Global Change Biology (2010), doi: 10.1111/j.1365-2486.2010.02315.x

Diagnosing and assessing uncertainties of terrestrial ecosystem models in a multimodel ensemble experiment: 2. Carbon balance

WEILE WANG*†, JENNIFER DUNGAN†, HIROFUMI HASHIMOTO*†, ANDREW R. MICHAELIS*†, CRISTINA MILESI*†, KAZUHITO ICHII‡ and RAMAKRISHNA R. NEMANI†

Q1 *California State University, Monterey Bay, Seaside, CA, USA, †NASA Ames Research Center, Moffett Field, CA, USA, ‡Faculty of Symbiotic Systems Science, Fukushima University, Japan

Abstract

This paper examines carbon stocks and their relative balance in terrestrial ecosystems simulated by Biome-BGC, LPJ, and CASA in an ensemble model experiment conducted using the Terrestrial Observation and Prediction System. We developed the Hierarchical Framework for Diagnosing Ecosystem Models to separate the simulated biogeochemistry into a cascade of functional tiers and examine their characteristics sequentially. The analyses indicate that the simulated biomass is usually two to three times higher in Biome-BGC than LPJ or CASA. Such a discrepancy is mainly induced by differences in model parameters and algorithms that regulate the rates of biomass turnover. The mean residence time of biomass in Biome-BGC is estimated to be 40–80 years in temperate/moist climate regions, while it mostly varies between 5 and 30 years in CASA and LPJ. A large range of values is also found in the simulated soil carbon. The mean residence time of soil carbon in Biome-BGC and LPJ is $\sim$200 years in cold regions, which decreases rapidly with increases of temperature at a rate of $\sim 10\,\mathrm{yr}\,^{\circ}\mathrm{C}^{-1}$. Because long-term soil carbon pool is not simulated in CASA, its corresponding mean residence time is only about 10–20 years and less sensitive to temperature. Another key factor that influences the carbon balance of the simulated ecosystem is disturbance caused by wildfire, for which the algorithms vary among the models. Because fire emissions are balanced by net ecosystem production (NEP) at steady states, magnitudes, and spatial patterns of NEP vary significantly as well. Slight carbon imbalance may be left by the spin-up algorithm of the models, which adds uncertainty to the estimated carbon sources or sinks. Although these results are only drawn on the tested model versions, the developed methodology has potential for other model exercises.

Keywords: biogeochemistry, carbon balance, carbon cycle, model intercomparison, model uncertainty, multimodel ensemble, terrestrial ecosystem model

Received 7 May 2010 and accepted 12 July 2010

UNCORRECTED PROOF

Introduction

As vital scientific tools for investigating the terrestrial carbon/water cycles, ecosystem models synthesize our current understanding of how terrestrial ecosystems function and how they interact with climate and other environmental systems (Waring & Running, 1998). Yet because ecosystems are highly complex, it is impossible to include all component processes into a numerical model and simplifications must be made one way or another. Different simplification schemes often lead to different model behaviors, generally referred to as structural uncertainty. Inherently, structural uncertainty cannot be evaluated by experiments that use one single model; instead, experiments that use an ensemble of multiple models, or multimodel ensemble (MME), must be conducted (Tebaldi & Knutti, 2007). Ideally, if the ensemble members are independent, their structural uncertainties are expected to cancel each other, and thus the combined results will be better than those of individual member models. For this reason, MME experiments are now widely accepted in carbon-cycle studies (VEMAP Members, 1995; Cramer *et al*., 1999, Q2 2001; Sitch *et al*., 2008; Randerson *et al*., 2009).

Challenges; however, exist in how to evaluate and combine ensemble model results. Although in general models are equally valid in their algorithms and structures, they rarely perform equally well in all aspects. Therefore, it is legitimate to screen outliers while weighting more heavily models that produce more realistic final estimates. Model outputs ultimately need to be evaluated against certain reference (usually

Correspondence: W. Wang, Mail Stop 242-4, NASA Ames Research Center, Moffett Field, CA 94035, USA. tel. + 1 650 604 6444, fax + 1 650 604 6569, e-mail: weile.wang@gmail.com

© 2010 Blackwell Publishing Ltd

(BWUK GCB 2315 Webpdf=08/31/2010 06:06:10 1253768 Bytes 12 PAGES n operator=) 8/31/2010 6:08:54 PM

	GCB	2315	B	Dispatch: 31.8.10 Journal: GCB	CE: Guna
	Journal Name	Manuscript No.		Author Received: No. of pages: 12	PE: Chris/Dheiran

GCB 2315

observation-based) datasets. For instance, Randerson *et al*. (2009) presented a framework to evaluate simulated terrestrial biogeochemistry in the Carbon-Land Model Intercomparison Project (C-LAMP), in which the outputs of ecosystem models are compared with satellite datasets [leaf area index (LAI) and net primary production (NPP)], flux tower measurements (carbon/energy fluxes), and some other data sources. However, this approach is not problem-free: there are measurement uncertainties and scale mismatch issues with the flux tower data (Baldocchi, 2003), while satellite products of LAI and NPP are strictly not observations but model results. Also, since numerical parameter optimization is now a common practice in ecosystem model applications (Raupach *et al*., 2005), there are potential risks that the models may be 'tuned' to some of the reference datasets. As such, a good match between a model and the reference data does not necessarily mean that the model will perform well in other regions or in other aspects, where it may not be well constrained.

This study proposes the Hierarchical Framework for Diagnosing Ecosystem Models (HFDEM) as an alternative approach to analyze and evaluate uncertainties of ecosystem models in MME experiments. Originally developed from our analysis of Biome-BGC (Wang *et al*., 2009), HFDEM decomposes the simulated biogeochemistry into a cascade of three functional tiers and examine them sequentially. The first tier analyzes the gross and net primary production (GPP/NPP) – the influx of carbon into the ecosystem – by checking their relationships with major climate drivers as well as vegetation leaf area. The second tier analyzes how carbon is allocated into different vegetation components and how biomass is accumulated through the balance between primary production and biomass turnover. The third tier analyzes the soil carbon balance between the influx from litter fall and the effluxes generated by heterotrophic respiration (HR). As such, HFDEM covers most important aspects of the simulated carbon cycle. Because the internal logic of the models is simplified with this framework, it not only helps evaluate structural differences between ecosystem models, but also how these differences originate.

We applied HFDEM to analyze and evaluate uncertainties among four public-domain ecosystem models in an MME experiment conducted using the Terrestrial Observation and Prediction System (TOPS; Nemani *et al*., 2009). The results are reported in a pair of papers: the companion paper (Wang *et al*., 2010; hereafter referred as Part 1) describes the experimental setup, the developed methodology, and reports the results centered on the primary production; this paper focuses on the simulated carbon pools [biomass and soil organic matters (SOM)] as well as the net ecosystem exchange (NEE). Below, we give a brief introduction to the model experiment and summarize the key findings of the companion paper.

Model experiment setup

We set up our model experiment with the TOPS, a data and modeling system developed at NASA Ames Research Center to perform ecological monitoring/forecasting and related research (Nemani *et al*., 2009). TOPS integrates various datasets from different sources, processes them into compatible formats, and uses them to drive application models. One of the key components of TOPS, the Surface Observation and Gridding System (SOGS), ingests daily meteorological observations from station networks and interpolates them to spatially continuous grids using the algorithms developed in Thornton *et al*. (1997) and Jolly *et al*. (2005).

For this study, we used SOGS to generate daily climate variables for the entire North American continent at 8 km resolution and for the time period of 1982–2006. The generated meteorological fields include minimum/maximum temperatures, precipitation, vapor pressure deficit, and incident solar radiation. The source data of these variables were obtained from the Global Summary of the Day (GSOD) and the Cooperative Summary of the Day (TD3200), both archived at the National Climatic Data Center (NCDC). When combined, the two datasets have about 8000 reporting stations over the continent at daily time steps. We used the natural-neighbor procedure (Sibson, 1981) to interpolate the station data to the defined domain. Climatologies of the generated annual mean temperature and total precipitation are shown in Part 1.

For models requiring satellite measured vegetation indices, we used a satellite dataset of LAI developed by Ganguly *et al*. (2008), which is generated from the multidecade records of the Advanced Very High Resolution Radiometer (AVHRR) sensors but with comparable quality to the Moderate Resolution Imaging Spectroradiometer (MODIS) LAI products. The MODIS land cover product (Friedl *et al*., 2002) is used to describe the distribution of plant functional types (PFTs). A global dataset of land surface parameters, ECOCLIMAP (Masson *et al*., 2003), is used to specify soil property (e.g., texture and depth) and other parameters (e.g., albedo).

Four public-domain ecosystem models are tested in this experiment: Biome-BGC (Thornton *et al*., 2002), LPJ (Sitch *et al*., 2003), CASA (Potter *et al*., 1993), and TOPS-BGC (White & Nemani, 2004). These models cover a range of different approaches of simulating ecological processes and have different complexity. In particular, Biome-BGC (version 4.1.2) and LPJ (version 3.1.1) fully

UNCORRECTED PROOF

(BWUK GCB 2315 Webpdf:=08/31/2010 06:06:10 1253768 Bytes 12 PAGES n operator=) 8/31/2010 6:08:54 PM

simulate the carbon cycle of terrestrial ecosystems such that they are able to project the states of ecosystems solely from climate/environmental forcing and specified initial conditions. On the other hand, CASA (version 4.2.2) and TOPS-BGC (version 1.0.2) do not simulate but require external inputs (usually satellite measured vegetation indices) to estimate structural variables of the vegetation canopy (e.g., LAI or FPAR). CASA does simulate the carbon flow of NPP into various biomass compartments and then into litter/soil pools, while TOPS-BGC does not. Characteristics of these models are summarized in Table 1 of Part 1 (for convenience, the same table is also given in the Supporting Information).

Spin-up runs are used to bring the states of the above models into equilibrium with the provided climate and ancillary datasets. This stage of the model experiment does not consider climate change scenarios; instead, a constant level of atmospheric CO_2 at 360 ppm is assumed in all the spin-up runs. No disturbances are externally specified in the simulations; however, fire is internally simulated by Biome-BGC and LPJ. Using the endpoints of the spin-up runs as initial conditions, we ran the models for the climate years from 1982 to 2006 as the normal model experiment. Standard model outputs include monthly carbon fluxes (e.g., GPP, NPP, NEP, and respiration), annual averaged carbon stocks (biomass and soil carbon pools), and annual maximum LAI (in the case of Biome-BGC and LPJ). For LPJ, because it allows multiple PFTs in a grid cell, the results represent the sum of all PFTs within the grid cells.

Method of diagnosis

In order to examine the internal logic of a model, we use diagrams that relate the model outputs with their major drivers. Because most ecological processes are driven by climate, climate-domain diagrams of carbon fluxes are frequently used in this study. A climate-domain diagram organizes model results by their climate context instead of their geographic location. Owing to their dominant importance to terrestrial ecosystems, temperature and precipitation are usually chosen as the two dimensions to define the climate domain (e.g., Churkina & Running, 1998; Schloss *et al.*, 1999).

The climate domain adopted in this study spans across annual mean temperature between −10 and 30 °C (as the *x*-axis) and annual total precipitation from 0 to 2000 mm yr^{-1} (as the *y*-axis; referenced as the *T–P* domain hereafter), which covers the climate conditions of most vegetated areas over North America. To generate 2D diagrams, we divide the *T–P* domain into 100 × 100 bins, each having a 'climate resolution' of 0.4 °C and 20 mm yr^{-1}. Geographic grid cells are projected into these climate bins according to their annual temperature and precipitation. Because the geographically defined (latitude–longitude) grid is not an equal-area projection, results from grid cells are weighted by the cosine of their latitude before they are summed or averaged in the *T–P* domain.

Additional features can be shown on the climate domain diagrams. To facilitate analysis, a boundary [The 'dry-moist' boundary adopted here is based on the Köppen climate classification system for the case where precipitation is evenly distributed throughout the year (Strahler & Strahler, 1978). For cases of precipitation concentrated in summer or winter, the offset of the boundary changes to 280 or 0 mm, respectively.] between dry and moist climate conditions is shown in the *T–P* diagrams that is defined by the line of $P = 20T + 140$ (where *T* is in °C and *P* is in millimeters). Another line we found helpful is the upward ('lower/left-to-upper/right') diagonal of the *T–P* diagrams, defined by $P = 50T + 500$ (Fig. 1). As described in Part 1, this line largely represents (local) optimal climate conditions for vegetation primary production in our model experiments. In addition, considering that forests and nonforests (grasses, crops, and shrubs) can have very different characteristics, we sometimes show *T–P* diagrams for these two types of PFTs separately.

In addition to the climate-domain diagrams, another tool adopted in this study is the density scatter plot, which is essentially a scatter plot with a color scheme to indicate the density of the points. It is mainly used to illustrate the relationship between two variables, for instance, leaf mass and total biomass.

Results and discussion

In Part 1, we report analyses of the simulated primary production (GPP/NPP) from the four tested models. The results indicate an overall agreement among the models. GPP tends to be optimal in climate regions where the relationship between annual temperature and precipitation is defined by $P = 50T + 500$. Along this line (and for temperatures below 20 °C), GPP typically increases by $\sim$70 g C yr^{-1} m^{-2} for forests and $\sim$30 g C yr^{-1} m^{-2} for nonforests per 1 °C increase in temperature (or 50 mm increase in precipitation). These broad patterns reflect the dominant control of climate variables on ecosystem production. In addition, linear relationships are found between annual GPP and annual mean LAI, suggesting that the latter is a good indicator of vegetation production. The patterns of NPP largely follow those of GPP, and the ratio between the two is close to 50% on average.

On the other hand, there are also considerable differences in the simulated GPP/NPP. For instance, Biome-BGC has the highest GPP for forests but the lowest GPP for nonforests among the tested models, while the corresponding contrast in GPP is much smaller in the other models. A major source of such differences is traced back to how leaf area (LAI) is simulated in the models. For Biome-BGC and LPJ, in particular, carbon assimilated through photosynthesis (NPP) is ultimately allocated to develop more leaves

UNCORRECTED PROOF

(BWUK GCB 2315 Webpdf:=08/31/2010 06:06:10 1253768 Bytes 12 PAGES n operator=) 8/31/2010 6:08:54 PM

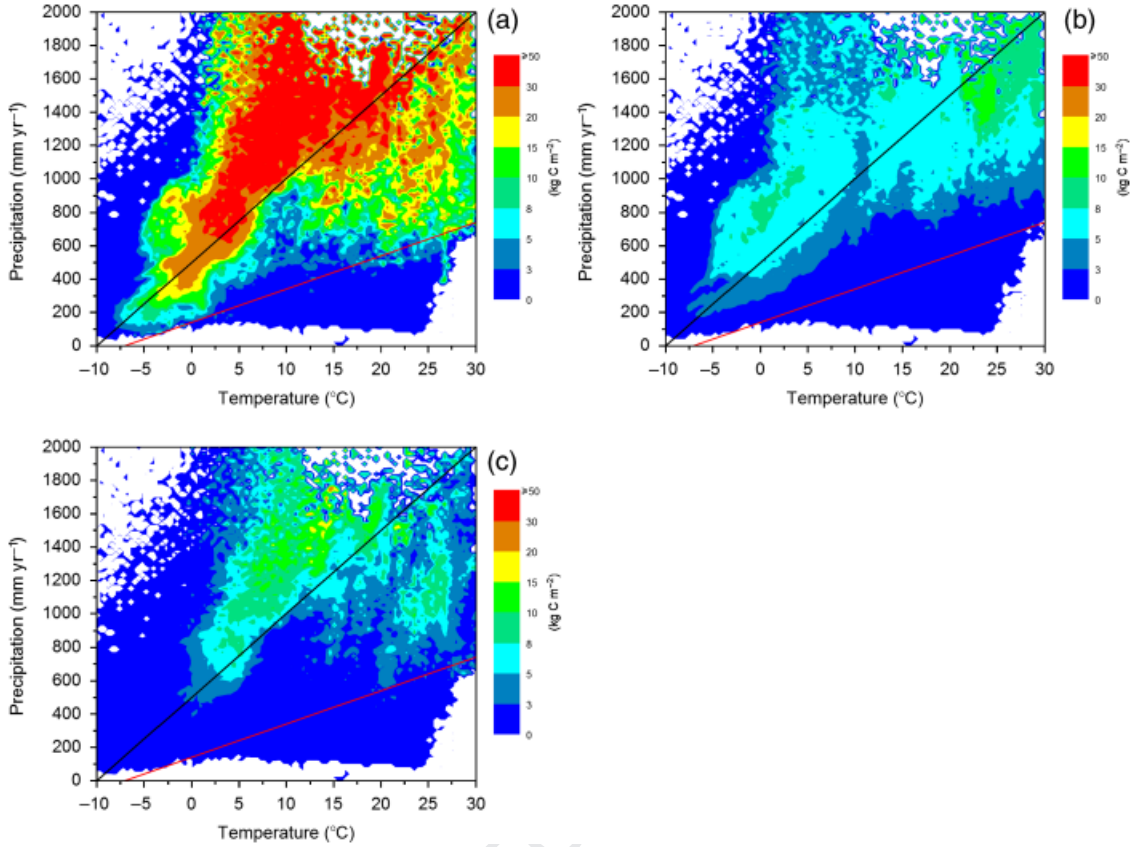

Fig. 1 Distributions of biomass (kg C m^{-2}) in the *T*–*P* domain: (a) Biome-BGC; (b) LPJ; and (c) CASA.

(and supporting tissues), and thus to enhance GPP/NPP subsequently. This constructs a positive feedback between primary production and leaf growth. Different strategies to constrain such a feedback lead to distinctive characteristics between the models (Part 1).

In this paper, we examine how carbon is allocated to and accumulated in biomass and how it is cycled through other carbon pools. Because TOPS-BGC does not simulate these processes, our discussions will focus on the remaining three models.

Biomass

The distributions of the simulated biomass in the *T*–*P* domain are shown in Fig. 1. For each model, higher biomass is generally found in climate regions along the diagonal of $P = 50T + 500$, which represents favorable climate conditions for primary production (Part 1). The centers of the distributions are slightly shifted towards regions with more rainfall, a result induced by the fact that in moist environments forests are more predominant than nonforests. Comparing the distribution patterns of biomass (Fig. 1) with those of NPP (Fig. S1) indicates some broad similarities, reflecting the fact that NPP is the input for biomass growth; however, such correlation is not always strong, for the accumulation process is also affected by other factors.

Despite the above stated similarities, however, the most striking feature of Fig. 1 is the different (absolute) magnitudes of biomass among the models. In particular, Biome-BGC is found to simulate much higher biomass than LPJ and CASA. In temperate (0–20 °C) and moist ($P \geq 50T + 500$) climate regions, biomass density in Biome-BGC is often above 30 kg C m^{-2} and sometimes even as high as 50 kg C m^{-2}; in contrast, the biomass density in LPJ and CASA is mostly below 10 kg C m^{-2} and rarely reaches 15 kg C m^{-2}. As such, the overall biomass simulated by Biome-BGC is about two to three times higher than that simulated by LPJ or CASA. For the latter two models, though their overall biomass magnitudes are similar, CASA simulates higher biomass in temperate regions while LPJ has higher biomass in the tropics (Fig. 1).

Because NPP is largely comparable among the three models (Fig. SA1), it is an unlikely cause for the discrepancy in biomass. To verify this point, we calculate ratios between biomass and corresponding annual NPP for all vegetated grid cells and represent them in

© 2010 Blackwell Publishing Ltd, *Global Change Biology*, doi: 10.1111/j.1365-2486.2010.02315.x

Colour

(BWUK GCB 2315 Webpdf:=08/31/2010 06:06:10 1253768 Bytes 12 PAGES n operator=) 8/31/2010 6:08:54 PM

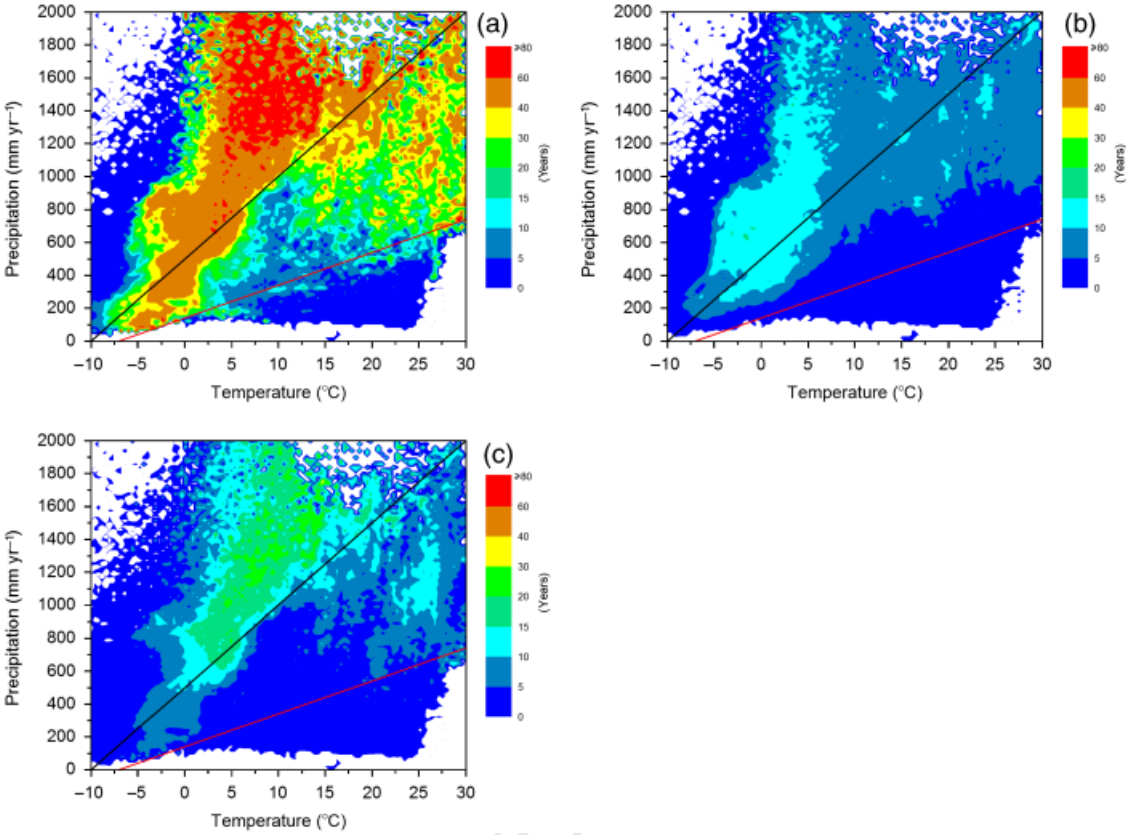

Fig. 2 Mean residence time of biomass (years), as estimated by the ratios between biomass and annual net primary production (NPP), in the *T*–*P* domain: (a) Biome-BGC; (b) LPJ; and (c) CASA.

the *T*–*P* domain (Fig. 2). Distributions of the biomass-to-NPP ratios have similar patterns as those of the biomass, with (relatively) high ratios located along (and above) $P = 50T + 500$. The magnitudes of the ratios are also very different among the models. In temperate and moist climate regions, the biomass-to-NPP ratios in Biome-BGC are generally between 40 and 80; the corresponding ratios in CASA range between 10 and 30, while those in LPJ are between 5 and 15 (Fig. 2). Therefore, when NPP is the same, Biome-BGC will have biomass three or four times as much as in LPJ or CASA, consistent with the results of Fig. 1.

The rate at which biomass is accumulated is determined by two factors: the gain of carbon as NPP and the loss of carbon through tissue turnover or whole plant mortality. At model equilibrium, the carbon influx must be balanced with the efflux so that biomass becomes steady. Therefore, steady-state NPP also reflects the total turnover (including mortality) rate of biomass. Following this line of thinking, the biomass-to-NPP ratios of Fig. 2 indeed indicate the mean residence time (or the '*e*-folding' time) of biomass in these models. As shown, the mean residence time of biomass in Biome-BGC is mostly between 40 and 80 years in temperate/moist climate regions, while it ranges from 10 to 30 and 5 to 15 years over the same regions in CASA and LPJ, respectively. Because biomass persists for a much longer time in Biome-BGC, it can accumulate to a much higher amount than in CASA or LPJ. Therefore, differences in model parameters and algorithms that regulate tissue turnover rates and plant mortality are mainly responsible for the discrepancies of biomass in Fig. 1.

Allometry

In addition to the amount of total biomass, we are also interested in how biomass is allocated to different plant components (e.g., leaves, stems, and roots). Because the definition of vegetation components varies among different models, here we check one particular aspect of the modeled allometry, namely, the relationship between the leaf and the whole plant. The results are illustrated in density scatter plots in Fig. 3. The *T*–*P* diagrams of annual maximum leaf mass (herein referred to as simply 'leaf mass' unless otherwise specified) are given in the Supporting Information (Fig. SA2).

© 2010 Blackwell Publishing Ltd, *Global Change Biology*, doi: 10.1111/j.1365-2486.2010.02315.x

Colour
UNCORREC
(BWUK GCB 2315 Webpdf:=08/31/2010 06:06:10 1253768 Bytes 12 PAGES n operator=) 8/31/2010 6:08:54 PM

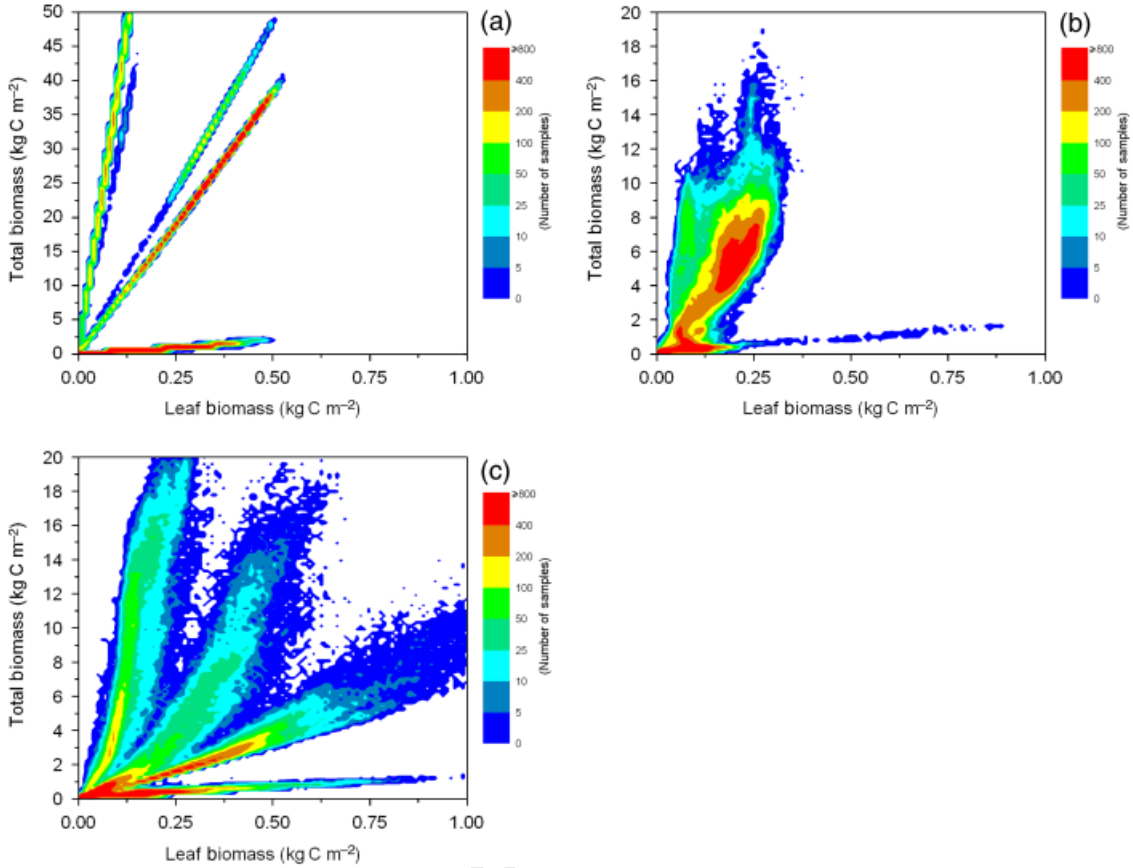

Fig. 3 Density scatter plots between annual maximum leaf mass (kg C m^{-2}) and total biomass (kg C m^{-2}): (a) Biome-BGC; (b) LPJ; and (c) CASA. The color scheme shows the number of samples represented by each pixel of the scatter plot (i.e., the density).

Overall, all the models show strong linear relationships between leaf mass and total biomass; in addition, such relationships are often comprised of multiple components that represent characteristics of different vegetation types (PFTs). The scatter plot of Biome-BGC (Fig. 3a) shows five nearly perfect linear components: two of them have the plant-to-leaf biomass ratios (i.e., the slopes) around 400, two components have the ratios around 100, and the last component has a ratio around 5. These components can be identified as representing deciduous (broadleaf/needleleaf) forests, evergreen (broadleaf/needleleaf) forests, and nonforest species (grasses and shrubs), respectively. The scatter plot from LPJ (Fig. 3b) shows two components: the one representing forests has a plant-to-leaf ratio about 50–100, and the one for nonforests has a ratio about 2. The scatter plot from CASA (Fig. 3c) shows four components that have plant-to-leaf biomass ratios around 80, 30, 10, and 1, respectively. The first three components of CASA represent forests in different regions: in particular, they represent temperate deciduous forests (mainly east-coast forests), tropical evergreen forests, and boreal forests (including marine west-coast forests), respectively. The last component represents herbaceous PFTs.

The range of leaf mass also varies between different vegetation types and/or models. For instance, leaf mass of deciduous forests in Biome-BGC is mostly below 0.12 kg C m^{-2}, whereas leaf mass of evergreen forests reaches 0.50 kg C m^{-2}. This difference is induced by the fact that the leaf mass density (leaf mass per unit area, or LMA) of deciduous species is usually about 1/3–1/2 of the evergreen species (White *et al*. 2000). When both leaf mass and the plant-to-leaf biomass ratio are considered, therefore, the total biomass density of deciduous forests in Biome-BGC is comparable to that of evergreen forests. In LPJ, the leaf mass of forests is generally below 0.30 kg C m^{-2}, which likely represents a mixture of deciduous and evergreen species. Leaf mass of nonforests in LPJ is mostly below 0.20 kg C m^{-2}; however, it can occasionally reach as high as 0.9 kg C m^{-2}. In CASA, leaf mass of deciduous forests and evergreen forests is generally below 0.25 and 0.50 kg C m^{-2}, respectively. However, the marine west-coast forests (the third component shown in Fig. 3c) can

© 2010 Blackwell Publishing Ltd, *Global Change Biology*, doi: 10.1111/j.1365-2486.2010.02315.x

Colour
Q3
(BWUK GCB 2315 Webpdf:=08/31/2010 06:06:10 1253768 Bytes 12 PAGES n operator=) 8/31/2010 6:08:54 PM

have leaf mass as high as 1.0 kg C m^{-2}, accounting for about 10% of their total biomass. Also, although the leaf mass of nonforests is under 0.5 kg C m^{-2} most of the time, it reaches 1.0 kg C m^{-2} in some situations.

Characteristics visible in of Fig. 3 may be explained by the biomass allocation and turnover algorithms of the models. In Biome-BGC, the allometric allocation scheme is defined by constant ratios relative to leaf growth; and the tissue turnover rates and various mortalities are also prescribed by constant parameters. Therefore, at system equilibrium the mass balance between leaf and other plant tissues is analytically guaranteed to be linear (Wang *et al*., 2009). In LPJ, in contrast, the allocation relationships are determined by solving a set of joint scaling equations that relate leaf area, canopy crown area, sapwood area, stem diameter, and tree heights together. In particular, the model constrains the growth of trees by imposing an upper limit on individual tree sizes or tree densities in a stand, and plants are forced to shed tissues or to die when their growth exceeds such thresholds. Such self-thinning processes induce a dynamical component to the overall biomass turnover rate of forests, and thus their corresponding allometric relationships are more dispersed (Fig. 3b). On the other hand, nonforests in LPJ are not subjected to such self-thinning processes. Therefore, their allometric relationships are more linear and their leaf mass (and leaf area) can become much higher than forests under certain conditions.

The carbon allocation algorithm in this version of CASA is based on local availability of resources such as light, water, and nutrients (Friedlingstein *et al*., 1999). Our model experiment shows that the simulated east-coast deciduous forests are mainly limited by solar radiation, and thus more carbon is allocated to stem; in contrast, the simulated boreal forests are mostly limited by soil water and nutrients, and thus more carbon is allocated to roots (Friedlingstein *et al*., 1999). These environmental factors cause forests in these regions to have different plant-to-leaf ratios (Fig. 3c). Another feature of CASA is that its leaf carbon is calculated independently from the input LAI. Therefore, one cannot relate the two by the means of LMA. For instance, marine west-coast forests can have leaf mass up to 1.0 kg C m^{-2}, twice as much as the peak leaf mass of boreal forests (Fig. 3c and Fig. SA2c), even though there is no substantial difference in LAI between the two regions (Part 1).

Soil carbon

Because SOM undergoes chemical degradation at different rates, in ecosystem models it is generally represented by cascading carbon pools that may have residence times ranging from a couple of years to thousands of years (Parton *et al*., 1987). Because the structure of SOM pools can vary significantly among different models, we only analyze the total SOM as a whole, referred to as 'soil carbon' for simplicity.

Figure 4 shows the *T*–*P* diagrams of the simulated soil carbon. For all models, (relatively) high magnitudes of soil carbon are found in cold and moist regions, reflecting its dependence on biological (e.g., biomass) and environmental factors (temperature, in particular). In Biome-BGC and LPJ, soil carbon reaches 30–40 kg C m^{-2} in regions along $P = 50T + 500$ with $T < 5\,^{\circ}$C; it rapidly decreases to < 10 kg C m^{-2} when T is above 10 °C (Fig. 4a and b). The link between soil carbon and biomass is more evident in CASA, in which highest soil carbon are found in temperate regions with temperature around 10 °C and precipitation around 1400 mm yr^{-1} (Fig. 4c), where biomass density is also the highest (Fig. 2c). However, the highest magnitude of soil carbon in CASA is only about 10 kg C m^{-2}, remarkably lower than that of Biome-BGC and LPJ.

Because NPP approximates the rate of litter generation at system equilibrium, we calculate the mass ratios between soil carbon and NPP in order to separate the influences of environmental factors on SOM dynamics from those of litter (Fig. 5). As in the case of biomass-to-NPP ratios, the SOM-to-NPP ratios also reflect the mean residence time of soil carbon in corresponding climate conditions. As shown, the SOM-to-NPP ratios in Biome-BGC and LPJ reach as high as 150–200 when T is below −5 °C, but become mostly below 20 when T is above 10 °C. The changes of the ratios (as indicated by the contour lines) are almost perpendicular to the axis of temperature, further confirming the dominant regulation of temperature on SOM decomposition. In these two models, therefore, organic matter can reside in the soil for centuries in cold climate regions, but it only takes a couple of decades for it to decompose in temperate/warm regions; the gradient (related to temperature) of the mean residence time of soil carbon is about −10 yr °C^{-1} on average.

In contrast to Biome-BGC and LPJ, the SOM-to-NPP ratios in CASA range between 10 and 20 in regions where T is between −10 and 5 °C; they decrease to below 10 when T is above 10 °C. Therefore, the mean residence time of soil carbon in CASA is much shorter than in the other two models. In addition, the gradient of the residence time is also much lower (about −1 yr °C^{-1}), suggesting the turnover of SOM is less sensitive to variations in temperature (Fig. 5c).

The above discrepancy between CASA and the other two models mainly reflects their different strategies in modeling soil carbon processes. In particular, soil carbon pools in CASA are only designed to describe SOM in the

UNCORRECTED PROOF

(BWUK GCB 2315 Webpdf:=08/31/2010 06:06:10 1253768 Bytes 12 PAGES n operator=) 8/31/2010 6:08:54 PM

© 2010 Blackwell Publishing Ltd, *Global Change Biology*, doi: 10.1111/j.1365-2486.2010.02315.x

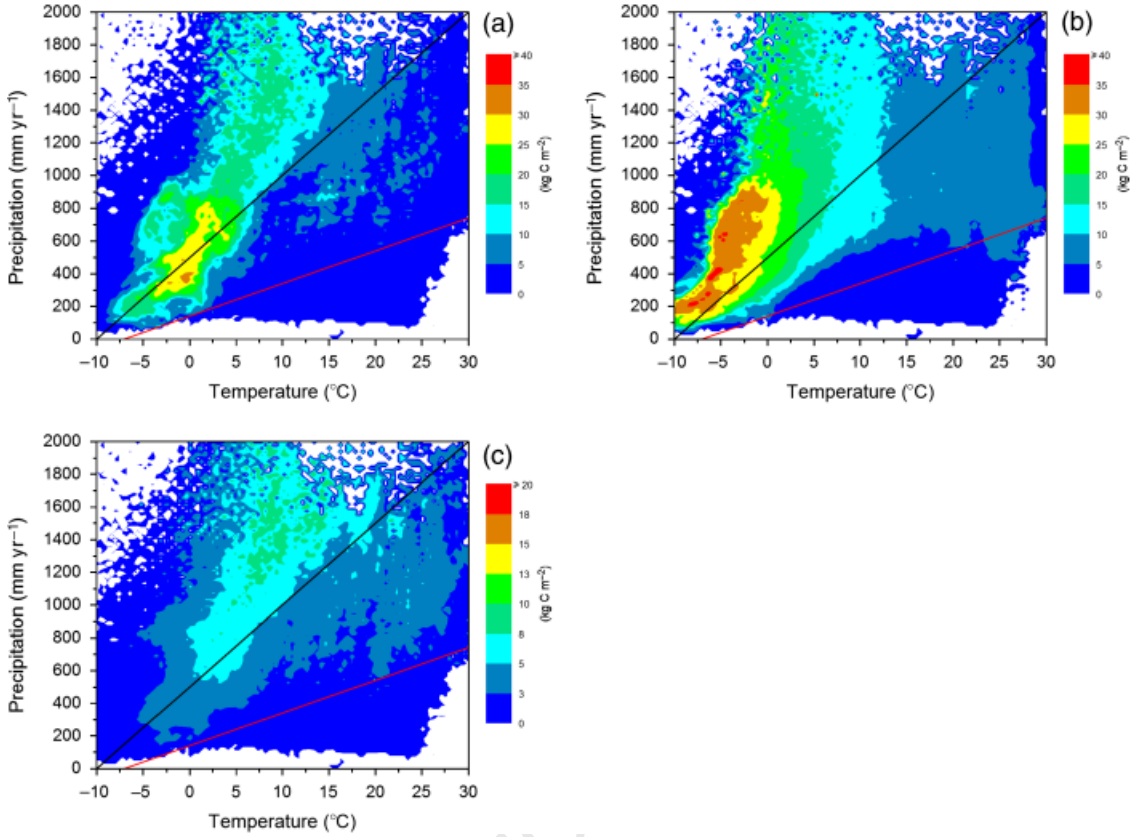

Fig. 4 Distributions of soil carbon (kg C m^{-2}) in the *T*–*P* domain: (a) Biome-BGC; (b) LPJ; and (c) CASA.

upper 0.3 m of soil; the cycling of carbon in the deeper soil is thus ignored (Potter *et al.*, 1993). On the other hand, both Biome-BGC and LPJ include the full set of SOM pools (Thornton *et al.*, 2002; Sitch *et al.*, 2003). As such, soil carbon simulated by CASA is not directly comparable to that simulated by Biome-BGC and LPJ. Nevertheless, by including CASA in the comparison, we simply want to demonstrate that such model structural differences can be captured by our diagnostic methods.

Fire emission and NEP

In ecosystem models, carbon fluxes released through the degradation of litter and SOM are summed up as HR, and the difference between NPP and HR represents net ecosystem production (NEP). When there are no other nonrespiratory CO_2 fluxes (e.g., fire emissions) simulated, NEP is the same as NEE but opposite in sign. This corresponds to the case of CASA in our experiment. In Biome-BGC and LPJ, on the other hand, fire is internally simulated, and thus NEE is defined by fire emissions minus NEP. At system equilibrium, the climatology of annual NEE should be zero, and the comparison of mean annual NEE among the models is thus trivial. Instead, we briefly compare NEP and fire emissions between Biome-BGC and LPJ.

Figure 6 shows the simulated fire emissions in the *T*–*P* domain. For Biome-BGC, highest fire emissions are found in temperate ($0\,°C < T < 10\,°C$) and moist (above $P = 50T + 500$) regions, where fire burns up 40–60 g C m^{-2} yr^{-1}; in the rest of the domain, they are mostly 20–40 g C m^{-2} yr^{-1} along $P = 50T + 500$, but gradually decrease to below 10 g C m^{-2} yr^{-1} in hot/dry or cold/wet conditions that are unfavorable for vegetation growth. Comparing the distribution patterns of fire emissions (Fig. 6a) with those of biomass (Fig. 1a) indicates a clear association between the two fields. Indeed, Biome-BGC assumes that for each PFT, fire occurs uniformly at a prescribed rate and with the same severity (vegetation experiences complete mortality when there is a fire). Therefore, fire emission in Biome-BGC is proportional to the amount of biomass. It can be inferred from Figs 1a and 6a that the rate of fire in Biome-BGC is about once in a century on average – note that the fire frequency can vary significantly between different PFTs. For nonforests like grasslands, for instance, fire is much more frequent and occurs once in a decade or so (Part 1).

© 2010 Blackwell Publishing Ltd, *Global Change Biology*, doi: 10.1111/j.1365-2486.2010.02315.x

Colour
UNCORRECTED
(BWUK GCB 2315 Webpdf:=08/31/2010 06:06:10 1253768 Bytes 12 PAGES n operator=) 8/31/2010 6:08:54 PM

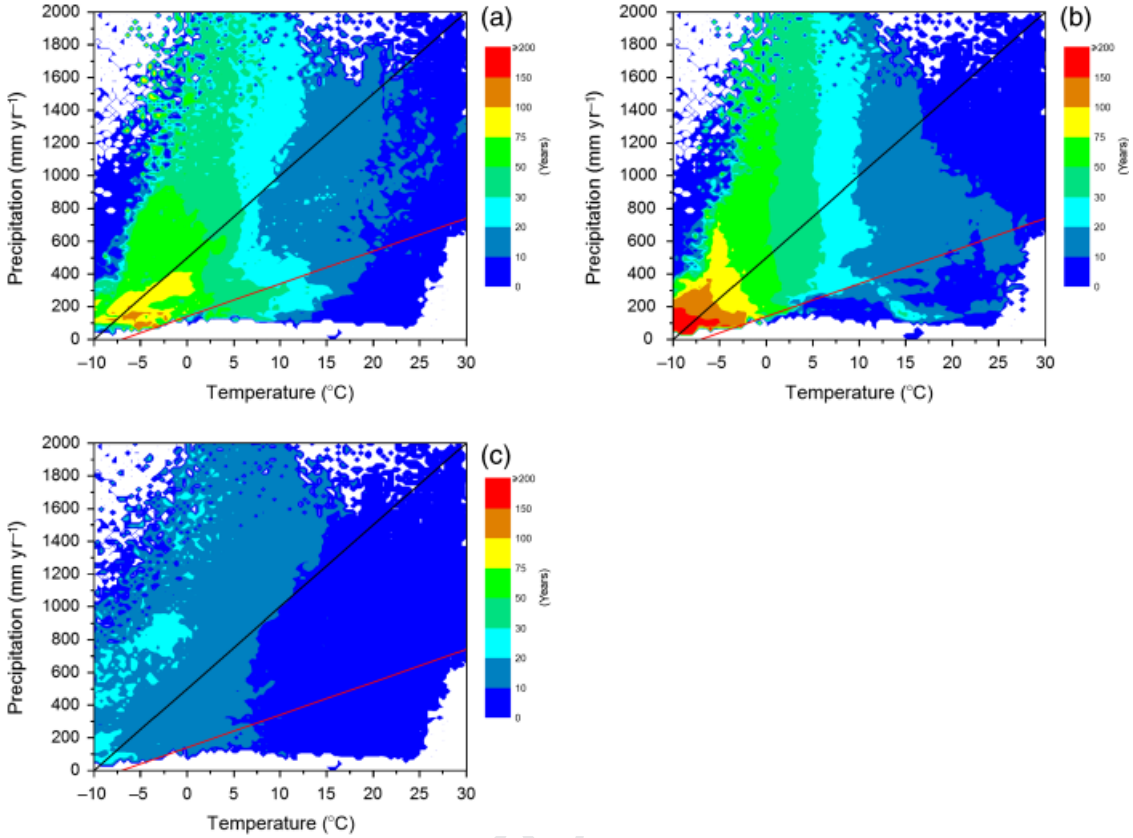

Fig. 5 Mean residence time of soil carbon (years), as estimated by ratios between soil carbon and annual mean net primary production (NPP), in the *T–P* domain: (a) Biome-BGC; (b) LPJ; and (c) CASA.

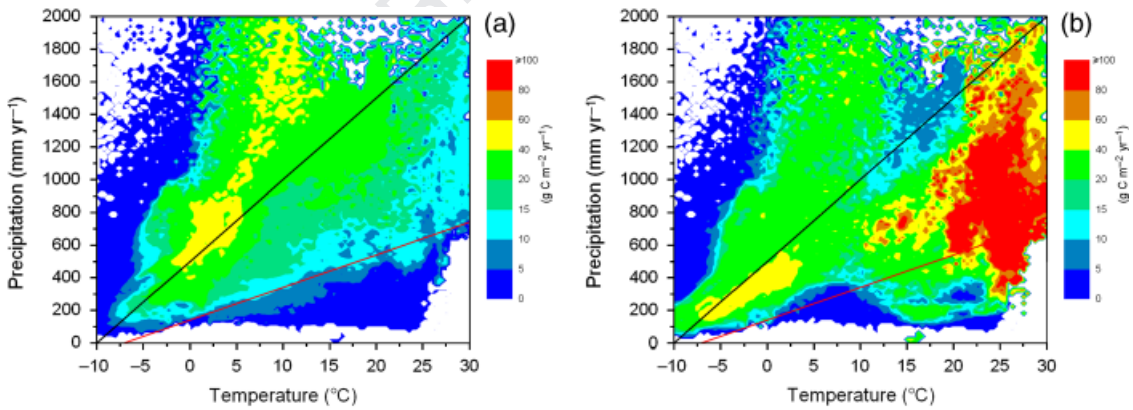

Fig. 6 Annual fire emissions (g C m^{-2} yr^{-1}) in the *T–P* domain: (a) Biome-BGC; (b) LPJ.

Compared with Biome-BGC, the fire algorithm in LPJ is more complex and calculates the probability and the extent of fire based on fuel load and litter moisture (Thonicke *et al.*, 2001; Sitch *et al.*, 2003). As a result, climatic regulations of fire emissions are more evident on the *T–P* diagram of LPJ (Fig. 6b). For instance, fire emissions are mostly below 40 g C m^{-2} yr^{-1} above the diagonal of $P = 50T + 500$; in contrast, they easily exceed 40 g C m^{-2} yr^{-1} under the diagonal and become dominantly 80–100 g C m^{-2} yr^{-1} when temperature is above 20 °C. Therefore, the impact of fire on the carbon cycle is proportionally much stronger in hot and (relatively) dry regions. Because the biomass density (fuel load) in LPJ is far less than in Biome-BGC (Fig. 1b), the fact that fire emissions in LPJ are comparable to or even higher than in Biome-BGC indicates that the fire rate in

© 2010 Blackwell Publishing Ltd, *Global Change Biology*, doi: 10.1111/j.1365-2486.2010.02315.x

Colour
Colour
Q4
(BWUK GCB 2315 Webpdf=08/31/2010 06:06:10 1253768 Bytes 12 PAGES n operator=) 8/31/2010 6:08:54 PM

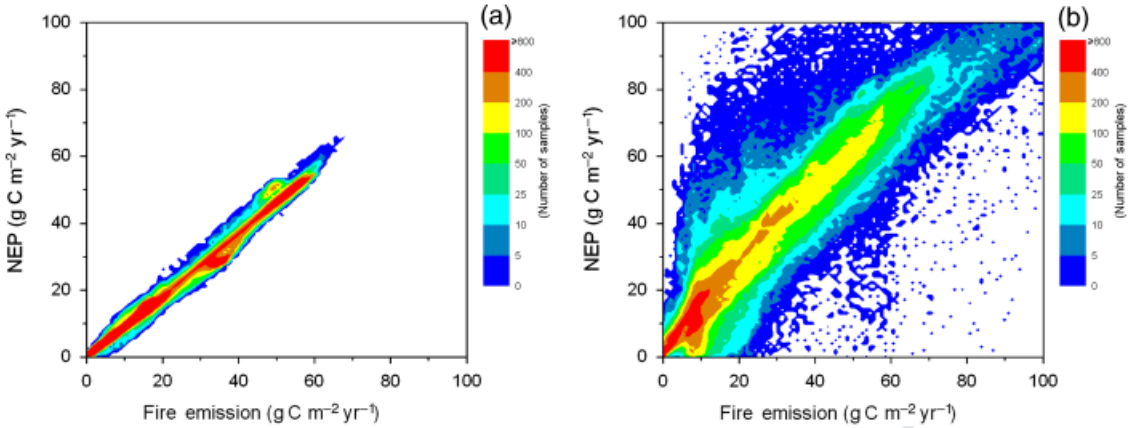

Fig. 7 Density scatter plots between annual fire emissions and NEP: (a) Biome-BGC; (b) LPJ. The color scheme shows the number of samples represented by each pixel of the scatter plot (i.e., the density).

LPJ is much higher than in Biome-BGC on average. In regions where annual mean temperature is above 20 °C, for instance, it may occur as frequently as once in every 5–10 years (Figs 1b and 6b).

Because both models are supposed to be at system equilibrium, fire emissions should be balanced by NEP. Two density scatter plots (Fig. 7) illustrate the balance or lack thereof. For Biome-BGC (Fig. 7a), almost all the points fall on the 1 : 1 line, indicating that the model does reach its equilibrium at most grid cells. On the other hand, the scatter plot for LPJ (Fig. 7b) is more dispersed, and its center is slightly shifted towards higher NEP than fire emissions. These features indicate that the model has not strictly reached its equilibrium, but is slightly sequestering carbon from the atmosphere. Note that this residual imbalance exists after the model is spun-up for one thousand more climate years than recommended by the default model settings (Part 1). It suggests that the spin-up algorithm of LPJ may need to be strengthened. At the same time, users of other models should also be aware of this issue, because it adds additional uncertainties to the simulated NEE.

It should be recognized that the results presented in this study are only drawn on our experiment of the tested model versions, and thus may not reflect the latest development of these models. Also, the tested model parameters were not calibrated particularly for applications in the experiment domain. Therefore, these results do not reflect the optimal performance and skills of these models in simulating carbon dynamics over North America.

Conclusions

This study reports results from an MME experiment designed to diagnose and assess structural uncertainties among terrestrial ecosystem models. The experiment was conducted under the TOPS, and the tested models include public versions of Biome-BGC, LPJ, TOPS-BGC, and CASA, all driven by gridded climate fields generated from observations of weather-station networks.

To systematically analyze the outputs from the model experiments, we developed a diagnostic framework, the HDFEMs, which breaks the simulated carbon cycle into three functional tiers (i.e., primary production, biomass, and soil carbon pools) and examines them sequentially. The results of the analysis are reported in two parts. This paper, being the second part, examines how carbon is allocated and accumulated in biomass and how it is cycled through soil carbon pools.

Comparing biomass simulated by the models indicates some similarity in the distribution patterns but highlights substantial differences in their magnitudes. In particular, biomass simulated by Biome-BGC is usually two to three times higher than that simulated by LPJ or CASA, despite the fact that NPP is largely comparable among the models. Such discrepant biomass magnitudes are mainly caused by different residence time of biomass in the models. For instance, the mean residence time of biomass in Biome-BGC is mostly 40–80 years in temperate/moist climate regions, whereas those in CASA and LPJ range 5–30 years. Therefore, biomass in Biome-BGC accumulates to higher amounts.

In addition to the total magnitudes, we also examined the allometric relationships of biomass in different plant components. Overall, strong linear relationships are found between leaf mass and total biomass, which are often comprised of multiple components that represent different vegetation types. However, the characteristics of these relationships vary among the models. In Biome-BGC, the relationships for all vegetation types

UNCORRECTED PROOF

Colour

(BWUK GCB 2315 Webpdf:=08/31/2010 06:06:10 1253768 Bytes 12 PAGES n operator=) 8/31/2010 6:08:54 PM

© 2010 Blackwell Publishing Ltd, *Global Change Biology*, doi: 10.1111/j.1365-2486.2010.02315.x

are strictly linear so that the knowledge of leaf mass can be used to accurately estimate the total biomass (and vice versa). In LPJ, the relationship for woody species becomes more dispersed while it remains very linear for nonwoody species. These different characteristics reflect differences in model parameters and algorithms to regulate carbon allocation and biomass turnover in these models, as discussed in Part 1.

Similarly, a large range is found in the simulated soil pools, where carbon stocks simulated by CASA are significantly less than those of Biome-BGC and LPJ. Also, such difference is mainly induced by differences in dynamic characteristics of soil carbon pools. The mean residence time of soil carbon in Biome-BGC and LPJ is around 200 years in regions where annual mean temperature is below $-5\,^{\circ}\mathrm{C}$; it decreases with increases in temperature at a rate of about $-10\,\mathrm{yr}\,^{\circ}\mathrm{C}^{-1}$ on average. In contrast, the mean residence time of soil carbon in CASA is only about 10–20 years, which is also less sensitive to changes in temperature. The reason for such apparent discrepancy is that long-term soil carbon pools are not simulated in CASA. We show that model structural differences of this kind can be readily captured by the developed diagnosing framework.

Dynamic characteristics of carbon pools are determined by model parameters (e.g., turnover rates and mortality) as well as ecological processes that regulate these variables. For soil carbon pools, temperature appears to be the most important climate regulator in all the models. For biomass; however, the regulating mechanism differs significantly between models, as reflected in the ratios between leaf mass and the total biomass of the tested models. In Biome-BGC, carbon allocation and turnover is based on prescribed allometric relationships and is constrained solely by nitrogen availability; therefore, each PFT has an almost perfect linear relationship between leaf mass and total biomass. In LPJ, on the other hand, the turnover rates and mortality of forests are also regulated by tree sizes and stand densities (e.g., the 'self-thinning' process); therefore, the leaf-to-biomass ratios are more dispersed in LPJ for forests. On the other hand, because such a constraint is not imposed on nonforests in LPJ, the relationship between leaf mass and biomass is much more linear. For the same reason, leaf mass (and leaf area) of nonforests in LPJ may grow far above those of forests in some conditions.

Disturbances such as wildfire are a key factor that influences the carbon balance of the simulated ecosystem. Fire occurs uniformly in Biome-BGC with a prescribed rate and a certain severity for each PFT; therefore, carbon emissions from fires are largely proportional to the simulated biomass. In contrast, fire frequency and fire severity are modeled as functions of soil moisture and biomass (the fuel load) in LPJ, and therefore, its fire emissions are significantly higher in warm and dry regions. When these models are in steady state, the simulated fire emissions are supposed to be balanced by NEP of the ecosystem. However, depending on how the models are spun-up, there may be slight imbalance in the models' states. Such residual imbalance, if not addressed properly, introduces additional uncertainty to the estimated carbon sources or sinks.

Overall, the results of this analysis indicate that the tested models are more consistent in estimating primary production than simulating biomass and soil carbon pools. Growing observation networks and satellite sensors allow for systematic measurements of photosynthetic carbon fluxes and related climatic/ecological variables and result in improved direct or indirect calibration of the photosynthesis algorithms and parameters in these models. On the other hand, consistent and systematic observations of biomass and soil carbon stocks are much harder to acquire, and our knowledge about the regulating processes are poor. Therefore, it is not surprising that the models differ more in these aspects. However, the revealed large uncertainties also suggest that the corresponding algorithms and the interactions among them may not be sufficiently analyzed and tested in these models. As such, they justify the necessity of a systematic diagnostic framework like the one developed in this study (i.e., HDFEM). Indeed, HDFEM can also be applied to analyze observational datasets and thus facilitate model-data comparisons. We shall report the new developments of the analysis in future publications.

Acknowledgements

The dataset GSOD is from NOAA NCDC (http://www.ncdc.noaa.gov). The dataset TD3200 is from the Research Data Archive (RDA), which is maintained by the Computational and Information Systems Laboratory (CISL) at the National Center for Atmospheric Research (NCAR). NCAR is sponsored by the National Science Foundation (NSF). The original data are available from the RDA (http://dss.ucar.edu) in dataset number ds510.0. This research was funded by grants from the NASA Earth Science program. We thank two anonymous reviewers for their insightful comments.

References

Baldocchi DD (2003) Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. *Global Change Biology*, **9**, 479–492.

Churkina G, Running SW (1998) Contrasting climatic controls on the estimated productivity of global terrestrial biomes. *Ecosystems*, **1**, 206–215.

Cramer W, Bondeau A, Woodward FI *et al*. (2001) Global response of terrestrial ecosystem structure and function to CO_2 and climate change: results from six dynamic global vegetation models. *Global Change Biology*, **7**, 357–373.

Cramer W, Kicklighter DW, Bondeau A *et al*. (1999) Comparing global models of terrestrial net primary productivity (NPP): overview and key results. *Global Change Biology*, **5** (Suppl. 1), 1–15.

Friedl MA, McIver DK, Hodges JCF *et al*. (2002) Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, **83**, 287–302.

UNCORRECTED PROOF

(BWUK GCB 2315 Webpdf:=08/31/2010 06:06:10 1253768 Bytes 12 PAGES n operator=) 8/31/2010 6:08:54 PM

© 2010 Blackwell Publishing Ltd, *Global Change Biology*, doi: 10.1111/j.1365-2486.2010.02315.x

Friedlingstein P, Joel G, Field CB, Fung IY (1999) Toward an allocation scheme for global terrestrial carbon models. *Global Change Biology*, **5**, 755–770.

Ganguly S, Mitchell AS, Samanta A *et al*. (2008) Generating vegetation leaf area index earth system data records from multiple sensors. Part 1: theory. *Remote Sensing of Environment*, **112**, 4333–4343.

Jolly WM, Graham JM, Michaelis A, Nemani RR, Running SW (2005) A flexible, integrated system for generating meteorological surfaces derived from point sources across multiple geographic scales. *Environmental Modeling and Software*, **20**, 873–882.

Masson V, Champeaux J-L, Chauvin F, Meriguet C, Lacaze R (2003) A global database of land surface parameters at 1-km resolution in meteorological and climate models. *Journal of Climate*, **16**, 1261–1282.

Nemani R, Hashimoto H, Votava P *et al*. (2009) Monitoring and forecasting protected area ecosystem dynamics using the terrestrial observation and prediction system (TOPS). *Remote Sensing of Environment*, **113**, 1497–1509.

Parton WJ, Schimel DS, Cole CV, Ojima DS (1987) Analysis of factors controlling soil organic matter levels in the Great Plains grasslands. *Soil Society of America Journal*, **51**, 1173–1179.

Potter CS, Randerson JT, Field CB, Matson PA, Vitousek PM, Mooney HA, Klooster SA (1993) Terrestrial ecosystem production: a process model based on global satellite and surface data. *Global Biogeochemical Cycles*, **7**, 811–841.

Randerson JT, Hoffman FM, Thornton PE *et al*. (2009) Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models. *Global Change Biology*, **15**, 2462–2484.

Raupach MR, Rayner PJ, Barrett DJ *et al*. (2005) Model-data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications. *Global Change Biology*, **11**, 378–397.

Schloss AL, Kicklighter DW, Kaduk J, Wittenberg Uthe participants of the Potsdam NPP Model Intercomparison (1999) Comparing global models of terrestrial net primary productivity (NPP): comparison of NPP to climate and the normalized difference vegetation index. *Global Change Biology*, **5** (Suppl. 1), 25–34.

Sibson R (1981) A brief description of natural neighbor interpolation. In: *Interpreting Multivariate Data* (ed. Barnett V), pp. 21–36 John Wiley, New York.

Sitch S, Smith B, Prentice IC *et al*. (2003) Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology*, **9**, 161–185.

Sitch S, Huntingford C, Gedney N *et al*. (2008) Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five dynamic global vegetation models (DGVMs). *Global Change Biology*, **14**, 2015–2039.

Strahler AN, Strahler AH (1978) *Modern Physical Geography*. John Wiley & Sons, New York.

Tebaldi C, Knutti R (2007) The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society A*, **365**, 2053–2075.

Thonicke K, Venevsky S, Sitch S, Cramer W (2001) The role of fire disturbance for global vegetation dynamics: coupling fire into a dynamic global vegetation model. *Global Ecology and Biogeography*, **10**, 661–677.

Thornton PE, Running SW, White MA (1997) Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal of Hydrology*, **190**, 214–251.

Thornton PE, Law BE, Gholz HL *et al*. (2002) Modeling and measuring the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests. *Agricultural and Forest Meteorology*, **113**, 185–222.

VEMAP Members (1995) Vegetation/ecosystem modeling and analysis project (VEMAP): comparing biogeography and biogeochemistry models in a continental-scale study of terrestrial ecosystem responses to climate change and CO_2 doubling. *Global Biogeochemical Cycles*, **9**, 407–437.

Wang W, Ichii K, Hashimoto H, Michaelis AR, Thornton PE, Law BE, Nemani RR (2009) A hierarchical analysis of terrestrial ecosystem model Biome-BGC: equilibrium analysis and model calibration. *Ecological Modelling*, **220**, 2009–2023.

Wang W, Dungan JD, Hashimoto H, Michaelis AR, Milesi C, Ichii K, Nemani RR (2010) Diagnosing and assessing uncertainties of terrestrial ecosystem models in multi-model ensemble experiment: 1. Primary production. *Global Change Biology* (in review). Q5

Waring RH, Running SW (1998) *Forest Ecosystems Analysis at Multiple Scales*, 2nd edn. Academic Press, New York.

White MA, Nemani RR (2004) Soil water forecasting in the continental United States: relative forcing by meteorology versus leaf area index and the effects of meteorological forecast errors. *Canadian Journal of Remote Sensing*, **30**, 717–730.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure SA1. Annual mean NPP ($g\,C\,m^{-2}\,yr^{-1}$) in the T-P domain: (a) Biome-BGC; (b) LPJ; and (c) CASA.

Figure SA2. Annual maximum leaf mass ($g\,C\,m^{-2}$) in the T-P domain: (a) Biome-BGC; (b) LPJ; and (c) CASA.

Table SA1. Selected characteristics and attributes of models used in the MME.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

UNCORRECTED PROOF

(BWUK GCB 2315 Webpdf:=08/31/2010 06:06:10 1253768 Bytes 12 PAGES n operator=) 8/31/2010 6:08:54 PM

© 2010 Blackwell Publishing Ltd, *Global Change Biology*, doi: 10.1111/j.1365-2486.2010.02315.x