New Phytologist

Confronting model predictions of carbon fluxes with measurements of Amazon forests subjected to experimental drought

Thomas L. Powell[1], David R. Galbraith[2,3], Bradley O. Christoffersen[4], Anna Harper[5,6], Hewlley M. A. Imbuzeiro[7], Lucy Rowland[8], Samuel Almeida[9], Paulo M. Brando[10], Antonio Carlos Lola da Costa[11], Marcos Heil Costa[7], Naomi M. Levine[1], Yadvinder Malhi[3], Scott R. Saleska[4], Eleneide Sotta[12], Mathew Williams[8], Patrick Meir[8] and Paul R. Moorcroft[1]

[1]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA; [2]School of Geography, University of Leeds, Leeds, LS2 9JT, UK; [3]Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford, OX1 3QY, UK; [4]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA; [5]College of Engineering, Mathematics, and Physical Science, University of Exeter, Exeter, EX4 4QF, UK; [6]Department of Atmospheric Science, Colorado State University, Fort Collins, CO 80523, USA; [7]Grupo de Pesquisas em Interação Atmosfera-Biosfera, Universidade Federal de Viçosa, Viçosa, CEP 36570-000, Minas Gerias, Brazil; [8]School of GeoSciences, University of Edinburgh, Edinburgh, EH8 9XP, UK; [9]Museu Paraense Emilio Goeldi, Belém, CEP 66077-530, Pará, Brazil; [10]Instituto de Pesquisa Ambiental da Amazônia, CEP 71503-505, Brasília, Distrito Federal, Brazil; [11]Centro de Geociências, Universidade Federal do Pará, Belém, CEP 66017-970, Pará, Brazil; [12]Embrapa Amapá, Macapá, CEP 68903-419, Amapá, Brazil

Author for correspondence:
Paul R. Moorcroft
Tel: +1 617 496 6744
Email: paul_moorcroft@harvard.edu

Received: *12 February 2013*
Accepted: *20 May 2013*

New Phytologist (2013) **200:** 350–364
doi: 10.1111/nph.12390

Key words: Amazon, carbon cycle, drought, terrestrial biosphere model, throughfall exclusion, tropical rainforest.

Summary

- Considerable uncertainty surrounds the fate of Amazon rainforests in response to climate change.
- Here, carbon (C) flux predictions of five terrestrial biosphere models (Community Land Model version 3.5 (CLM3.5), Ecosystem Demography model version 2.1 (ED2), Integrated BIosphere Simulator version 2.6.4 (IBIS), Joint UK Land Environment Simulator version 2.1 (JULES) and Simple Biosphere model version 3 (SiB3)) and a hydrodynamic terrestrial ecosystem model (the Soil–Plant–Atmosphere (SPA) model) were evaluated against measurements from two large-scale Amazon drought experiments.
- Model predictions agreed with the observed C fluxes in the control plots of both experiments, but poorly replicated the responses to the drought treatments. Most notably, with the exception of ED2, the models predicted negligible reductions in aboveground biomass in response to the drought treatments, which was in contrast to an observed *c*. 20% reduction at both sites. For ED2, the timing of the decline in aboveground biomass was accurate, but the magnitude was too high for one site and too low for the other.
- Three key findings indicate critical areas for future research and model development. First, the models predicted declines in autotrophic respiration under prolonged drought in contrast to measured increases at one of the sites. Secondly, models lacking a phenological response to drought introduced bias in the sensitivity of canopy productivity and respiration to drought. Thirdly, the phenomenological water-stress functions used by the terrestrial biosphere models to represent the effects of soil moisture on stomatal conductance yielded unrealistic diurnal and seasonal responses to drought.

Introduction

Changes in precipitation patterns are projected to be one of the biggest consequences for the Amazon rainforest as global climate change intensifies over this century. Predicted shifts in precipitation include: an increase in the frequency of extremely wet or dry months (Lintner *et al.*, 2012), regional increases or decreases in dry season length and intensity (Malhi *et al.*, 2008; Costa & Pires, 2010; Good *et al.*, 2013; Joetzjer *et al.*, 2013), and either increased precipitation or chronic drying across large regions of the basin (Cox *et al.*, 2000; Li *et al.*, 2006). Although the spatial and temporal patterns of predicted shifts in precipitation vary considerably between climate models (Jupp *et al.*, 2010), there is increasing consensus toward drying and longer dry seasons (Joetzjer *et al.*, 2013). However, it is presently unclear how resilient forests in different regions will be to a drier climate.

Process-based terrestrial biosphere models are key tools for assessing ecosystem resilience to climate change because of their ability to mechanistically predict ecosystem responses to novel environmental conditions. However, it is unclear whether current

model formulations can accurately capture the impacts of chronic drought on Amazon forest ecosystems. Several modeling studies have been conducted to evaluate the importance of different ecosystem processes and physiological mechanisms that either modify Amazon carbon fluxes or confer tolerance during periods of water stress (e.g. Fisher *et al.*, 2007, 2010; Baker *et al.*, 2008; Sakaguchi *et al.*, 2011). However, these studies examined the predictions of single models using different meteorological forcing data and different representations of the soil properties; therefore, it is difficult to understand how contrasting terrestrial biosphere formulations and parameterizations affect predictions of the responses of different Amazon forests to severe water limitation. In this analysis, we performed a detailed evaluation of the ability of five terrestrial biosphere models (Community Land Model version 3.5 (CLM3.5), Ecosystem Demography model version 2.1 (ED2), Integrated BIosphere Simulator version 2.6.4 (IBIS), Joint UK Land Environment Simulator version 2.1 (JULES) and Simple Biosphere model version 3 (SiB3)) and a site-specific ecosystem model (the Soil–Plant–Atmosphere (SPA) model) to correctly capture the effects of water limitation on carbon fluxes of two Amazon forests. Model predictions were compared against observations from the two throughfall exclusion (TFE) drought experiments located in the Caixuanã (CAX) and Tapajós (TNF) National Forests in the eastern Brazilian Amazon (Nepstad *et al.*, 2002; Fisher *et al.*, 2007; Meir *et al.*, 2009). All model simulations used standardized initial spin-up conditions, soil physics, and local meteorological forcings.

The TFE drought experiments are ideal for evaluating the model-specific soil water-stress responses as they have prevented *c.* 50% of the precipitation from entering the soil without altering atmospheric conditions (Nepstad *et al.*, 2002). Moreover, the TFEs serve as useful benchmarks for vegetation models as the simulated droughts cover a broader range of drying than is currently predicted by most climate models, thus ensuring conservative parameterizations. Also, unlike glasshouse-based drought manipulations, they evaluate ecosystem-level drought responses. Although soil type and water table depth are considerably different between the two sites (see the Materials and Methods section), the two drought experiments had similar responses involving reductions in wood production and elevated mortality of dominant trees in the treatment plots (Nepstad *et al.*, 2007; Brando *et al.*, 2008; da Costa *et al.*, 2010).

In this study, the carbon dynamics of the TNF and CAX forests were simulated under observed precipitation (0% reduction), and three drought levels, classified as substantial (30%), severe (50%, also TFE treatment level), and catastrophic (80%) reductions in precipitation. The six models were evaluated for their ability to predict reported carbon fluxes at the control and treatment levels. The models were compared to determine the level of agreement in the timing and magnitude of the response of ecosystem carbon fluxes to different levels of drought. Finally, the various formulations associated with soil water stress were evaluated to determine the dominant mechanisms necessary for inclusion in terrestrial ecosystem models in order to provide useful information about the fate of the Amazon rainforest under future climate change.

Materials and Methods

Study sites

Two TFE experiments were initiated in the Tapajós (TNF; 2.897ºS, 54.952ºW) and Caxiuanã (CAX; 1.737ºS, 51.458ºW) National Forests, Pará, Brazil, in 1999 and 2001, respectively, to assess whole-ecosystem responses to drought. Mean annual precipitation at TNF is 2000 mm (Nepstad *et al.*, 2002) with a wet season from December to mid-June, while at CAX mean annual precipitation is 2272 mm (Fisher *et al.*, 2007) with a wet season from December to mid-July. Except for the below-average rainfall at TNF during the 2003 wet season, precipitation rates at the two sites during the experiments were typical (Supporting Information Fig. S1) (Rosolem *et al.*, 2008). The soils at both sites are Oxisols, but they differ in texture and depth: TNF is comprised of 60% clay and 38% sand with no hardpan layers in the top 12 m and a water table at > 80 m (Nepstad *et al.*, 2002). CAX is 15% clay and 78% sand with a stony/laterite layer 3–4 m deep, and a water table at *c.* 10 m during the wet season (Fisher *et al.*, 2007).

Aboveground biomass at the beginning of the TNF drought experiment – estimated for trees ≥ 10 cm diameter at breast height (dbh) using allometric equations from Chambers *et al.* (2001) – was *c.* 15.0 kg C m^{-2} (Nepstad *et al.*, 2002). The TNF plot had a relatively rough canopy that ranged from 18 to 40 m in height with some emergent trees reaching 55 m. Aboveground biomass (trees ≥ 10 cm dbh) at the beginning of the CAX drought experiment – estimated using the average of eight published allometric equations – was *c.* 21.4 kg C m^{-2} (see table 1 in da Costa *et al.*, 2010). The CAX canopy was comparatively smooth with a mean height of 30 m.

A brief description of the experimental designs is given in Notes S3, and they are described in greater detail in Nepstad *et al.*, (2002) and Fisher *et al.*, (2007). The observations against which the models are evaluated are listed in Table 2 and descriptions of the measurement methodologies can be found in Notes S3.

Model descriptions

Five terrestrial biosphere models and one terrestrial ecosystem model were analyzed in this study. All the models had been parameterized before this study. The five biosphere models CLM3.5, ED2, IBIS, JULES and SiB3 used existing regional or global-scale parameterizations, while the SPA model had been parameterized for the CAX site as part of an earlier study (see references in Tables 1, S3, S4). The definitions of model variables and parameters are given in Table S2. A brief description of the model formulations relevant for understanding how the modeled ecosystem responds to drought is given below. The terrestrial biosphere models represented the biological response to water stress with schemes relating to atmospheric demand and soil moisture supply. The former regulates carbon assimilation and evapotranspiration through biophysical processes that link stomatal conductance to atmospheric humidity and the surface energy budget

Table 1 Summary of the six models in this study

Model name	Dynamic vegetation	Hydrodynamic	Canopy layers	Reference
CLM3.5 Community Land Model version 3.5 Dynamic Global Vegetation Model	Yes	No	2 : 1 sun and 1 shade	Bonan *et al.* (2003); Levis *et al.* (2004); Oleson *et al.* (2008)
ED2 Ecosystem Demography model version 2.1 (rv76)	Yes	No	Spatially variable	Medvigy *et al.* (2009)
IBIS Integrated BIosphere Simulatorversion 2.6.4	Yes	No	2 : 1 sun and 1 shade	Foley *et al.* (1996); Kucharik *et al.* (2000)
JULES Joint UK Land Environment Simulator version 2.1	Yes	No	10	Best *et al.* (2011); Clark *et al.* (2011)
SiB3 Simple Biosphere model version 3	No	No	1	Sellers *et al.* (1996); Baker *et al.* (2008)
SPA Soil–Plant–Atmosphere model	No	Yes	3 layers, each with sun and shade	Williams *et al.* (1996, 2005); Fisher *et al.* (2007)

Table 2 List of observations and associated references

Definition	Symbol	Units	Site and source
Aboveground biomass	AGB	kg C m^{-2}	CAX: da Costa *et al.* (2010)TNF: Brando *et al.* (2008)
Gross primary production of carbon	GPP	kg C m^{-2} yr^{-1}	TNF: Hutyra *et al.* (2007)
Leaf area index	LAI	m^2 m^{-2}	CAX: Fisher *et al.* (2007)TNF: Nepstad & Moutinho (2008)
Litter production		kg C m^{-2} yr^{-1}	TNF: Brando *et al.* (2008)
Net ecosystem production of carbon	NEP	kg C m^{-2} yr^{-1}	TNF: Hutyra *et al.* (2007)
Net primary production of carbon in wood	NPP_w	kg C m^{-2} yr^{-1}	CAX: da Costa *et al.* (2010)TNF: Brando *et al.* (2008)
Autotrophic respiration	R_a	kg C m^{-2} yr^{-1}	CAX: Metcalfe *et al.* (2010a)
Whole-ecosystem respiration	R_e	kg C m^{-2} yr^{-1}	CAX: Metcalfe *et al.* (2010a) TNF: Hutyra *et al.* (2007)
Heterotrophic respiration	R_h	kg C m^{-2} yr^{-1}	CAX: Metcalfe *et al.* (2010a)
Leaf respiration	R_{lf}	kg C m^{-2} yr^{-1}	CAX: Metcalfe *et al.* (2010a,b)
Root respiration	R_r	kg C m^{-2} yr^{-1}	CAX: Metcalfe *et al.* (2010a)
Soil respiration	R_s	kg C m^{-2} yr^{-1}	CAX: Sotta *et al.* (2007); Metcalfe *et al.* (2010a)TNF: Davidson *et al.* (2008)
Wood respiration	R_w	kg C m^{-2} yr^{-1}	CAX: Metcalfe *et al.* (2010a)

CAX, Caixuanã; TNF, Tapajós.

(Farquhar *et al.*, 1980; Farquhar & Sharkey, 1982; Collatz *et al.*, 1991), while the latter regulates responses to soil water stress and differed across all the models.

Four of the terrestrial biosphere models, CLM3.5, IBIS, JULES and SiB3, are so-called 'big leaf' models in which the plant canopy is horizontally aggregated (see references in Table 1). In SiB3, the composition of the plant canopy is prescribed as a single plant functional type (PFT) with a single canopy layer parameterized for tropical trees. In CLM3.5, IBIS and JULES, the canopy is comprised of different PFTs competing for available resources within the grid cell and the relative success of each PFT determines its fractional coverage. CLM3.5 and IBIS have two canopy layers, parameterized for sun and shade leaves, while JULES has 10 canopy layers, each with its own nitrogen content that both increases with height and is used to scale the apparent maximum photosynthetic rate (V_{cmax}). In CLM3.5, IBIS, JULES and SiB3, leaf photosynthesis declines as soil water stress increases through cumulative distribution functions of available soil moisture (see Table S3 and references therein). The specific functions vary between the models, yielding differing sensitivities of the plant canopy to decreasing soil moisture.

ED2 differs from the other models by being explicitly formulated at the scale of individual plants and using a system of size- and age-structured partial differential equations to dynamically track the horizontally and vertically heterogeneous ensemble of individual trees growing within a grid cell (see references in Table 1). In addition, ED2 has three tropical tree PFTs (defined as early-, mid- and late-successional tree species) that differ in their photosynthesis, water use, energy exchange, carbon allocation, and mortality (see Moorcroft *et al.*, 2001; Medvigy *et al.*, 2009; M. Longo, unpublished). Reflecting its individual-based nature, ED2 also explicitly represents mortality as a process distinct from other forms of tissue turnover, with *per capita* mortality rates varying as a function of the carbon balance of the individual plants. Soil water stress scales the maximum photosynthetic rate through a soil moisture supply versus transpiration demand function (Table S3). The demand function includes a leaf biomass term that allows for leaf drop when soil moisture becomes limiting.

The SPA model (see references in Table 1) is a terrestrial ecosystem model that uses a hydrodynamic formulation to mechanistically simulate changes in water potential and storage from the soil, through the stem, to leaves in each canopy layer. SPA

simulated a single tropical PFT with three canopy layers, each partitioned into an average sun or shade leaf fraction. Compared with the terrestrial biosphere models, SPA's hydrodynamic formulation more mechanistically represents the genesis of plant water stress. However, it has a much simpler formulation of carbon fluxes, assuming, for example, that autotrophic respiration (R_a) was a fixed fraction of gross primary production (GPP) (Table S4). To facilitate comparison with the other models, SPA's R_a was subdivided into leaf (R_{lf}), root (R_r) and wood (R_w) respiration components by using the C : N ratios reported in Williams *et al.* (2002) to estimate the relative size of each respiring pool, with 10% of the wood pool assumed to be actively respiring.

Simulation protocol and meteorological drivers

In order to isolate modeled biological responses to drought, the physical representation of the soil was standardized across all the models (see Notes S4). The models were run off-line using site-level meteorological measurements made above the forest canopy at nearby weather stations. The TNF meteorological measurements covered 2002–2004 (N. Restrepo-Coupe, unpublished data) and were recycled sequentially over the 8-yr simulation period from 1999 to 2006. The CAX meteorological measurements covered the entire 2001–2008 simulation period (da Costa *et al.*, 2010). Shortwave radiation was split into 68% direct and 32% diffuse and then further split into 43% visible and 57% near-IR for direct, and 52% visible and 48% near-IR for diffuse (Goudriaan, 1977).

All model simulations followed a standardized initialization, spin-up, and drought simulation protocol. The models were initialized with a near-bare-ground initial condition and then forced with sequentially recycled site-level meteorological drivers until aboveground biomass and soil carbon pools reached equilibrium under preindustrial atmospheric CO_2 concentrations (278 ppm). The models were then brought up to present-day atmospheric CO_2 concentrations (380 ppm) following the exponential increase in CO_2 since 1750. The drought simulations were initialized from these spin-ups. The simulations were run with one baseline year followed by 7 yr of reduced precipitation for each site (TNF: baseline1999, TFE 2000–2006 and CAX: baseline 2001, TFE 2002–2008). The site years were selected to coincide with the actual TFE experiments. During the experimental periods, precipitation was reduced by 30, 50 or 80%. These are denoted throughout this paper as d30, d50 and d80, and with d0 identifying the control simulation. Consistent with the field experiment protocols, precipitation in the TNF simulations was reduced only during the wet season, while CAX precipitation was reduced all year, and the other meteorological variables (e.g. humidity) were not manipulated.

Data analysis and presentation

The most detailed carbon accounting for either experiment is reported for the fourth year (2005) at CAX (see table 2 of Metcalfe *et al.*, 2010a), and so modeled whole-ecosystem (R_e) and component respiratory (R) fluxes were evaluated against these reported values. Similar data were not available for TNF; however, soil respiration (R_s) (Davidson *et al.*, 2008) and net ecosystem production (NEP), GPP and R_e from a neighboring flux tower site (Hutyra *et al.*, 2007) were available for the fourth year (2003).

In the plots, the effect of drought on carbon fluxes is presented as the change (Δ) in magnitude with respect to the control (d0) simulation, with negative Δ values indicating reductions in flux caused by drought. The ensemble median values of the model simulations were calculated rather than mean values so as not to over-weight individual model outliers. In the time series plots, 95% confidence intervals (CIs) are shown for the ensemble means of the models and are a measure of model agreement about the associated flux. The 95% CI is not, however, a measure of whether or not the process is represented correctly; the latter is evaluated through the data–model comparison. Observations are given as within-plot means with 95% CIs indicating spatial variability. Errors were propagated by summing in quadrature absolute errors for addition and subtraction and relative errors for products and quotients (Taylor, 1997), assuming that reported observation errors were independent and random.

Results

As this analysis is principally concerned with agreement between the predictions of the individual models and the observations at the two drought experiments, model ensemble median and mean predictions with 95% CIs are shown, but are not emphasized. The overall ecosystem responses to the two drought treatments predicted by the models are summarized in Fig. 1, which shows the dynamics of aboveground biomass (AGB) (see also Table S1). As indicated by the AGB dynamics of the control (d0) simulations, the CLM3.5, ED2, IBIS and JULES terrestrial biosphere models predicted equilibrium AGB values similar to the observed value of 14.2 kg C m^2 at TNF (Fig. 1a). At CAX, the terrestrial biosphere models had higher AGB predictions, yet all were lower than an observed value of 21.4 kg C m^2 (Fig. 1b). SiB3 did not track AGB. By contrast, the SPA ecosystem model prediction for the CAX control plot was close to the observed value, but at TNF it was more than double the observed value.

The models exhibited divergent predictions of AGB under increasing drought treatment levels (Figs 1c,d, S2a,b, Table S1). The observed reductions in AGB at the d50 treatment level at TNF and CAX were −3.2 and −3.8 kg C m^2; however, CLM3.5, IBIS, JULES and SPA all exhibited little or no response at either site (Fig. 1c,d, Table S1). By contrast, ED2 predicted a marked reduction in AGB at TNF (−5.4 kg C m^2), and a small reduction at CAX (−1.4 kg C m^2). Although the magnitude of ED2's predicted decrease in AGB at TNF was too large, the timing of its predicted decrease in the third year of the drought treatment agreed with the observations. At the d80 treatment level, both CLM3.5 and ED2 predicted almost a complete loss of AGB, while IBIS, JULES and SPA still predicted only marginal losses in AGB by the end of the experiment (Fig. S2a,b).

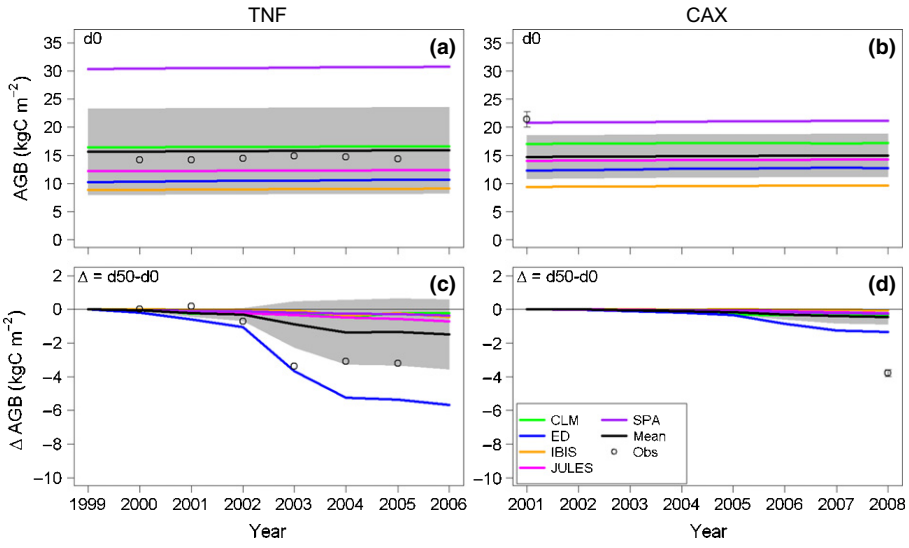

Fig. 1 Annual aboveground biomass (AGB; kg C m^{-2}) predicted for Tapajós (TNF; left side) and Caxiuanã (CAX; right side) National Forests. Colored lines are individual model predictions and the black line is the five-model ensemble mean. The shaded area is the 95% confidence interval (CI) of the models. Open symbols are published observations (mean ± 95% CI (when reported); TNF: Brando *et al.*, 2008; CAX: da Costa *et al.*, 2010). d0 and d50 are drought levels indicating a 0 and 50% reduction in precipitation, respectively. The Δ plots (c, d) show the amount the control (a, b) AGB was altered by the 50% drought treatment.

Net ecosystem productivity and its constituents

Individual model predictions of control plot NEP ranged between a small carbon source and a moderately strong carbon sink at both sites (Fig. 2a,c, Table S1). JULES and SiB3 simulations of NEP agreed with the flux tower measurements at TNF in year 4 (-0.09 ± 0.05 kg C m^{-2} yr^{-1}), predicting the control plot to be a weak carbon source (Fig. 2a, Table S1). However, CLM3.5, ED2, IBIS and SPA predicted TNF control plot NEP to be a carbon sink (Fig. 2a, Table S1). At CAX, CLM3.5 predicted the control plot NEP to be a weak carbon source, while ED2, IBIS, JULES, SiB3 and SPA predicted a carbon sink (Fig. 2c, Table S1).

At both sites, model predictions of GPP and component *R* fluxes in the control plots generally agreed well with the available measurements (Fig. 2b,d, Table S1). All models except SiB3

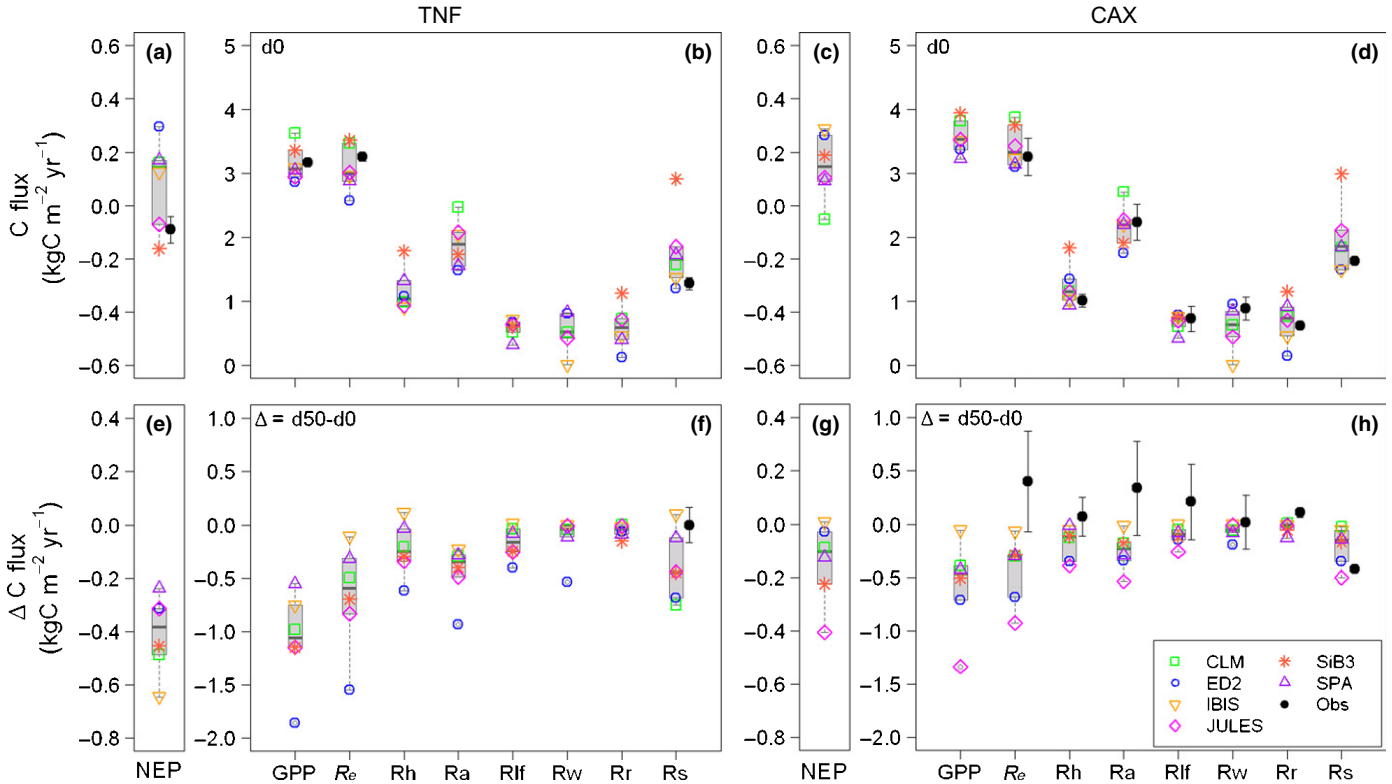

Fig. 2 Net and component ecosystem carbon fluxes (kg C m^{-2} yr^{-1}) in the fourth year of the experiment for Tapajós (TNF; left side) and Caxiuanã (CAX; right side) National Forests. Carbon flux definitions and observation sources are given in Table 2. Colored symbols are model predictions and black symbols are observations (mean ± 95% confidence interval (CI)). The control plot (d0) carbon fluxes are shown in panels (a) to (d) and the drought treatment plot (d50) fluxes are shown in panels (e) to (h). The Δ indicates the amount the control (d0) fluxes were altered by the 50% drought treatment.

correctly predicted that autotrophic respiration (R_a) should exceed heterotrophic respiration (R_h) as observed at CAX.

When precipitation was reduced by 50%, all models, except IBIS at CAX, predicted considerable reductions in NEP at both sites (Fig. 2e,g, Table S1). Although measurements of overall NEP were unavailable for comparison, model predictions of the annual component R fluxes compared poorly with observations from CAX (Fig. 2h, Table S1). The response of individual model predictions of annual R_e and its components to d50 ranged from neutral to negative, while the observed responses were neutral to positive. Almost the entire mismatch between observed R_e and model predictions arose from the disagreement within the components of R_a, in particular R_{lf} (Fig. 2h, Table S1). For comparative purposes, the TNF component R fluxes for the fourth year are also provided in Fig. 2(f) and Table S1 despite absence of validation data at this site. The patterns and magnitudes of the model predictions for TNF (Fig. 2f) were generally similar to those at CAX except for ED2, whose predictions changed substantially, reflecting the predicted reduction in AGB at this site (Fig. 1c, Table S1).

Soil respiration

A direct comparison of the magnitudes of observed and modeled R_s was only possible for CAX, because the meteorological drivers for the simulations were concurrent with the R_s measurements between November 2001 and 2003 (Fig. S3). In the control plot, predictions of R_s by CLM3.5, IBIS and SPA were consistently in agreement with the magnitude and seasonality of the observations, while the magnitude of JULES and SiB3, and seasonal dynamics of ED2 and SiB3 exceeded the observations (Fig. S3a). At the treatment level, however, there was good agreement between model predictions and measurements of the R_s response to d50 only during the wet season, and not during the dry season (Fig. S3b).

The sensitivity of R_s to volumetric soil water content (θ_s) in the observations was significantly different at the two sites, where a dependence was only observed at CAX (Fig. 3). Accordingly, the absence of a θ_s dependence in the IBIS and SPA formulas was more realistic at TNF (Fig. 3a), while the approximately parabolic relationship between R_s and θ_s found in JULES was more realistic for CAX, but incorrectly parameterized to match the observations (Fig. 3b). Predictions of R_s by ED2 and SiB3 were excessively sensitive to soil moisture, but in contrasting directions relative to the observations. For ED2, R_s declined well below the observations at low θ_s, but rapidly exceeded them at high θ_s (Fig. 3). Interestingly, at the annual time-scale, the low R_s predictions by ED2 for the control plots were caused by relatively low R_r rather than R_h (Fig. 2b,d, Table S1). However, the parameterization of R_s in SiB3 produced fluxes equal to the observations at low θ_s, which then increased linearly to an optimal θ_s level (Fig. 3). Hence, SiB3's R_s was generally higher than in the other models, particularly at TNF, where θ_s was often above the parameterized optimal level (*c.* 0.25 $m^3\ m^{-3}$) (Figs 2b,d, 3, Table S1). Moreover, SiB3's overestimation of R_s in both control plots (e.g. Fig. 2b,d) was also attributable to its R_r formulation. Unlike CLM3.5, ED2, IBIS and JULES, growth respiration (R_g) is not explicitly modeled in SiB3; rather, a relatively high, nondimensional scalar (0.5) is used in its R_r calculation to achieve plant carbon balance closure (Table S4).

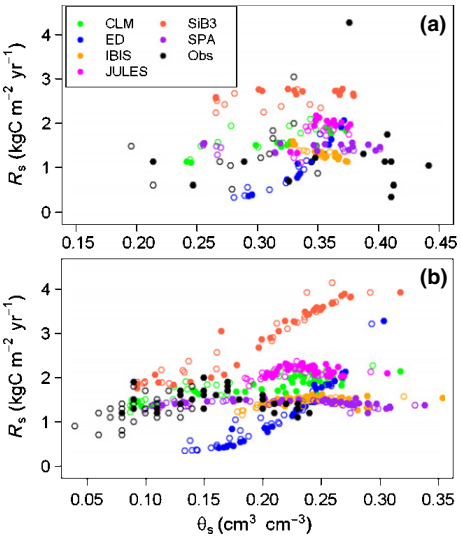

Fig. 3 Periodic measurements of soil respiration (R_s; mean ± 95% confidence interval (CI); kg C m^{-2} yr^{-1}; black symbols) as a function of volumetric soil water content (θ_s; cm^3 cm^{-3}) for (a) Tapajós (TNF; Davidson *et al.*, 2008) and (b) Caxiuanã (CAX; Sotta *et al.*, 2007) National Forests. Concurrent R_s predictions are given for each model (colored symbols). Closed symbols are for the control plots (d0), and open symbols are for the treatment plots (d50).

Net primary production and litter fluxes

Model predictions of annual woody biomass increment (net primary production of carbon in wood (NPP_w)) in the control plots of both sites were systematically higher than the observations, with the exception of JULES at TNF (Fig. 4a,b, Table S1). The model predictions of NPP_w under the drought treatment at TNF were, however, realistic in the sense that the models generally captured the observed decline in NPP_w and the accompanying pattern of interannual variability (Fig. 4c,d). By contrast, NPP_w was poorly captured under the drought treatment at CAX, where CLM3.5, ED2 and JULES all systematically over-predicted the changes in observed NPP_w, while IBIS and SPA under-predicted the observed reductions in NPP_w in 2003, 2007 and 2008 (Fig. 4d, Table S1).

The predictions of litter fluxes for the control plots of both sites had a very large range (Fig. 5a,b, Table S1). IBIS's litterfall predictions for the TNF control plot agreed well with the observations (*c.* 0.30 kg C m^{-2} yr^{-1}), while ED and CLM3.5 both under-predicted observations, and JULES and SPA had litter fluxes that were significantly lower and higher, respectively, than the observations. At CAX, each model's prediction of annual litterfall in the control plots was of a similar magnitude to its TNF prediction, except for a 35% lower prediction by SPA. Except for JULES, the models generally predicted a long-term decline in litterfall under the d50 drought treatment, which contrasted the initial increase and then subsequent decline found in the observations at TNF (Fig. 5c,d).

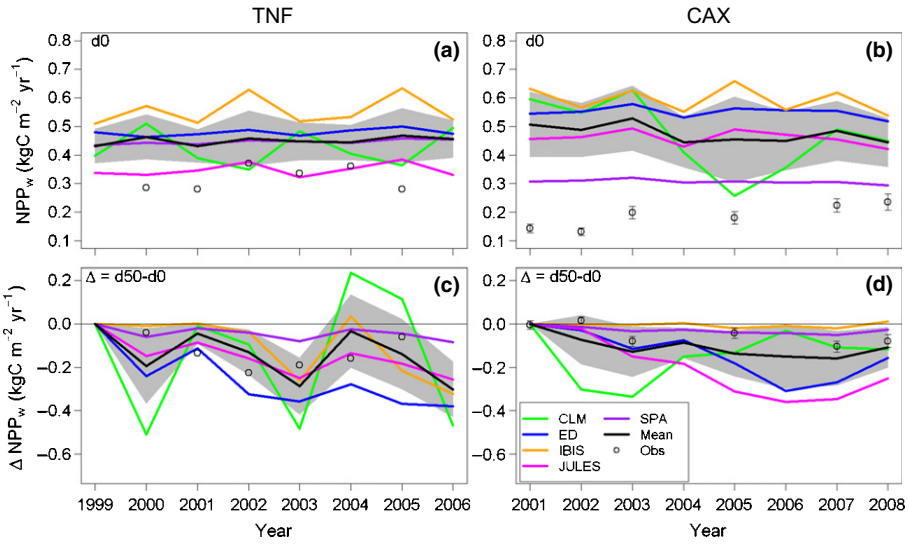

Fig. 4 Annual net primary production of wood (NPP_w, kg C m^{-2} yr^{-1}) for Tapajós (TNF; left side) and Caxiuanã (CAX; right side) National Forests. Colored lines are individual model predictions and the black line is the five-model ensemble mean. The shaded area is the 95% confidence interval (CI) of the models. Open symbols are published observations (mean ± 95% CI (when reported); TNF: Brando *et al.*, 2008; CAX: da Costa *et al.*, 2010). d0 and d50 are drought levels indicating a 0 and 50% reduction in precipitation, respectively. The Δ plots (c, d) show the amount the control (a, b) NPP_w was altered by the 50% drought treatment.

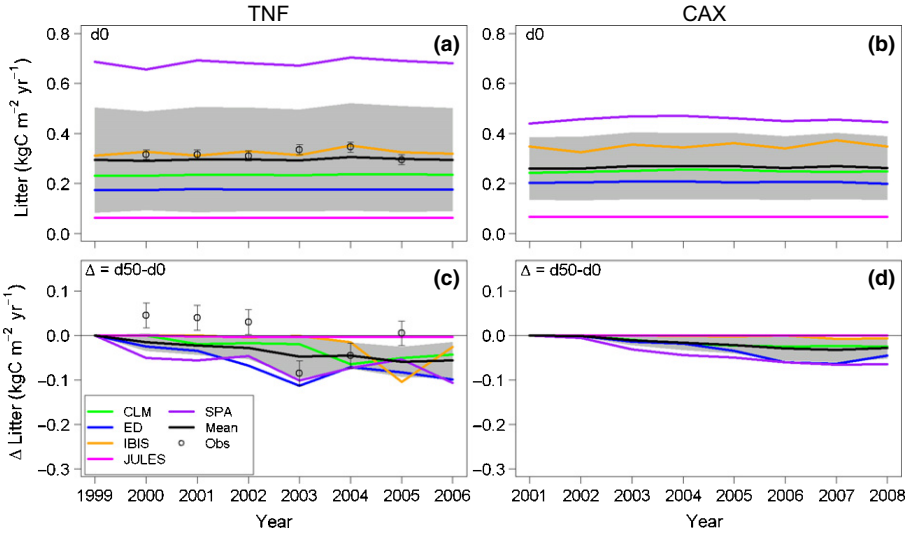

Fig. 5 Annual litter production (kg C m^{-2} yr^{-1}) for Tapajós (TNF; left side) and Caxiuanã (CAX; right side) National Forests. Colored lines are individual model predictions and the black line is the five-model ensemble mean. The shaded area is the 95% confidence interval (CI) of the models. Open symbols are published observations (mean ± 95% CI; TNF: Brando *et al.*, 2008). d0 and d50 are drought levels indicating a 0 and 50% reduction in precipitation, respectively. The Δ plots (c, d) show the amount the control (a, b) litter production was altered by the 50% drought treatment.

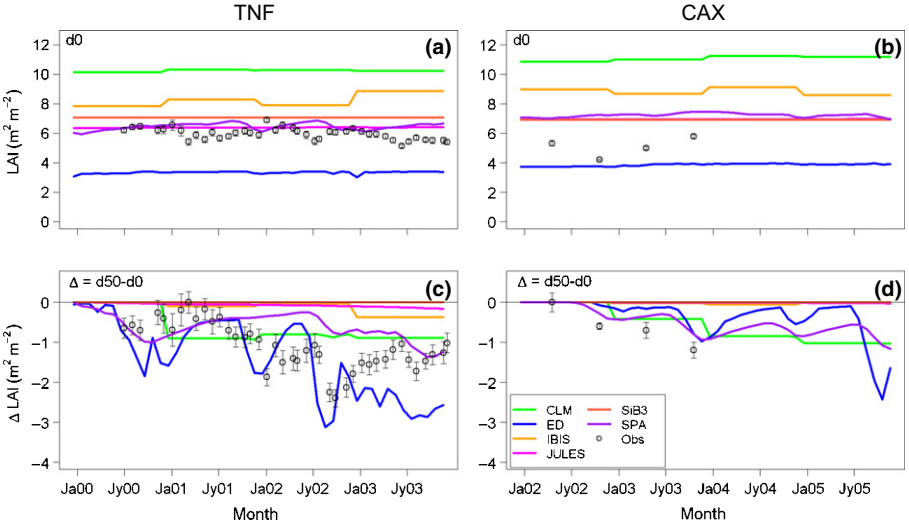

Fig. 6 Leaf area index (LAI; m^2 m^{-2}) for Tapajós (TNF; left side) and Caxiuanã (CAX; right side) National Forests over years 1–4 of the experiment. Colored lines are individual model predictions. Open symbols are published observations (mean ± 95% confidence interval (CI); TNF: Nepstad & Moutinho, 2008; CAX: Fisher *et al.*, 2007). d0 and d50 are drought levels indicating a 0 and 50% reduction in precipitation, respectively. The Δ plots (c, d) show the amount the control (a, b) LAI was altered by the 50% drought treatment.

Leaf area index (LAI) predictions for the control simulations at both sites varied significantly between models (Fig. 6a,b). The LAI predictions of JULES, SIB3 and SPA were in agreement with observed LAI values of 6 m^2 m^{-2} measured at TNF (Fig. 6a). By contrast, CLM3.5 and IBIS over-predicted TNF LAI by 4 and 2 m^2 m^{-2}, respectively, while ED2 under-predicted it by

$3\ m^2\ m^{-2}$. The LAI observations for the control plot at CAX were intermediate between the ED2 prediction of $4\ m^2\ m^{-2}$ and the JULES, SIB3 and SPA predictions *c.* $7\ m^2\ m^{-2}$ (Fig. 6b). In both treatment plots there was an overall decline in LAI by 20–30% over the first 4 yr (Fig. 6c,d), and at TNF, the observed seasonality of LAI increased markedly (Fig. 6c). At CAX, observed LAI in the treatment plot remained suppressed by *c.* 20% in the sixth year (Metcalfe *et al.*, 2010a). Although the leaf area formulations of CLM3.5, ED2, IBIS and SPA all contained soil moisture dependencies, their predictions of LAI displayed contrasting sensitivities to the d50 treatment (Fig. 6c,d). ED2 and SPA replicated the observed trends of LAI in the treatment plots. However, by the third year, the leaf-drop formulation in ED2 resulted in a recurring 90% loss of canopy foliage each dry season, whereas SPA predicted a 20% relative reduction in LAI. In CLM3.5, LAI declined in a stepwise fashion that was similar in magnitude to the observations; but the relative reduction in LAI predicted by CLM3.5 was only half of relative reduction observed between the control and treatment plots. IBIS predicted a 30% reduction in LAI after the fourth year for the TNF treatment plot (not shown, as TNF LAI was not measured after 2003), but no response at all in the CAX treatment plot. The leaf area formulations for JULES and SiB3 were not dependent on soil moisture, and thus LAI did not display a significant drought response over the 7-yr experimental period.

Carbon fluxes as a function of drought level

To better understand the drought sensitivities of the models, we examined their responses to three levels of drought: 30, 50 and 80% reductions in precipitation. The range among the six models in cumulative NEP after 7 yr was large at the control level, with SiB3 at one end predicting both sites to be carbon neutral, and ED2 at the other end predicting both sites to be strong carbon sinks (data not shown). There was also considerable disagreement among models about the interannual variation of NEP, where the magnitudes and signs were often opposing (Fig. 7a,b). As drought severity increased from d30 to d80, model agreement about the magnitude and interannual variability of ΔNEP decreased (Figs 7c–h, 8).

The contrasting trends in NEP under ambient and drought conditions reflected differences among the six models in their hypotheses about the balance between GPP and *R* for the two ecosystems. After 7 yr under the control simulation, modeled GPP ranged between 20.4 and 24.9 $kg\ C\ m^{-2}\ 7\ yr^{-1}$ at TNF and between 22.5 and 27.4 $kg\ C\ m^{-2}\ 7\ yr^{-1}$ at CAX. CLM3.5

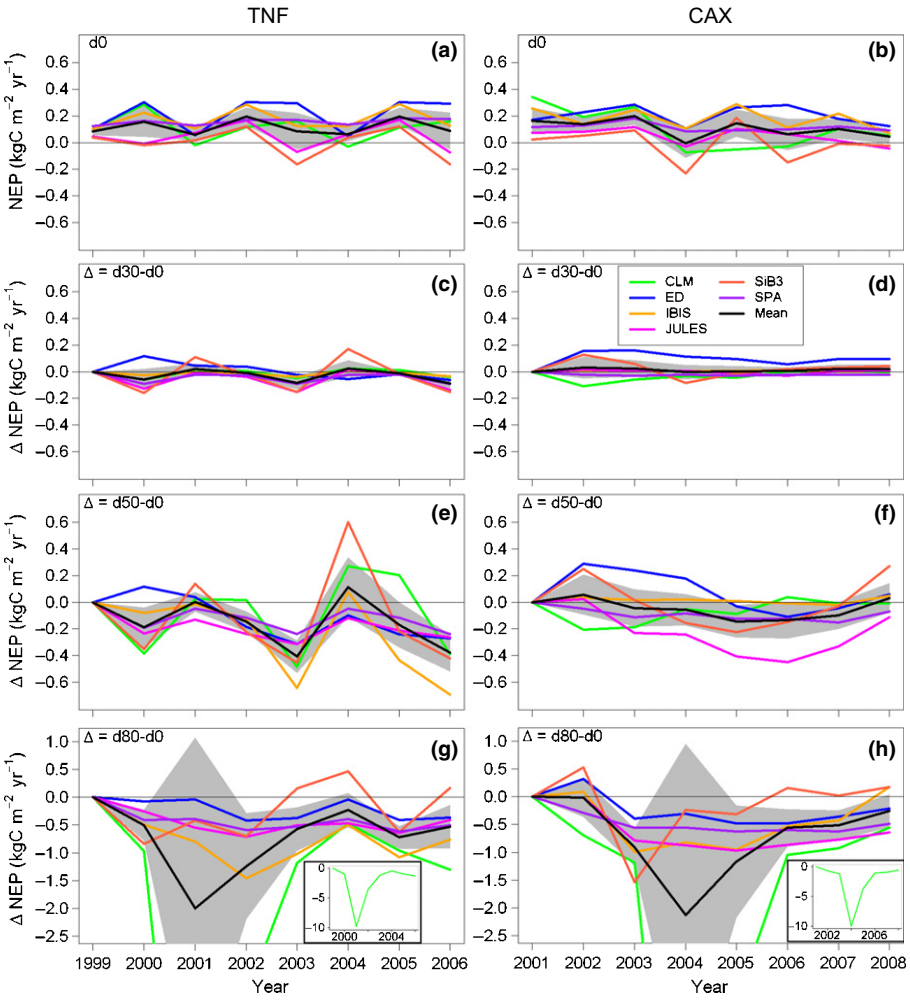

Fig. 7 Annual net ecosystem production of carbon (NEP; $kg\ C\ m^{-2}\ yr^{-1}$) for Tapajós (TNF; left side) and Caxiuanã (CAX; right side) National Forests. Colored lines are individual model predictions and the black line is the six-model ensemble mean. The shaded area is the 95% confidence interval (CI) of the models. NEP for the control plots (d0) are shown in panels (a) and (b). Drought levels are indicated by (c, d) d30, (e, f) d50 and (g, h) d80, which are, respectively, 30, 50 and 80% reductions in precipitation. The Δ indicates the amount the d0 carbon fluxes were altered by the indicated drought level. Insets in (g) and (h) show the full range for CLM3.5.

and SiB3 occupied the higher end of the range of GPP at both sites, while ED2, IBIS, JULES and SPA occupied the lower end. As drought intensity increased, there was increasing disagreement in both the cumulative magnitude (Fig. 8) and interannual variation (Fig. S4) of the model GPP predictions at both sites. Among the biosphere models, IBIS's predictions of GPP were generally the least sensitive to the increasing levels of drought at both sites, while the greatest reductions were predicted by ED2 at the d30 and d50 levels for TNF (Figs 8b,f, S4c,e), by CLM3.5 at the d30 and d80 levels at CAX and TNF, respectively (Fig. 8d,j), and by JULES at the d50 and d80 levels at CAX (Figs 8h,i, S4f,h). The reductions in GPP predicted by the SPA ecosystem model under the d50 and d80 drought levels at both sites were also comparatively small (Fig. 8f,h,j,l).

These differences in predicted GPP could in part be accounted for by the contrasting formulations associated with soil water stress. Canopy aggregated GPP declines under soil water deficits through a combination of two sets of mechanisms, each operating differently among all six models. The first set modifies canopy leaf area. The substantially lower predictions of GPP by ED2 at the d30 and d50 levels at TNF (Fig. 8b,f) were predominately caused by a reduction in LAI, first through leaf-shedding (Fig. 6c) and then through a loss of AGB (Fig. 1a).

GPP also decreases through plant responses to atmospheric demand, that is, vapor pressure deficit (VPD), and reductions in soil-water supply. In the terrestrial biosphere models, the photosynthesis routine accounts for the effects of VPD on stomatal conductance. The effects of supply are incorporated via a

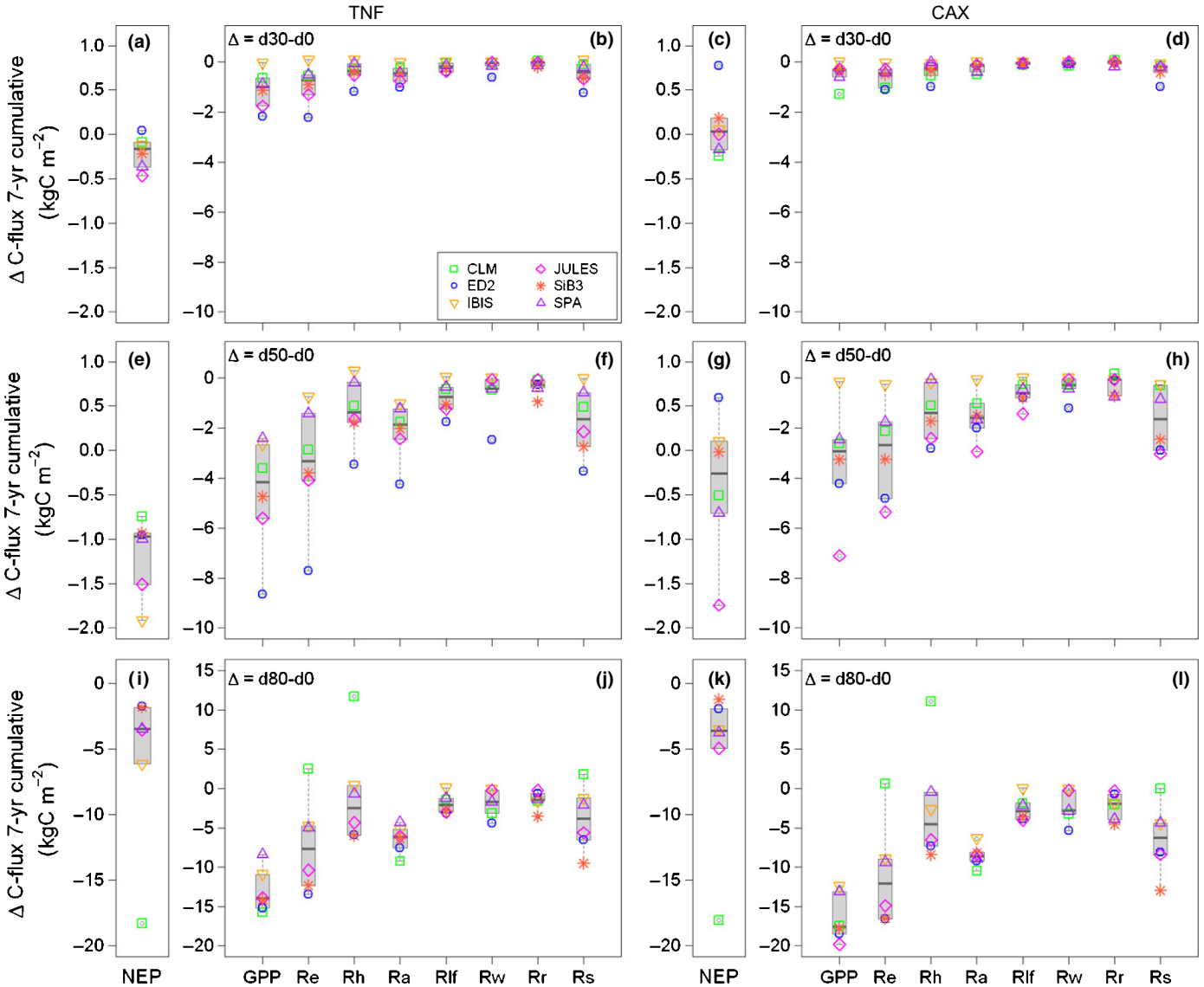

Fig. 8 Change in the 7-yr cumulative net and component ecosystem carbon fluxes (kg C m^{-2} yr^{-1}) predicted by each model for Tapajós (TNF; left side) and Caxiuanã (CAX; right side) National Forests. The line in the box is the median value for the models. Carbon flux definitions are given in Table 2. Drought levels are indicated by (a–d) d30, (e–h) d50 and (i–l) d80, which are, respectively, 30, 50 and 80% reductions in precipitation. The Δ indicates the amount the d0 carbon fluxes were altered by the indicated drought level.

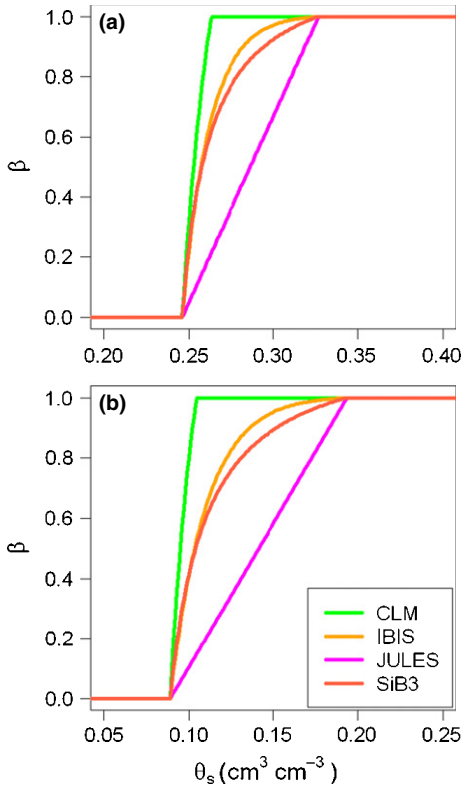

Fig. 9 Relationships between the soil water-stress factors (β) and volumetric soil moisture (θ_s; cm^3 cm^{-3}) for CLM3.5, IBIS, JULES and SiB3 for (a) Tapajós (TNF) and (b) Caxiuanã (CAX) National Forests. The function for each model is given in Supporting Information Table S3 and the parameter values used for each model are given in Table S2.

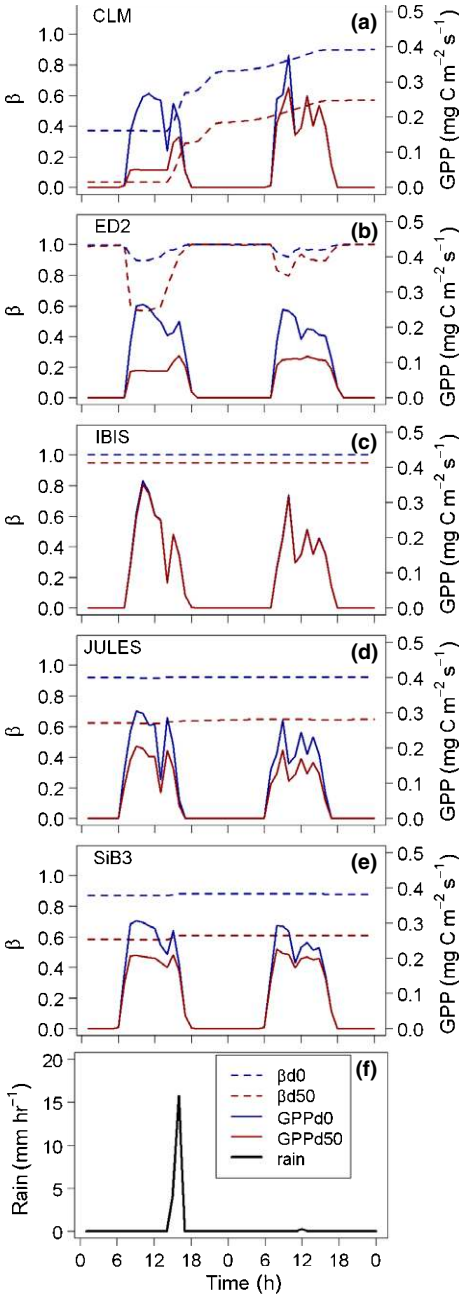

Fig. 10 Diurnal trends of hourly ecosystem water-stress factor (β) and ecosystem gross primary production (GPP; mg C m^{-2} s^{-1}) at the Tapajós (TNF) National Forest for (a) CLM3.5, (b) ED2, (c) IBIS, (d) JULES, and (e) SiB3 for 2 d, 13–14 October 2000. (f) The panel shows a concurrent rain (mm h^{-1}) event. d0 and d50 indicate 0 and 50% reductions in precipitation, respectively.

phenomenological function (β) that down-regulates stomatal conductance (g_s), and as a consequence GPP, with decreasing soil moisture. (Throughout this paper the symbol g_s includes β.) β varies between 0 and 1, where 1 implies that stomata are only VPD-regulated and 0 causes full stomatal closure (Fig. 9, Notes S1). Across all models, β increasingly trended toward 0 with increasing drought intensity and duration over the 7-yr simulation (Fig. S5). β was also generally lower at each drought level at TNF compared with CAX (Fig. S5). Examination of diurnal-scale variation of β revealed contrasting dynamics among the models (Fig. 10a–e): β had a diurnal cycle in ED2, while in CLM3.5, IBIS, JULES and SiB3 it did not (Fig 10a–e). In addition, β in CLM3.5 and ED2 was responsive to individual rain events (Fig. 10a,b,f), whereas in IBIS, JULES and SiB3 changes in β were more gradual and occurred over seasonal time-scales (Figs 10c–f, S5e–j). Finally, in ED2, IBIS and JULES, β directly scaled net photosynthesis, while in CLM3.5 and SiB3 β scaled V_{cmax} (Table S3), dampening the effect of β on GPP following rain events (Fig. 10a,e). In contrast to the biosphere models, in SPA stomatal function is mechanistically linked to both the supply of soil moisture and atmospheric demand through a porous medium pipe-model formulation.

Fig. 11 shows the predicted distributions of hourly midday (11:00–14:00 h) g_s values at both sites spanning a 6-month period between the wet and dry seasons (May–November during 2000 at TNF and during 2003 at CAX). In the control simulations at both sites, JULES had very broad distributions of g_s that peaked at *c.* 400 mmol CO_2 m^{-2} leaf s^{-1}, while the distributions of g_s in ED2 and IBIS were intermediate and peaked at *c.* 250 and 500 mmol CO_2 m^{-2} leaf s^{-1}, respectively, and the distributions of g_s in CLM3.5 and SiB3 were relatively narrow and peaked at *c.* 40 and 300 mmol CO_2 m^{-2} leaf s^{-1}, respectively (Fig. 11a–j). SPA, in contrast, predicted a bimodal distribution, with the larger of the two peaks switching between TNF and CAX (Fig. 11k,l). Under the drought treatment, the β-correction caused a

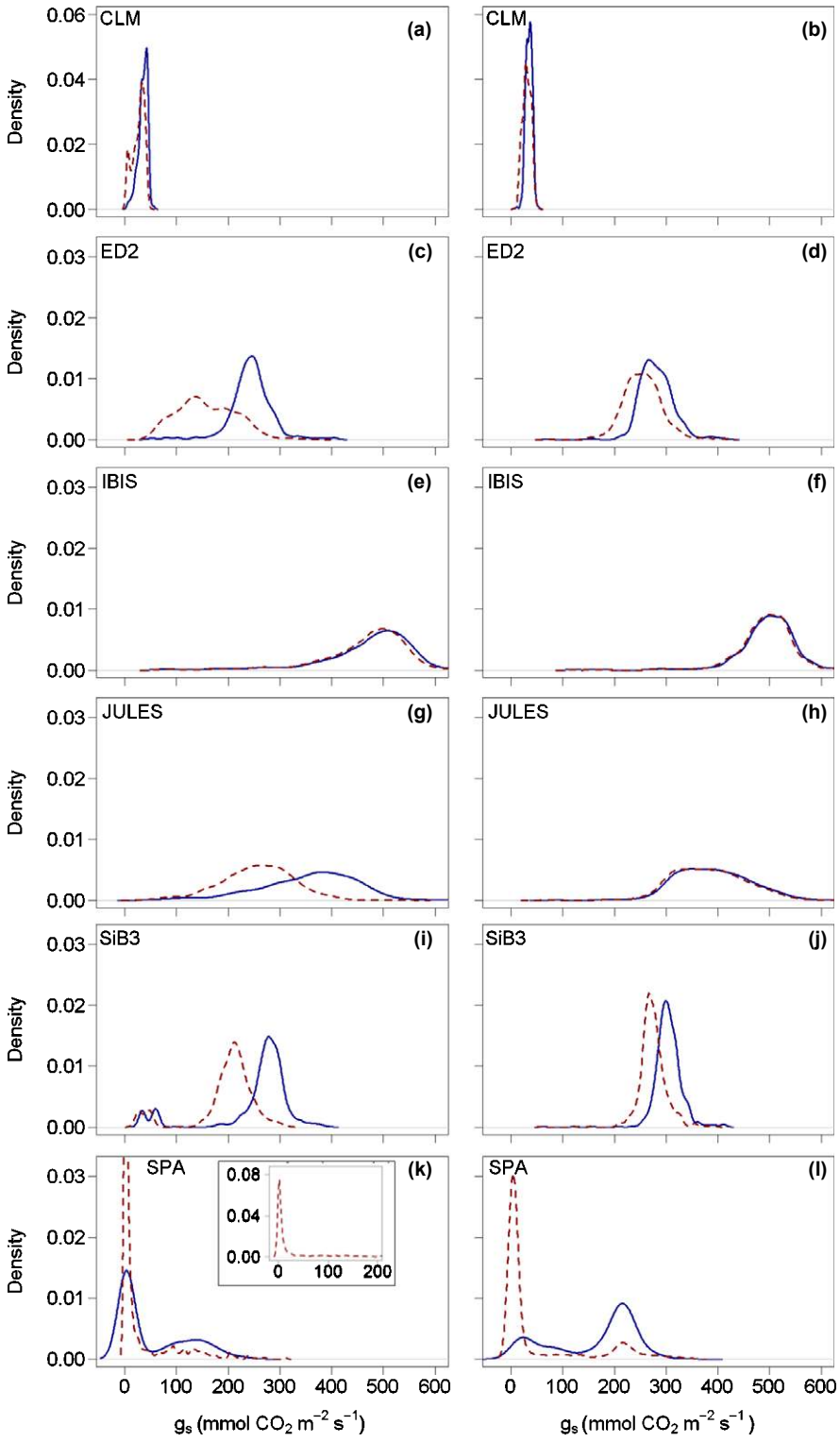

Fig. 11 Density plots of modeled hourly stomatal conductance (g_s, mmol CO_2 m^{-2} leaf s^{-1}) during midday (11:00–14:00 h) over a 6-month period, May–November, in 2000 for Tapajós (TNF; left side) and 2003 for Caxiuanã (CAX; right side) National Forests. Panels (a, b) show CLM, (c, d) ED2, (e, f) IBIS, (g, h) JULES, (i, j) SiB3, and (k, l) SPA. d0 (solid line) and d50 (dashed line) indicate 0 and 50% reductions in precipitation, respectively. The inset in (k) shows the full range of g_s at the d50 level for the SPA model. Note the scale difference on the *y*-axis for CLM (a, b).

downward shift by 50–100 mmol CO_2 m^{-2} leaf s^{-1} in the distributions of g_s in ED2 and SiB3 at both sites and JULES at TNF, but not in CLM3.5 and IBIS at both sites and JULES at CAX (Fig. 11a–j). In SPA, the drought treatment at both sites resulted in a very narrow concentration of g_s values just above 0 mmol CO_2 m^{-2} leaf $s^{-1}$with a long positive tail, which was a substantially greater change compared to the other models (Fig. 11a–l).

There were considerable differences among the six models in the relative contribution of each constituent flux to R_e in the control simulations for both sites. For all models except SiB3,

R_a was the dominant component of R_e, but by varying magnitudes (Figs S6a,b, S7a,b). Accordingly, CLM3.5 was at one end of the range, estimating R_a to be more than double R_h, and SiB3 was at the other end, with R_a and R_h contributing equally to R_e. Also, the patterns of interannual variation of R_h and R_a were often opposing between models (Figs S6a,b, S7a,b). Of the constituent fluxes of R_a, there was generally strong agreement in model predictions of R_{lf} and R_w, and considerable uncertainty in R_r, which accounted for most of the uncertainty in R_a (not shown, but similar pattern visible in Fig. 2b,d). One notable exception was IBIS's significantly low predictions of R_w (see also Fig. 2b,d).

Variation between model predictions of respiration fluxes also increased substantially with drought intensity, yet some clear patterns of individual models emerged (Fig. 8). ED2, JULES and SiB3 predicted increasingly the highest reductions in R_h as drought levels intensified as a consequence of their decomposition formulations including a soil moisture factor. In ED2, the soil moisture factor regulating R_h was so influential that it resulted in ΔNEP being positive at the d30 and d50 drought levels (Figs 7c–f, 8a,c,g). IBIS was the least sensitive to all drought levels, predicting only small increases in R_h and reductions in R_a at TNF (Fig. 8b,f,j), and virtually no response of its respiratory fluxes at CAX (Fig. 8d,h,l). While the cumulative 7-yr response of IBIS was small, its prediction of d50 R_a at TNF came into alignment with the other models by the seventh year (Fig. S7e), implying that its respiration did eventually respond to the drought treatment. At the d80 level, the R_h prediction of CLM3.5 was a significant positive outlier (Figs 8j,l, S6g,h), thus reflecting the respiration of the additional dead biomass, which was not captured by ED2 even though it too experienced a high mortality flux (Fig. S2) These substantial increases in R_h predicted by CLM3.5 occurred only in the second and third years at TNF and in the third and fourth years at CAX; otherwise the R_h predictions were in alignment with the other models (Fig. S6). Finally, for CLM3.5 at the d80 level, its predictions of R_a had much more interannual variation relative to the similar behavior of the other models (Fig. S7).

Discussion

Under current climate conditions, the six models analyzed here had a reasonably high degree of skill in replicating ecosystem carbon fluxes (Fig. 2a,b,d, Table S1). The models also had some success in capturing the responses to the experimental drought, generally capturing the observed reductions in NPP_w (Fig. 4c,d), and ED2 correctly predicted the timing of the reduction in AGB (Fig. 1c), but the models failed to accurately predict the magnitude of the drought-induced reductions in AGB and accompanying component carbon fluxes (R_e, R_a, R_r and R_s at CAX and R_s at TNF) (Figs 1c,d, 2f,h, Table S1).

One explanation for this pattern of having high skill under current climate conditions and poor skill under drier climate conditions is that the models are poorly parameterized because they have been developed to replicate past observations of rainforests, which are rarely limited by water (Huete *et al.*, 2006; Hutyra *et al.*, 2007). Data-model assimilation using the observations from the two drought experiments may help to correct this problem. However, the results of this analysis more strongly suggest that the models require more realistic representations of key photosynthetic and respiratory drought response mechanisms. These are discussed in more detail in each section below.

Photosynthesis and soil water stress

In all models except SPA, the stomatal response to decreasing soil moisture is represented by down-regulation imposed through a phenomenological function (β). The nature of the β function differs across these models (Fig. 9, Table S3), and thus each represents a contrasting hypothesis about how plant productivity is affected by available soil moisture. Any bias associated with β translates directly into a corresponding bias in the magnitude and timing of the decline in GPP as θ_s declines. For IBIS, JULES and SiB3, β changed gradually as available soil moisture was drawn down or replenished (Fig. S5), but was effectively static over diurnal periods (Fig. 10c–e). In CLM3.5, β was very responsive to rain events (Fig. 10a,f) as it was parameterized based on soil matrix potential (ψ_s) instead of θ_s (Table S3), but was static during dry days. By contrast, in ED2 β had a clear diurnal cycle (Fig. 10b) because its value is affected by evaporative demand at the leaf, which changes significantly over the daily photoperiod. It is also important to note that in CLM3.5, ED2, IBIS and SPA the level of soil water supply is impacted by changes in demand that are modulated through leaf area dynamics (Fig. 6). We return to this issue in the 'Changes in plant carbon pools' section below.

The mechanistic hydrodynamic formulation in SPA has been shown to be effective in capturing both the seasonal and diurnal dynamics of stomatal control over GPP (Fisher *et al.*, 2006) and thus obviates the need for a β-type correction factor. However, the reduction in the mean density of ecosystem-averaged leaf-level g_s predicted by SPA for CAX in this study is considerably greater than the reduction of *c.* 40 mmol CO_2 m^{-2} leaf s^{-1} reported for four canopy trees, and simulated by SPA in a local CAX parameterization using detailed soil hydraulic measurements (Fisher *et al.*, 2006). In this study, using generic soil hydraulics parameters, the sensitivity of g_s to drought predicted by SPA increased. However, the drought-induced reduction in g_s predicted by ED2 and SiB3 was similar in magnitude to the individual tree measurements reported by Fisher *et al.* (2006). These results indicate the challenge of generating at a regional scale the local hydraulic parameters required for a SPA-style mechanistic formulation of plant water use. Therefore, for regional applications, where soil hydraulics data are limited, the β formulations of ED2 and SiB3 are relatively robust in terms of down-regulating g_s.

Autotrophic respiration

After the drought treatment was initiated, leaf dark respiration at CAX increased in terms of both unit leaf area and unit mass (Metcalfe *et al.*, 2010b). At the same time, observed LAI

decreased, but the decline did not offset the increase in dark leaf respiration rate, resulting in an overall increase in R_{lf} (Fig. 2h). The predicted reductions in R_{lf} under the experimental drought directly conflict with these observations (Fig. 2h, Table S1). Of particular note, JULES and SiB3 explicitly down-regulate R_{lf} by β (Table S4), with this effect being stronger in JULES (Fig. 8h) as its decrease in β begins at a higher θ_s relative to SiB3 (Fig. 9).

The results of this study suggest that respiration formulations need to be modified to test hypotheses about increases in maintenance respiration when soil moisture becomes severely limiting. For example, it is possible that R_{lf} increases under drought as a consequence of futile respiratory cycles or because of demand for additional energy to maintain solute gradients, repair damaged tissue (Würth *et al.*, 2005; Meir *et al.*, 2008; Metcalfe *et al.*, 2008), or repair leaf embolisms (Brodribb & Holbrook, 2003). Furthermore, all six models evaluated in this study follow the paradigm, either explicitly (SPA only) or heuristically, that growth declines when drought induces stomatal closure and down-regulates GPP. However, many studies across a range of plant forms have shown that under water stress growth declines before photosynthesis, such that they become uncoupled and carbon accumulates within the plant (Würth *et al.*, 2005; reviewed by Muller *et al.*, 2011). In many cases, the excess carbon is converted to carbon-rich compounds that require extra energy to generate. Inclusion in future model iterations of formulations that up-regulate respiration during water stress to account for all of these discussed processes is critical for correctly computing the plant carbon balance during drought.

Changes in plant carbon pools

Another critical component of capturing how GPP and R_a respond to soil water-stress is correctly characterizing how carbon stocks change under drought. Accordingly, properly constraining seasonal leaf-area dynamics is imperative for reducing substantial biases in predicted GPP and R_a (Richardson *et al.*, 2012). Yet, all of the models failed to capture both the initial increase in litter production observed in the drought treatment plots at TNF (Fig. 5a) and, except for CLM3.5, ED2 and SPA, the reductions in LAI observed at both sites (Fig. 6c,d). While the mechanisms controlling leaf-shedding and flushing by tropical trees are not well understood (Kim *et al.*, 2012), they appear to be related to seasonal cumulative soil water deficit (Nepstad *et al.*, 2007; Brando *et al.*, 2010). The seasonally static leaf area of CLM3.5, IBIS, JULES and SiB3 introduced additional biases in their predictions of GPP and R_{lf} (Fig. 6), while the phenological formulations of ED2 and SPA, which depend on soil moisture, appear to be more reasonable approximations for canopy leaf-area dynamics.

Similarly, correctly predicting R_a under drought conditions requires not only the per unit tissue R_a to be accurate, but also the changes in the plant's carbon stocks. In this regard, the insensitivity of plant biomass to increasing water stress seen in Fig. 1 is particularly problematic for the dynamic vegetation models. For IBIS, JULES, SPA and also SiB3 in terms of LAI, even the d80 drought level did not trigger a reduction in biomass (Fig. S2), which implies that, even though they contain mechanisms to modify plant carbon balance under extreme drought, they still do not contain a mechanism that translates a carbon *imbalance* into a loss of biomass (D. Galbraith, unpublished). Without a drought-related mortality mechanism, these models will not be able to accurately predict carbon fluxes under drought because the fluxes will be derived from incorrect levels of biomass or leaf area. In CLM3.5 and ED2, significant changes in carbon stocks occurred (Figs 1, 6) after chronically negative plant carbon balance occurred. Accordingly, for ED2, the 50% reduction in AGB and 90% reduction in LAI at the d50 level at TNF caused reductions in GPP and component R fluxes (except R_r) to be considerably greater than predicted by the other models (Fig. 8f). Also, even for CLM3.5 and ED2, mortality attributable to negative carbon balance was incorrectly tuned to replicate the observed losses in AGB in the treatment plots (Figs 1, S2).

Net ecosystem carbon balance

Accurate predictions of changes in NEP are dependent not only on correctly quantifying changes in plant carbon stocks, but also on correctly characterizing how rates of decomposition change in response to drought. Accordingly, CLM3.5 and ED2 reflect contrasting hypotheses about the rate at which soil carbon is respired under drought, which becomes evident in their contrasting predictions of R_h (Fig. S6) and resulting NEP (Figs 7, 8) following their predicted AGB declines (Figs 1c, S2). CLM3.5 predicted R_h to dramatically increase for the 2 yr following its predicted AGB loss (Fig. S6g,h). Meanwhile, ED2 predicted a general decline in R_h under all drought levels (Fig. S6c–h), even though there was an increase in its soil carbon pool (not shown) following its predicted AGB loss (Figs 1c, S2) As a result, CLM3.5's predicted NEP was significantly lower than those of the other models (Figs 7g,h, 8i,k), whereas ED2's NEP prediction was similar to those of the other models (Figs 7e,h, 8e,k).

Overall the models compared well to the site-specific empirically derived R_s annual estimates (Sotta *et al.*, 2007; Davidson *et al.*, 2008) at the d0 level (Fig. 2b,d), but poorly under the d50 conditions (Fig. 2f,h, Table S1). The inability of the models to correctly capture drought effects on R_s stems from the compounding effect of two related problems. First, in an adjacent forest at CAX, measurements of R_r accounted for 42% and 61% of R_s during the wet and dry seasons, respectively (Metcalfe *et al.*, 2007), while all models except SPA predicted R_h to be the dominant component of R_s (Fig. 2d). Secondly, inaccurate soil moisture dependencies of R_h are apparent from the poor agreement between the modeled dependencies of R_s on θ_s and the observations (Fig. 3). In particular, the dependencies in ED2 and SiB3 resulted in extreme excursions from the R_s observations at high (ED2 and SiB3) and low (ED2) θ_s values (Fig. 3). While no direct measurements of NEP for the drought plots were available for evaluation, the predictions of R_h (Fig. S6c–h), and consequently NEP (Fig. 7c–h), under drought by ED2 and SiB3 were probably unrealistic given the mismatch between their predictions and the observations of R_s (Fig. S3). In order to make correct estimates of NEP under drought,

future model development must focus on correcting both the relative contribution of R_h to R_s and the parameterization of the dependence of R_h on θ_s.

Conclusions

To our knowledge, this is the first time a series of terrestrial biosphere models run under a standardized protocol have been evaluated for their ability to predict how chronic drought affects plant and ecosystem carbon balances. This study demonstrates that terrestrial biosphere models are competent at predicting plant and ecosystem carbon fluxes under the present climate, but still require substantial development for predicting the consequences of severe drought scenarios. Model development should be focused on testing hypotheses associated with enhanced R_a under severe water stress, controls on leaf phenology, and drought-induced mortality.

Acknowledgements

This research was funded by a grant from the Andes-Amazon Initiative of The Gordon and Betty Moore Foundation. T.L.P., B.O.C. and S.R.S. gratefully acknowledge support from the National Science Foundation Partnership for International Research and Education in Amazon Climate Interactions grant (NSF award #OISE-0730305). P.M. was supported by NERC award NE/J011002/1 and ARC award FT110100457. The LBA-DMIP project (NASA award #NNX09AL52G) provided the TNF meteorological data. We are grateful to the Museu Paraense Emilio Goeldi, the LBA office in Santarém, and the dedicated researchers who collected the data that made this model evaluation possible. We thank Eric Davidson for furnishing us with their published original TNF R_s data.

References

Baker IT, Prihodko L, Denning AS, Goulden M, Miller S, da Rocha HR. 2008. Seasonal drought stress in the Amazon: reconciling models and observations. *Journal of Geophysical Research* **113**, G00B01.

Best MJ, Pryor M, Clark DB, Rooney GG, Essery RLH, Ménard CB, Edwards JM, Hendry MA, Porson A, Gedney N *et al.* 2011. The Joint UK Land Environment Simulator (JULES), Model description-Part I: energy and water fluxes. *Geoscientific Model Development* **4**: 677–699.

Bonan GB, Levis S, Sitch S, Vertenstein M, Oleson KW. 2003. A dynamic global vegetation model for use with climate models: concepts and description of simulated vegetation dynamics. *Global Change Biology* **9**: 1543–1566.

Brando PM, Goetz SJ, Baccini A, Nepstad DC, Beck PSA, Christman MC. 2010. Seasonal and interannual variability of climate and vegetation indices across the Amazon. *Proceedings of the National Academy of Sciences, USA* **107**: 14685–14690.

Brando PM, Nepstad DC, Davidson EA, Trumbore SE, Ray D, Camargo P. 2008. Drought effects on litterfall, wood production and belowground carbon cycling in an Amazon forest: results of a throughfall reduction experiment. *Philosophical Transactions of the Royal Society, B* **363**: 1839–1848.

Brodribb TJ, Holbrook NM. 2003. Stomatal closure during leaf dehydration, correlation with other leaf physiological traits. *Plant Physiology* **132**: 2166–2173.

Chambers JQ, dos Santos J, Ribeiro RJ, Higuchi N. 2001. Tree damage, allometric relationships, and above-ground net primary production in central Amazon forest. *Forest Ecology and Management* **152**: 73–84.

Clark DB, Mercado LM, Sitch S, Jones CD, Gedney N, Best MJ, Pryor M, Rooney GG, Essery RLH, Blyth E *et al.* 2011. The Joint UK Land Environment Simulator (JULES), model description – Part 2: carbon fluxes and vegetation dynamics. *Geoscientific Model Development* **4**: 701–722.

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

da Costa ACL, Galbraith D, Almeida S, Portela BTT, da Costa M, de Athaydes Silva Junior J, Braga AP, de Gonçalves PHL, de Oliveira AAR, Fisher R *et al.* 2010. Effect of 7 yr of experimental drought on vegetation dynamics and biomass storage of an eastern Amazonian rainforest. *New Phytologist* **187**: 579–591.

Costa MH, Pires GF. 2010. Effects of Amazon and Central Brazil deforestation scenarios on the duration of the dry season in the arc of deforestation. *International Journal of Climatology* **30**: 1970–1979.

Cox PM, Betts RA, Jones CD, Spall SA, Totterdell IJ. 2000. Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature* **408**: 184–187.

Davidson EA, Nepstad DC, Ishida FY, Brando PM. 2008. Effects of an experimental drought and recovery on soil emissions of carbon dioxide, methane, nitrous oxide, and nitric oxide in a moist tropical forest. *Global Change Biology* **14**: 2582–2590.

Farquhar GD, Sharkey TD. 1982. Stomatal conductance and photosynthesis. *Annual Review of Plant Physiology* **33**: 317–345.

Farquhar GD, von Caemmerer S, Berry JA. 1980. A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species. *Planta* **149**: 78–90.

Fisher R, McDowell N, Purves D, Moorcroft P, Sitch S, Cox P, Huntingford C, Meir P, Woodward FI. 2010. Assessing uncertainties in a second-generation dynamic vegetation model caused by ecological scale limitations. *New Phytologist* **187**: 666–681.

Fisher RA, Williams M, da Costa AL, Malhi Y, da Costa RF, Almeida S, Meir P. 2007. The response of an Eastern Amazonian rain forest to drought stress: results and modelling analyses from a throughfall exclusion experiment. *Global Change Biology* **13**: 2361–2378.

Fisher RA, Williams M, Do Vale RL, Da Costa AL, Meir P. 2006. Evidence from Amazonian forest is consistent with isohydric control of leaf water potential. *Plant, Cell & Environment* **29**: 151–165.

Foley JA, Prentice IC, Ramankutty N, Levis S, Pollard D, Sitch S, Haxeltine A. 1996. An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. *Global Biogeochemical Cycles* **10**: 603–628.

Good P, Jones C, Lowe J, Betts R, Gedney N. 2013. Comparing tropical forest projections from two generations of Hadley Centre Earth System models, HadGEM2-ES and HadCM3LC. *Journal of Climate* **26**: 495–511.

Goudriaan J. 1977. *Crop micrometeorology: a simulation study.* Wageningen, the Netherlands: Center for Agricultural Publishing and Documentation.

Huete AR, Didan K, Shimabukuro YE, Ratana P, Saleska SR, Hutyra LR, Yang W, Nemani RR, Myneni R. 2006. Amazon rainforests green-up with sunlight in dry season. *Geophysical Research Letters* **33**: L06405.

Hutyra LR, Munger W, Saleska SR, Gottlieb E, Daube BC, Dunn AL, Amaral DF, de Camargo PB, Wofsy SC. 2007. Seasonal controls on the exchange of carbon and water in an Amazonian rain forest. *Journal of Geophysical Research* **112**: G03008.

Joetzjer E, Douville H, Delire C, Ciais P. 2013. Present-day and future Amazonian precipitation in global climate models: CMIP5 versus CMIP3. *Climate Dynamics*. doi:10.1007/s00382-012-1644-1.

Jupp TE, Cox PM, Rammig A, Thonicke K, Lucht W, Cramer W. 2010. Development of probability density functions for future South American rainfall. *New Phytologist* **187**: 682–693.

Kim Y, Knox RG, Longo M, Medvigy D, Hutyra LR, Pyle EH, Wofsy SC, Bras RL, Moorcroft PR. 2012. Seasonal carbon dynamics and water fluxes in an Amazon rainforest. *Global Change Biology* **18**: 1322–1334.

Kucharik CJ, Foley JA, Delire C, Fisher VA, Coe MT, Lenters JD, Young-Molling C, Ramankutty N, Norman JM, Gower ST. 2000. Testing the performance of a Dynamic Global Ecosystem Model: water balance,

carbon balance, and vegetation structure. *Global Biogeochemical Cycles* **14**: 795–825.

Levis S, Bonan GB, Vertenstein M, Oleson KW. 2004. *The Community Land Model Dynamic Global Vegetation Model (CLM-DGVM): technical description and user's guide*. Boulder, CO, USA: National Center for Atmospheric Research.

Li W, Fu R, Dickinson RE. 2006. Rainfall and its seasonality over the Amazon in the 21st century as assessed by the coupled models for the IPCC AR4. *Journal of Geophysical Research* **111**: D02111.

Lintner B, Biasutti M, Diffenbaugh NS, Lee J-E, Niznik MJ, Findell KL. 2012. Amplification of wet and dry month occurrence over tropical land regions in response to global warming. *Journal of Geophysical Research* **117**: D11106.

Malhi Y, Roberts JT, Betts RA, Killeen TJ, Li W, Nobre CA. 2008. Climate change, deforestation, and the fate of the Amazon. *Science* **319**: 169.

Medvigy D, Wofsy SC, Munger JW, Hollinger DY, Moorcroft PR. 2009. Mechanistic scaling of ecosystem function and dynamics in space and time: Ecosystem Demography model version 2. *Journal of Geophysical Research* **114**: G01002.

Meir P, Brando PM, Nepstad D, Vasconcelos S, Costa ACL, Davidson E, Almeida S, Fisher RA, Sotta ED, Zarin D *et al.* 2009. The effects of drought on Amazonian rain forests. In: Keller M, Bustamante M, Gash J, Silva Dias P, eds. *Amazonia and global change, Geophysical Monograph Series* **186**: 429–449.

Meir P, Metcalfe DB, Costa ACL, Fisher RA. 2008. The fate of assimilated carbon during drought: impacts on respiration in Amazon rain forests. *Philosophical Transactions of the Royal Society, London B. Biological Sciences* **363**: 1849–1855.

Metcalfe DB, Lobo-do-Vale R, Chaves MM, Maroco JP, Aragão LEOC, Malhi Y, Da Costa AL, Braga A, Gonçalves PL, De Athaydes J *et al.* 2010b. Impacts of experimentally imposed drought on leaf respiration and morphology in an Amazon rainforest. *Functional Ecology* **24**: 524–533.

Metcalfe DB, Meir P, Aragão LEOC, Da Costa ACL, Braga AP, Gonçalves PHL, De Athaydes Silva J Jr, de Almeida SS, Dawson LA, Malhi Y *et al.* 2008. The effects of water availability on root growth and morphology in an Amazon rainforest. *Plant and Soil* **311**: 189–199.

Metcalfe DB, Meir P, Aragão LEOC, Lobo-do-Vale R, Galbraith D, Fisher RA, Chaves MM, Maroco JP, da Costa ACL, de Almeida SS *et al.* 2010a. Shifts in plant respiration and carbon use efficiency at a large-scale drought experiment in the eastern Amazon. *New Phytologist* **187**: 608–621.

Metcalfe DB, Meir P, Aragão LEOC, Malhi Y, da Costa ACL, Braga A, Gonçalves PHL, de Athaydes J, de Almeida SS, Williams M. 2007. Factors controlling spatio-temporal variation in carbon dioxide efflux from surface litter, roots, and soil organic matter at four rain forest sites in the eastern Amazon. *Journal of Geophysical Research – Biogeosciences* **112**: G04001.

Moorcroft PR, Hurtt GC, Pacala SW. 2001. A method for scaling vegetation dynamics: the Ecosystem Demography model (ED). *Ecological Monographs* **71**: 557–587.

Muller B, Pantin F, Génard M, Turc O, Freixes S, Piques M, Gibon Y. 2011. Water deficits uncouple growth from photosynthesis, increase C content, and modify the relationship between C and sink organs. *Journal of Experimental Botany* **62**: 1715–1729.

Nepstad DC, Moutinho PR. 2008. *LBA-ECO LC-14 Rainfall Exclusion Experiment, LAI, Gap Fraction, TNF, Brazil: 2000-05*. Data set. [WWW document] URL http://lba.cptec.inpe.br/ from LBA Data and Information System, National Institute for Space Research (INPE/CPTEC), Cachoeira Paulista, Sao Paulo, Brazil [accessed 17 November 2012].

Nepstad DC, Moutinho P, Dias-Filho MB, Davidson E, Cardinot G, Markewitz D, Figueiredo R, Vianna N, Chambers J, Ray D *et al.* 2002. The effects of partial throughfall exclusion on canopy processes, aboveground production, and biogeochemistry of an Amazon forest. *Journal of Geophysical Research* **107**(D20): 8085.

Nepstad DC, Tohver IM, Ray D, Moutinho P, Cardinot G. 2007. Mortality of large trees and lianas following experimental drought in an Amazon forest. *Ecology* **88**: 2259–2269.

Oleson KW, Niu G-Y, Yang Z-L, Lawrence DM, Thornton PE, Lawrence PJ, Stöckli R, Dickinson RE, Bonan GB, Levis S *et al.* 2008. Improvements to the Community Land Model and their impact on the hydrological cycle. *Journal of Geophysical Research* **113**: G01021.

Richardson AD, Anderson RS, Altafarain M, Barr AG, Bohrer G, Chen G, Chen JM, Ciais P, Davis KJ, Desai A *et al.* 2012. Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis. *Global Change Biology* **18**: 566–584.

Rosolem R, Shuttleworth WJ, Gonçalves LGG. 2008. Is the data collection period of the Large-Scale Biosphere-Atmosphere Experiment in Amazonia representative of long-term climatology? *Journal of Geophysical Research* **113**: G00B09.

Sakaguchi K, Zeng X, Christoffersen BJ, Restrepo-Coupe N, Saleska SR, Brando PM. 2011. Natural and drought scenarios in an east central Amazon forest: fidelity of the Community Land Model 3.5 with three biogeochemical models. *Journal of Geophysical Research* **116**: G01029.

Sellers PJ, Randall DA, Collatz GJ, Berry JA, Field CB, Dazlich DA, Zhang C, Collelo GD, Bounoua L. 1996. A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: Model formulation. *Journal of Climate* **9**: 676–705.

Sotta ED, Veldkamp E, Schwendenmann L, Guimarães BR, Paixão RK, Ruivo MLP, da Costa ACL, Meir P. 2007. Effects of an induced drought on soil carbon dioxide (CO_2) efflux and soil CO2 production in an eastern Amazonian rainforest, Brazil. *Global Change Biology* **13**: 2218–2229.

Taylor JR. 1997. *An introduction to error analysis: the study of uncertainties in physical measurements, 2nd edn*. Sausalito, CA, USA: University Science Books, 49–61.

Williams M, Rastetter EB, Fernandes DN, Goulden ML, Wofsy SC, Shaver GR, Melillo JM, Munger JW, Fan S-M, Nadelhoffer KJ. 1996. Modelling the soil–plant–atmosphere continuum in a *Quercus-Acer* stand at Harvard Forest: the regulation of stomatal conductance by light, nitrogen and soil/plant hydraulic properties. *Plant, Cell & Environment* **19**: 911–927.

Williams M, Schwarz PA, Law BE, Irvine J, Kurpius MR. 2005. An improved analysis of forest carbon dynamics using data assimilation. *Global Change Biology* **11**: 89–105.

Williams M, Shimabukuro YE, Herbert DA, Pardi Lacruz S, Renno C, Rastetter EB. 2002. Heterogeneity of soils and vegetation in an eastern Amazonian rain forest: Implications for scaling up biomass and production. *Ecosystems* **5**: 692–704.

Würth MKR, Peláe-Riedl S, Wright SJ, Körner C. 2005. Non-structural carbohydrate pools in a tropical forest. *Oecologia* **143**: 11–24.

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 Monthly rainfall measured at TNF from 2002 to 2004 and CAX from 2001 to 2008.

Fig. S2 Change in aboveground biomass (AGB; kg C m^{-2}) with an 80% reduction (d80) in precipitation relative to the control (d0) simulations shown in Fig. 1.

Fig. S3 Time series of published periodic measurements of soil respiration (R_s) for CAX (Sotta *et al.*, 2007) and concurrent individual model estimates for the control (d0) and treatment (d50) plots of CAX.

Fig. S4 Annual gross primary production (GPP; kg C m^{-2} yr^{-1}) for TNF and CAX.

Fig. S5 Full 7-yr time series of the monthly mean water-stress factor (β) for CLM3.5, ED2, IBIS, JULES and SiB3 at each treatment level for TNF and CAX.