New Coupled Model Used Inversely for Reconstructing Past Terrestrial Carbon Storage from Pollen Data: Validation of Model Using Modern Data

Haibin Wu [1, 2, 3] Joel Guiot [2] Changhui Peng [3,4*] Zhengtang Guo[5]

([1]State Key Laboratory of Loess and Quaternary Geology, Institute of Earth Environment, Chinese Academy of Sciences, Xi'an 710075, China.

[2]CEREGE, UMR 6635, CNRS/ Aix-Marseille Université, BP 80, 13545, Aix-en-Provence cedex 4, France.

[3]Institut des sciences de l´environnement, Département des sciences biologiques, Université du Québec à Montréal, Montréal H3C 3P8, Canada.

[4]Guest Professor of Ecology Research Section, Central-South University of Forestry & Technology, Changsha, Hunan 410004, China

[5]Institute of Geology and Geophysics, Chinese Academy of Sciences, P.O. Box 9825, Beijing 100029, China)

"Global Change Biology" in press

Running title: *Terrestrial carbon stock reconstruction by means of inverse process*

*To whom correspondence should be addressed: Institut des sciences de l´environnement, Université du Québec à Montréal, Case postale 8888, succ Centre-Ville, Montréal (QC), Canada H3C 3P8.
Tel: (514) 987-3000 ext. 1056, Fax: (514) 987-4718
E-mail: peng.changhui@uqam.ca

Abstract

The knowledge of potential impacts in climate change on terrestrial vegetation is crucial to understand long-term global carbon cycle development. Discrepancy in data has long existed between past carbon storage reconstructions since the Last Glacial Maximum by way of pollen, carbon isotopes, and general circulation model (GCM) analysis. This may be due to the fact that these methods do not synthetically take into account significant differences in climate distribution between modern and past conditions as well as the effects of atmospheric CO_2 concentrations on vegetation. In this study, a new method to estimate past biospheric carbon stocks is reported utilizing a new integrated ecosystem model (PCM) built on a physiological process vegetation model (BIOME4) coupled with a process-based biospheric carbon model (DEMETER). The PCM was constrained to fit pollen data to obtain realistic estimates. It was estimated that the probability distribution of climatic parameters, as simulated by BIOME4 in an inverse process, was compatible with pollen data while DEMETER successfully simulated carbon storage values with corresponding outputs of BIOME4. The carbon model was validated with present day observations of vegetation biomes and soil carbon, and the inversion scheme was tested against 1491 surface pollen spectra sample sites procured in Africa and Eurasia. Results show that this method can successfully simulate biomes and related climates at most selected pollen sites providing a coefficient of determination (R) of 0.83 to 0.97 between the observed and reconstructed climates, while also showing a consensus with an R value of 0.90 to 0.96 between the simulated biome average terrestrial carbon variables and the available observations. The results demonstrate the reliability and feasibility of the climate reconstruction method and its potential efficiency in reconstructing past terrestrial carbon storage.

***Keywords*:** terrestrial carbon storage, biome model, inverse model, BIOME6000, pollen biome scores

1. Introduction

Ice core measurements reveal that CO_2 concentrations in earth's atmosphere have exhibited large variations over glacial-interglacial cycles (Siegenthaler *et al*., 2005). There is little doubt that oceans were primarily responsible for reducing CO_2 levels during glacial phases (Siegenthaler and Wenk, 1984; Martin, 1990; François *et al*., 1998; Joos *et al*., 2004). However, paleoecological data has shown that the distribution and composition of terrestrial ecosystems during the Last Glacial Maximum (LGM) were significantly different from current conditions, and changes in global vegetation patterns continued throughout deglaciation periods and into the Holocene epoch (Prentice *et al*., 2000). Changes to terrestrial biospheric carbon stocks and those fluxes related to them may have played a role in determining atmospheric CO_2 concentrations during the past 21 ka (Adam *et al*., 1990; Prentice *et al*. 1993; Indermuhle *et al*., 1999; Kaplan *et al*., 2002; Joos *et al*., 2004). Therefore, quantitative estimates of terrestrial ecosystem carbon stocks since the LGM are required to narrow uncertainties in the global carbon cycle inventory from this time period (Adam *et al*., 1990; Prentice *et al*. 1993; François *et al*., 1998; Peng *et al*., 1998a; Indermuhle *et al*., 1999; Kaplan *et al*., 2002; Joos *et al*., 2004).

There have been many attempts to estimate terrestrial carbon storage variations between glacial and interglacial environmental conditions. One method uses palynological, pedological, and sedimentological proxy data to map the distribution of vegetation types and to estimate stocks by assuming that the average carbon density in each biome is the same as observed conditions today (Van Campo *et al*., 1993; Crowley, 1995; Adam *et al*., 1990, 1998). This estimate is solely dependent upon vegetation or biome type and does not vary geographically for specific types under different climatic conditions and atmospheric CO_2 levels, which may lead to substantial errors during glacial-interglacial periods (Esser and Lautenschlager, 1994; François *et al*., 1998; Kaplan *et al*., 2002; Joos *et al*., 2004).

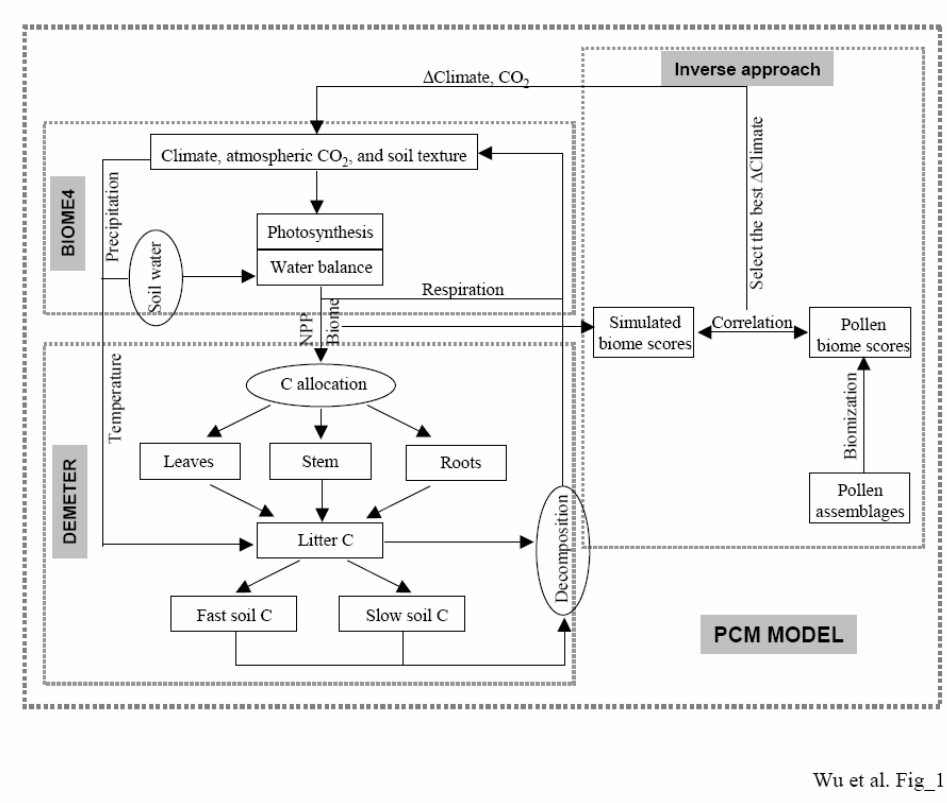
Based on statistical paleoclimatic reconstructions from pollen (Guiot *et al*., 1993; Cheddadi *et al*., 1997) and paleoclimatic simulations by GCMs, an improved method of past carbon storage estimates was developed by using a simple biospheric model (Esser and Lautenschlager, 1994; Peng *et al*., 1995a; 1995b; 1998b) to take into account the variations in carbon density applied to biomes. However, climatic reconstructions from pollen have been built upon the assumption that plant-climate interactions remain the same throughout time, and implicitly assume that these interactions are independent on changes in atmospheric CO_2 (Cowling and Sykes, 1999; Guiot *et al*., 2000; Jackson and Williams, 2004). This assumption may lead to considerable bias (Jolly and Haxeltine, 1997; Street-Perrott *et al*., 1997; Cowling and Sykes, 1999; Wu *et al*., 2007a) since physiological data has demonstrated that the processes that modify carbon and water uptake in plants are highly dependent on CO_2 concentrations (Polley *et al*., 1993; Farquhar, 1997). Polar ice core records, for example, show that atmospheric CO_2 concentrations have fluctuated significantly over at least the last 740,000 years (EPICA community members, 2004). To date, paleoclimatic simulations using GCMs are not precise enough (François *et al*., 1998; Joussaume and Taylor, 2000; Ramstein *et al*., 2007) to provide unbiased carbon storage estimates. A solution that solves this problem is to use an ecophysiological process model in an inverse mode (Guiot *et al*., 2000; Wu *et al*., 2007a; Wu *et al*., 2007b) to calculate model inputs (e.g., climate) when model outputs are constrained by pollen data and CO_2 concentrations are set to their observed value.

In this paper, a new past terrestrial carbon storage model (PCM) constrained within the BIOME6000 database (Prentice *et al*., 2000) is presented, combining the process-based biospheric models BIOME4 (Kaplan, 2001) and DEMETER (Foley, 1995) into an inverse model (Guiot et al., 2000; Wu *et al*., 2007a) where the reconstructed climate and carbon storage PCM is tested against the modern observable global climate and carbon data.

2. Model structure

Figure 1 provides a schematic of the applied method based upon vegetation, carbon models, and an inversion algorithm. The first model, BIOME4 (Kaplan *et al*., 2001), is a physiological process-based global model that is particularly useful when working with paleovegetation since it operates with a limited number of inputs (monthly temperature, precipitation, sunshine, absolute minimum temperature, soil granulometry, and atmospheric CO_2 concentrations) that are easily available. However, its equilibrium design makes it impossible to directly simulate terrestrial carbon stocks. The second model, DEMETER (Foley, 1995), provides a better simulation of global and continental-scale biospheric carbon storage within vegetation, litter, and soil, but offers only a simple potential vegetation submodel derived from the BIOME1

model (Prentice *et al*., 1992). To overcome the shortcomings of both approaches, these models are coupled with biome types and NPP that are calculated by BIOME4 and then used as inputs for the DEMETER model.



Wu et al. Fig_1

The data reveals the biome types of the pollen sites but not the corresponding climate, a key model input. To work around this, each model output was matched with the pollen data and model inputs were deduced. Typically, this is an inversion problem (Mosegaard and Tarantola, 1995) already applied to the pollen data (Guiot *et al*., 2000, Wu *et al*., 2007a; Wu *et al*., 2007b).

2.1 The vegetation submodel

The vegetation submodel is based upon the BIOME4 vegetation model developed from BIOME3 (Haxeltine and Prentice, 1996) and modified and validated by Kaplan (2001). BIOME4 predicts the global steady state of vegetation distribution, structure, and biogeochemical processes by simulating the direct effects of CO_2 on photosynthesis, stomata conductance, and the leaf area index (LAI) by means of the interactively coupled carbon and water flux model.

BIOME4 includes 12 plant functional types (PFT) defined by a set of bioclimatic limits and physiological parameters (Kaplan, 2001). These PFTs represent broad and physiologically distinct classes ranging from cushion forbs to tropical rain forest trees. For a given site, ecophysiological constraints determine the potential occurrence of specific PFTs. A coupled carbon and water flux scheme for each PFT is then used to calculate the seasonal maximum LAI that maximizes net primary production (NPP), based on a daily time step simulation of soil water balance and monthly process-based calculations of canopy conductance, photosynthesis, respiration, and phonological state. Competition between PFTs is simulated by using the optimal NPP of each PFT as an index of competitiveness.

To identify the biome of a given site, the model ranks both woody and grass PFTs

that were calculated for the site. The ranking is defined according to a set of rules based on biogeochemical variables including LAI, NPP, and mean annual soil moisture. The ranked combination of PFTs is classified into one of 27 biome types which are then sequentially used as inputs to DEMETER.

2.2 The carbon submodel

Carbon simulation follows the approaches of BIOME4 (Kaplan, 2001) and DEMETER (Foley, 1995). The photosynthetic simulation was adapted from BIOME4 since it is a process-based model better equipped at coupling carbon and water flux simulations than DEMETER, but without explicitly modeling the nitrogen cycle (Haxeltine and Prentice, 1996). Plant respiration is also calculated following the approach of BIOME4. The carbon pool simulation in vegetation, litter, and soil was adapted from Foley's terrestrial biospheric carbon model DEMETER (1995). For each plant type, annual NPP was partitioned among three parts: leaf, stem, and root (Table 1). Soil organic carbon was divided into fast and slow carbon pools (Table 1) to a depth of 1m, a simplification of the Century model developed for organic carbon material in soils (Parton *et al*., 1993) with the exception of the nitrogen cycle (see the appendix for a detail description of the model).

2.3 Inverse modeling procedure

An innovative approach of the PCM model in this study is the inversion technique used for climatic reconstruction and terrestrial carbon storage estimates. The principle behind the method is to estimate the input of BIOME4 (e.g., monthly climate) given the known data related to the output of the model, that is, biome scores (Prentice et al., 1996) (see section 3.2) derived, in this case, from pollen. The reconstructed climate is then used as an input for DEMETER to deduce the terrestrial vegetation carbon cycle (Figure 1). However, this is not an analytical inversion since it cannot be calculated mathematically. Instead, an iterative approach is used in order to find a representative set of climate scenarios compatible with vegetation records by exploring an input space defined by input parameters, here represented by monthly climatic values.

The inversion process consists of finding all combinations of climatic factors that could support a biome similar to the one observed at a given site. The main input parameters driving vegetation are temperature, precipitation, and atmospheric CO_2 concentration. To limit the number of parameters, model outputs were fitted to the observed data by altering January and July temperatures and precipitation. The other monthly parameters were deduced, including monthly sunshine, from these four parameters using empirical equations (Guiot *et al*., 2000). This procedure implies that the maximum anomalies are found either in January or July, although this may not be necessarily true in practice (Bartlein *et al*., 1998). The absolute minimum temperature (T_{min}) is based upon the relationship between T_{min} and the mean temperature of the coldest month (MTCO) deduced from global climatic data (Spangler and Jenne, 1988; Leemans and Cramer, 1991): $T_{\min} = 1.1MTCO - 21.3$ ($R^2 = 0.92$).

Since no full compatibility exists between the biome typology of BIOME4 and the biome typology of pollen data, a transfer matrix was therefore defined (Table 2). These values were set empirically by examining the modern pollen biome score data and by taking into account the theoretical definition of each biome based upon modern vegetation maps (Prentice *et al*., 1992).

For a given pollen site, the ensuing procedure was followed: (1) Selection of a four-dimensional vector of climatic anomalies (ΔT: temperature, ΔP: precipitation),

that is, the difference between past and modern values using a uniform random generator within prescribed ranges. (2) Estimation of additional monthly components of the climate using empirical equations (Eqn.36 to 38 in the appendix). (3) Addition of anomalies to the modern climate before running it through BIOME4. (4) Application of a transfer matrix (Table 2) to convert the BIOME4 biome into the pollen biome scores, and comparing the simulated biome scores to the observed scores using a Euclidian distance between the observed and the simulated biome scores (Eqn.33) and then calculating the maximum likelihood statistic LH (Eqn.32). (5) Acceptation or rejection of this climate vector based upon its LH, relative to criterion C (Eqn.35). (6) Random selection of another climate anomaly vector before returning back to procedure 1 above. This iterative process is terminated when a sufficient number of valid scenarios to calculate the *a posteriori* probability distributions are obtained, typically 200 to 300. Finally, the most probable climatic and carbon storage scenarios, together with their confidence intervals, are calculated from the *a posteriori* probabilities.

3. Model input and test data

3.1 Climate and soil

Modern monthly mean climatic variables, including temperature, precipitation, and cloudiness have been spatially interpolated to each modern pollen site using a global climatic dataset (Leemans and Cramer, 1991). The absolute minimum temperature is interpolated from the dataset compiled by Spangler and Jenne (1988). A two-layer backpropagation (BP) artificial neural network technique was used as described by Guiot *et al*. (1996) for the interpolation. Modern CO_2 concentrations were set to 340 ppmv, the concentration of the period when most of the modern pollen samples were collected (between 1970 and 1990). Soil properties were derived from the FAO digital soil map of the world (FAO, 1995).

3.2 Pollen data

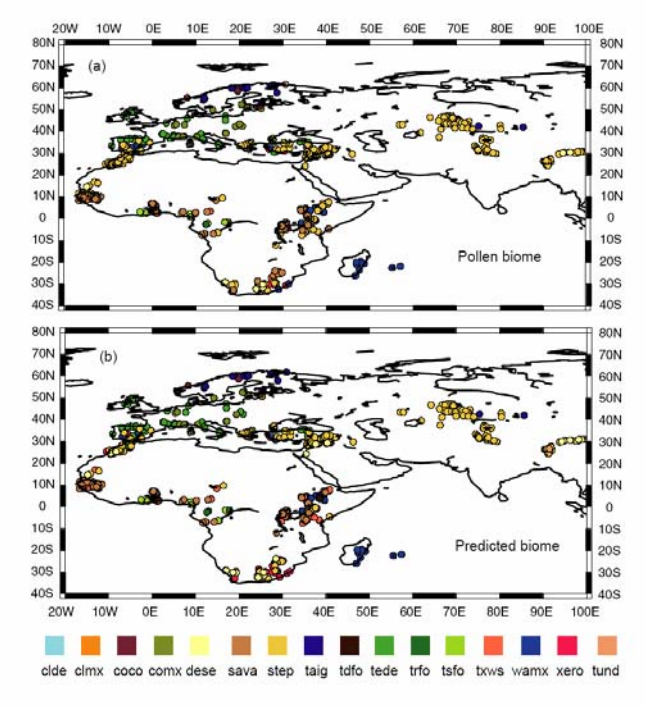
Pollen data compiled by the BIOME6000 project (Prentice *et al*., 2000) concerns three key periods: 0 k, 6 k and 21 k cal ^{14}C BP. A project goal was to classify pollen assemblages into a set of vegetation types. For this test, the modern dataset containing 1491 sample sites from Africa and Eurasia were used (Prentice *et al*., 1996, Jolly *et al*., 1998, Tarasov *et al*, 1998) covering most biome types throughout the world.

3.3 Carbon data

Several global observational datasets exist that describe the potential natural state of terrestrial carbon stocks (Schlesinger, 1977; Post *et al*., 1982; Matthews, 1983; Zinke *et al*., 1984; Olson *et al*., 1985; Batjes, 1996). For this study, the vegetation biomass and soil carbon density data complied by Olson *et al*. (1985) and Zinke *et al*. (1984) for vegetation biomass and soil carbon, respectively, were used.

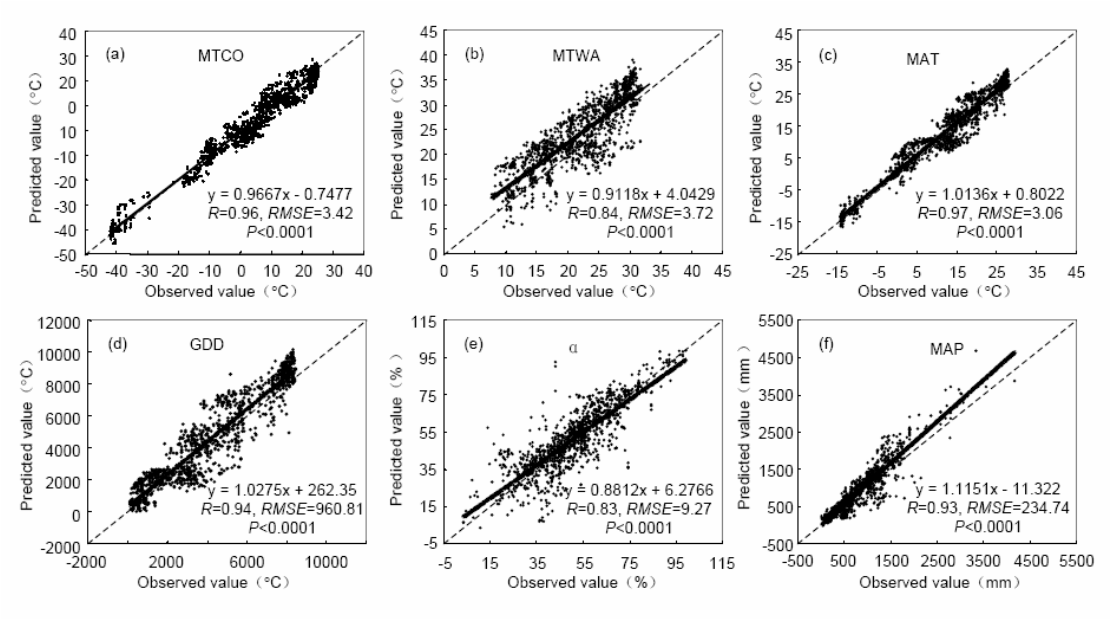
4. Validation applying measured data

To evaluate the reliability of the method, 1491 modern pollen sites in Eurasia and Africa were used for the validation test. Using the same inverse procedure discussed in section 2.3 (Inverse modeling procedure), the interpolated climatic variables from each pollen site discussed in section 3.1 (Climate and soil) were used and added with the anomalies to the modern climate by randomly browsing the climatic space as shown in Table 3. The biomes for the pollen sites were then predicated while the present day climatic and carbon densities were reconstructed.



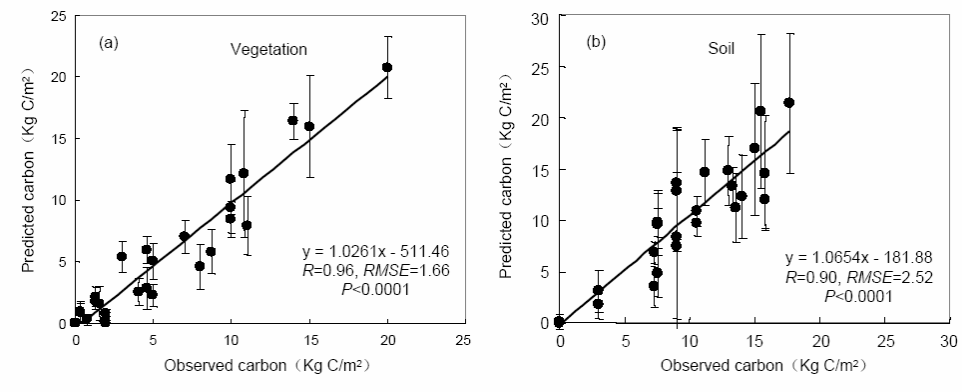
Wu et al., Fig. 2

The reconstructed biomes for the modern pollen sites had to be first compared (Figure 2a-b). No systematic regional errors exist between pollen and predicted biomes. In total, 61% of the biomes were correctly predicted (Table 4). The method worked particularly well for arboreal biomes, which correctly predicted 57% to 100% of the sites. A total of 61% signifies that the algorithm is not always able to converge to the observed biome; however, if climatically contiguous biomes are acceptable (e.g., STEP/DESE or SAVA/STEP or TUND/TAIG), it increases to 91%, considered a good fit. For the other 9% of sites, disagreement can be explained (1) partly by human impact on modern vegetation (e.g., deforestation, irrigation, planting new species in populated regions) since pollen biomes do not reflect potential vegetation; (2) partly by an incorrect estimation of the modern climate in mountainous regions where there are fewer weather stations; or (3) partly by the uncertainties in pollen-based biomization itself (e.g., some pollen samples may have two or more dominant biomes that have the same biome score).



Wu et al., Fig. 3

Second, statistical correlations between actual and reconstructed climatic variables at pollen sample sites were examined. Since climatic values reconstructed by the model are given as anomalies, the reconstructed climates are the sum of the observed climates and predicted climatic anomalies. These predicted climatic anomalies should equal zero if the observed climate and predicted climate are perfectly consistent. In the climatic validation (Figure 3), the predicted climates (in which the reconstructed climatic anomalies were added to the modern climatic values) were used to compare against the observed values, the same modern climatic values in the modern validation in this case. Correlations (R) between the observed and estimated parameters were significantly high, that is, from 0.83 to 0.97 (Figure 3a-f), based on the acceptable biomes used for climate reconstruction. Although the acceptable biomes cover a larger spatial climatic space than the same biomes, increasing the range of climatic reconstructions, the acceptable biomes cover most of the pollen sites. If a straight line is drawn between estimations and observations, an intercept of 0 and a slope of 1 are expected. The slope is slightly biased for mean temperature variables for the warmest month (MTWA) and mean annual precipitation (MAP) as well as the ratio of actual to potential evapotranspiration (α). The biases on the intercepts show a tendency to overestimate MTWA, MAP, and the growing degree for daytime temperatures above 5°C (GDD).



Wu et al., Fig. 4

Third, terrestrial carbon between model reconstructions and measurements was compared. Since the measurement data of vegetation and soil carbon densities were unavailable from the same location as the pollen sites and that these pollen sites were restricted to Africa and Eurasia, another strategy was used for validation. The PCM model was run for all global climatic data from Leemans and Cramer (1991), and the carbon simulations were averaged by biome. These averages were compared to the available data, also averaged by biome (Table 5). The coefficient of determination (R) was 0.96 for vegetation C (Figure 4a) and 0.90 for soil C (Figure 4b). The simulated biome averaged carbon densities are therefore in strong agreement with the vegetation and soil carbon data presented by Olson *et al*. (1985) and Zinke *et al*. (1984).

On a global scale, the potential vegetation carbon storage in preindustrial times simulated by the PCM based on global climatic data was approximately 907 Pg C (Table 5). Soil carbon up to 1 m deep was simulated to be 1284 Pg C. Tropical rainforests and tropical seasonal forests had the highest vegetation carbon levels, approximately 188 and 80 Pg C, respectively; whereas the largest soil carbon pool (181 Pg C) was found within cold evergreen needleleaf forests due of its large expanse.

5. Discussion and conclusion

5.1 Comparison to previous approaches

The method described in this paper is a new concept for past carbon storage simulations, allowing physiological processes constrained by pollen data. The key innovative point, in comparison to previous methods, is that this double-data model approach in which the model is constrained by data provides more realistic simulations.

For vegetation simulations, BIOME4, a physiological process-based global model, was used to improve global biome simulations. This also provides an important advantage in comparison with previous simulations based on the simple biogeographical model BIOME1 (Prentice *et al*., 1992). Table 4 shows that the simulated biomes in this study closely match the observed values.

Concerning climatic reconstruction, the results of this study on the whole are comparable to or better than previous statistical methods since the correlation between the modern observed and reconstructed variables range (1) from 0.83 (GDD) to 0.90 (MTCO) using the taxa-based best modern analogues technique in Europe (Cheddadi *et al*., 1997); (2) from 0.70 (MAP) to 0.96 (MAT) using the PFT-based best modern analogues technique in Eurasia (Tarasov *et al*., 1999); and (3) from 0.87 (MTWA) to 0.91 (MTCO) in Europe using the PFT-based best modern analogues technique (Davis *et al*., 2003). In comparison, the numbers obtained in this study range from 0.83 (α) to 0.97 (MAT).

One major difference between this study and other statistical approaches is that previous climatic reconstruction methods were built on the assumption that plant-climate interactions remain constant through time as well as the fact that calibration is carried out on modern data with the implicit assumption that these interactions are independent on changes in atmospheric CO_2. The process-based model used here does not require such hypotheses since the climate and CO_2 concentration inputs are model parameters. Even though both approaches display similar results for modern data, the process-based approach is significantly better suited for past periods where vegetation changes are potentially led by lowered CO_2

levels, especially during glacial periods (Guiot *et al*., 1999; Wu *et al*., 2007a; Wu *et al*, 2007b). This can result in significant consequences on the simulation of terrestrial carbon storage dynamics during the Quaternary period. Similarly, the inverse approach can also take into account climatologies that do not exist in modern times as increased seasonality has occurred during the Holocene, possibly resulting in a lack of modern analogues for the past.

Concerning the simulation of carbon stocks, results from this study are consistent with those obtained by DEMETER (Foley, 1995). The PCM, therefore, has the advantage of the biospheric carbon storage simulation of DEMETER. The global potential of vegetation carbon storage (907 Pg C) during the preindustrial period is also comparable to previous estimates, ranging from 560 to 1100 Pg C (Whittaker and Likens, 1973; Atjay *et al*., 1979; Olson *et al*., 1985). Soil carbon simulation (1284 Pg C) is consistent with the measured data of 1220 Pg C by Sombroek *et al*. (1993) and 1309 Pg C by Zinke *et al*. (1984), but appears to underestimate the range of estimates (1456 Pg C) from Schlesinger (1977), (1395 Pg C) Post *et al*. (1982), (1457 Pg C) Meentemeyer *et al*. (1985), and (1462 to 1548 Pg C) Batjes (1996).

This new approach is, in some ways, similar to another approaches based on integrating paleoecological data and biospheric models (Peng *et al*., 1995a; 1995b; 1998b), but differs in two major ways: (1) The input climate of the biospheric carbon models from previous studies was based on climatic reconstructions using statistical methods, whereas the climatic reconstruction carried out for this study takes into account the atmospheric CO_2 effect; (2) previous carbon storage estimates were based on statistical models or an empirical biospheric model (OBM) whereas estimates in this study were based on a process-based model. Even though the models used in this study are relatively simple, they are reliable enough to estimate past carbon storage levels since changes driving vegetation are significant. Moreover, most of the recent sophisticated global dynamic vegetation models cannot run under paleoconditions due to the lack of input information.

The validation of the PCM with modern data shows that the method can successfully simulate most pollen biomes, modern climates, and biome-averaged terrestrial carbon variables. It can be applied to past pollen data in the future, particularly to data between periods from 6 ka and 21 ka BP.

5.2 Model validation and future improvements

A major challenge for all process-based models is validation. The results presented in this study demonstrate that the PCM is able to provide reasonable reconstructions of climatic and carbon density for different vegetation types from pollen data and modern climates. However, the PCM approach is not a panacea. Since it is a model-based approach, it is highly dependent upon the quality of the vegetation model itself. Additional verification should be used when applying vegetation models.

Terrestrial carbon accumulation may be constrained by nutrients, particularly nitrogen (Oren *et al*., 2001; Finzi *et al*., 2002; Luo *et al.*, 2004), through mechanisms that are not explicitly modeled in this study. Moreover, the PCM requires a great amount of computation time that will become problematic if using more sophisticated dynamic models that include more precise simulations of NPP, belowground biomass, soil carbon, nitrogen, and water dynamics.

The MCMC algorithm used in the PCM does not guarantee convergence towards an optimal solution, which is further complicated by the fact that simulated biome

types in BIOME4 are not directly comparable to pollen biomes, adding a source of uncertainty to the analysis that is difficult to quantify.

In future versions of the PCM, these shortcomings will be minimized, and the carbon simulation will be scaled up to the regional or global levels by a greater use of the BIOME6000 pollen database. Changes to global or regional terrestrial carbon dynamics since the LGM will therefore be estimated with less uncertainties. It remains important, however, to use this approach in parallel with other methods. The comparison of different past carbon storage reconstructions can improve our understanding of dynamic processes of global vegetation and carbon cycling during glacial-interglacial periods.

Acknowledgements

This research was supported by funding by the National Basic Research Program of China (No. 2004CB720203), the National Natural Science Foundation of China (No. 40730104, 40599422), the NSFC for Innovation Group project (N0.40121303), a grant from the French Ministry of Research to Haibin Wu as well as from the Furong Scholar Program, the Natural Sciences and Engineering Research Council of Canada Discovery Grant and Canada Research Chair Program, the PICC project funded by the French National Agency for Research (ANR) and the DECVEG project funded by the French CNRS and the European Science Foundation.

Reference:

Adams JM, Faure H (1998) A new estimate of changing carbon storage on land since the last glacial maximum, based on global land ecosystem reconstruction. *Global and Planetary Change*, **16-17**, 3-24.

Adams JM, Faure H, Faure-Denard L, McGlade JM, Woodward FI. (1990) Increase in terrestrial carbon storage from the last glacial maximum to the present. *Nature*, **348**, 711-714.

Atjay GL, Ketner P, Duvingeaud P (1979) Terrestrial primary production and phytomass. In: *The global carbon cycle*, (eds Bolin B, Degens ET, Kempe S, Ketner P), pp. 129-181. New York, Wiley.

Bartlein PJ, Anderson KH, Anderson PM, *et al*. (1998) Palaeoclimate simulations for North America over the past 21,000 years: features of the simulated climate and comparisons with paleoenvironmental data. *Quaternary Science Review*, **17**, 547-585.

Batjes NH (1996) Total C and N in soils of the world, *Europen Journal of Soil Science*, **47**, 151–163.

Cheddadi R, Yu G, Guiot J, Harrison SP, Prentice IC (1997) The climate of Europe 6000 ago. *Climate Dynamics*, **13**, 1-9.

Cowling SA, Sykes M (1999) Physiological significance of low atmospheric CO_2 for plant-climate interactions. *Quaternary Research*, **52**, 237-242.

Crowley TJ (1995) Ice age terrestrial carbon change revised. *Global Biogeochemical Cycles*, **9**, 377-389.

Davis BAS, Brewer S, Stevenson AC, Guiot J, Data Contributors (2003) The

temperature of Europe during the Holocene reconstructed from pollen data. *Quaternary Science Review*, **22**: 1701-1716.

EPICA community members (2004) Eight glacial cycles from an Antarctic ice core. *Nature,* **429**, 623-628.

Esser G, Lautenschlager M (1994) Estimating the change of carbon in the terrestrial biosphere from 18,000 BP to present using a carbon cycle model. *Environmental Pollution*, **83**, 45-53.

Farquhar GD (1997) Carbon dioxide and vegetation. *Science*, **278**, 1411.

Finzi AC, DeLucia EH, Hamilton JG, Richter DD, Schlesinger WH (2002) The nitrogen budget of a pine forest under free air CO_2 enrichment. *Oecologia*, **132**, 567-578.

Foley JA (1995) An equilibrium model of the terrestrial carbon budget. *Tellus*, **47B**, 310-319.

Food and Agriculture Organization (FAO) (1995) *Soil Map of the World 1:5,000,000*. U.N. Educ., Sci., and Cult. Organ., Paries.

François LM, Delire C, Warnant P (1998) Modelling the glacial-interglacial changes in the continental biosphere. *Global and Planetary Change*, **16-17**, 37-52.

Guiot J, Harrison SP, Prentice IC (1993) Reconstruction of Holocene precipitation patterns in Europe using pollen and lake-level data. *Quaternary Research*, **40**, 139-149.

Guiot J, Cheddadi R, Prentice IC, Jolly D (1996) A method of biome and land surface mapping from pollen data: application to Europe 6000 years ago. *Palaeoclimates: data model*, **1**, 311-324.

Guiot J, Torre F, Jolly D, Peyron O, Boreux JJ, Cheddadi R (2000) Inverse vegetation modeling by Monte Carlo sampling to reconstruct palaeoclimate under changed precipitation seasonality and CO_2 conditions: application to glacial climate in Mediterranean region. *Ecology Modelling*, **127**, 119-140.

Haxeltine A, Prentice IC (1996) BIOME3: An equilibrium terrestrial biosphere model based on ecophysiological constraints, resource availability, and competition among plant functional types. *Global Biogeochemical Cycles*, **10**, 693-709.

Indermuhle A, Stocker T F, Joos F, *et al*. (1999) Holocene carbon-cycle dynamics based on CO_2 trapped in ice at Taylor Dome, Antarctica. *Nature*, **398**, 121-126.

Jackson ST, Williams JW (2004) Modern analogs in quaternary paleoecology- Here Today, Gone Yesterday, Gone Tomorrow? *Annual Review of Earth Planetary Science*, **32**, 495-537.

Jolly D, Haxeltine A (1997) Effect of low glacial atmospheric CO_2 on tropical African montane vegetation. *Science,* **276**, 786-788.

Jolly D, Prentice IC, Bonnefille R, *et al*. (1998) Biome reconstruction from pollen and plant macrofossil data for Africa and the Arabian Peninsula at 0 and 6 ka. *Journal of Biogeography*, **25**, 1007-1028.

Joos F, Gerber S, Prentice IC, Otto-Bliesner BL, Valdes PJ (2004) Transient simulations of Holocene atmospheric carbon dioxide and terrestrial carbon since the Last Glacial Maximum. *Global Biogeochemical Cycles*, **18**, doi: 10.1029/2003GB002156.

Joussaume S, Taylor KE (2000) The paleoclimate modeling Intercomparsion project. In: *Paleoclimate modeling intercomparsion project (PMIP)* (eds. Braconnot P). Proc Third PMIP Workshop, Canada, 4-8 October 1999, vol WCRP-111, WMO/TD-no.1007, pp 9-24.

Kaplan JO, Prentice IC, Knorr W, Valdes PJ (2002) Modeling the dynamics of terrestrial carbon storage since the Last Glacial Maximum. *Geophysical Research Letters*, **29**, 2074-2077.

Kaplan JO (2001) *Geophysical applications of vegetation modeling*. Lund University, Sweden.

Leemans R, Cramer W (1991) *The IIASA Climate Database for mean monthly values of temperature, precipitation and cloudiness on a global terrestrial grid*. RR-91-81, International Institute of Applied Systems Analysis, Laxenburg.

Luo YQ, Su B, Currie WS, *et al*. (2004) Progressive nitrogen limitation of ecosystem responses to rising atmospheric carbon dioxide. *BioScience*, **54**, 731-739.

Martin JH (1990) Glacial-interglacial CO_2 changes: the iron hypothesis. *Paleoceanography*, **5**, 1-13.

Matthews E (1983) Global vegetation and land use: new high resolution data bases for climate studies. *Journal of Climatology and Applied Meteorology*, **22**, 474-487.

Mosegaard K. Tarantola A (1995) Monte Carlo sampling of solutions to inverse problems. *Journal of geophysical Research*, **100 B7**, 12431-12447.

Olson JS, Watts JA, Allison LJ (1985) *Major world ecosystem complexes ranked by carbon in live vegetation*. A Database. NPD-017. Carbon Dioxide Information Center. Oak Ridge National Laboratory. Oak Ridge, Tennessee.

Oren R, Ellsworth DS, Johnsen KH, *et al*. Soil fertility limits carbon sequestration by forest ecosystem in a CO_2 enriched atmosphere. *Nature*, **411**, 469-472.

Parton WJ, Scurlock JM, Ojima DS, *et al*. (1993) Observations and modelling of biomass and soil organic matter dynamics for the grassland biome worldwide. *Global Biogeochemical Cycles*, **7**, 785-809.

Peng CH, Guiot J, Campo EV (1998a) Estimating changes in terrestrial vegetation and carbon storage: using palaeoecological data and models. *Quaternary Science Reviews*, **17**, 719-735.

Peng CH, Guiot J, Van Campo E (1998b) Past and future carbon balance of European ecosystem from pollen data and climatic models simulations. *Global and Planetary Change*, **18**, 189-200.

Peng CH, Guiot J, Van Campo E, Cheddadi R (1995a) Reconstruction of the past terrestrial carbon storage of the Northern Hemisphere from the Osnabrűck Biosphere Model and palaeodata. *Climate Research*, **5,** 107-118.

Peng CH, Guiot J, Van Camp E, Cheddadi R (1995b) The variation of terrestrial carbon storage at 6000 yr BP in Europe: reconstruction from pollen data using two empirical biosphere models. *Journal of Biogeography*, **22**, 863-873.

Polley HW, Johnson HB, Marino B, Mayeux HS (1993) Increase in C3 plant water-use efficiency and biomass over Glacial to present CO_2 concentrations. *Nature*, **361**, 61-64.

Post WM, Emanual WR, Zinke PJ, Stangenberger AG (1982) Soil carbon pools and

world life zones. *Nature*, **298**, 156-159.

Prentice IC, Cramer W, Harrison SP, Leemans R, Monserud RA, Solomon AM (1992) A global biome model based on plant physiology and dominance, soil properties and climate. *Journal of Biogeography*, **19**, 117-134.

Prentice IC, Guiot J, Huntley B, Jolly D, Cheddadi R (1996) Reconstructing biomes from palaeoecological data: a general method and its application to European pollen data at 0 and 6 ka. *Climate Dynamics*, **12**, 185-194.

Prentice IC, Jolly D, BIOME 6000 P (2000) Mid-Holocene and glacial-maximum vegetation geography of the northern continents and Africa. *Journal of Biogeography*, **27**, 507-519.

Prentice IC, Sykes MT, Lautenschlager M, Harrison SP, Dennissenko O, Bartlein PJ (1993) Modeling global vegetation patterns and terrestrial carbon storage at the last glacial maximum. *Global Ecology and Biogeography Letters*, **3**, 67-76.

Ramstein G, Kageyama M, Guiot J, Wu H, Hely C, Krinner G, Brewer S (2007) How cold was Europe at the Last Glacial Maximum? A synthesis of the progress achieved since the first PMIP model-data comparison. *Climate of Past Discussion*, **3**, 197-220.

Schlesinger WH (1977) Carbon balance in terrestrial detritus. *Annual Review of Ecology and Systematics*, **8**, 51-81.

Siegenthaler U, Wenk T (1984) Rapid atmospheric CO_2 variations and ocean circulation. *Nature*, **308**, 624-626.

Siegenthaler U, Stocher TF, Monnion E, *et al* (2005) Stable carbon cycle-climate relationship during the late Pleistocene. *Science*, **310**, 1313-1317.

Spangler WML, Jenne RL (1988) *World monthly surface station climatology*. NCAR, Scientific computing division, report 6p+CDROM.

Street-Perrott FA, Huang YS, Perrott A, Eglinton G, Barker P, Khelifa LB, Harkness DD, Olago DO (1997) Impact of lower atmospheric carbon dioxide on tropical mountain ecosystems. *Science,* **278**, 1422-1426.

Tarasov PE, Webb TIII, Andreev AA, *et al*. (1998) Present day and mid-Holocene biomes reconstructed from pollen and plant macrofossil data from the former Soviet Union and Mongolia. *Journal of Biogeography,* **25**, 1029-1053.

Tarasov PE, Peyron O, Guiot J, *et al*. (1999) Last Glacial Maximum climate of the former Soviet Union and Mongolia reconstructed from pollen and plant macrofossil data. *Climate Dynamics*, **15**, 227-240.

Van Campo E, Guiot J, Peng CH (1993) A data-based re-appraisal of the terrestrial carbon budget at the Last Glacial Maximum. *Global and Planetary Change*, **8**, 189-201.

Whittaker RH, Likens GE (1973) Carbon in the biota. In: *Carbon and the biosphere* (eds.Wordwell GM, Pecan EV), pp. 281-302, Washington, DC: National Technical Servie Inforamtion Service.

Wu HB, Guiot J, Brewer S, Guo ZT (2007a) Climatic changes in Eurasia and Africa at the Last Glacial Maximum and mid-Holocene: reconstruction from pollen data using inverse vegetation modelling, *Climate Dynamics*, **29**, 211-229, doi: 10.1007/s00382-007-0231-3.

Wu HB, Guiot J, Brewer S, Guo ZT, Peng CH (2007b), Dominant factors controlling glacial and interglacial variations in the treeline elevation in tropical Africa. *Proceeding of the National Academy of Science of the United States of America*, 104: 9720-9724.

Zinke PJ, Stangenberger AG, Post WM, Emanuel WR, Olson JS (1984) *Worldwide organic soil carbon and nitrogen data (No. ORNL/TM-8857)*. Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Figure Captions

Figure 1: Schematic diagram of the inverse vegetation modeling approach to past carbon storage reconstruction.

Figure 2: Comparison of each site between pollen-based and simulated biomes in Eurasia and Africa at 0 ka BP. (a) Observed pollen biome. (b) Predicted biome by model. See caption of Table 1 for biome code.

Figure 3: Correlation between observed climates and predicted values by the model in Eurasia and Africa at 0 ka BP. (a) Mean temperature of the coldest month (MTCO); (b) mean temperature of the warmest month (MTWA); (c) growing degree-days above 5°C (GDD); (d) ratio of actual to potential evapotranspiration (α); (e) mean annual temperature (MAT); and (f) mean annual precipitation (MAP). The solid line is the least-squares linear regression, and the dashed line is the 1:1 line. R is the correlation coefficient, and RMSE is the root-mean-square error of the residuals.

Figure 4: Correlation between averaged PCM carbon simulations and averaged observations: (a) vegetation carbon density; and (b) soil carbon density. The error bars relate to the 95% confidence intervals. R and RMSE are the correlation coefficient and the root-mean-square error of the residuals, respectively.

Table 1: Parameters of vegetation carbon simulation

Vegetation type	p_{leaf}	p_{stem}	p_{root}	t_{leaf} (yr)	t_{stem} (yr)	t_{root} (yr)	f_{resp}	P_{fast}	P_{slow}
TrEgFo	0.40	0.20	0.40	1	50	10	0.65	0.980	0.020
TrSeDeFo	0.40	0.20	0.40	1	50	10	0.65	0.980	0.020
TrDeFo	0.40	0.20	0.40	1	25	5	0.65	0.980	0.020
TeDeFo	0.40	0.20	0.40	1	50	10	0.70	0.985	0.015
TeCoFo	0.40	0.20	0.40	2	60	10	0.70	0.985	0.015
WaMxFo	0.40	0.20	0.40	1	50	10	0.70	0.985	0.015
CoMxFo	0.40	0.20	0.40	2	60	10	0.70	0.985	0.015
CoCoFo	0.40	0.20	0.40	2	80	10	0.70	0.985	0.015
ClMxFo	0.40	0.20	0.40	2	80	10	0.70	0.985	0.015
EgTaig	0.40	0.20	0.40	2	80	10	0.70	0.985	0.015
DeTaig	0.40	0.20	0.40	2	80	10	0.70	0.985	0.015
TrSav	0.50	0.10	0.40	2	30	5	0.70	0.985	0.015
TrXsSl	0.50	0.10	0.40	2	40	5	0.70	0.985	0.015
TeXsSl	0.50	0.10	0.40	1	40	5	0.70	0.985	0.015
TeScWo	0.50	0.10	0.40	1	40	5	0.70	0.985	0.015
TeBlSav	0.50	0.10	0.40	1	40	5	0.70	0.985	0.015
OpCoWo	0.50	0.10	0.40	1	40	5	0.70	0.985	0.015
BoPrkl	0.50	0.10	0.40	1	40	5	0.70	0.985	0.015
TrGrl	0.70	0.10	0.20	1	30	2	0.70	0.985	0.015
TeGrlc	0.70	0.10	0.20	1	30	2	0.70	0.985	0.015
TeGrlw	0.70	0.10	0.20	1	30	2	0.70	0.985	0.015
HotDesert	0.50	0.20	0.30	1	30	2	0.70	0.985	0.015
Desert	0.60	0.10	0.30	1	30	2	0.70	0.985	0.015
ShTund	0.60	0.10	0.30	1	80	5	0.70	0.985	0.015
DShTund	0.60	0.10	0.30	1	80	5	0.70	0.985	0.015
PsShTund	0.60	0.10	0.30	1	80	5	0.70	0.985	0.015
FoLiMoss	0.60	0.10	0.30	1	1	1	0.70	0.985	0.015
Barren	0.00	0.00	0.00	0	0	0	0.00	0.000	0.000
LIce	0.00	0.00	0.00	0	0	0	0.00	0.000	0.000

Vegetation (BIOME4) types: Barren, barren land; BoPrkl, boreal parkland; ClMxFo, cold mixed forest; CoCoFo, cool evergreen needleleaf forest; CoMxFo, cool mixed forest; Desert, desert; DeTaig, cold deciduous forest; DshTund, erect dwarf-shrub tundra; EgTaig, cold evergreen needleleaf forest; FoLiMoss, cushion-forb, lichen, and moss tundra; HotDesert, hot desert; LIce, land ice; OpCoWo; temperate evergreen needleleaf open woodland; PsShTund, prostrate dwarf-shrub tundra; ShTund, low and high shrub tundra; TeBlSav, temperate deciduous broadleaved savanna; TeCoFo, temperate evergreen needleleaf forest; TeDeFo, temperate deciduous broadleaf forest; TeGrlc, cool temperate grassland; TeGrlw, warm temperate grassland; TeScWo, temperate sclerophyll woodland and shrubland; TeXsSl, temperate xerophytic shrubland; TrDeFo, tropical deciduous broadleaf forest and woodland; TrEgFo, tropical evergreen broadleaf forest; TrGrl, tropical grassland; TrSav, tropical savanna; TrSeDeFo, tropical semi-evergreen broadleaf forest; TrXsSl, tropical xerophytic shrubland; WaMxFo, warm-temperate evergreen broadleaf and mixed forest.

p_{leaf}, p_{stem}, and p_{root}t represent the allocation of NPP into the biomass compartments. t_{leaf}, t_{stem}, and t_{root} are the average lift span (years) of plant compartments of vegetations. f_{resp} is the fraction of litter carbon released as CO_2 during decomposition. P_{fast} and P_{slow} are the fractions of humus sent to the fast and slow pools, respectively.

Table 2: Transfer matrix from BIOME4 typology to the pollen biome scores

BIOME4 type	Pollen biome type																
	CLDE	CLMX	COCO	COMX	DESE	STEP	TAIG	TEDE	TUND	XERO	HODE	SAVA	TDFO	TRFO	TSFO	WAMX	TXWS
TrEgFo	0	0	0	0	0	0	0	0	0	0	0	0	5	15	10	0	0
TrSeDeFo	0	0	0	0	0	0	0	0	0	0	0	0	10	10	15	0	5
TrDeFo	0	0	0	0	0	0	0	0	0	0	0	5	15	5	10	0	0
TeDeFo	0	5	5	10	0	0	0	15	0	0	0	0	0	0	0	10	0
TeCoFo	0	0	15	10	0	0	0	5	0	0	0	0	0	0	0	0	0
WaMxFo	0	0	0	0	0	0	0	10	0	10	0	0	0	0	0	15	0
CoMxFo	0	0	10	15	0	0	0	10	0	0	0	0	0	0	0	0	0
CoCoFo	0	0	15	10	0	0	5	0	0	0	0	0	0	0	0	0	0
ClMxFo	10	15	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0
EgTaig	5	10	5	0	0	0	15	0	0	0	0	0	0	0	0	0	0
DeTaig	10	5	0	0	0	0	15	0	5	0	0	0	0	0	0	0	0
TrSav	0	0	0	0	0	5	0	0	0	0	0	15	5	0	0	0	10
TrXsSl	0	0	0	0	0	10	0	0	0	0	0	5	0	0	0	0	15
TeXsSl	0	0	0	0	0	5	0	0	0	15	0	0	0	0	0	5	0
TeScWo	0	0	0	0	0	5	0	0	0	15	0	5	0	0	0	10	0
TeBlSav	0	0	0	0	0	5	0	5	0	5	0	15	0	0	0	5	0
OpCoWo	0	0	10	0	0	5	0	0	0	0	0	0	0	0	0	0	0
BoPrkl	0	0	5	0	0	10	10	0	0	5	0	0	0	0	0	0	0
TrGrl	0	0	0	0	0	15	0	0	0	0	5	5	0	0	0	0	10
TeGrlc	0	0	0	0	5	15	0	0	5	0	0	0	0	0	0	0	0
TeGrlw	0	0	0	0	5	15	0	0	0	5	0	5	0	0	0	0	0

HotDesert	0	0	0	0	0	10	0	0	0	0	15	0	0	0	0	0	0
Desert	0	0	0	0	15	10	0	0	0	0	0	0	0	0	0	0	0
ShTund	5	0	0	0	0	14	5	0	15	0	0	0	0	0	0	0	0
DShTund	0	0	0	0	0	5	0	0	15	0	0	0	0	0	0	0	0
PsShTund	0	0	0	0	0	5	0	0	15	0	0	0	0	0	0	0	0
FoLiMoss	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0
Barren	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
LIce	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Pollen biome types: CLDE, cold deciduous forest; CLMX, cold mixed forest; COCO, cool coniferous forest; COMX, cool mixed forest; DESE, desert; HODE, hot desert; SAVA, savanna; STEP, steppe; TAIG, taiga; TDFO, tropical dry forest; TEDE, temperate deciduous forest; TRFO, tropical rain forest; TSFO, tropical seasonal forest; TUND, tundra; TXWS, tropical xerophytic woods/scrub; WAMX, broadleaved evergreen/warm mixed forest; XERO, xerophytic woods/scrub. BIOME4 type codes are given in Table 1. Temperate grassland were divided into cool temperate grassland and warm temperate grassland, and desert into cold desert and hot desert, based upon the minimum temperature (22°C) of the mean temperature of the warmest month (Prentice *et al*., 1992).

Table 3: Modern ranges of input parameters for simulations.

Parameter	Modern
ΔT_{jan}	[-10, 10]°C
ΔT_{jul}	[-10, 10]°C
ΔP_{jan}	[-90, 100]%
ΔP_{jul}	[-90, 100]%
CO_2	340ppmv
Number of iterations	2000

Ranges are given in anomalies for modern values (deviation for temperature and percentages for precipitation).

Table 4: Numerical comparison of each site between pollen-derived ('p') and simulated ('s') biomes at modern sites in Eurasia and Africa

Biome	$CLDE_s$	$CLMX_s$	$COCO_s$	$COMX_s$	$DESE_s$	$SAVA_s$	$STEP_s$	$TAIG_s$	$TDFO_s$	$TEDE_s$	$TRFO_s$	$TSFO_s$	$TUND_s$	$TXWS_s$	$WAMX_s$	$XERO_s$	N	Excellent (%)	Acceptable (%)
$CLDE_p$	**0**	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
$COCO_p$	0	0	**4**	0	0	0	0	0	0	0	0	0	0	0	0	0	4	100	100
$COMX_p$	0	0	0	**24**	0	0	0	3	0	**5**	0	0	0	0	0	0	32	75	91
$DESE_p$	0	0	0	0	**12**	0	0	0	0	0	0	0	0	0	0	0	12	100	100
$SAVA_p$	0	0	0	0	0	**49**	**1**	0	1	0	0	0	0	**41**	1	**10**	103	48	98
$STEP_p$	0	1	0	1	**35**	**41**	**270**	6	2	24	0	2	**4**	**133**	15	**57**	581	46	92
$TAIG_p$	0	0	0	1	0	0	0	**106**	0	1	0	0	**1**	0	0	0	108	97	98
$TDFO_p$	0	0	0	0	0	1	0	0	**4**	0	0	0	0	1	0	1	7	57	57
$TEDE_p$	0	0	0	0	0	0	0	0	0	**113**	0	0	0	0	**1**	0	114	98	100
$TRFO_p$	0	0	0	0	0	0	0	0	**1**	0	**8**	0	0	0	0	0	9	89	100
$TSFO_p$	0	0	0	0	0	0	0	0	**4**	0	**13**	**36**	0	1	1	0	55	65	96
$TUND_p$	0	0	0	0	1	0	**12**	**37**	0	6	0	0	**22**	0	1	1	80	28	89
$TXWS_p$	0	0	0	0	0	0	0	0	0	0	0	0	0	**4**	0	0	4	100	100
$WAMX_p$	0	0	0	0	0	0	0	0	1	**22**	1	3	0	0	**169**	**3**	199	85	98
$XERO_p$	0	0	7	1	2	0	**3**	2	0	49	0	0	0	2	**26**	**90**	182	49	65
Average																		61%	91%
Total SN																	1491	910	1350

N: number of sites; Excellent: the same biomes between observation and simulation; Acceptable: the same and climatically contiguous biomes between observation and simulation. Bold value: the same and climatically contiguous biomes (i.e., excellent and acceptable biomes). Value in cell (i,j) provides the number of pollen biomes i simulated as biome j. Biome codes are given in Table 2.

Table 5: Potential global terrestrial carbon storage during preindustrial periods

Biome type	Area	Carbon density (Kg C/m^2)			Carbon storage (Pg C)			
	(10^{10} m^2)	Vegetation	Litter	Soil	Vegetation	Litter	Soil	Total
TrEgFo	906.1	20.7	1.0	10.9	187.7	9.4	98.6	295.8
TrSeDeFo	489.5	16.4	1.0	9.7	80.1	4.8	47.7	132.6
TrDeFo	663.8	7.0	1.1	9.8	46.8	7.4	64.9	119.0
TeDeFo	594.3	11.7	3.0	16.9	69.4	18.0	100.7	188.1
TeCoFo	313.8	12.1	2.2	12.3	38.0	6.8	38.6	83.5
WaMxFo	870.3	15.9	1.7	11.2	138.8	15.2	97.5	251.4
CoMxFo	475.4	9.4	2.7	14.8	44.7	12.6	70.6	127.9
CoCoFo	429.0	8.4	2.4	13.3	36.1	10.3	57.0	103.4
ClMxFo	119.2	7.9	2.3	20.6	9.4	5.2	24.5	39.1
EgTaig	1506.7	5.8	2.2	12.1	87.3	33.9	181.7	302.9
DeTaig	379.1	4.6	2.9	14.6	17.3	10.9	55.3	83.5
TrSav	678.7	5.9	0.8	6.8	40.3	5.5	46.3	92.1
TrXsSl	1252.2	2.8	0.5	3.5	35.2	5.7	43.8	84.7
TeXsSl	557.4	2.5	0.7	4.9	14.1	3.8	27.1	45.1
TeScWo	341.6	5.0	1.4	9.8	17.2	4.8	33.4	55.4
TeBlSav	155.7	5.4	2.2	14.7	8.4	3.4	22.9	34.6
OpCoWo	156.6	2.2	1.7	9.6	3.5	2.6	15.0	21.2
BoPrkl	17.7	1.6	1.4	7.5	0.3	0.2	1.3	1.9
TrGrl	238.3	1.8	2.6	13.7	4.2	6.1	32.5	42.9
TeGrl	496.1	2.0	2.5	13.2	9.7	12.3	65.7	87.7
Desert	1818.4	0.9	0.3	1.9	16.4	6.2	34.8	57.5
Tund	1379.3	0.4	0.6	11.3	5.7	8.3	156.2	170.2
Barren	101.5	0.0	0.0	0.1	0.0	0.0	0.1	0.1
Global total					**906.5**	**187.5**	**1283.7**	**2377.7**

Biome codes are given in Table 2.

Appendix:

1. Carbon submodel

The carbon simulation follows the approaches of BIOME4 and DEMETER. Although a full description of the model is provided by Haxeltine and Prentice (1996) and Foley (1995), a brief overview of the carbon cycle model is presented below.

1.1 Photosynthesis

The photosynthetic scheme is based on the Farquhar photosynthetic model as simplified by Collatz *et al*. (1991, 1992). The optimality hypothesis is assumed to apply since nitrogen content and Rubisco activity of leaves are assumed to vary both seasonally and with variation in canopy position in a way that maximizes net assimilation at leaf level. The hypothesis states that this is due to light use efficiencies and photosynthetic rates that are reduced in natural ecosystems by temperature, water, and nutrient stress.

Plant photosynthesis adopting the C3 versus C4 biochemical pathways is modeled in slightly different ways. For C3 plants assimilation, A_{nd} is given by:

$$A_{nd} = \frac{J_E + J_C - \sqrt{(J_E + J_C)^2 - 4J_E J_C}}{2\theta} - R_d \quad (1)$$

where

$$J_E = C1_{C3} APAR \quad (2)$$

$$C1_{C3} = \Phi_C \Phi_{TC3} C_{mass} \alpha_a \alpha_{C3} \frac{p_i - \Gamma_*}{p_i + \Gamma_*} \quad (3)$$

$$J_C = C2_{C3} V_m \quad (4)$$

$$C2_{C3} = \frac{p_i - \Gamma_*}{p_i + K_C \frac{[O_2]}{Ko}} \quad (5)$$

where A_{nd} is the daily net photosynthetic uptake; R_d the daily leaf respiration rate; θ = 0.7; APRA the total absorbed PAR; Φ_C the PFT specific parameter; Φ_{TC3} the effect of low temperatures on C_3 photosynthesis; C_{mass} the molar mass of carbon; α_a the scaling parameter for α ; α_{C3} the intrinsic quantum efficiency for CO_2 uptake; p_i the internal partial pressure of CO_2; Γ_* the CO_2 compensation point; K_C and K_O the Michalis constant for CO_2 and O_2; and $[O_2]$ the partial pressure O_2.

$$p_i = \lambda p_a \tag{6}$$

$$\Gamma_* = \frac{[O_2]}{2\tau} \tag{7}$$

$$\Phi_{TC3} = \frac{1}{1 + e^{[0.2(10 - T_C)]}} \tag{8}$$

$$R_d = b_{C3} V_m \tag{9}$$

where p_a is the ambient partial of CO_2; λ a parameter; τ the kinetic parameter; b_{C3} = 0.015; T_c and V_m the monthly temperatures and the maximum daily rate of net photosynthesis.

For C4 photosynthesis, the functions are calculated as:

$$C1_{C4} = \Phi_{pi} \Phi_{TC4} C_{mass} \alpha_a \alpha_{C4} \tag{10}$$

$$C2_{C3} = 1 \tag{11}$$

$$\Phi_{pi} = \frac{\lambda}{\lambda_{mC4}} \tag{12}$$

$$\Phi_{TC4} = \frac{1}{(1 + e^{0.3(13 - T_C)})(1 + e^{0.3(T_C - 36)})} \tag{13}$$

where α_{C4} is the intrinsic quantum efficiency for C4 photosynthesis; Φ_{pi} the effect of reduced p_i on C4 photosynthesis; Φ_{TC4} the response of C4 plants to extreme temperature; and λ_{mC4} the value of λ that C4 plants maintain under non-water stress conditions.

The daytime assimilation rate A_{dt} is related to p_i through the CO_2 diffusion gradient between the atmosphere and intercellular air spaces.

$$g_C = g_{\min} + \frac{1.6 A_{dt}}{C_\alpha (1 - \lambda)} \tag{14}$$

where g_c is the average daytime canopy conductance; g_{min} the PFT-specific minimum canopy conductance; and C_a the ambient mole fraction of CO_2. A_{dt} is obtained from A_{nd} by the addition of nighttime respiration.

Under non-water stress conditions the maximum values of λ (λ_{mC3} and λ_{mC4} for C3 and C4 plants) allow the maximum potential photosynthesis (Eqn.1) and maximum potential conductance (g_p, Eqn.14). Water stress results in a lower canopy

conductance, in which case the water balance calculation provides the actual canopy conductance. Annual GPP is obtained by adding daytime respiration to A_{dt} and summing the result over a one year period.

1.2 Respiration

Plant respiration is calculated as:

$$R_{plant} = R_{leaf} + R_{transt} + R_{fine\_root} + R_{growth} \quad (15)$$

where R_{plant} is the total annual plant respiration cost; R_{leaf} the annual leaf respiration that is calculated in a monthly time step by the photosynthetic model; R_{transt} the annual maintenance respiration cost for transport tissues; R_{fine_root} and R_{growth} the fine root respiration and annual growth respiration; and R_{growth} the estimate of 20% of the gross photosynthesis remaining after all other respiration costs have been removed (Ryan, 1991)

$$R_{\text{transt}} = \sum_{\text{m=1}}^{\text{m=12}} K_r C_s e^{E_o f(T_c)} \quad (16)$$

$$f(T_C) = \frac{1}{T_{ref} - T_o} - \frac{1}{T_C - T_o} \quad (17)$$

$$C_S = LAI \cdot C_n \quad (18)$$

$$R_{fine\_root} = aL_f \quad (19)$$

where m is the number of the month and T_c is the mean monthly temperature; K_r the sapwood respiration rate of the reference temperature (T_{ref}) of 10ºC; E_o = 308.56K; T_o = -46.02ºC; C_s the total sapwood carbon content; LAI the leaf area index; C_n a parameter for the sapwood carbon content per unit LAI; α = 1; and L_f the total annual leaf litterfall carbon. Therefore, the net primary production (NPP) is calculated:

$$NPP = GPP - R_{plant} \quad (20)$$

1.3 Vegetation Carbon

Carbon was simulated in vegetation using an adaptation of the terrestrial biosphere carbon model by Foley (1995) and King *et al*. (1997). The steady state carbon density of each plant compartment (C_k) is:

$$C_k = NPP \cdot p_{jk} t_{jk} \quad (21)$$

where NPP is the equilibrium steady state annual net primary production; p_{jk} the

partitioning coefficient for the *k*th plant compartment of the vegetation type *j*; and t_{jk} the average lifespan of plant part *k* of vegetation type *j*. The parameters (Table 1) for each of the vegetation types of the BIOME4 model were defined by Foley (1995) and King *et al*. (1997).

1.4 Litter Carbon

Like vegetation carbon, litter carbon has an associated three components: leaf, stem, and root litter. Under equilibrium conditions, the litter production *L* for each plant *k* is:

$$L_k = \frac{C_k}{t_{jk}} \tag{22}$$

where C_k is the steady state vegetation carbon density; and t_{jk} the average lifespan of plant part *k* of vegetation type *j* (Table 1).

Carbon stored in litter type *j* (C_{lj}) is:

$$C_{lj} = \frac{L_j}{k_{lj}} \tag{23}$$

where k_{lj} is the decomposition coefficient for the litter. Each vegetation type has an associated above and below-ground litter pool. The rate of litter decay is controlled by climate (temperature and moisture) and the quality of the vegetation material (Meetenmeyer, 1978, 1984; Vogt *et al*., 1986; Dyer *et al*., 1990).

$$k_{lj} = \frac{g(T) f(W_1) \varepsilon_j}{\tau_{10}} \tag{24}$$

$$g(T) = e^{[\frac{308.56}{56.02} - (\frac{1}{T} + 273.0 - 227.13)]} \tag{25}$$

$$f(W_1) = 0.25 + 0.75 W_1 \tag{26}$$

where W_1 is the average moisture status in the upper soil layer. The empirical soil moisture relationship ($f(W_1)$) is from Foley (1995). The effect of temperature (T) on the decomposition rate is adopted from Lloyd and Taylor (1994); ε_j is the relative ability of litter types to decay (Foley, 1995). Above-ground litter decomposition is dependent on air temperature, whereas below-ground litter decomposition is related to soil temperature.

1.5 Soil Organic Carbon

Carbon pool obeys the following formula:

$$\frac{\partial C_{s,k}}{\partial t} = (1 - f_{resp}) P_k \sum_{j=1}^{2} (k_{l,j} C_{l,j} - k_{s,k} C_{s,k}) \quad (27)$$

where k is an index for the fast (k=1) and slow (k=2) pools of soil; $C_{s,k}$ the organic carbon of soil; and $k_{s,k}$ the decomposition constant. The fraction of decomposed litter carbon lost to the atmosphere is represented by f_{resp} (Table 1), with variations between biomes (Parton *et al*., 1987; 1992). The remaining carbon (1- f_{resp}) enters the organic pools of soil. P_k is the partitioning coefficients for carbon flowing into fast and slow soil pools, with approximately 98.5% and 1.5% of the remainder entering fast and slow soil pools, respectively. Under equilibrium conditions, the organic carbon of soil is:

$$C_{s,k} = \frac{(1 - f_{resp}) \sum_{j=1}^{2} k_{l,j} C_{l,j}}{k_{s,k}} \quad (28)$$

$$k_{s,k} = \frac{g(T) f(W_1)}{\tau_{10}} \quad (29)$$

The control of the climate in the turnover of soil carbon is expressed by a simple relationship between the organic decomposition rate of soil, temperature, and soil moisture (Parton *et al*., 1992).

2. Inverse process

The crudest approach is exhaustive sampling, requiring the greatest calculation time, where all points in a dense grid covering the input space are applied. This method is not recommended if the number of parameters is high. An alternative method, used for this study, is the Bayesian approach (Gelman *et al*., 1995) that uses "*a priori* information" $B(x)$ for the input parameter vector x. This information is then combined with information provided by a comparison of the model output to the observations $p = (p_i, i = 1, 2, ..., m)$ in order to define a probability distribution representing the *a posteriori* information $\beta(x)$ of the parameter vector x:

$$\beta(x) = k \cdot B(x) \cdot L(x) \quad (30)$$

where k is an appropriate normalization constant; and $L(x)$ the likelihood function that roughly measures the fit between the observed data (p_i, i =1, 2, ..., m) and the predicted data ($\hat{p}(x) = (\hat{p}_1(x),..., \hat{p}_m(x))$) by the model. If model errors are assumed to be independent and Gaussian, $L(x)$ can be written as follows:

$$L(x) = -\frac{n}{2} \log \sqrt{2\pi S^2} - \frac{1}{2S^2} \sum_{i=1}^{m} (p_i - \hat{p}_i(x))^2 \quad (31)$$

where S^2 is the estimated error variance. Monte Carlo sampling of the Markov Chains (MCMC) is the method used to calculate $\beta(x)$ for the given observational vector p. In this study, $L(x)$ is approximated by LH as follows:

$$LH(p) = -\frac{n}{2}\log\sqrt{2\pi S^2} - \frac{1}{2S^2}\sum_{i=1}^{n}(D_i) \tag{32}$$

$$D_i = \sum_{j=1}^{m}(\hat{p}_{ij} - p_{ij})^2 \tag{33}$$

$$S^2 = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{\ 2} \tag{34}$$

where p_{ij} is the jth biome score of the ith pollen sample, $\hat{p}_{ij}$ is the simulated biome score of jth to sample i. D_i measures the similarity between the observed and simulated biome scores. S^2 can be considered as the variance of the biome scores and is used to standardize the LH function. Finally, LH/n is used instead of LH to facilitate the comparison of different runs with different numbers of samples.

Let us consider a multidimensional domain where each dimension represents a parameter range, and where a vector of parameters is an element of the multidimensional domain. An iterative method of the Metropolis-Hastings (MH) algorithm is used, defined in Bayesian statistics context (Gelman *et al*., 1995), for browsing the domain of the parameters according to an acceptance-rejection rule. $x_{ACTUAL}^{(k)}$ is denoted as the actual value of the parameter vector at interaction (k) and $x_{CAND}^{(k)}$ is a candidate for the next position. If the data collection $p=(p_j, j=1, 2, \ldots, m)$ is known, the acceptance-rejection rule is based on the criteria C (Fahmy, 1997):

$$\mathrm{C}(x_{ACTUAL}^{(k)}, x_{CAND}^{k}) = \frac{P(x_{CAND}^{k}|p)}{P(x_{ACTUAL}^{(k)}|p)}$$

$$= \frac{LH_{CAND}(p)B(x_{CAND}^{k})}{LH_{ACTUAL}(p)B(x_{ACTUAL}^{k})} \tag{35}$$

where $P(./p)$ is the posterior distribution; B the prior distribution; and $LH(p)$ the LH function of the (actual/candidate) parameter vector for dataset p. If $\mathrm{C}(x_{ACTUAL}^{(k)}, x_{CAND}^{k}) > u$, then the $x_{ACTUAL}^{(k)}$ value is accepted; u is randomly chosen, without preference, between 0 and 1 at each interaction. In practice, the parameter vector is initiated according to its prior distribution. A candidate is then chosen

around the previous value using a multivariate Gaussian distribution. The variance-covariance matrix is updated at every *k* iteration with regard to the observed variance-covariance matrix of the last *k* iteration. The MH algorithm was applied to the LH function defined by Eqn.3 with a multivariate uniform distribution as a prior of the hyper-parameter.

The empirical equations by Guiot *et al*. (2000) are applied to the changes of temperature (ΔT) and precipitation (ΔP) for January and July as well as monthly sunshine (S_j).

$$\Delta T_j = \Delta T_{Jan} + (\Delta T_{Jul} - \Delta T_{Jan}) \sin[\pi(j-1)/12] \quad (36)$$

$$\Delta P_j = 1 + \Delta P_{Jan} + (\Delta P_{Jul} - \Delta P_{Jan}) \sin\left[\pi(j-1)/12\right] \quad (37)$$

$$S_j = a_j P_j + b_j T_j + c_j \quad (38)$$

where a_j and b_j are the slopes of monthly precipitation and temperature; c_j the monthly intercepts; and *j* the month from January to December. ΔT_j is applied additively and ΔP_j multiplicatively to modern values.

References:

Collatz GJ, Ball JT, Grivet C, Berry JA (1991) Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminer boundary layer. *Agricultural and Forest Meteorology*, **54**, 107-136.

Collatz GJ, Ribas-Carb M, Berry JA (1992) Coupled photosynthesis-stomatal conductance model for leaves of C4 plants. *Australian Journal of Plant Physiology*, **19**, 519-538.

Dyer ML, Meentemeyer V, Berg B (1990) Apparent controls of mass loss rate of leaf litter on a regional scale: litter quality versus climate. *Scandianvian Journal of Forest Research*, **5**, 311-323.

Fahmy T (1997) *Modélisation de la qualité bactériologique de l'eau potable et optimization des procedures de contrôle. Doctorat*. ENGREF, Paris, 224pp.

Gelman A, Carlin JB, Stein HS, Rubin DB (1995) *Bayesian Data Analysis*. Chapman and Hall, New York.

King AW, Post WM, Wullschleger SD (1997) The potental reponse of terrestrial carbon storage to changes in climate and atmospheric CO_2. *Climatic Change*, **35**, 199-227.

Lloyd J, Taylor JA (1994) On the temperature dependence of soil respiration. *Functional Ecology*, **8**, 315-323.

Meentemeyer V (1978) Macroclimate and lignin control of litter decomposition rates.

Ecology, **59**, 465-472.
Meentemeyer V (1984) The geography of organic decomposition rates. *Annals of the Association of American Geographers*, **74**, 551-560.
Parton WJ, Running S, Walker B (1992) A toy terrestrial carbon flow model. In: *Modelling the earth system boulder* (eds. Ojima D), OIES/UCAR.
Parton WJ, Schimel DS, Cole CV, Ojima DS (1987) Analysis of factors controlling soil organic matter levels in Great Plains grasslands. *Soil Science Society of American Journal*, 51, 1173-1179.
Ryan MG (1991) Effects of climatic change on plant respiration. *Ecological Application*, **1**, 157-167.
Vogt KA, Grier CC, Vogt DJ (1986) Production, turnover, and nutrient dynamics of above- and belowground detritus in world forests. *Advances in Ecological Research*, **15**, 303-377.