, climate parameters are often interrelated. Thus, it has been necessary to develop methods taking into account



Correspondence to: J. Guiot
(guiot@cerege.fr)

the ecological complexity of species and assemblages, and of their relationships with climatic factors. A relatively ancient evolution was to work with several climatic variables and several species (Iversen, 1944; Atkinson et al., 1987).

One of the power of the pollen counts is to give information partly related on species abundances, making possible to develop response models where the abundance of the species is expressed as a function of the climate (Bartlein and Prentice, 1986). These statistical models are only valid on climatic niches presently realized and their extrapolation to the past could be problematic, when the past is too different from the present (Guiot et al., 2008). Moreover, to reconstruct climate, it is necessary to inverse these response models, but this cannot be achieved directly. Usually one calculates backward statistical relationships between climate and species, which are usually called *transfer functions* by the palaeoclimatologists. They are based on a few assumptions:

- (1) Climate is the ultimate cause of changes in the paleobiological data.
- (2) The ecological properties of the species considered has not changed between the period analyzed and the present time, and the relationship between the species and the climate is thus uniform through time.
- (3) The modern observations contain all the necessary information to interpret the fossil data.

The second and third assumptions originate from the uniformitarian principle that the same scientific laws and processes are constant throughout space and time (this theory has been proposed by James Hutton in 1795 and popularised by Charles Lyell in 1830 “Amid all the revolutions of the globe the economy of nature has been uniform and her laws are the only things that have resisted the general movement”). Without the second assumption, the reconstruction of past environments becomes impossible. To satisfy the third assumption, it is necessary to collect a large diversity of modern samples to optimise the chance to cover all the possible situations of the period studied. But sometimes, non climatic forcings are so different today that there is no true modern analogues. An example can easily be given for vegetation. A number of physiological and palaeoecological studies (e.g. Jolly and Haxeltine, 1997; Cowling and Sykes, 1999) have proved that plant-climate interactions are sensitive to the atmospheric CO₂ concentration, and we know, from ice cores (EPICA, 2004), that this concentration is presently much higher than ever during the past 740 000 yr. Consequently modern samples collected under high CO₂ concentration are hardly good analogues for low CO₂ periods. Moreover, pollen assemblages are noisy and sometimes biased records of the climate variables, because (1) pollen productivity is not equal to vegetation productivity, (2) pollen assemblages are disturbed by pollen grain transportation, (3) a pollen

taxon is not a univocal species, (4) and the species are not affected by a single climatic variables.

All these problems make difficult the use of statistical methods based on the reference modern data. We synthesise in this paper recent progresses achieved in the last years to relax these constraints. One way was the use of mechanistic vegetation models together with pollen data (a similar approach can followed with other proxies if such models are available). Another complementary way is the use of several proxies measured on the same samples. We will show how to combine both approaches when adequate models are available and what are the perspectives of what is called model inversion. The purpose of this paper is not to detail the mathematics behind the methods, but to give an intuitive flavor of the concepts involved by them. The reader can find in the cited papers more details to satisfy his curiosity.

2 The methods

Even if the inversion of statistical response models should be the “natural” way to proceed to reconstruction past climate, the large majority of works published in the three last decades were based on a very simple “one-step” concept. A very popular method is the modern analogue technique.

2.1 Modern analogues technique (MAT)

MAT is illustrated by the schema of Fig. 1a. This schema does not reflect the exact way in which the algorithm is built, but it facilitates the comparison with the other methods. The caption of the figure explains the five steps. To implement it, it is necessary to define a distance index. Usually a the Euclidian distance of the square-root of the pollen frequencies (chord distance) is used (Overpeck et al., 1985). The number of analogues depends on a threshold above which the similarity is considered as too poor. The reconstructed climate is provided as a weighted mean of the climate of the analogues (according to the inverse of the distance index). It is accompanied by an error bar based on the climatic range of the analogues. This error bar cannot be considered as a confidence interval *sensu stricto* as it depends on the number of good analogues available and not directly on the tolerance of vegetation to a climatic range nor on the noise in the data. Advantages and limits of the method are discussed in Guiot and DeVernal (2007).

2.2 Vegetation modelling

A pollen assemblage (or spectrum) is assumed to reflect the composition and structure of the regional vegetation. It is composed by a large number of taxa which can be grouped into what is usually called plant functional types (PFT, i.e. groups of plant species of similar characteristics and responding in a similar way to climate). This has the advantage

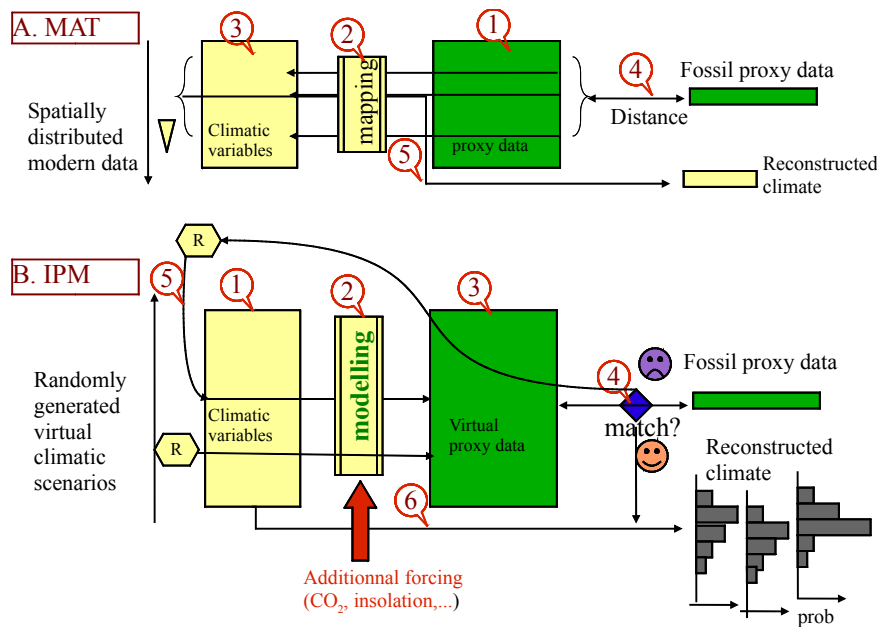


Fig. 1. Sketch of the modern analogues methods (MAT) and inverse proxy modelling (IPM). A. MAT starts from (1) a matrix of proxy assemblages and a fossil assemblage (2) which is compared to all the modern proxy assemblages using a distance index; the few most similar ones are identified, they are called the best analogues, (3) they are located on a map, (4) the corresponding climatic variables are selected among the climate database and (5) averaged to provide the reconstructed climate variables. B. IPM starts (1) from a climatic scenario (a vector of climatic variables), randomly generated (*R*), which is (2) introduced into the proxy model, and produces (3) a simulated proxy assemblage; (4) the fossil assemblage is compared to simulated pollen assemblages; if the matching is acceptable, the climatic scenario is kept, if not acceptable, it is rejected; (5) a new climatic scenario is randomly selected and the procedure (1 to 4) is repeated; (6) when a sufficient number of virtual climatic scenarios is obtained, the procedure is stopped and distribution histograms of the scenarios retained are build. It is possible to change other inputs of the model, such as the CO₂ or insolation, and to study the proxy model sensitivity to that variable. For both approaches, the steps are idealised to facilitate the their intercomparison. The practical algorithms operate generally in slightly different way.

to reduce the size of the assemblages and overall to be coherent with vegetation model outputs, according to the work of Prentice et al. (1996).

There exists a large variety of vegetation models. Some of them need a fine knowledge of climate to estimate vegetation. They are hardly usable at a continental scale where often monthly climatic records are available. This explain why paleostudies have used relative simple biogeochemical models. The most popular model was BIOME3 (Haxeltine and Prentice, 1996) or a modified version BIOME4 (Kaplan et al., 2003). It is a process-based terrestrial biosphere model which includes a photosynthesis scheme that simulates acclimation of plants to changed atmospheric CO₂ by optimisation of nitrogen allocation to foliage and by accounting for the effects of CO₂ on net assimilation, stomatal conductance, leaf area index (LAI) and ecosystem water balance. It assumes that there is no nitrogen limitation. The inputs of the model are soil texture, CO₂ rate, absolute minimum temperature (*T*_{min}), monthly mean temperature (*T*), monthly total precipitation (*P*) and monthly mean sunshsine (*S*), i.e. the ratio between the actual number of hours with sunshine over the potential number (with no clouds). From these in-

put variables, the model computes bioclimatic variables, and from them, the maximum sustainable leaf area index and the net primary production (NPP, in kg m⁻² yr⁻¹) for the PFT's able to live in this input climate. Competition among PFT's is simulated by using the optimal NPP of each PFT as an index of competitiveness. The most important PFT's in Europe are: temperate broadleaved evergreen trees (tbe), temperate summergreen trees (ts), temperate evergreen conifer trees (tc), boreal evergreen trees (bec), boreal deciduous trees (bs), temperate grass (tg), woody desert plant type (wd), tundra shrub type (tus), cold herbaceous type (clg), lichen/forb type (lf). The pollen PFT's are sometimes more precise and pollen information is sufficient to recognize several varieties of the same model PFT, for example pollen is able to separate warm and cool ts. The use of such models in the paleoclimatological context and the simulation of the CO₂ effect on ecosystems are particularly well reviewed in Prentice and Harrison (2009).

BIOME3 and BIOME4 are equilibrium models. LPJ-GUESS is a noticeable improvement as the dynamics of the vegetation stands are taken into account (Smith et al., 2001). While, in the equilibrium models, two runs with the

same climate gives always the same vegetation output, in a dynamic model, random processes as competition between species, and mortality introduce stochasticity in the outputs. In LPJ-GUESS, cohorts of trees of different species, age and structure compete for light and soil resources on a number of replicated patches of plants. Either PFT (Sitch et al., 2003) or species (Hickler et al., 2004) may be simulated. Garreta et al. (2009) used the species version which includes 18 species. LPJ-GUESS has standard inputs, i.e. with monthly values of precipitation, temperature and cloudiness. For each study site, past and present, precipitation and temperature chronologies were interpolated from the CRU TS 1.2 dataset (New et al., 2002), which has a spatial resolution of $10'$. For cloudiness, they fitted a relationship between monthly cloudiness and both monthly precipitation and temperature per site.

2.3 Inversion modelling and Bayesian approach

As indicated by Fig. 1a, the statistical method to estimate climate starts from pollen assemblages and goes back to climate. Vegetation models start from climate and go to vegetation. The idea proposed by Guiot et al. (2000) is then to use massive computation algorithms to “inverse” the model, starting from vegetation and going back to climate. It is not an analytical inversion, but an iterative procedure where one converges progressively towards the climate which has produced the observed vegetation (Fig. 1b). The caption of the figure explains the steps of the method. The climatic space is randomly sampled to produce a large variety of climatic scenarios which are introduced in the vegetation model to simulate the corresponding vegetation composition and productivity. The simulated pollen assemblages are compared to the fossil assemblage and those matching reasonably well are retained. The corresponding climatic scenarios are then considered to be compatible with the observed vegetation. They are used to build histograms, which are estimates of probability distribution functions of a climate able to generate such a vegetation. The outputs of the model do not correspond exactly to the pollen assemblages. A transformation is necessary and it is a major tricky point of the method. Several tested approaches are presented in the following sections. This transformation is assimilated to the model box in the figure.

The second tricky point is the number of input climatic parameters. The used vegetation models use 36 monthly climate inputs above described, which define the scenario. One has to modify them randomly to browse the climatic space, but its size is too high to converge to the true solution. So, we decided to reduce them to a small number of representative variables (Tjan, Tjul, Pjan, Pjul), from which all the other climatic variables are deduced. A sine function is adjusted to the two temperature variables and another to the two precipitation variables, enabling an interpolation of the missing months. The sunshine percentage of each month is

estimated by a linear regression from the temperature and precipitation of the same month. See Guiot et al. (2000) for more details. To provide a comparison between sites and time periods, climate variables are expressed as “anomalies” or Δ climate, i.e. differences between proposed climate and the modern climate at the considered site.

The third tricky point is then to define what is considered as a matching. Bayesian theory provides a framework for such a definition (Robert and Casella, 1999). In this context, it uses main concepts of *prior* and *posterior*. The *prior* is the information, summarised under the form of a distribution, which is available prior to the data analysis. The *posterior* is the information that we will deduce from the data and a hierarchical model. In that respect, the hierarchical model is not restricted to the vegetation model, but it the function which relates the *prior* to the *posterior*. In statistical terms, it is the probability of pollen assemblage Y conditional on climate C . It is noted $p(Y|C) = \int p(Y|V)p(V|C)dV$ where $p(V|C)$ is the vegetation model which links vegetation V to climate C and $p(Y|V)$ is the function which links pollen to vegetation. The *prior* is an initial guess of the probability distribution of the climate. It can be given by the knowledge we have from other paleoclimatic data, or from, if nothing is available, from the knowledge which has been accumulated in that science. The distribution law is then an uniform law defined on that range.

Bayesian statistics have been conceptually introduced in paleoclimatology by Korhola et al. (2002) and Haslett et al. (2006), but without any reference to a mechanistic model. They underlined that such an approach is slow despite making unreasonable compromises on the models employed. With a mechanistic model, it is even slower. The reason is that, to draw the *posterior*, one has to use Monte-Carlo algorithms which need thousands of iterations. These algorithms – coherently with the Bayesian inference – provide an integration over the climate parameter space instead of an optimisation. A popular type of such algorithms is known as Monte Carlo Markov Chain (MCMC) algorithm. Let us consider a multi-dimensional mathematical space where each dimension represents a climatic variable. A vector of parameters is an element of the multi-dimensional climate space. The Metropolis-Hastings algorithm is an iterative method which browses the climate space according to an acceptance-rejection rule (Metropolis et al., 1953; Hastings, 1970). The output of this algorithm is a “path” or “chain” of climate parameters describing the posterior distribution of climate parameter. The MCMC algorithm can be considered as an equilibrium inversion method, compatible with equilibrium vegetation models as BIOME3.

To realise the temporal inversion of the dynamic model LPJ-GUESS, a statistical framework has been developed around a temporal hierarchical model and a Sequential Monte Carlo (SMC or particle filter Doucet et al. (2001)) inference algorithm, because (1) the random character of the vegetation simulated by LPJ-GUESS prohibits the use of the

“static” MCMC approach and, (2) the need of dimension reduction in the reconstructed climate space, which is equal to the number of samples in the reconstructed climatic time-series.

3 Applications

The method is illustrated starting, in one side from the equilibrium vegetation model BIOME3 towards the dynamic vegetation model LPJ-GUESS and, in the other side, from a single proxy (pollen) towards a double proxy constraint. This is sketched in Fig. 2 with a double axis. The second proxy is, in one case, lake levels data and, in the other one, isotopic data, each proxy giving information on different aspects of the climate. Finally, we will conclude on the directions that palaeoclimatology should follow to fully exploit the increasingly diverse and improved set of archives and proxies.

All these results concern annual temperature and precipitation. Even if these variables are not those which are the most determinant for vegetation, they are the average of the input variables and are then the most synthetic. Moreover, annual precipitation is that one which concerns both lake-levels and vegetation. We have thought that it was better to present them instead of more bioclimatic variables (growing degree-days, water availability ...), even if for a better interpretation of the results, it is necessary to look also the bioclimatic variables.

3.1 Application A: Europe at the Last Glacial Maximum

The first application uses BIOME3 constrained by pollen data (application A in Fig. 2) for the Last Glacial Maximum (LGM, 21 ± 2 ka BP) in Europe. The data and the method are fully described in Guiot et al. (2000). The model outputs are transformed into pollen PFT's scores by an Artificial Neural Network (ANN) calibrated on a modern dataset (Tarasov et al., 1998). Unlike standard transfer function, the relationship is not calculated between climate and pollen, but between vegetation and pollen, the bridge between climate and vegetation being given by BIOME3. The measure of fit between the vegetation model outputs (NPP) and the observations (pollen PFT scores) is a likelihood index. It assumes a probability model for the simulation “errors”, here a Gaussian model. It is then proportional to the sum of square discrepancies between ANN-transformed NPP and observed pollen PFT scores. The *priors* are given by an uniform distribution law on $[-30, +5^\circ\text{C}]$ for temperature anomalies and $[-60, 60\%]$ for precipitation relative anomalies.

A dataset of 15 LGM samples is considered. We present two experiments. The first experiment is done with a high level of CO_2 (340 ppmv) close to the atmospheric concentration existing during the modern data sampling. The second experiment with a low level of CO_2 (200 ppmv), such is measured in the ice cores for the LGM (Petit et al.,

1999). The results are presented according to the latitude (Fig. 3). Annual temperature shows an increased gradient from the southernmost site (about 35°N) towards the northernmost (48°N), and, for the annual precipitation, a decreased gradient. It is $0.81\pm 0.35^\circ\text{C}/^\circ\text{latitude}$ for temperature and $-29\pm 12\text{ mm}/^\circ\text{latitude}$ for precipitation. It means that a temperature decrease larger in the south than in the north is necessary to transform forest into steppes and, in the north, a stronger precipitation decrease is necessary. When the LGM CO_2 level is applied, the gradients become insignificant for both variables. So the CO_2 lowering is large enough to reduce forest extent: under a high CO_2 level, temperature must fall sufficiently to reduce the growing season under a certain level, and under a low CO_2 level, the forest reduction is due to both temperature lowering and carbon limitation. There is then a real bias in ignoring the true level of CO_2 for climate reconstruction (when statistical methods are used instead of mechanistic models). This bias reaches 3°C in southernmost sites but not more than 1°C in northernmost ones, meaning that CO_2 becomes more limiting than temperature far away from the ice cap.

3.2 Application B: Eurasia and Africa at the LGM

Wu et al. (2007a) have improved the method. First BIOME3 has been replaced by BIOME4 (Kaplan et al., 2003). Second, the ANN-relationship between NPP simulations of the model PFT's and pollen PFT scores has been replaced by a correspondence matrix between the model biomes and the biome scores calculated from pollen. This matrix is an empirical result based on modern data and theoretical definition of the biomes (see the original paper for more information). The method has been applied to LGM of Eurasia and Africa (application B in Fig. 2).

The estimated anomalies of the climatic parameters for the LGM period are shown in Fig. 4. The left part of the each graphic concerns Africa. There is a large dispersion which can mainly be explained by a large dispersion of the elevations. Wu et al. (2007a) have shown a strong altitudinal gradient of precipitation. For the modern level of CO_2 , one cannot fit a linear relationship of temperature to latitude in Africa, but yet in Europe, the relationship is negative: high latitude sites had a temperature anomaly of about -12°C and southern sites anomalies of -10 to -5°C . The gradient is negative while it was positive in Fig. 3. It is likely due to the better ability of BIOME4 to simulate the LGM vegetation, which is intermediate between cool steppes and tundra. A biome called steppe-tundra was introduced in the most recent version of the model, which fits then much better to the data. Even if that biome does not exist explicitly in the pollen data, it exists cryptically when tundra and steppic scores are of the same magnitude. The reconstructed anomalies under low CO_2 concentration are not significantly different from the reconstructed anomalies under high CO_2 concentration. Wu et al. (2007a) found a clear bias for winter temperature, i.e.

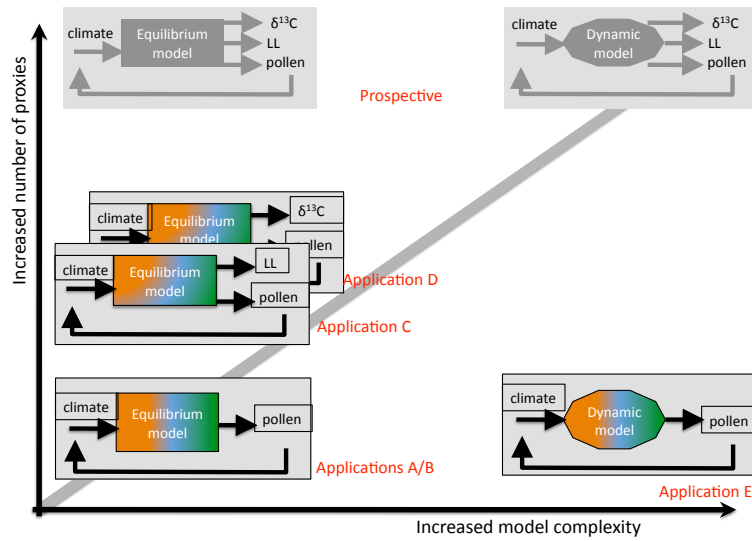


Fig. 2. Evolution of the inverse modelling method according to directions: complexification of the proxy model and increased number of proxies. The applications A to E, corresponding to the sub-sections labelled in the text, refer to examples of increased complexity using vegetation models. Applications A/B are equilibrium model inversion constrained by pollen alone (low complexity); applications C and D involves an equilibrium model with two proxies (pollen with lake levels (LL), pollen with $\delta^{13}C$); application E involves a dynamic model constrained by pollen alone; the two icons corresponding to prospective represent two cases not yet implemented: a simple equilibrium model with three proxies and a more complex dynamic model with also three proxies.

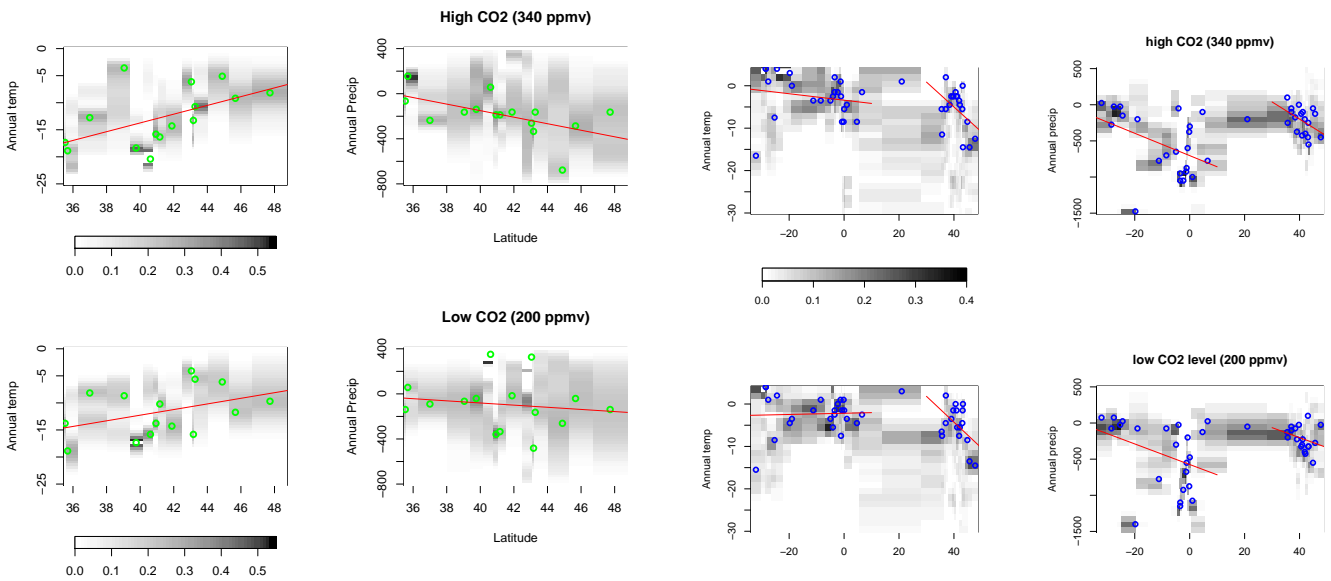


Fig. 3. Annual temperature and precipitation anomalies (i.e. deviations form present value) at the Last Glacial Maximum (21 ka \pm 2 kaBP) in function of latitude. The region covered is Europe. Two experiments are defined, one with quasi-modern CO₂ concentration (340 ppm), one with LGM CO₂ concentration (200 ppm). The grey scale indicates the probability distribution, the blue circles show the mode of the distribution and the red line the linear relationship between the mode and the latitude.

Fig. 4. Annual temperature and precipitation anomalies (i.e. deviations form present value) at the Last Glacial Maximum (21 ka \pm 2 kaBP) in function of latitude. The regions covered are Europe and Africa (42 sites). The grey scale indicates the probability distribution, the blue circles show the mode of the distribution and the red line the linear relationships between these modes and the latitude of the sites, one for sites south of 10° N and one for sites north of 30° N.

about 10°C colder under higher CO₂, but nothing for summer temperature. The only bias that is found for annual temperature concerns Mediterranean sites, with an annual temperature rather lower under high CO₂.

In Africa, temperature was not very different from present values. Precipitation shows a much more structured profile. Under high CO₂, anomalies were close to zero in South Africa and between −1000 and 0 mm/yr at equator, depending on the elevation. Under low CO₂ concentration, the reconstruction in the southern part of the continent was similar, and in the central part, the dispersion was higher: −1200 to 0 mm/yr. If we focus on high elevation sites (>1500 m), the precipitation mode, for 340 ppmv, was at about 1000 mm/year and is replaced, with 200 ppmv, by a large double peak from 1100 to 700 mm/year. In fact, Wu et al. (2007b) have shown that the disappearance of forest above 2000 m elevation can be explained partly by a precipitation decrease and partly by a CO₂ lowering. Wu et al. (2007a) analysed the water stress variable α , which is the ratio of actual and potential evapotranspiration and is closely related to the stomatal area and the water use efficiency. They found that its maximum probability ranges within [−40, −28%] for high CO₂ and within [−40, −8%] for low CO₂. There is then an oversetting of CO₂, if we use a high CO₂ concentration, inducing an overestimation of the water stress.

Several solutions are possible for the LGM climate in regions where a mixture of steppes and tundra existed. As these biomes have no clear analogues today, a reconstruction based on statistical methods will tend to choose the least poor matching, or fail to find a matching (Peyron et al., 1998; Jost et al., 2005). These analogues were taken in tundra or very cold steppes, resulting in very low reconstructed temperatures. By using a mechanistic model and probability distributions, the results are multi-modal and the most probable mode is different according to the CO₂ concentration. All possible solutions at LGM CO₂ levels can be explored. Complementary proxies are, in this case, of great help to precise the best suitable solution.

3.3 Application C: lake levels and an equilibrium vegetation model

The third example is a single site application with a core covering a part of the Holocene and Younger Dryas (YD) for which pollen assemblages and lake-levels data are available. This application illustrates the effect expected from the use of a second proxy to precise climate components not optimally accessible from pollen data alone (application C in Fig. 2).

The palaeo-lake Le Locle (47°03' N, 6°43' E) has been dried at the last century. It is located at 915 m a.s.l. in the high Swiss Jura. The pollen and lake-level data used in this study were obtained as described in Magny et al. (2001). The lake level status curve indicates that the YD was characterized by a trend toward a lake-level lowering and strong instability (Fig. 5). The early Holocene had three major phases of low

levels, before 10 ka BP, between 9 and 8.5 ka BP, and after 7 ka BP. Concerning the vegetation history (represented by the deciduous trees curve and the total of tree pollen (Fig. 5), the Younger Dryas was characterized by rather large percentages of trees (*Pinus*, *Betula*) together with about 10% of *Artemisia*. Early Holocene was characterized by an increase of *Corylus*, then *Ulmus* and *Quercus*. Nothing in the vegetation history can be related to the rise in lake level at ca. 8400–8300 cal yr BP.

First, we use the method as defined in Sect. 3.1, pollen being used alone to constrain the model and CO₂ assumed to be constant and equal to the pre-industrial value 280 ppmv. The *priors* for January and July temperature are assumed to be uniform between −8 and +4°C (in anomalies) and for precipitation, between −40 and +40% of modern conditions. It is called the “pollen experiment” (ExP) (Fig. 6). The YD was characterised by a temperature lower than present by 8°C. Annual precipitation did not seem to have any trend across the whole studied period. The second experiment (pollen-CO₂ experiment, ExPC) is obtained by providing to the model the atmospheric CO₂ as reconstructed from the Taylor Dome ice core (Indermühle et al., 1999) (Fig. 6). It has the largest effect on the reconstruction of temperature – an anomaly of 5°C instead an anomaly of 8°C with ExP – when its concentration is the lowest. This is enlightened by the differences ExPC-ExP between probability distributions of ExPC and that of ExP: the modes of ExP (in blue) are systematically lower than the modes of ExPC (in red). As for the LGM, this shows that, when the true value of CO₂ is not taken into account, there is a bias in the temperature reconstruction, the effect being maximum during the YD, when the CO₂ was the lowest. The effect on precipitation seems to be negligible (the blue and red distributions being flat and not contrasted).

The last experiment (pollen-CO₂-lakes experiment, ExPCL) is obtained by constraining the model with pollen, CO₂ and lake-levels (Fig. 6). The integration of lake-levels is not straightforward. A solution has been proposed by Cheddadi et al. (1996), called the constrained analogue method. The lake-levels were compared, for each iteration, to the precipitation minus evapotranspiration (P–E), closely related to run-off. Both quantities are subtracted by their modern value at the study site. we call ΔL , the anomaly of lake-level and $\Delta(P-E)$ the anomaly of P–E. Even if the matching between simulated and observed pollen assemblages are acceptable, the iterations where

$$\begin{aligned} &(|\Delta L| \leq 0.5) \text{ and } (|\Delta(P-E)| > 200 \text{ mm}) \\ &(\Delta L > 0.5) \text{ and } (\Delta(P-E) < -100 \text{ mm}) \\ &(\Delta L < -0.5) \text{ and } (\Delta(P-E) > 100 \text{ mm}) \end{aligned} \quad (1)$$

are eliminated. The thresholds used in that equation are in some way arbitrary and obtained by trials and errors. Cheddadi et al. (1996) found that the results were not too much

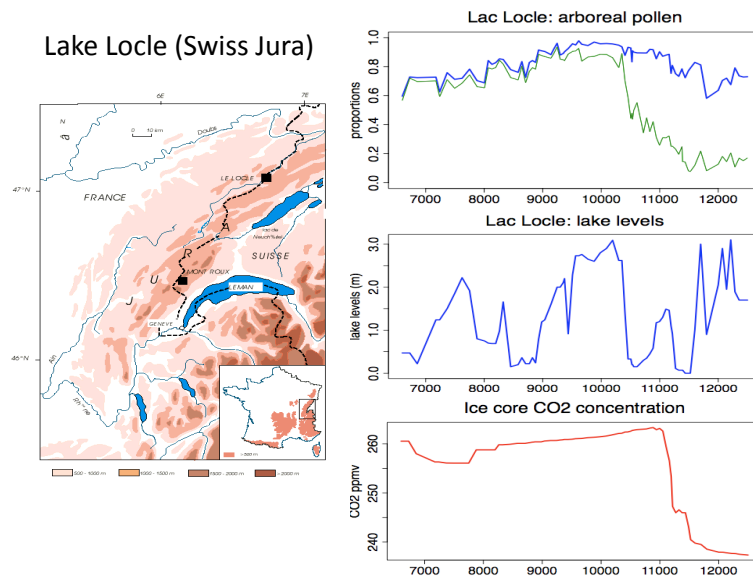


Fig. 5. Location of Lake Le Locle in Swiss Jura. The upper right graphic represents the proportion of tree pollen and the proportion of deciduous tree pollen in the pollen diagram. The middle right graphic represents the lake-levels. The lower right graphic represents the CO₂ concentration in the ice core bubbles of Taylor Dome (Indermühle et al., 1999). Time scales are in calibrated years BP.

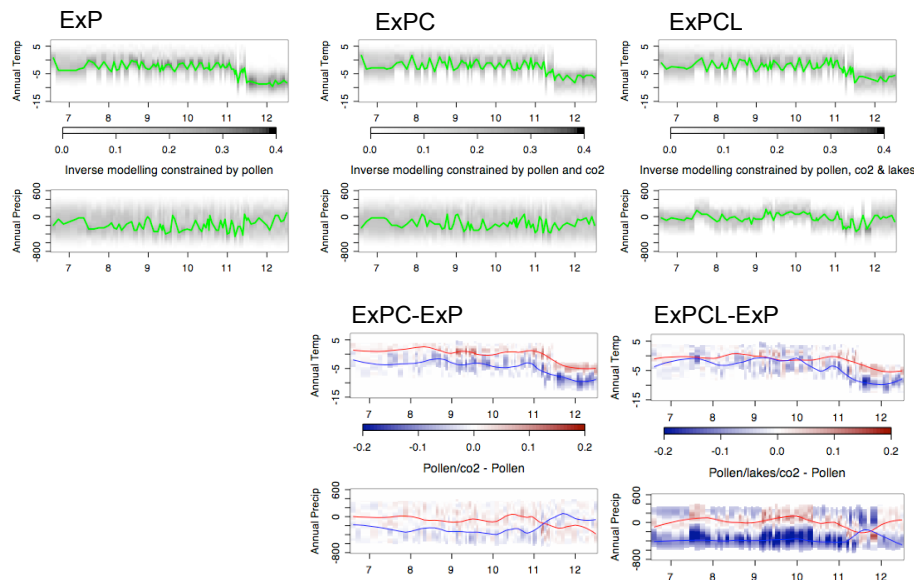


Fig. 6. Results of three experiments on Lake Le Locle. Two variables are constructed, annual temperature and precipitation. Grey sale indicates the probability distributions. The green curve indicates the modal curves. ExP shows the results when pollen alone is used in the inverse modelling; ExPC, when pollen is used with variable CO₂ concentration (Indermühle et al., 1999); ExPCL, when pollen is used with variable CO₂ and with the lake levels constraints. The “blue/red” graphic represent the difference between probability distribution of two experiments. The curves represents the modal curves.

sensitive to the choice of these values. Fig. 6 shows that the reconstructed variations of temperature do not change, but those of precipitation follow much better those of the lake levels, with also a decrease of the uncertainties (indicated by a narrowing of the probability distribution). The probability distribution differences (ExPCL-ExP) shows that ExPCL

distributions are narrower than the ExP modes (blue areas on both sides of the red area indicate large distributions). So when pollen is used alone, the precipitation reconstruction have much larger uncertainties.

These experiments prove again then that CO₂ must be taken into account at least during periods where it is low.

Another point is that precipitation, in temperate regions (at least), can not be inferred with a sufficient confidence from vegetation proxies only. Vegetation uses a part of precipitation falling on the ecosystems, a significant part runs off and consequently, a complementary proxy is needed to infer correctly the total amount of water available within the ecosystem.

3.4 Application D: $\delta^{13}\text{C}$ proxy and equilibrium vegetation model

We present now a another single site application with a core approximately covering the Eemian warm period (128 to 100 ka BP) for which pollen diagram and $\delta^{13}\text{C}$ of organic matter are available. This application illustrates the effect expected from the use of a second proxy of vegetation to decrease uncertainties of pollen data alone. It corresponds to application D of Fig. 2. The procedure used is based on BIOME4 model (as in Sect. 3.2). The likelihood function LH assumes a Gaussian probability distribution for the errors of $\delta^{13}\text{C}$.

$$\text{LH} = -\frac{(\delta^{13}\text{C}_o - \delta^{13}\text{C}_s)^2}{S^2} \quad (2)$$

where subscripts o and s correspond to target and simulated values respectively and where $1/S^2$ is the whole model precision, the inverse of the model error variance. It is an adjustable number which measures the quality of the fit between model outputs and data (Hatté and Guiot, 2005). When pollen data are also available, it is possible to use biome assignment to the sample to make an additional selection of the runs. If the simulated biome matches with the biome obtained from pollen, the iteration is kept, if not, it is rejected. Hatté et al. (2009) have compared the results obtained with biome alone (which is a single pollen approach) and with carbon isotopes constrained by pollen biomes. The method is validated in Hatté et al. (2009).

We reproduce here the results obtained for La Grande Pile sequence. This site is located at $47^\circ 44' \text{N}$, $6^\circ 30' \text{E}$, 330 m a.s.l. with annual precipitation of 1080 mm, a mean annual temperature of 9.5°C , and a seasonal range of about 18°C between the warmest and the coldest months. The data are presented in Rousseau et al. (2006).

For each sample of the La Grande Pile core, an input vector is defined and composed by (1) the $\delta^{13}\text{C}$ of the sample, (2) the target biomes as the two with the highest scores achieved by the biomisation procedure (further information in Rousseau et al., 2006), (3) the atmospheric CO_2 concentration based on Petit et al. (1999) record interpolated at La Grande Pile time-scale and (4) soil type and texture. The reconstructed annual temperature and precipitation are based on iteration with value of LH higher than -0.5 , corresponding to an accepted error of maximum 0.7‰ for $\delta^{13}\text{C}$.

Mean annual temperature and annual precipitation reconstructed by inverse modelling constrained by both pollen

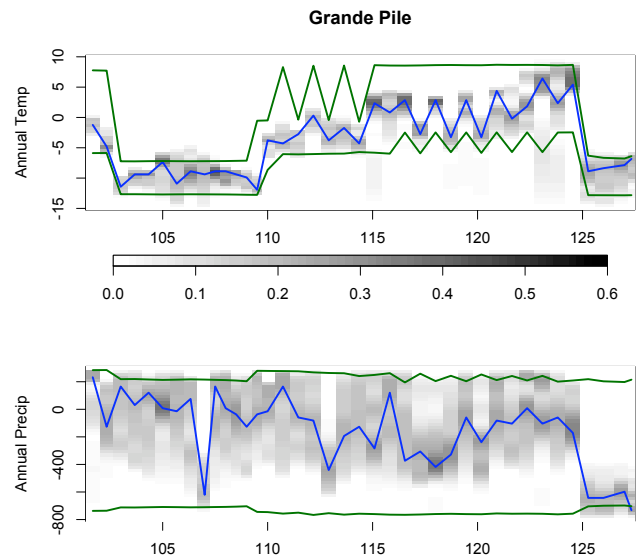


Fig. 7. Temperature and Precipitation reconstruction at La Grande Pile during Eemian period. Mean annual temperature and annual precipitation reconstructed by biome(s) and $\delta^{13}\text{C}$ inverse modelling are represented by a grey scale color for the probability distribution and its modal curve (in blue). They are bracketed (in green) by the domain that encompasses both potential climatic niches of both most likely biomes. Modified from Hatté et al. (2009).

biomes and $\delta^{13}\text{C}$ are bracketed by the ranges which should be obtained by pollen biome constrains alone (Fig. 7). The added-value of double constraints is particularly clear for precipitation reconstruction: single constraint infers a constant value with large uncertainties ($[-600, +200 \text{ mm/yr}]$ in precipitation anomaly) and a double constraint decrease uncertainty by 2 to 4. Furthermore, reconstructed temperature ranges are often decreased by a factor 2. This confirms the conclusion of previous section that pollen alone cannot give a sufficiently precise reconstruction of precipitation. This shows also that the use of two proxies decrease the uncertainty on reconstruction of both variables and inverse modelling is an elegant way to integrate several proxies related to vegetation. Nevertheless, we must note that the uncertainty provided by pollen biome is higher than uncertainty provided by the whole PFT assemblage, as in the previous subsections.

3.5 Application E: dynamic vegetation model

This section intends to illustrate the use of a dynamic vegetation model, LPJ-GUESS, with a single proxy, i.e. pollen assemblages. As the model is dynamic, this application deals with the temporal characteristics of the data, such as already suggested by Haslett et al. (2006). Vegetation is not only assumed to be dependent on the contemporaneous climate but also on the previous vegetation. Autocorrelation in the time-series is considered as an important information. Moreover, the dynamic model is not a deterministic model (two

runs with the same inputs do not produce exactly the same results). MCMC algorithm are then not applicable. Garreta et al. (2009) have proposed to use a particle filter technique more adapted to time-series and stochastic processes.

We do not use here PFT's scores but a restricted vector of 14 arboreal pollen taxa (*Abies*, *Alnus*, *Betula*, *Carpinus*, *Corylus*, *Fagus*, *Fraxinus*, *Picea*, *Pinus*, evergreen *Quercus*, deciduous *Quercus*, *Tilia*, *Ulmus* and *Populus*) and a 15th herbaceous taxon, summing all the herbaceous taxa. This choice has the maximum of coherency with the 18 species defined in LPJ-GUESS. Garreta et al. (2009) applied their method to a fossil core (Meerfelder maar) (Litt et al., 2009), but we just present here the validation of the method with modern samples. The monthly temperature and precipitation were deduced from a 6-dimensional climate parameter vector: $C=(T_{\text{Jan}}, T_{\text{Jul}}, P_{\text{win}}, P_{\text{spr}}, P_{\text{sum}}, P_{\text{aut}})$, which is slightly different from what has been done in the previous sections. The first two variables are temperature anomalies (in °C) from January and July for 1901–2000. The four other ones were seasonal (winter, spring, summer and autumn) precipitation relative anomalies (in %).

To simulate vegetation at time $t_j > t_i$, with t_i and t_j consecutive time periods (corresponding to the resolution of the core), the vegetation model starts with V_{t_i} and runs for $t_j - t_i$ years. If $t_j - t_i$ is short, vegetation simulated at t_j is strongly forced by vegetation V_{t_i} and then, implicitly, by climate C_{t_i} . This constraint gives a time-coherence to vegetation and then to reconstructed climate, and helps to produce “histories” or “dynamics” or joint distributions of vegetation and climate. This constraint can be seen as a smoother of the local bias within independent reconstructions.

A key element of the inversion model is the relationship between simulated vegetation and pollen data. In the previous sections, this has been calculated either with a statistical non linear relationship or with a correspondence matrix. Here it is approximated by a kernel surface (or a response surface) where the pollen taxon is expressed as a function of the taxonomically closest model taxon. This kernel was calibrated on a dataset of 1209 surface samples covering Europe and North Africa. It is illustrated for *Alnus* (Fig. 8) where the maximum weight is found where the coherency is best between data and model (here, in the region of low pollen and NPP values and in the region of mean pollen values and NPP around 0.02). Where pollen values are high (>4), model is unable to simulate high NPP.

The results of the method are shown as (smoothed) posterior distributions of each climatic variable. It is illustrated for an Andalusian site (Fig. 9). Mean discrepancies between posterior medians and expected values of the 6 reconstructed parameters are negligible by comparison with interval widths: the differences between the modes are <5°C for temperature and close to 0% for precipitation. There is then a bias for temperature. But, this kind of analysis has not really a sense for a single site. To really evaluate the biases, it is necessary to repeat this validation for several sites. It has

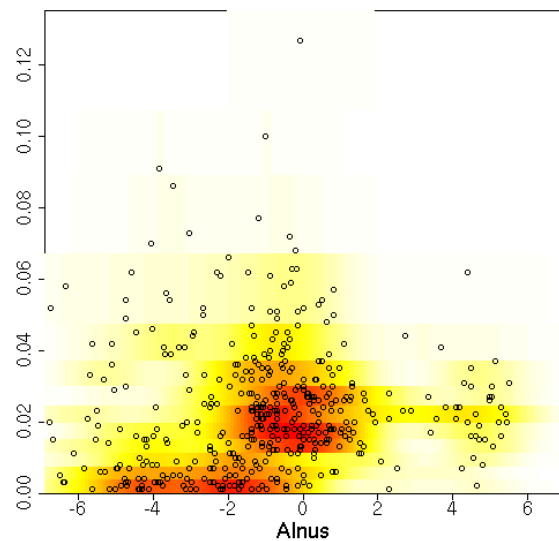


Fig. 8. *Alnus* distribution: the points are the modern sites, the x-axis is the χ^2 transform of pollen percents= $\log(P_{\text{alnus}}/P_{\text{GrSh}})$ (P for percent), the y-axis is the simulated *Alnus* using LPJ-GUESS (in $\text{kg carbon m}^{-2} \text{ year}^{-1}$) from CRUTS-1.2 (New et al., 2000) climate interpolated at each site; the color scale represents the surface fitted to the density of sites (red meaning maximum density).

been done for 30 sites in Europa and Garreta et al. (2009) have shown that the mean bias was <1°C and 3% in absolute value. Thus, the method seems to be unbiased.

To provide a valuable information, the posterior distributions must be narrower than the prior ones. It is the case for temperature where the lower limit of temperature distribution goes from -15°C to -5°C in January and from -10°C to -6°C in July (Fig. 9). Precipitation posteriors are not narrower than their priors, a result which shows that improvements, in both the vegetation model and the inversion scheme, are still necessary. Some of them concern a better modelling of the relationship between pollen dispersion and plant productivity.

4 The main results relevant to palaeoclimatology

This paper has shown the progresses which have been made in the last ten years by introducing more mechanisms in the climate reconstructions. The hypotheses behind classical approaches say that we may find in the modern world, similarities for the past and then explain the past in one location as a realization of a present situation somewhere else in the world. This is clearly the basis of the analogue approaches, but also of all statistical approaches based on a modern dataset considered as a training dataset (regression based methods, artificial neural networks ...). Mechanistic models, able to simulate a proxy in function of climate, give us the chance to work around this hypothesis at the condition to replace similarity of data by uniformity of processes.

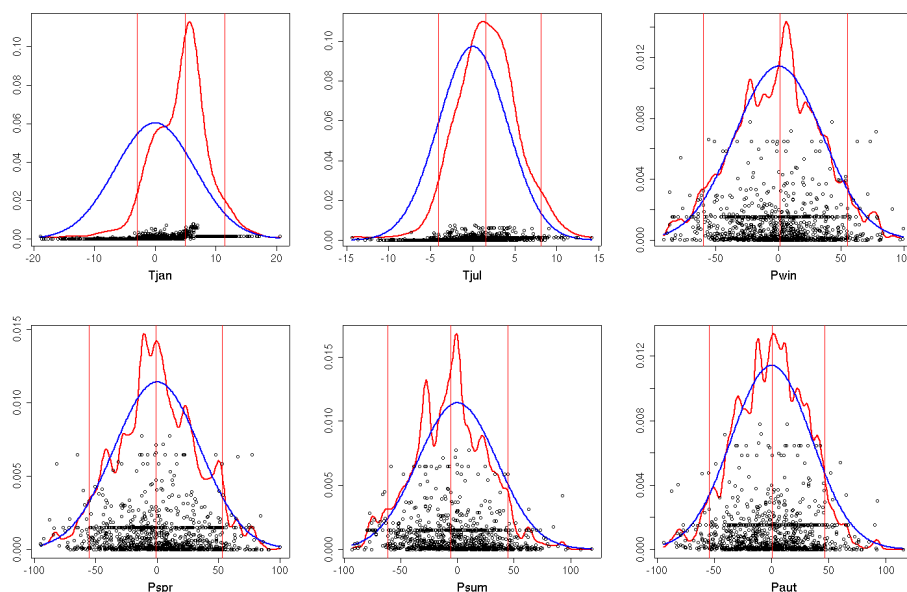


Fig. 9. Verification of the method on modern pollen data of an Andalucian site. Prior (blue lines) and posterior (red lines) univariate distributions of the 6 climate variable weighted with the particle importance and smoothed: monthly January and July temperature (in °C of anomaly), seasonal winter, spring, summer and autumn total precipitation (in % of anomaly).

It implies that such models were strongly based on mechanisms and not only a set of linear or non linear equations calibrated on modern datasets. Vegetation models is among this category.

Pollen data have the chance that vegetation models based on physiological laws have been developed more than fifteen years ago (Prentice et al., 1992). Having such model available is not the only condition. These models must also be enough simple to work with accessible inputs (climate, soil structure ...). It has been the leading mind of most of the vegetation models developed since this pioneer work. This paper has shown how to go from a relatively simple equilibrium model (BIOME3) to a dynamic model as LPJ-GUESS. These models give the possibility to work under conditions very different from the modern ones. It is clearly the case for the atmospheric CO₂ concentration often lower than the continuously increasing present one (200 ppmv during the glacial periods, around 280 ppmv during the interglacials and more than 370 ppmv today). The seasonality changes are also an interesting point. It is induced by variations of earth orbit around the sun (see the pioneer work of Berger, 1978). This feature is implicitly partly taken into account by the inversion procedure, through its effect on temperature and precipitation, as different priors are set for winter and summer. But solar radiation influences also directly photosynthesis and this should also been taken into account in the future.

Our results enable to draw several important points:

- 1) there may be a significant bias in not taking into account the difference of CO₂ between modern and past time periods. Particularly, during the glacial periods where the

difference is maximum, CO₂ fall is partly responsible of the destruction of forest in Mediterranean area. Not taking it into account, the results tend to attribute it to a too important temperature fall. The tundra-steppe vegetation of central and southern Europe is interpreted as a tundra vegetation when statistical methods are used, while a mechanistic model as BIOME4 interpret it as a cool steppe, less cold than the tundra, especially in summer. Some biases can also exist during less cold periods (Younger Dryas and even Holocene).

- 2) The use of lake-levels to constrain the reconstruction from pollen data reduces the uncertainty associated with the fact that pollen in temperate zones is a temperature indicator rather a precipitation proxy. The results, using lake Le Locle pollen data and lake levels proxies, have shown that, not only uncertainty is reduced but also larger variations are reconstructed across the Holocene.
- 3) $\delta^{13}\text{C}$ is another proxy strongly related to precipitation. The results on the Grande Pile Eemian have confirmed that the joined use of pollen and carbon isotopes reduces also the uncertainties on precipitation reconstruction.
- 4) The use of a dynamics model confirms the main role of temperature in the vegetation shifts in Europe. This approach is still in development and some improvements are necessary to make the method fully operationnal. A first result here, which maybe confirms points 2 and 3 above, is that the effect of precipitation seems to be underestimated in LPJ-GUESS or BIOME4.

5 Conclusions

Vegetation models are an elegant solution to integrate several proxies. They simulate quantities which may be related to pollen data. They simulate also fractionation of $\delta^{13}\text{C}$ in the plant which can be compared with isotopic measurements in the sediment bulk. They simulate also water absorbed by the plant and water running off. The run-off, represented by precipitation minus evapotranspiration, can be directly compared to lake-levels data when the core is lacustrine. As often lake recharge is done in winter and water useful for vegetation must be available in the growing season, the use of both proxies give a complementary enlightening of two complementary aspects of the climate, which enables also to study various seasonalities.

Despite these points, this inverse vegetation modelling approach is not the panacea. First, because it is a model-based approach, it is highly dependent on the quality of the proxy model. Second, it requires a great deal of computation time, which will increasingly become a problem in adapting the technique to more sophisticated models. Third, the outputs of the model are not directly comparable with the pollen data without a pollen dispersion modelling. Further verification is required by adapting this approach to other vegetation models. It remains important, however, to use this approach in parallel with classical statistical approaches. The comparison of results is a major key in understanding relationships between paleoclimates and palaeovegetation.

Finally, it is expected a lot in building an integrated model of the pollen accumulation in the core: this model should include all the processes such as vegetation development, pollen dispersion, catchment basin erosion, sediment accumulation, diagenesis, chronology uncertainties... A lot of work is still to be done.

Dedication: André Berger has not only been a pioneer in the theory of paleoclimates. He has stimulated a lot of scientists around him. He early understood the necessity to obtain quantified information on the past climates. Concerning the first author of this paper, he was his PhD supervisor and he pushed him to develop such quantified approaches for terrestrial archives. The first author dedicates this article to an inspiring colleague and friend.

Acknowledgements. This research has been partly funded by a grant of the French Ministry of Research to the second author, the 5th PCRD EU project MOTIF (EVK2-CT-2002-00153), by the European Science Foundation, EUROCLIMATE/DECVEG and the French ANR PICC. This is an LSCE contribution number 4000. The paper has been strongly improved thanks to the comments of C. Buck and two anonymous reviewers.

Edited by: M. Crucifix



The publication of this article is financed by CNRS-INSU.

References

- Atkinson, T. C., Briffa, K. R., and Coope, G. R.: Seasonal temperatures in Britain during the past 22,000 years, reconstructed using beetle remains, *Nature*, 325(6105), 587–592, 1987.
- Bartlein, P. J. and Prentice, I. C.: Climatic response surfaces from pollen data for some eastern North American taxa, *J. Biogeogr.*, 13, 35–57, 1986.
- Berger, A.: Long-term variations of caloric insolation resulting from the earth's orbital elements, *Quaternary Res.*, 9, 139–167, 1978.
- Cheddadi, R., Yu, G., Guiot, J., Harrison, S. P., and Prentice, I. C.: The climate of Europe 6000 years ago, *Clim. Dynam.*, 13, 1–9, 1996.
- Cowling, S. and Sykes, M.: Physiological significance of low atmospheric CO₂ for plant-climate interactions, *Quaternary Res.*, 52, 237–242, 1999.
- Doucet, A., De Freitas, N., and Gordon, N.: *Sequential Monte Carlo Methods in Practice, Statistics for Engineering and Information Science*, Springer, New York, 2001.
- EPICA: Eight Glacial cycles from an Antarctic ice core, *Nature*, 429, 623–628, 2004.
- Garreta, V., Miller, P., Guiot, J., Hely, C., Brewer, S., Sykes, M. T., and Litt, T.: A method for climate and vegetation reconstruction through the inversion of a dynamic vegetation model, *Clim. Dynam.*, doi:10.1007/s00382-009-0629-1, in press, 2009.
- Guiot, J. and DeVernal, A.: Transfer Functions: Methods for Quantitative Paleoceanography Based on Microfossils, in: *Developments in Marine Geology, Volume 1*, edited by: Hillaire-Marcel, C. and DeVernal, A., Elsevier, Amsterdam, 523–563, 2007.
- Guiot, J., Torre, F., Jolly, D., Peyron, O., Borreux, J. J., and Cheddadi, R.: Inverse vegetation modeling by Monte Carlo sampling to reconstruct paleoclimate under changed precipitation seasonality and CO₂ conditions: application to glacial climate in Mediterranean region, *Ecol. Model.*, 1, 119–140, 2000.
- Guiot, J., Hély-Alleau, C., Wu, H., and Gaucherel, C.: Interactions between vegetation and climate variability: what are the lessons of models and paleovegetation data, *C.R. Geoscience*, 340, 595–601, 2008.
- Haslett, J., Whitley, M., Bhattacharya, S., Salter-Townshend, M., Wilson, S., Allen, J. R. M., Huntley, B., and Mitchell, F. J. G.: Bayesian Paleoclimate Reconstruction, *J. Roy. Stat. Soc. A Sta.*, 169, 395–438, 2006.
- Hastings, W.: Monte Carlo sampling methods using Markov chains and their application, *Biometrika*, 57, 97–109, 1970.
- Hatté, C. and Guiot, J.: Paleoprecipitation reconstruction by inverse modelling using the isotopic signal of loess matter: application to the NuBloch loess sequence (Rhine Valley, Germany), *Clim. Dynam.*, 25, 315–327, 2005.

- Hatté, C., Rousseau, D.-D., and Guiot, J.: Climate reconstruction from pollen and $\delta^{13}\text{C}$ records using inverse vegetation modeling – Implication for past and future climates, *Clim. Past*, 5, 147–156, 2009, <http://www.clim-past.net/5/147/2009/>.
- Haxeltine, A. and Prentice, I.: BIOME3: an equilibrium terrestrial biosphere model based on ecophysiological constraints, resource availability and competition among plant functional types, *Global Biogeochem. Cy.*, 10, 693–709, 1996.
- Hickler, T., Smith, B., Sykes, M., Davis, M., Sugita, S., and Walker, K.: Using a generalized vegetation model to simulate vegetation dynamics in northeastern USA, *Ecology*, 85, 519–530, 2004.
- Indermühle, A., Stocker, T., Joos, F., Fisher, H., Smith, H., Wahlen, M., Deck, B., Mastropianni, D., Tshumi, J., Blunier, T., Meyer, R., and Stauffer, B.: Holocene carbon-cycle dynamics based on CO_2 trapped in ice at Taylor Dome, Antarctica, *Nature*, 398, 121–126, 1999.
- Iversen, J.: *Viscum, Hedera and Ilex as climatic indicators*, *Geologiska Föreningen Förhandlingar*, 66, 463–483, 1944.
- Jolly, D. and Haxeltine, A.: Effect of low glacial atmospheric CO_2 on tropical African montane vegetation, *Science*, 276, 786–788, 1997.
- Jost, A., Lunt, D., Kageyama, M., Abe-Ouchi, A., Peyron, O., Valdes, P. J., and Ramstein, G.: High-resolution simulations of the last glacial maximum climate over Europe: a solution to discrepancies with continental palaeoclimatic reconstructions?, *Clim. Dynam.*, 24, 577–590, doi:10.1007/s00382-005-0009-4, 2005.
- Kaplan, J. O., Bigelow, N. H., Prentice, I. C., Harrison, S. P., Bartlein, P. J., Christensen, T. R., Cramer, W., Matveyeva, N. V., McGuire, A. D., Murray, D. F., Razzhivin, V. Y., Smith, B., Walker, D. A., Anderson, P. M., Andreev, A. A., Brubaker, L. B., Edwards, M. E., and Lozhkin, A. V.: Climate change and Arctic ecosystems: 2. Modeling, paleodata-model comparisons, and future projections, *J. Geophys. Res.-Atmos.*, 108, 8171, doi:10.1029/2002JD002559, 2003.
- Korhola, A., Vasko, K., Toivonen, H., and Olandor, H.: Holocene temperature changes in northern Fennoscandia reconstructed from chironomids using Bayesian modelling, *Quaternary Sci. Rev.*, 21, 1841–1860, 2002.
- Litt, T., Schölzel, C., Kühl, N., and Brauer, A.: Vegetation and climate history in the Westeifel Volcanic Field (Germany) during the past 11 000 years based on annually laminated lacustrine maar sediments, *Boreas*, doi:10.1111/j.1502-3885.2009.00096.x., in press, 2009.
- Magny, M., Guiot, J., and Schoellammer, P.: Quantitative reconstruction of Younger Dryas to mid-Holocene paleoclimates at Le Locle, Swiss Jura, using pollen and lake-level data, *Quaternary Res.*, 56, 170–180, 2001.
- Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and E. T.: Equations of state calculations by fast computing machines, *J. Chem. Phys.*, 21, 1087–1093, 1953.
- New, M., Hulme, M., and Jones, P. D.: Representing twentieth century space-time climate variability. Part 2: development of 1901–96 monthly grids of terrestrial surface climate, *J. Climate*, 13, 2217–2238, 2000.
- New, M., Lister, D., Hulme, M., and Makin, I.: A high-resolution data set of surface climate over global land areas, *Climate Res.*, 21, 1–25, 2002.
- Overpeck, J. T., Prentice, I. C., and Webb, T. I.: Quantitative interpretation of fossil pollen spectra: dissimilarity coefficients and the method of modern analogs, *Quaternary Res.*, 23, 87–108, 1985.
- Petit, J., Jouzel, J., Raynaud, D., Barkov, N. I., Barnola, J., Basile, I., Bender, M., Chappellaz, J., Davis, M., Delaygue, G., Delmotte, M., Kotlyakov, V. M., Legrand, M., Lipenkov, V., Lorius, C., Pepin, L., Ritz, C., Saltzman, E., and Stievenard, M.: Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica, *Nature*, 399, 429–436, 1999.
- Peyron, O., Guiot, J., Cheddadi, R., Tarasov, P., Reille, M., De Beaulieu, J., Bottema, S., and Andrieu, V.: Climatic reconstruction in Europe for 18,000 yr BP. from pollen data, *Quaternary Res.*, 49, 183–196, 1998.
- Prentice, I. and Harrison, S.: Ecosystem effects of CO_2 concentration: evidence from past climates, *Climate of the Past*, 5, 297–307, 2009.
- Prentice, I. C., Cramer, W., Harrison, S. P., Leemans, R., Monserud, R. A., and Solomon, A. M.: A global biome model based on plant physiology and dominance, soil properties and climate, *J. Biogeogr.*, 19, 117–134, 1992.
- Prentice, I. C., Guiot, J., Huntley, B., Jolly, D., and Cheddadi, R.: Reconstructing biomes from palaeoecological data: A general method and its application to European pollen data at 0 and 6 ka, *Clim. Dynam.*, 12, 185–194, 1996.
- Robert, C. and Casella, G.: *Monte Carlo Statistical Methods*, Springer Texts in Statistics, Springer-Verlag, New York, 1999.
- Rousseau, D., Hatté, C., Guiot, D., Duzer, D., Schevin, P., and Kukla, G.: Reconstruction of the Grande Pile Eemian using inverse modelling of biomes and $\delta^{13}\text{C}$, *Quaternary Sci. Rev.*, 25, 2808–2819, 2006.
- Sitch, S., Smith, B., Prentice, I., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J., Levis, S., Lucht, W., Sykes, M., Thonicke, K., and Venevsky, S.: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ Dynamic Global Vegetation Model, *Glob. Change Biol.*, 9, 161–185, 2003.
- Smith, B., Prentice, I., and Sykes, M.: Representation of vegetation dynamics in modelling of terrestrial ecosystems: comparing two contrasting approaches within European climate space, *Global Ecol. Biogeogr.*, 10, 621–637, 2001.
- Tarasov, P., Cheddadi, R., Guiot, J., Bottema, S., Peyron, O., Belmonte, J., Ruiz-Sanchez, V., Saadi, F., and Brewer, S.: A method to determine warm and cool steppe biomes from pollen data; application to the Mediterranean and Kazakhstan regions, *J. Quaternary Sci.*, 13, 335–344, 1998.
- Wu, H., Guiot, J., Brewer, S., and Guo, Z.: Climatic changes in Eurasia and Africa at the Last Glacial Maximum and mid-Holocene: reconstruction from pollen data using inverse vegetation modelling, *Clim. Dynam.*, 29, 211–229, 2007a.
- Wu, H., Guiot, J., Brewer, S., Guo, Z., and Peng, C.: Dominant Factors Controlling Glacial and Interglacial Variations in the Tree-line Elevation in Tropical Africa, *P. Natl. Acad. Sci.*, 104, 9720–9724, 2007b.