North American Carbon Program (NACP) regional interim synthesis: Terrestrial biospheric model intercomparison

D.N. Huntzinger [a,*], W.M. Post [b], Y. Wei [b], A.M. Michalak [c], T.O. West [d], A.R. Jacobson [e,f], I.T. Baker [g], J.M. Chen [h], K.J. Davis [i], D.J. Hayes [b], F.M. Hoffman [b], A.K. Jain [j], S. Liu [k], A.D. McGuire [l], R.P. Neilson [m], Chris Potter [n], B. Poulter [o], David Price [p], B.M. Raczka [i], H.Q. Tian [q], P. Thornton [b], E. Tomelleri [r], N. Viovy [o], J. Xiao [s], W. Yuan [t], N. Zeng [u], M. Zhao [v], R. Cook [b]

a School of Earth Science and Environmental Sustainability, Northern Arizona University, P.O. Box 5694, Flagstaff, AZ 86011-5694, United States
b Earth Science Division, Oak Ridge National Laboratory, Oak Ridge, TN, United States
c Department of Global Ecology, Carnegie Institute for Science, Stanford, CA, United States
d Joint Global Change Research Institute, College Park, MD, United States
e NOAA Earth System Research Lab Global Monitoring Division, Boulder, CO, United States
f Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, United States
g Department of Atmospheric Sciences, Colorado State University, Fort Collins, CO, United States
h Department of Geography and Program in Planning, University of Toronto, Toronto, Ontario, Canada
i Department of Meteorology, The Pennsylvania State University, University Park, PA, United States
j Atmospheric Sciences, University of Illinois, Urbana Champaign, Urbana, IL, United States
k United States Geologic Survey National Center for EROS, Sioux Falls, SD, United States
l U.S. Geological Survey, Alaska Cooperative Fish and Wildlife Research Unit, University of Alaska Fairbanks, Fairbanks, AK, United States
m Department of Botany and Plant Pathology, University of Utah, Salt Lake City, UT, United States
n NASA Ames Research Center, Moffett Field, CA, United States
o Laboratoire des Sciences du Climat et de l'Environnement, LSCE, Gif sur Yvette, France
p Northern Forestry Centre, Natural Resources Canada, Edmonton, Alberta, Canada
q Ecosystem Dynamics and Global Ecology Laboratory, Auburn University, Auburn, AL, United States
r Max Planck Institute for Biogeochemistry, Jena, Germany
s Earth Systems Research Center, Institute for the Study of Earth, Oceans, and Space, University of New Hampshire, Durham, NH, United States
t College of Global Change and Earth System Science, Beijing Normal University, Beijing, China
u Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD, United States
v Numerical Terradynamics Simulation Group, University of Montana, Missoula, MT, United States

ARTICLE INFO

ABSTRACT

Understanding of carbon exchange between terrestrial ecosystems and the atmosphere can be improved through direct observations and experiments, as well as through modeling activities. Terrestrial biosphere models (TBMs) have become an integral tool for extrapolating local observations and understanding to much larger terrestrial regions. Although models vary in their specific goals and approaches, their central role within carbon cycle science is to provide a better understanding of the mechanisms currently controlling carbon exchange. Recently, the North American Carbon Program (NACP) organized several interim-synthesis activities to evaluate and inter-compare models and observations at local to continental scales for the years 2000–2005. Here, we compare the results from the TBMs collected as part of the regional and continental interim-synthesis (RCIS) activities. The primary objective of this work is to synthesize and compare the 19 participating TBMs to assess current understanding of the terrestrial carbon cycle in North America. Thus, the RCIS focuses on model simulations available from analyses that have been completed by ongoing NACP projects and other recently published studies. The TBM flux estimates are compared and evaluated over different spatial ($1° \times 1°$ and spatially aggregated to different regions) and temporal (monthly and annually) scales. The range in model estimates of net ecosystem productivity (NEP) for North America is much narrower than estimates of productivity or respiration, with estimates of NEP varying between −0.7 and 2.2 PgC yr^{-1}, while gross primary productivity and heterotrophic respiration vary between 12.2 and 32.9 PgC yr^{-1} and 5.6 and 13.2 PgC yr^{-1}, respectively. The range in estimates from the models appears to be driven by a combination of factors, including the representation of photosynthesis, the source and of environmental driver data and the temporal

* Corresponding author. Tel.: +1 928 523 1669; fax: +1 928 523 7423.
E-mail address: deborah.huntzinger@nau.edu (D.N. Huntzinger).

variability of those data, as well as whether nutrient limitation is considered in soil carbon decomposition. The disagreement in current estimates of carbon flux across North America, including whether North America is a net biospheric carbon source or sink, highlights the need for further analysis through the use of model runs following a common simulation protocol, in order to isolate the influences of model formulation, structure, and assumptions on flux estimates.

1. Introduction

North America has been identified as both a significant source (e.g., fossil fuel emissions) and biospheric sink of atmospheric carbon dioxide (CO_2) (Gurney et al., 2002; CCSP, 2007; Prentice, 2001). However, as summarized in the State of the Carbon Cycle Report (SOCCR; CCSP, 2007), estimates of the North American biosphere carbon sink vary widely, ranging from less than $0.1\,PgC\,yr^{-1}$ to over $2.0\,PgC\,yr^{-1}$. While some of the mechanisms responsible for this sink are understood (e.g., forest regrowth), the current and future role of other mechanisms, such as extreme weather events (Jentsch et al., 2007), changes in land-use, CO_2 and nitrogen fertilization, natural disturbances (e.g., Kurz et al., 2007; Bond-Lamberty et al., 2007), and other carbon-climate feedbacks (Friedlingstein et al., 2006; Pan et al., 1998) in controlling the North American carbon cycle are highly uncertain (CCSP, 2007). Thus, a basic goal of carbon cycle studies has been to address key scientific questions ranging from carbon flux diagnosis (What are net carbon sources and sinks, and how do they change with time?), to attribution (What are the processes controlling flux variability?), and prediction (How might changes in climate and other factors alter future fluxes?). Understanding the sources and sinks of carbon and their distribution across North America is critical for the successful management of the carbon cycle (CCSP, 2007) and for useful predictions of its future evolution, and requires a strong understanding of carbon dynamics. Providing useful information about the carbon cycle and projecting future CO_2 concentrations is also urgently needed for informing policies addressing fossil fuel emissions.

Understanding of carbon exchange between terrestrial ecosystems and the atmosphere can be improved through direct observations and experiments, as well as through modeling activities. Terrestrial biosphere models (TBMs), sometimes called forward models, have become an integral tool for extrapolating local observations and understanding to much larger terrestrial regions (Waring and Running, 2007; Davis, 2008), as well as for testing hypotheses about how ecosystems will respond to changes in climate and nutrient availability. Although TBMs vary in their specific goals and approaches, their central role within carbon cycle research is to provide a better understanding of the mechanisms currently controlling carbon exchange. This understanding is then used as the basis of prediction and, ultimately to inform the development of any potential carbon management plans (Schimel et al., 2000).

The ultimate objective is to model all the processes that result in the net carbon exchange between the terrestrial system and the atmosphere, called the net ecosystem exchange (NEE). This includes many processes, most importantly gross primary production (GPP), autotrophic and heterotrophic respiration (Ra and Rh respectively, which together add up to ecosystem respiration, Re), and losses due to fire and other disturbance processes (herbivory, insects, disease, physical disturbance from storms, etc.) Therefore, understanding how TBM estimates of ecosystem photosynthesis, respiration, and net carbon exchange vary spatially and temporally is of great importance, not only for improving TBMs, but also for understanding their contribution to uncertainty in global climate simulations. By extension, it is also important to know why different TBMs product different estimates, even when forced with the same driving conditions. The former can be examined by bringing together existing model results and comparing them within a consistent framework, while the later requires a substantial, formal intercomparison effort.

Individual TBMs are often based on different simplifying assumptions, use different environmental driving data and initial conditions, and formulate the processes controlling carbon exchange in different ways. Thus, there is diversity in both the complexity of the model structure and formulation, as well as model estimates of regional net carbon exchange. Each TBM, therefore, is a complex combination of scientific hypotheses and choices, and their estimates depend on these inherent assumptions (Beer et al., 2010). Available observations of carbon flux components, as well as our current understanding of the processes controlling carbon exchange over regional scales, however, are not sufficient to rank models in terms of which is "best" at representing current fluxes or predicting carbon exchange under future climate conditions (Melillo et al., 1995). Therefore, in order to move towards more robust estimates of carbon cycle dynamics, we must first compare estimates from a variety of model types, as well as evaluate estimates against those measurements that are available (Cramer et al., 1999; Melillo et al., 1995; Beer et al., 2010).

Recently, the North American Carbon Program (NACP) (Denning et al., 2005; Wofsy and Harriss, 2002) organized several interim-synthesis activities to evaluate and inter-compare models and observations at local to continental scales for the time period of 2000 through 2005. These interim synthesis activities include three companion studies, each conducted on different spatial scales: (1) site-level analyses that examine process-based model estimates and observations at over 30 AmeriFlux[1] and Fluxnet-Canada[2] tower sites across North America; (2) a regional, mid-continent intensive study centered in the agricultural regions of the United States and focused on comparing inventory-based estimates of net carbon exchange with those from atmospheric inversions; and (3) a regional and continental synthesis evaluating model estimates against each other and available inventory-based estimates across North America. A number of other interim syntheses are underway, including ones focusing on non-CO_2 greenhouse gases, the impact of disturbance on carbon exchange, and coastal carbon dynamics.

Here, we compare the model estimates from the regional and continental interim-synthesis (RCIS) activities. The primary objective of this work is to synthesize and compare TBMs to assess current understanding of the terrestrial carbon cycle in North America. Thus, the RCIS focuses on "off-the-shelf" model simulations, i.e., existing model results currently available from analyses that have been completed by ongoing NACP projects and other recently published studies. Although there is a challenge in interpreting existing results compared to prescribing new simulations designed for the controlled comparison of different modeling system, there is also great value in using independent estimates to assess the overall spread or variability in model results. While it is necessary to limit variability between models (by, for example, prescribing consistent driver data and a detailed simulation protocol) in order to better understand what is driving the differences among model estimates, this approach provides an unrealistic assessment

---

[1] http://public.ornl.gov/ameriflux/.
[2] http://www.fluxnet-canada.ca/.

of the true uncertainty in our ability to model land-atmosphere carbon exchange. Models differ structurally in how they represent the processes controlling carbon exchange between the land and atmosphere, in their input or driver data (land cover, climate), and in the parameter values used within their varying process descriptions. These varying approaches to modeling terrestrial carbon exchange result in a large degree of variability in the land-atmosphere flux estimates. Thus, this work provides a valuable assessment of the current status of terrestrial carbon modeling in NA by bringing together model estimates that incorporate a wide range of modeling choices and input data. This work also serves as a starting point for analyses that compare these model results to different observational data products. Specifically, Raczka and Davis (personal communication) evaluated flux estimates of RCIS models against observations from 30 flux towers across a wide range of NA ecosystems. In addition, Hayes et al. (2012) has assembled and analyzed available agricultural and forest biomass inventory-based data for NA and compared them alongside estimates from TBM and inverse approaches available from the RCIS. In addition, ongoing work is comparing TBM estimates of net ecosystem exchange to flux estimates derived from atmospheric inversions. Flux estimates from atmospheric inverse models are more comprehensive, in the sense that all ecosystem sources and sinks, fossil fuel emissions, and any other processes emitting or absorbing CO_2 are, in principle, captured in the atmospheric signal (GCP, 2010). Combined, the comparison of TBM estimates to different observational data products and modeling approaches can provide further insight into our ability to model land-atmosphere carbon dynamics. This manuscript provides the foundation for these types of comparisons.

2. Overview of participating models

TBMs represent processes controlling carbon cycle dynamics; however, the level of detail with which processes are represented varies across models. Whereas some models are empirically or statistically-based with relatively simple relationships between driver variables and flux, others are more complex, simulating the coupled carbon, nutrient, and water cycles in terrestrial ecosystems. Models also differ in their representation of soil properties, vegetation type, and environmental forcings, as well as how carbon pools are initialized. Here we compare carbon flux estimates over North America (NA) for the 19 TBMs that participated in the RCIS. Key features of the models participating in this study in terms of how they represent photosynthesis, autotrophic respiration, decomposition, and other processes affecting carbon fluxes are summarized in Tables 1–3 (see Supplemental Material for additional model descriptions). The TBM flux estimates are evaluated over different land cover regions of NA, and with respect to photosynthetic formulation, soil carbon dynamics, and whether they explicitly account for the impact of fire disturbances on carbon pools and stocks.

TBMs can be divided into two general classes: diagnostic and prognostic models. In order to specify the internal (time-varying) state of the system, diagnostic models rely on forcing data (e.g., leaf area) provided directly or indirectly from satellite or other external sources. In contrast, the internal states of the system in prognostic models are computed as part of the system equations. Therefore, in principle, prognostic models can be used to predict future conditions using external climate forcing alone, in addition to being used for diagnostic analyses (e.g., reproducing past or measured fluxes).

The distinction between diagnostic and prognostic models is important. Diagnostic models frequently use observed leaf area index (LAI) as a specified driving variable, along with empirical algorithms of varying complexity, to estimate fluxes over regional or global domains and changes in carbon pool over time (Table 2, models: BEPS, CASA, NASA-CASA, CASA GFEDv2, EC-MOD, EC-LUE, ISAM, MODIS, MOD17+). Conversely, prognostic models determine the amount of leaf area as the result of carbon allocation and water balance dynamics within the model. As a result, they can project or estimate carbon cycle dynamics into the future under changing environmental conditions (Can-IBIS, CLM-CASA′, CLM-CN, DLEM, LPJ-wsl, MC1, ORCHIDEE, SiB3.1, TEM6, VEGAS2). In addition, some prognostic models also contain dynamic algorithms to estimate vegetation distribution over time (Can-IBIS, LPJ-wsl, MC1, ORCHIDEE, and VEGAS2). Although prognostic models can be used for future predictions, they are much less constrained by observations than diagnostic models. As a result, one would expect their results to be more variable (and perhaps less reliable) even when used in a diagnostic mode.

The model results submitted to the interim synthesis activity also vary in terms of the processes included, the choice of driving data, and the types of algorithms employed to represent these processes (Tables 1–3, Supplementary Information). For example, eight of the nineteen models represent photosynthesis using an enzyme kinetic formulation (Farquhar et al., 1980), normally at a sub-daily time step, while nine of the models use a light-use efficiency calculation at daily to monthly time steps. The models also differ in how they model soil carbon decomposition. Five of the models use a zero-order calculation, where decomposition is a function of temperature and moisture only. Two of the models omitted soil carbon decomposition altogether, and the remainder of the models represent decomposition through first-order kinetics, where decomposition depends on the magnitude of soil carbon stocks in addition to environmental drivers, and interactive processes such as N dynamics. In addition, models differ in the types of disturbance considered (e.g., wind or storm, fire, disease) and how these disturbances are included within the model (e.g., explicitly described or implicitly accounted for through vegetation indices). Most of the models in this study do not directly account for the impacts of fire, disease, or storm events on carbon fluxes or pools. In addition, those that do include the impact of fire disturbances (e.g., Can-IBIS, TEM6, MC1, LPJ-wsl) do so in varying ways (refer to Table 3 and Supplementary Information).

This diversity in model structure and process representation makes evaluation and comparison of model performance challenging. However, information on model differences helps to inform the analysis and was used here to define subsets or groups of models based on specific defining characteristics, and aid in the interpretation of observed differences.

3. Methods for comparison

Prior to analysis, all model output was processed, as necessary, to a spatial resolution of one-degree by one-degree, temporally aggregated to monthly fluxes, and placed on a grid with a spatial extent of 10–84° North, and 50–120° West. Fluxes are compared for the six years covering the period of 2000 through 2005.

3.1. Regional analysis of TBM output

Several of the model estimates lack full spatial coverage of North America (Fig. 1); therefore, in order to better compare net flux across models, $1° \times 1°$ flux estimates were spatially aggregated to regions defined by the TransCom intercomparison study (Gurney et al., 2002) and the Global Land Cover classification for 2000 (GLC2000; Latifovic et al., 2004; NRCan and USGS, 2003). The aggregation of fluxes to large contiguous regions, with similar land cover or biome types and climatic conditions, allows for the examination of regional differences between the models. This approach is similar

Table 1
Terrestrial biospheric models participating in the NACP regional interim synthesis.

Model	Spatial range	Native spatial resolution	Native temporal resolution	Fluxes submitted	Temporal range	Selected references
Can-IBIS	Canada and U.S.	–	30 min	GPP, NEE, NEP, NPP, Ra, Rh	2000–2005	Wang et al. (2011), Kucharik et al. (2000), and Foley et al. (1996)
CLM-CASA′	Global	2.8°	20 min	GPP, NPP, Rh, NEE, NEP	2000–2004	Randerson et al. (2009)
CLM-CN	Global	2.8°	20 min	GPP, NPP, Rh, NEE, NEP	2000–2004	Thornton et al. (2009) and Randerson et al. (2009)
DLEM	N. America	32 km	Daily	GPP, NEE, NPP, Ra, Rh	2000–2005	Tian et al. (2010)
ISAM	N. America	1°	Weekly	NEE, Rh, NPP	2000–2005	Jain and Yang (2005) and Yang et al. (2009)
LPJ-wsl	N. America	0.5°	Daily	GPP, NPP, Rh, NEE, CFire, NEEF	2000–2005	Bondeau et al. (2007) and Sitch et al. (2003)
MC1	Global, Continental U.S.	0.5°	Monthly	NPP, Rh, NEE, CFire, NEEF	2000–2005	Bachelet et al. (2000), Daly et al. (2000), and Lenihan et al. (2008)
ORCHIDEE	Global	0.5°	30 min	GPP, NPP, Rh, NEE, CO2Flux	2000–2005	Krinner et al. (2005) and Viovy et al. (2000)
SiB3	Global	1°	Hourly	NEE, GPP, Reco	2000–2005	Baker et al. (2008)
TEM6	N. A. > 45°N	0.5°	Monthly	GPP, NPP, Rh, NEE, CFire, NECB	2000–2005	McGuire et al. (2010) and Hayes et al. (2011)
VEGAS2	N. America	1°	Daily	GPP, NPP, Ra, Rh, NEE, CFire	2000–2005	Zeng (2003) and Zeng et al. (2004, 2005)
BEPS	N. America	1°	Hourly	GPP, NEE, NEP, NPP, Rh	2000–2004	Chen et al. (1999) and Ju et al. (2006)
CASA	Global	1°	Monthly	NEE	2002–2003	Randerson et al. (1997)
NASA CASA	Continental U.S.	8 km	Monhly	NPP, Rh, NEE, NEP	2001–2004	Potter et al. (2007)
CASA GFEDv2	Global	1°	Monthly	GPP, NPP, Rh, CFire, NEE	2000–2005	van der Werf et al. (2004, 2006)
EC-LUE	N. America	1°	Weekly	GPP	2004–2005	Yuan et al. (2007)
EC-MOD	N. America	1°	8-Day	GPP, NEE	2000–2006	Xiao et al. (2008, 2010, 2011)
MODIS	N. America	–	8-Day	GPP, annual NPP	2000–2005	Heinsch et al. (2003) and Running et al. (2004)
MOD17+	Global	0.5°	Daily	GPP, NEE, Reco	2000–2004	Reichstein et al. (2005)

Gross primary productivity (GPP); net ecosystem exchange (NEE); net ecosystem productivity (NEP); net primary productivity (NPP), autotrophic respiration (Ra); heterotrophic respiration (Rh); carbon emissions from fires (CFire); net ecosystem exchange including fire emissions (NEEF); net carbon flux including fire and disturbance (CO2Flux); ecosystem respiration (Reco); net ecosystem carbon balance (NECB).

to that used by Kicklighter et al. (1999), where net primary productivity (NPP) estimates were averaged across global biomes defined by the potential natural vegetation map developed by Melillo et al. (1993). The choice of land cover classification for defining spatially contiguous regions is somewhat subjective. As with the Potsdam model intercomparison study (e.g., Cramer et al., 1999; Kicklighter et al., 1999), landcover classification is used here solely as a mask for flux aggregation to smaller regions in order to examine regional differences among models.

The models used (or prognostically generated) different vegetation maps with varying classification schemes. Therefore, the choice of land cover scheme applied in this analysis does not reflect how well a model predicts flux for a particular biome type, but rather how predicted fluxes compare over large, spatially contiguous regions with similar land cover or climatic conditions. To avoid comparing models with limited spatial coverage in a region, only those models with at least 80% representation (i.e., those that estimate fluxes for at least 80% of the cells) in a given land region were included in the comparison within that region.

3.2. Subsetting models based on model formulation

In addition to comparing aggregated carbon fluxes, flux estimates were also compared by grouping models by their photosynthetic formulation and treatment of soil carbon dynamics (Table 2). Both the spread in model estimates and the across-model average for these different subsets were evaluated and compared. As mentioned above, the models in this study can be divided into two predominant photosynthetic formulation classes: light-use

efficiency (LUE) and enzyme kinetic (EK). Light-use efficiency models estimate productivity by quantifying the fraction of photosynthetically active radiation (fPAR) absorbed by the vegetation and then adjust the conversion of solar energy to photosynthesis or biomass production through climatological and physiological restrictions (e.g., temperature, moisture). Thus, carbon fixation is a strong function of solar radiation and leaf area index (LAI), or a proxy such as normalized vegetative difference index (NDVI). In contrast, models with enzyme kinetic formulations are more physiologically based, simulating photosynthesis using equations that represent biochemical/biophysical reactions driven by absorbed PAR, atmospheric CO_2 concentration, leaf temperature, and leaf water status (Farquhar et al., 1980). Thus, EK models quantify photosynthesis by emphasizing the light and enzyme limiting rates that affect photosynthesis. In addition to LUE and EK formulations, some models employ more statistical or regression-based approaches, modeling productivity as an empirical function of different environmental drivers. Photosynthetic formulation controls, to some extent, estimates of carbon uptake or productivity predicted by the models. Photosynthesis can also be influenced by other factors including driving meteorology, atmospheric CO_2 concentration, nutrient availability, and moisture and temperature limitations.

In addition to photosynthesis, models were grouped based on their treatment of soil carbon dynamics and decomposition. The CO_2 released (i.e., heterotrophic respiration, Rh) from the decomposition of above and below-ground dead organic matter is controlled by three factors, including: substrate quality and quantity, moisture availability, and temperature (Waring and Running, 2007). Thus, the degree to which these limitations are accounted for in the model

Table 2

Comparison of environmental drivers, vegetation and soil distribution, phenology, compartments, and photosynthetic and soil carbon decomposition formulations among models.

Model[a]	Vegetation distribution	Soil distribution	Weather/climate data	Phenology	# PFTs	# Veg pools	# Soil pools	Photo-synthetic formulation[b]	Soil carbon decomposition
Can-IBIS	Dynamic	CSL (Canada), STATSGO (Alaska), VEMAP (cont. U.S.)	Canadian Forest Services (CFS)	Prognostic	12	3	7	EK	1st Order
CLM-CASA'	MODIS	IGBP-DIS (GSDTG, 2000)	NCEP reanalysis	Prognostic	15	3	5	EK	1st Order
CLM-CN	MODIS	IGBP-DIS (GSDTG, 2000)	NCEP reanalysis	Prognostic	15	4	7	EK	1st Order, with N
DLEM	Multiple sources (Tian et al., 2010)	Zobler (1986)/FAO (1995/2003)	NARR and PRISM	Prognostic	21 + 10	7	3	EK	1st Order, with N
ISAM	Loveland and Belward (1997) and Haxeltine and Prentice (1996)	Zobler (1986)/FAO (1995/2003)	Mitchell et al. (2005)	–	13	5	8	LUE	1st Order, with N
LPJ-wsl	Dynamic	Zobler (1986)/FAO (1995/2003)	CRU TS 3.0	Prognostic	9	3	2	EK	1st Order
MC1	Dynamic	STATSGO	PRISM	Prognostic	6	7	6	Statistical	1st Order, with N
ORCHIDEE	Dynamic	Zobler (1986)/FAO (1995/2003)	CRU05 and NCEP reanalysis	Prognostic	12	8	8	EK	1st Order, with N
SiB3	IGBP	IGBP-DIS (GSDTG, 2000)	NARR	MODIS LAI	14	1	0	EK	Zero Order
TEM6	Loveland et al. (2000) and Hurtt et al. (2006)	IGBP-DIS (GSDTG, 2000)	CRU05 and NCEP reanalysis	Prognostic	23	1	3	EK	1st Order, with N
VEGAS2	Dynamic	Related to vegetation	CRU05 and NCEP reanalysis	Prognostic	4	3	6	LUE	1st Order
BEPS	GLC2000	STATSGO (SSS, 2011)	NCEP reanalysis	VGETATION LAI	6	4	9	EK	1st Order, with N
CASA	DeFries and Townshend (1994)	Zobler (1986)/FAO (1995/2003)	Leemans and Cramer (1991) and Hansen et al. (1999)	GIMMS NDVI derived LAI	11	3	5	LUE	1st Order
NASA CASA	MODIS	STATSGO (SSS, 2011)	NCEP reanalysis	MODIS EVI	11	3	5	LUE	1st Order, with N
CASA GFEDv2	MODIS	Batjes (1996)	IISAS, GISSTEMP, and GPCPv2	GIMMS NDVI derived LAI	3	3	5	LUE	1st Order
EC-LUE	–	–	GMAO/DAO	MODIS NDVI	–	–	–	LUE	–
EC-MOD	MODIS	–	–	MODIS EVI, LAI	7	0	0	statistical	Zero Order
MODIS	MODIS	–	DAO	MODIS LAI	–	0	–	LUE	–
MOD17+	SYNMAP, Jung et al. (2006)	–	ERA-Interim reanalysis	MODIS LAI	10	0	0	LUE	Zero Order

Shaded boxes refer to model components that are not considered or needed within the model.

[a] Model acronyms are defined and additional model information is provided in Supplementary Information.

[b] Enzyme kinetic (EK) and light-use efficiency (LUE).

will likely impact their estimations of Rh and overall net carbon dynamics.

Some models lack soil carbon pools/layers altogether, and heterotrophic respiration is thus not explicitly calculated. Others calculate soil respiration as an empirical function of moisture and temperature conditions (e.g., zero-order). In most models, however, soil organic matter decomposition is based on first-order kinetics, where the rate of decomposition is a function of the size of the soil carbon pool (e.g., amount of carbon), a simple decomposition constant, as well as temperature and moisture limitations (Reichstein and Beer, 2008). The influence of nitrogen (N) dynamics and cycling on soil carbon decomposition may or may not be considered by the model (Table 2). In this analysis, two soil carbon dynamics classifications are used: models with (1) dynamic soil carbon pools, with first-order soil carbon decomposition rates and (2) dynamic soil carbon pools that include nitrogen cycling and limitations, with first-order soil carbon decomposition rates. A few of the models consider zero-order soil decomposition, and therefore lack soil carbon pools altogether and were not included in the comparison of heterotrophic respiration.

Models were also classified by other factors that affect their dynamics, including whether they consider fire disturbances and land-use change; and whether transient CO_2, or the combination of transient CO_2 and N deposition forcings are included within the

model (Tables 1 and 3). Although many of these classifications are not mutually exclusive (e.g., many prognostic models use an EK photosynthetic formulation), their use of in model evaluation helps to identify potential sources of variability that drive differences in GPP and Rh, which translate into differences in net ecosystem productivity (NEP).

4. Results and discussion

4.1. Magnitude and distribution of carbon sources and sinks

The carbon flux that all the models submitted to the RCIS have in common is net ecosystem production (NEP), where NEP is the difference between GPP and the sum of autotrophic and heterotrophic respiration (Chapin et al., 2006). NEP does not include direct disturbance-induced carbon fluxes, which many models in this study do not consider. If a model does consider disturbances (Table 3), however, this can alter carbon pools, and as a result, impact both NPP and Rh. In some models, such as Can-IBIS, the effects of disturbances on NEP are only accounted for at year's end. As a result, if NEP is compared over the summer months (June, July, August), the flux estimates from these months will not account for losses due to disturbance. Instead, disturbances will cause additions

Table 3
Components and processes (including disturbance events) influencing the estimation of net ecosystem productivity by each model.

Model[a]	NEP[b]	Land-use/land cover change	Fire[c]	Insect, storm damage[d]	Transient forcings[e]	DIC, DOC, PC losses[f]
Can-IBIS	GPP − (Ra+Rh)	–	Prognostic	–	CO_2, Ndep	DOC
CLM-CASA'	GPP − (Ra+Rh)	Prescribed land-use	–	–	CO_2	–
CLM-CN	GPP − (Ra+Rh)	Prescribed land-use	Prognostic	–	CO_2, Ndep	–
DLEM	GPP − (Ra+Rh)	Prescribed land-use	–	–	CO_2, Ndep	CH_4 loss
ISAM	NPP − Rh	Prescribed land-use	–	–	CO_2, Ndep	–
LPJ-wsl	GPP − (Ra+Rh)	–	Prognostic	–	CO_2	–
MC1	NPP − Rh	Prescribed land-use, prognostic forest harvest	Prognostic	–	CO_2, Ndep	–
ORCHIDEE	GPP − (Ra+Rh) − crop harvest	No land-use/land-cover change, 40% of cropland biomass is harvested	–	–	CO_2	–
SiB3.1	GPP − (Ra+Rh)	–	–	–	CO_2	–
TEM6	GPP − (Ra+Rh)	Prescribed land-use, and forest harvest	Prescribed	–	CO_2, Ndep	DOC
VEGAS2	GPP − (Ra+Rh)	–	–	Constant background mortality rate from cold and drought stress	CO_2	–
BEPS	GPP − (Ra+Rh)	–	–	–	CO_2	–
CASA	NPP − Rh	–	–	–	–	–
NASA CASA	NPP − Rh	–	–	–	CO_2, Ndep	–
CASA GFEDv2	NPP − Rh	–	Prescribed	–	–	–
EC-LUE	GPP only	–	–	–	–	–
EC-MOD	-NEE	–	–	–	–	–
MOD17+	GPP − Re	–	–	–	–	–

Shaded boxes refer to processes that are not included or considered in the model.

[a] Model acronyms are defined and additional model information is provided in Supplementary Information.

[b] Net ecosystem productivity (NEP), gross primary productivity (GPP), heterotrophic respiration (Rh), autotrophic respiration (Ra).

[c] Models without prognostic or prescribed.

[d] Transient atmospheric carbon dioxide concentration (CO_2), transient nitrogen deposition (Ndep).

[f] Dissolved inorganic carbon (DIC), dissolved organic carbon (DOC), particulate carbon (PC).

to litter pools and removals of live vegetation at year end, which will affect the NEP in the following (and subsequent) years.

The spatial distribution of average summer (June, July, August) NEP predicted by the models is shown in Fig. 1. Table 3 provides a list of processes or factors that influence each model's estimate of productivity. Although, as mentioned above, the direct and indirect effects of fires influence some model estimates of carbon flux and pools, direct CO_2 emissions from forest fires are not included in model NEP estimates. Throughout the following discussion a positive (+) sign on NEP indicates net uptake of carbon from the atmosphere by the land, while a negative (−) sign signifies a net release of carbon from the land back to the atmosphere. During the growing season, the magnitude and spatial distribution of fluxes vary substantially among the models (Fig. 1). Some models show strong carbon sources in the Midwest and Southeast portions of the U.S. (e.g., MC1, LPJ-wsl), Central Plains, West, and Southwest (LPJ-wsl, MOD17+, DLEM), while others estimate large sinks particularly in the Southeast (e.g., BEPS, EC-MOD, NASA-CASA, Can-IBIS). In the boreal regions of North America, however, there appears to be more consistency among the models. In these northern regions, most models show an overall sink of carbon during the summer months, although the strength of that sink varies across models (Fig. 1).

The overall similarities and differences among modeled mean NEP estimates were quantified for each one-degree cell by calculating the across-model standard deviation in estimated flux. During the summer months of June, July, and August, the largest differences between NEP estimates are located in the Midwestern and Southeast regions of the continental U.S. (Fig. 2). Much of the across-model spread in summertime NEP in the southeast is driven by differences in predicted GPP (Fig. 2). Overall, as expected, the greatest difference in model estimates occurs in areas of larger flux magnitude.

When fluxes are spatially aggregated to all of North America, the TBMs predict annual NEP ranging from −0.7 to +1.7 PgC yr^{-1} for prognostic models and −0.3 to +2.2 PgC yr^{-1} for diagnostic models, with an overall model average of +0.65 PgC yr^{-1} for the North American continent (Table 4). This model average is consistent with previous estimates of the strength of the North American sink of 0.35–0.75 PgC yr^{-1} (Goodale et al., 2002; Houghton et al., 1999; CCSP, 2007; Pacala et al., 2001; Xiao et al., 2011). Much of the spread in NEP estimates comes from the range in model estimates of photosynthesis or GPP, because the majority of models scale autotrophic respiration (Ra) based on their estimates of photosynthesis. TBM estimates of GPP and heterotrophic respiration for North America vary considerably between 12.2 and 32.9 PgC yr^{-1} and 5.6 and 13.2 PgC yr^{-1}, respectively (Table 4). Overall, prognostic models exhibit greater across-model spread or variability in their net GPP estimates relative to diagnostic models. Prognostic models also estimate a larger net GPP or uptake across North America compared to diagnostic models.
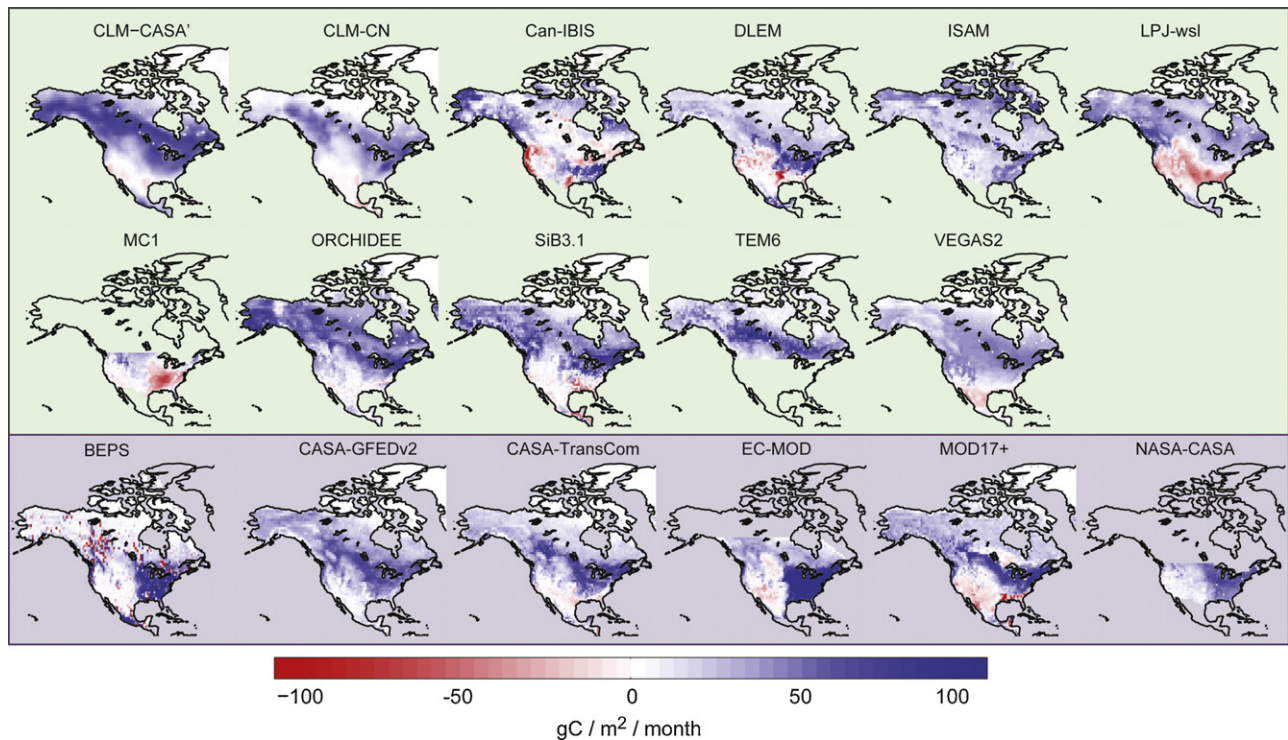
Fig. 1. Long-term mean summer (June, July, August) net ecosystem productivity by model (2000–2005). A positive sign indicates net terrestrial carbon uptake from the atmosphere, while a negative sign signifies net carbon release to the atmosphere. Prognostic models are shown above with a green background; diagnostic models are below with a purple background.

One potential reason for the narrower spread in GPP among the diagnostic models is that several of the diagnostic models (EC-LUE, EC-MOD, MOD17+) presented in this study are calibrated to flux tower data and use similar satellite observations for providing LAI and fPAR. As a result, their flux estimates tend to be more similar among themselves relative to the differences among prognostic models. However, only three of the eight diagnostic models explicitly calibrate their models using flux tower data, so this is un-likely to be the only cause of similarly among the diagnostic models.

It is surprising that diagnostic models have a greater range and standard deviation in NEP than prognostic models, given that diagnostic models have smaller ranges in the component fluxes GPP and Rh (Table 4). This indicates that the production and respiration components are less correlated within diagnostic models.

Fluxes were also spatially aggregated to Boreal and Temperate North America; regions defined by the TransCom inverse model intercomparison (Gurney et al., 2003). The TransCom regions were chosen for comparison because they cover a majority of North America (minus Greenland, the Northern Queen Elizabeth Islands,
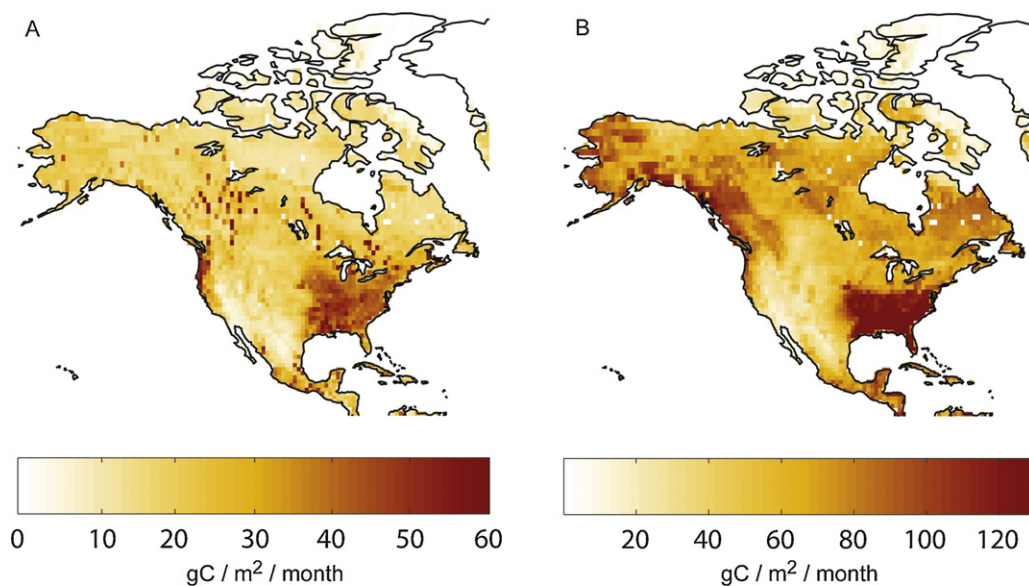

Fig. 2. Across-model standard deviation in long-term mean (2000–2005) summer (June, July, August) model estimates of (A) net ecosystem productivity and (B) gross primary productivity.
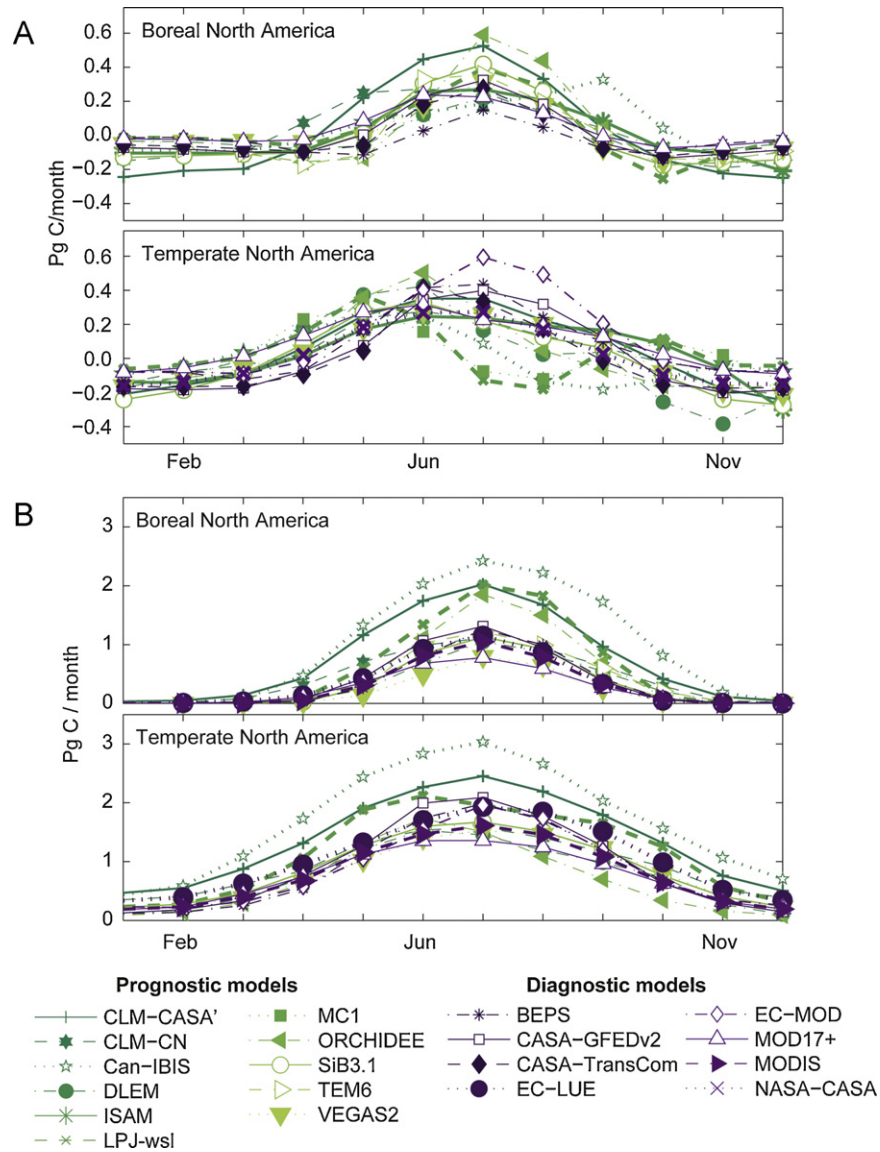
Fig. 3. Model estimates of the long-term mean (2000–2005) seasonal cycle of (A) net ecosystem productivity and (B) gross primary productivity for boreal and temperate North America.
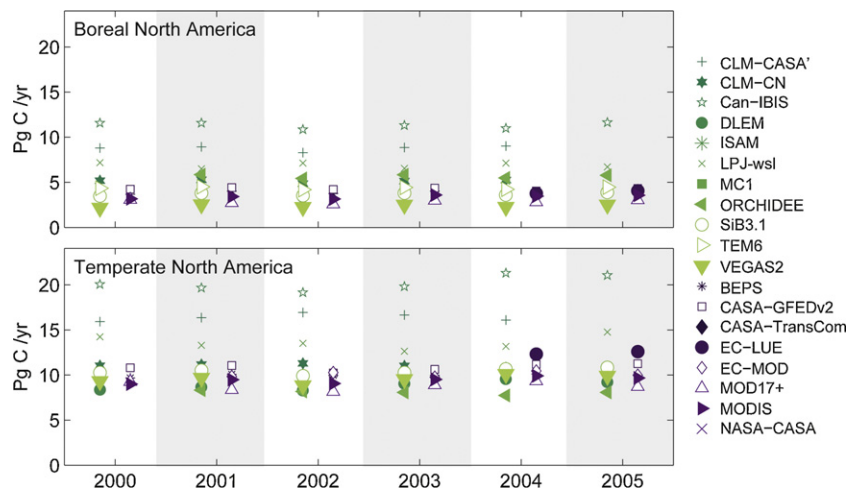

Fig. 4. Model estimates of annual gross primary productivity (GPP) for 2000 through 2005 for Boreal and Temperate North America. Prognostic models are shown in shades of green; diagnostic models are shown in purple.

Table 4

Long-term mean (2000–2005) net ecosystem productivity, gross primary productivity, and heterotrophic respiration estimated by the models in PgC yr^{-1} for North America. Not all models submitted all three fluxes (NEP, GPP, and Rh). To avoid comparing models with limited spatial coverage in a region, only those models with at least 80% representation (i.e., those that estimate fluxes for at least 80% of the cells) in a given land region were included in the comparison within that region.

	Prognostic models				Diagnostic models			
	Number of models	(min, max)	Mean	Std dev	Number of models	(min, max)	Mean	Std dev
Net ecosystem productivity ($n = 17$)								
North America	9	(−0.7, 1.7)	0.4	0.4	6	(−0.3, 2.2)	0.9	0.7
Boreal NA	10	(−0.2, 0.7)	0.1	0.2	4	(−0.4, 0.6)	0.1	0.3
Temperate NA	10	(−0.5, 1.1)	0.2	0.3	6	(−0.1, 1.6)	0.7	0.6
Gross primary productivity ($n = 15$)								
North America	8	(12.2, 32.9)	20.0	6.6	6	(12.2, 18.7)	14.8	1.9
Boreal NA	9	(2.2, 11.6)	5.7	2.7	5	(2.6, 4.4)	3.6	0.6
Temperate NA	8	(7.7, 21.3)	12.3	4.0	6	(8.2, 12.6)	10.0	1.0
Heterotrophic respiration ($n = 13$)								
North America	8	(5.6, 13.2)	8.2	2.3	2	(7.4, 8.6)	8.2	–
Boreal NA	9	(1.3, 4.6)	2.6	1.1	2	(2.1, 2.9)	2.4	–
Temperate NA	9	(3.4, 7.5)	4.8	1.3	3	(2.4, 5.6)	4.5	–

Central America, and parts of southern Mexico). Estimates of NEP and GPP by prognostic versus diagnostic models differ considerably in both the depth and timing of the seasonal cycle, with prognostic models estimating greater overall productivity during the summer months compared to diagnostic models (Fig. 3). These seasonal cycle differences translate into large variability in net annual estimates of NEP for 2000–2005, ranging from −0.4 to 0.7 PgC yr^{-1} (Boreal NA) and −0.5 to 1.6 PgC yr^{-1} (Temperate NA) (Table 4).

The differences among TBMs are even more apparent when comparing GPP over similar land regions. Overall, prognostic models exhibit a significantly greater across model variability in net annual uptake than diagnostic models (Table 4 and Fig. 4). In order

to examine regional differences among the models that may be contributing to variation in their estimates of North American net annual flux, model estimates of NEP, GPP, and Rh were compared across biomes (Fig. 5). To better compare model estimates, aggregated fluxes were normalized by the total land area covered by a given model for a given land cover region, and therefore the results are presented as gC m^{-2} yr^{-1}. Recall, that to be included in the comparison for a given biome, a model must have at least 80% spatial coverage within that region.

Model estimates vary considerably in their net annual estimates of flux with the greatest discrepancies occurring in more productive regions (e.g., mixed and deciduous forest, cultivated and
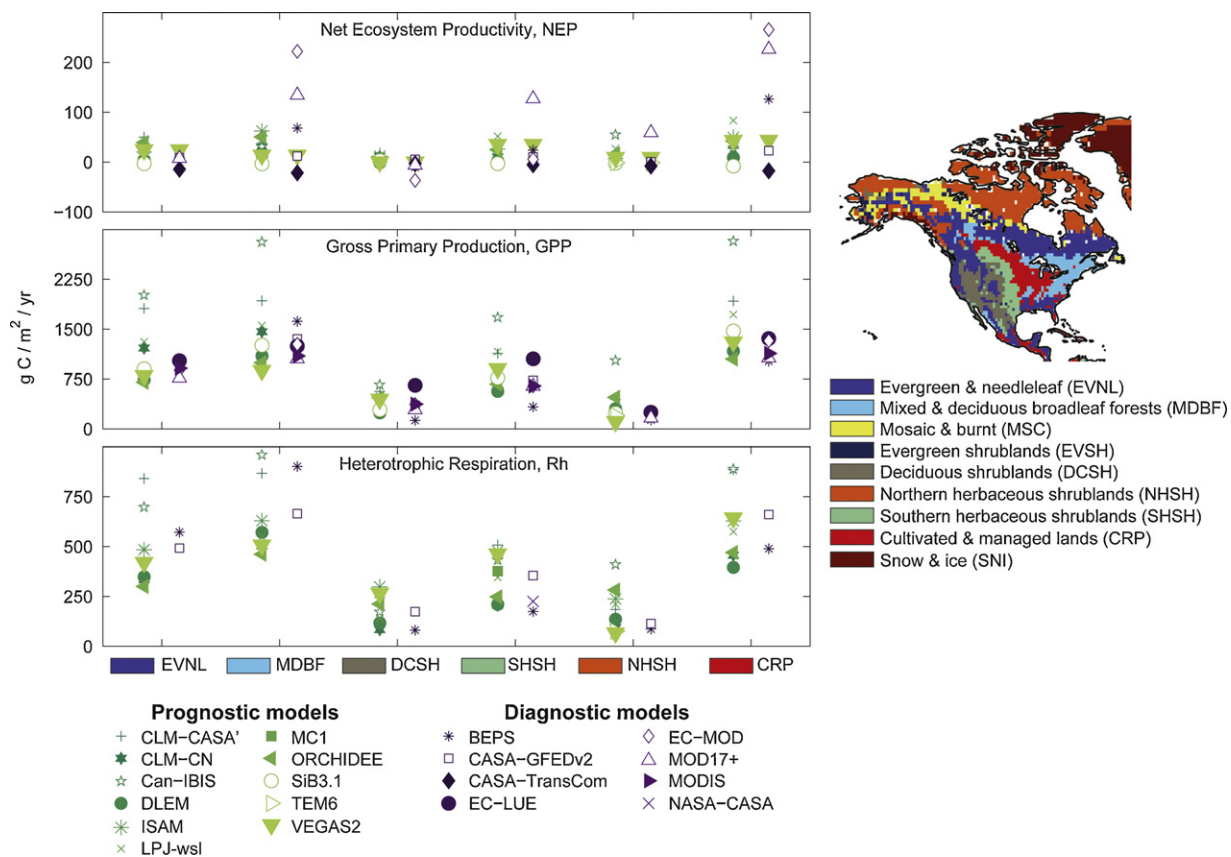

Fig. 5. Model estimates of the long-term mean (2000–2005) net ecosystem productivity (NEP), gross primary productivity (GPP) and heterotrophic respiration (Rh) by biome. Biome or vegetative cover classification based on the Global Land Cover 2000 classification scheme.

managed lands; Fig. 5). Model estimates of the long-term mean annual NEP in mixed and deciduous forested regions varies from about $-25\,\mathrm{gC\,m^{-2}\,yr^{-1}}$ to $+250\,\mathrm{gC\,m^{-2}\,yr^{-1}}$. One explanation for this difference is that models, and their estimates of GPP and Rh, have varying sensitivities to limitations, such as water availability and temperature. In low productivity systems (e.g., shrublands), limitations are likely strong regardless of a given model's sensitivity to these limitations. In more productive systems (e.g., forests and cultivated lands), however, a model's sensitivity to limiting factors (e.g. water availability) will have a much larger effect, and slight differences in the sensitivity of GPP and Rh to these limitations could result in more divergent NEP estimates. In addition, from examination of model estimates of long-term mean seasonal cycle at the biome level, it appears that across-model differences in growing season net uptake may be driving some of the average annual NEP and GPP variability among models. Conversely, a similar range in estimated NEP is seen in areas of cultivated and managed lands. For most models, NEP is calculated as the difference between GPP and ecosystem respiration (Rh + Ra). Model estimates of GPP and Rh vary considerably across biomes. However, in more productive areas (e.g., deciduous shrublands, evergreen and needleleaf), the larger productivity results in more decomposable substrate. As a result, Rh tends to be highly correlated with GPP, which yields relatively similar estimates of NEP across models compared to other regions (variability Rh and GPP somewhat cancel each other out).

The potential factors driving the differences seen across models are examined further below by subsetting models based on shared model attributes.

4.2. Attribution of intermodel differences to model formulation and driver data

Attribution of intermodel differences in net flux and the long-term mean seasonal cycle of NEP can best be examined through the component fluxes of GPP (photosynthetic uptake) and respiratory release of carbon (Rh). Thus, in order to identify potential drivers of differences between models, we compare estimates of component fluxes (e.g., GPP and Rh) by subsetting models based on differences in their photosynthetic and soil carbon decomposition formulations, as well as their treatment of fire disturbance, land cover change and external forcings, such as time-varying CO_2 and N deposition.

4.2.1. Differences in gross primary productivity

It is generally assumed that the physiology of photosynthesis and the kinetics of Rubisco are relatively well understood at the leaf-level (Collatz et al., 1992; Dai et al., 2004; Farquhar and von Caemmerer, 1982). However, there is a great deal of uncertainty as to how to scale leaf-level processes up to the canopy or ecosystem level (Chen et al., 1999; Baldocchi and Amthor, 2001). In addition, there are uncertainties concerning the exact influence of factors such as nitrogen content, nitrogen allocation, and radiative transfer on productivity. These processes must be parameterized in models, and can lead to a potentially large spread in GPP estimates across a collection of models. The complications in modeling productivity leads to significant disagreement among the model estimates of GPP, with peak growing season differences of greater than $2\,\mathrm{PgC\,month^{-1}}$ in both Temperate and Boreal NA TransCom regions (Fig. 3), and over $1000\,\mathrm{gC\,m^{-2}\,yr^{-1}}$ in regions of mixed and deciduous broadleaf forests and cultivated and managed lands (Fig. 5).

Overall, models with photosynthetic formulations based on enzyme kinetics predict a greater mean annual GPP with a larger range in estimates than light-use efficiency-based models (Fig. 6). Whether photosynthetic formulation is the driving cause of variability in modeled GPP is not clear. For example,

Medvigy et al. (2010) found that high-frequency meteorological data profoundly impacts simulated terrestrial carbon dynamics. Using the Ecosystem Demography model version 2 (ED2) forced with observed meteorology, as well as reanalysis weather, this study found that over an 8-year period, differences in climatic driver data alone resulted in a 10% difference in net GPP and 25% difference in NEP. This work suggests that precipitation and radiation data with higher temporal variability yield lower overall GPP and cumulative above ground biomass, due to non-linearities in the photosynthetic functions. Conversely, climate drivers with lower variability, e.g., from reanalysis weather products, may lead to higher GPP (Medvigy et al., 2010). Model estimates of GPP and NEP are also highly sensitive to biases in solar radiation (e.g., Ricciuto et al., in prep, Poulter et al., 2011; Zhao et al., 2011). Finally, many of the EK models examined in this study also model phenology prognostically, which could also explain much of the spread in GPP (Figs. 4 and 6) among the prognostic models. Therefore, much of the spread in GPP estimates in this study is likely to be driven by a combination of differences, including the source of driver data, the temporal variability of meteorological data, prognostic representation of phenology, and/or how changes in sunlight and precipitation affect productivity through the models' choice of photosynthetic formulation.

Disturbances can have a significant and immediate influence on ecosystems by redistributing stocks among live and dead organic matter pools and, in the case of fire, the atmosphere. Disturbances can also greatly alter the natural community (e.g., succession), which can influence biogeochemical cycling long after the direct impacts of a disturbance event have passed. To examine the potential impacts of a model's treatment of disturbance on GPP, models were grouped based on how they account for fire disturbances. Some models explicitly account for the effect of fire either prognostically or diagnostically (refer to Supplemental Information). However, a majority of the models in this study do not directly account for fire disturbances or do so implicitly through the use of satellite-based vegetative indices such as LAI or fPAR, which are themselves impacted by fire disturbance.

Overall, models that explicitly account for fire disturbances, and their associated impact on carbon pools, predict a greater mean annual GPP with a larger range in flux estimates than models without disturbance included (Fig. 6). The impacts of fire on a given ecosystem depend on a number of factors including the ecosystem type (e.g., ponderosa pine forest versus grasslands), fire intensity and type (i.e., stand replacing), and overall scale. For example, a large, stand-replacing fire would likely result in suppressed productivity (and GPP) for several years following the fire. Conversely, given the right conditions, a fire event could make more nitrogen available for growth (and thereby increase production of leaf tissue) and/or for photosynthesis (through higher leaf tissue N in the form of Rubisco). This, however, is balanced by any losses in leaf area during the fire. Many of the models that directly account for fire also employ an enzyme kinetic approach in their formulation of photosynthesis. Although, how a model accounts for disturbances (including fire) impacts their estimates of carbon pools and stocks, it is not likely the dominant driver for the differences in GPP seen among the participating models in this study.

There are limited datasets with which to compare modeled GPP. Although MODIS-derived estimates of GPP (Heinsch et al., 2006; Running et al., 2004; Zhao et al., 2005) have been favorably compared to flux tower measurements, tower-by-tower comparisons still show significant residuals. MODIS GPP is fundamentally a modeled product, not a direct observation. The MODIS product and other LUE-based models are similar in their estimates of net uptake, and generally predict lower productivity than models in which photosynthesis is more physiologically based (Figs. 4 and 6 and Table 3). For example, when totalled over the growing season and annually,
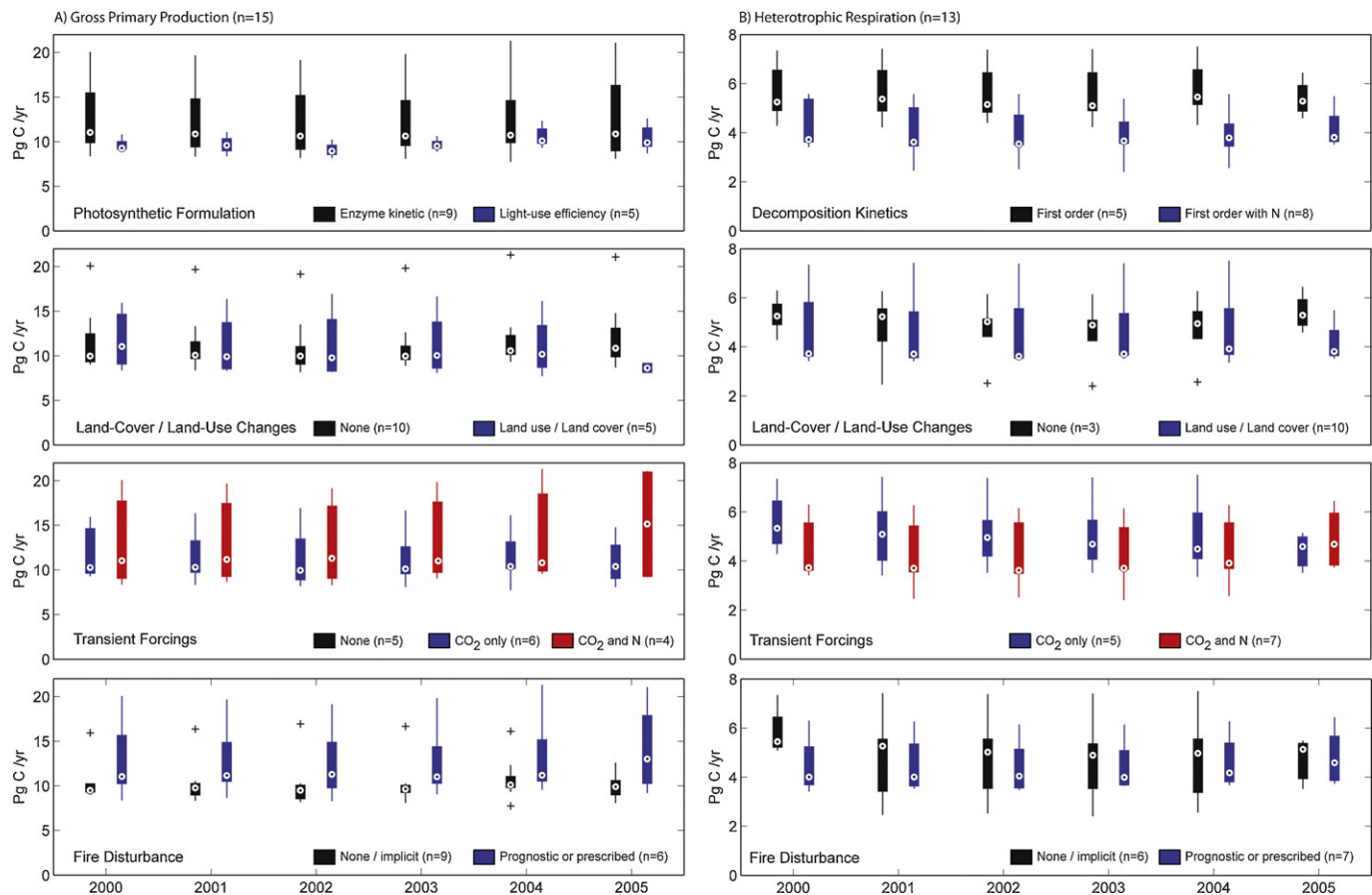
Fig. 6. Model estimates of (A) gross primary productivity (GPP) and (B) heterotrophic respiration (Rh) for Temperate North America, grouped by decomposition kinetics, photosynthetic formulation (enzyme kinetic versus light-use efficiency), and whether fire disturbance, land-cover/land-use changes, and transient forcings were considered by the models. See Tables 2 and 3 for more information.

many of the prognostic models in this study estimate 1.2–2 times the GPP predicted by the diagnostic or light-use efficiency based models. Razcka and Davis (personal communications) compared the TBM estimates in this study to flux tower measurements. They found that the mean GPP and ecosystem respiration (Ra + Rh) from the models is about 30–40% greater in most biomes (not including deciduous broadleaf forests) compared to those derived from eddy-covariance (EC) measurements at flux tower sites. As a result, although similarities exist between the lower end of the model-based GPP estimates and those derived from EC measurements, it is difficult to say whether these lower GPP estimates are more correct.

In addition to the influence of environmental drivers discussed above, whether a model accounts for time-varying CO_2 and/or nitrogen deposition could contribute to the differences in net carbon uptake simulated by the models (Fig. 6). Friedlingstein et al. (2006), for example, showed greater carbon uptake by ecosystems in uncoupled TBMs as a result of increased atmospheric CO_2 concentration.

4.2.2. Variability in heterotrophic respiration

Heterotrophic respiration is also difficult to model at a fundamental scale due to its dependence on poorly understood, complex processes, as well as the need to track diverse carbon pools of varying decomposability (Jastrow, 1996; Oades, 1988; Parton et al., 1987). While the overall magnitude in Rh is smaller than that of GPP, the variation among models is still large, with estimates differing by 50–600 gC m^{-2} yr^{-1} (Fig. 5). Models that estimate soil carbon decomposition based on zero-order kinetics (i.e., decomposition

rate independent of concentration) do not explicitly calculate Rh, and they are therefore not included in Fig. 6. Estimates of Rh from models with both first-order soil carbon decomposition rates, which also include nitrogen cycling, tend to exhibit a shallower seasonal cycle and less overall soil C release than models without N cycling. Nitrogen limitations on microbial decomposition could result in slower decomposition rates (Thornton et al., 2007; Yang et al., 2009). However, this in turn would reduce the rate of N availability for plant growth. The models that consider nitrogen deposition (in addition to CO_2) do not have lower GPP and may have a slightly larger GPP than the models that do not include N deposition (Fig. 6A). This added N from atmospheric deposition may, at least for North America, be enough to compensate for the reduction in N from decomposition, thus supplying the N required for GPP.

Overall, the differences in modeled GPP and Rh do not translate into large differences in the long-term mean seasonal cycle of NEP (Fig. 3), in part, because within many models respiration is highly correlated to GPP. This is also observed in other studies (e.g., Poulter et al., 2011) where modeled Rh tends to respond proportionally to changes in GPP or productivity, resulting in a smaller net range in absolute NEP among the models (Table 4).

5. Conclusions

This study brings together estimates of land-atmosphere carbon exchange from nineteen prognostic and diagnostic TBMs, in order to assess the current understanding of the terrestrial carbon cycle in North America. The models differ substantially in their estimates of net ecosystem productivity, as well as gross primary productivity

and respiration. Prognostic models exhibit greater overall range in their estimates and predict larger net uptake of carbon over North America relative to diagnostic models.

Photosynthetic formulation, the source and variability of climatic driver data, and how phenology is described all appear to influence the across-model difference in estimated fluxes, and the magnitude of overall carbon uptake predicted by the models. Much of the variability in modeled Rh is likely driven by variability in GPP, because the majority of models scale respiration based on their estimates of photosynthesis. While this type of scaling may be appropriate for forested regions where GPP and Rh are closely linked, this assumption is probably not appropriate for more managed lands (e.g., agricultural lands and forest plantations in the U.S. Southeast), where harvest, lateral transport, and other management activities can impact where carbon is respired.

For many biome types (e.g., evergreen and needleleaf, deciduous and herbaceous shrublands), there is a large range in both GPP and Rh, but a relatively small range in model-estimated NEP. This trend in simulation results is consistent with the work of Raczka and Davis (2011, personal communication), which compares model derived estimates of GPP and respiration to those inferred from flux tower observations. Thus, models that overestimate (or underestimate) GPP and Rh can still predict plausible values for NEP, but for the wrong reasons. For example, models that are calibrated to flux tower observations may be "tuned" to NEP, particularly when GPP and Rh observations are scarce. The flux tower records can help to interpret the cause of model difference, and suggest that the lower range of GPP in this collection of TBM models may be closer to tower-based observations. What we cannot tell from comparisons with observations, however, whether the model estimates reproduce observations for the right reasons (i.e., whether processes accurately are represented in the model).

Overall, flux estimates are a function not only of model algorithmic formulation, but also how models were calibrated (or tuned), initial conditions (e.g., soil properties, vegetation, and land-use), driver data (e.g., weather, CO_2 concentration), and their treatment of disturbances (e.g., fire, wind, disease). The entire modeling framework contributes to the results, and therefore all of the components require evaluation. The study reveals the large variation in TBM estimates of long-term mean net ecosystem productivity, as well as discrepancies in the magnitude and timing of the seasonal cycle. The results also provide a sobering picture of the current lack of consensus among model estimates of land-atmosphere carbon exchange across North America. Attributing the cross-model variability to differences in modeling approaches and driving data is difficult, however, given the focus on existing results from models run using a wide range of assumptions and inputs. Developing, improving, and evaluating TBMs such that they can provide useable forecasts (and past diagnoses) at near-term, inter-annual, decadal, and century timescales requires developments in quantitative model evaluation and rigorous benchmark development. While we were able to attribute some of this variation to model structure and aspects of model driver data, a more formal model-data comparison is required to more definitively quantify the impact of model formulation and supporting and driver data on the accuracy of the simulation outputs. Such efforts require substantial technical support for model participation, the development of consistent and optimal environmental driver datasets, a unified intercomparison protocol, as well as coordination of the intercomparison effort across research groups. These types of efforts are underway, including several projects working to understand how model formulation and model choices impact overall model performance through the use of detailed simulation protocol and controlled input environmental driver data (e.g., Schwalm et al., 2010) and the Multi-Scale Synthesis and Terrestrial Model Intercomparison Project (MsTMIP), which directly builds of the NACP regional interim synthesis present here.

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ecolmodel.2012.02.004.

References

Bachelet, D., Lenihan, J.M., Daly, C., Neilson, R.P., 2000. Interactions between fire, grazing and climate change at Wind Cave National Park, S.D. Ecological Modelling 134 (2–3), 229–244.

Baker, I.T., Prihodko, L., Denning, A.S., Goulden, M., Miller, S., da Rocha, H.A., 2008. Seasonal drought stress in the Amazon: reconciling models and observations. Journal of Geophysical Research-Biogeosciences, 113.

Baldocchi, D.D., Amthor, J.S., 2001. Canopy Photosynthesis: History, Measurements, and Models, Terrestrial Global Productivity. Academic Press, San Diego, ISBN 978-0-12-505290-0, doi:10.1016/B978-012505290-0/50003-X, pp. 9–31.

Batjes, N., 1996. Total carbon and nitrogen in the soils of the world. European Journal of Soil Science 47, doi:10.1111/j.1365-2389.1996.tb01386.x.

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., et al., 2010. Terrestrial gross cabon dioxide uptake: global distribution and covariation with climate. Science 329 (5993), 834–838, doi:10.1126/science.1184984.

Bond-Lamberty, B., Peckham, S.D., Ahl, D.E., Gower, S.T., 2007. Fire as the dominant driver of central Canadian boreal forest carbon balance. Nature 450, 89–92.

Bondeau, A., SmithPC, Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., Gerten, D., 2007. Modelling the role of agriculture for the 20th century global terrestrial carbon balance. Global Change Biology 13 (3), 679–706.

Carbon Cycle Science Program (CCSP), 2007. In: King, A.W., Dilling, L., Zimmerman, G.P., Fairman, D.M., Houghton, R.A., Marland, G., et al. (Eds.), The First State of the Carbon Cycle Report (SOCCR): The North American Carbon Budget and Implications for the Global Carbon Cycle. A Report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research,. National Oceanic and Atmospheric Administration, National Climatic Data Center, Asheville, NC, USA, p. 242.

Chapin, F.S., Woodwell, G.M., Randerson, J.T., Rastetter, E.B., Lovett, G.M., Baldocchi, D.D., et al., 2006. Reconciling carbon-cycle concepts, terminology, and methods. Ecosystems 9, 1041–1050, doi:10.1007/s10021-005-0105-7.

Chen, J.M., Liu, J., Cihlar, J., Goulden, M.L., 1999. Daily canopy photosynthesis model through temporal and spatial scaling for remote sensing applications. Ecological Modelling 124 (2–3), 99–119.

Collatz, G.J., Ribas-Carbo, M., Berry, J.A., 1992. Coupled photosynthesis-stomatal conductance model for leaves of c4 plants. Australian Journal of Plant Physiology 19 (5), 519–538.

Cramer, W., Kicklighter, D.W., Bondeau, A., Moore, B., Churkina, C., Nemry, B., et al., 1999. Comparing global models of terrestrial net primary productivity (NPP): overview and key results. Global Change Biology 5, 1–15.

Dai, Y.J., Dickinson, R.E., Wang, Y.P., 2004. A two-big-leaf model for canopy temperature, photosynthesis, and stomatal conductance. Journal of Climate 17 (12), 2281–2299.

Daly, C., Bachelet, D., Lenihan, J.M., Neilson, R.P., Parton, W., Ojima, D., 2000. Dynamic simulation of tree–grass interactions for global change studies. Ecological Applications 10 (2), 449–469.

Davis, K.J., 2008. Integrating field measurements with flux tower and remote sensing data. In: Hoover, C.M. (Ed.), Field Measurements For Landscape-Scale Forest Carbon Monitoring, XVIII. , ISBN 978-1-4020-8505-5, p. 242.

DeFries, R. S., Townshend, J.R.G., 1994. 1 Degree Global Land Cover Data Set Derived from AVHRR. Available on-line [http://glcf.umiacs.umd.edu/

data/landcover/index.shtml] from the Global Land Cover Facility, University of Maryland Institute for Advanced Computer Studies, College Park, Maryland, U.S.A.

Denning, A.S., et al., 2005. Science implementation strategy for the North American Carbon Program. In: Report of the NACP Implementation Strategy Group of the U.S. Carbon Cycle Interagency Working Group. U.S. Carbon Cycle Science Program, Washington, DC, p. 68.

FAO, 1995/2003. The Digitized Soil Map of the World and Derived Soil Properties (Version 3.5). FAO Land and Water Digital Media Series 1. FAO, Rome.

Farquhar, G.D., von Caemmerer, S., 1982. Modeling of photosynthetic response to environmental conditions. In: Lange, O.L., Nobel, P.S., Osmond, C.B., Zeigler, H. (Eds.), Physiological Plant Ecology. II. Water Relations and Carbon Assimilation. Springer-Verlag, New York.

Farquhar, G.D., Caemmerer, S.V., Berry, J.A., 1980. A biochemical-model of photosynthetic CO_2 assimilation in leaves of C-3 species. Planta 149 (1), 78–90.

Foley, J.A., Prentice, I.C., 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 (4), 603–628.

Friedlingstein, P., et al., 2006. Climate-carbon cycle feedback analysis: results from the (CMIP)-M-4 model intercomparison. Journal of Climate 19 (14), 3337–3353.

Global Change Project (GCP), 2010. REgional Carbon Cycle Assessment and Processes (RECCAP) Soft Protocol, Version 4. Global Carbon Project.

Global Soil Data Task Group, 2000. Global Gridded Surfaces of Selected Soil Characteristics (IGBP-DIS). Global Gridded Surfaces of Selected Soil Characteristics (International Geosphere-Biosphere Programme-Data and Information System). Data Set. Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA, doi:10.3334/ORNLDAAC/569, Available on-line http://www.daac.ornl.gov.

Goodale, C.L., et al., 2002. Forest carbon sinks in the Northern Hemisphere. Ecological Applications 12 (3), 891–899.

Gurney, K.R., Law, R.M., Denning, A.S., Rayner, P.J., Baker, D., Bousquet, P., et al., 2002. Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models. Nature 415 (6872), 626–630.

Gurney, K.R., Law, R.M., Denning, A.S., Rayner, P.J., Baker, D., Bousquet, P., et al., 2003. Transcom 3 CO_2 inversion intercomparison. 1. Annual mean control results and sensitivity to transport and prior flux information. Tellus 55B, 555–579.

Hansen, J., Ruedy, R., Glascoe, J., Sato, M., 1999. GISS analysis of surface temperature change. Journal of Geophysical Research 104, 30997–31022, doi:10.1029/1999JD900835.

Haxeltine, A., Prentice, I.C., 1996. BIOME3: An equilibrium terrestrial biosphere model based on ecophysiological constraints, resource availability, and competition among plant functional types. Global Biogeochemical Cycles 10 (4), 693–709.

Hayes, D.J., McGuire, A.D., Kicklighter, D.W., Gurney, K.R., Burnside, T.J., Melillo, J.M., 2011. Is the northern high latitude land-based CO_2 sink weakening? Global Biogeochemical Cycles 25 (3), GB3018, doi:10.1029/2010gb003813.

Hayes, D.J., Turner, D.P., Stinson, G., McGuire, A.D., Wei, Y., West, T.O., et al., 2012. Reconciling estimates of the contemporary North American carbon balance among inventory-based approaches, terrestrial biosphere models and atmospheric inversions. Global Change Biology 18 (3), doi:10.1111/j.1365-2486.2011.02627.x.

Heinsch, F.A., et al., 2003. User's Guide GPP and NPP (MOD17A2/A3) Products NASA MODIS Land Algorithm.

Heinsch, F.A., et al., 2006. Evaluation of remote sensing based terrestrial productivity from MODIS using regional tower eddy flux network observations. IEEE Transactions on Geoscience and Remote Sensing 44 (7), 1908–1925.

Houghton, R.A., Hackler, J.L., Lawrence, K.T., 1999. The US carbon budget: contributions from land-use change. Science 285 (5427), 574–578.

Jain, A.K., Yang, J.K., 2005. Modeling the effects of two different land cover change data sets on the carbon stocks of plants and soils in concert with CO_2 and climate change. Global Biogeochemical Cycles 19 (2).

Jastrow, J.D., 1996. Soil aggregate formation and the accrual of particulate and mineral-associated organic matter. Soil Biology & Biochemistry 28 (4–5), 665–676.

Jentsch, A., Kreyling, J., Beierkuhnlein, C., 2007. A new generation of climate-change experiments: events, not trends. Frontiers in Ecology and the Environment 5 (7), 365–374.

Ju, W.M., Chen, J.M., Black, T.A., Barr, A.G., Liu, J., Chen, B.Z., 2006. Modelling multiyear coupled carbon and water fluxes in a boreal aspen forest. Agricultural and Forest Meteorology 140 (1–4), 136–151.

Jung, M., Henkel, K., Herold, M., Churkina, G., 2006. Exploiting synergies of global land cover products for carbon cycle modeling. Remote Sensing of Environment 101, 534–553.

Kicklighter, D.W., Bondeau, A., Schloss, A.L., Kaduk, J., McGuire, A.D., et al., 1999. Comparing global models of terrestrial net primary productivity (NPP): global pattern and differentiation by major biomes. Global Change Biology 5, 16–24.

Krinner, G., Viovy, N., de Noblet-Ducoudre, N., Ogee, J., Polcher, J., Friedlingstein, P., et al., 2005. A dynamic global vegetation model for studies of the coupled atmosphere–biosphere system. Global Biogeochemical Cycles 19 (1).

Kucharik, C.J., Foley, J.A., Delire, C., Fisher, V.A., Coe, M.T., Lenters, J.D., et al., 2000. Testing the performance of a Dynamic Global Ecosystem Model: water balance, carbon balance, and vegetation structure. Global Biogeochemical Cycles 14 (3), 795–825.

Kurz, W.A., Stinson, G., Rampley, G., 2007. Could increased boreal forest ecosystem productivity offset carbon losses from increased disturbances? Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, doi:10.1098/rstb.2007.2198.

Latifovic, R., Zhu, Z.L., Cihlar, J., Giri, C., Olthof, I., 2004. Land cover mapping of north and central America – Global Land Cover 2000. Remote Sensing of Environment 89 (1), 116–127.

Leemans, R., Cramer, W., 1991. The IIASA Database for Mean Monthly Values of Temperature, Precipitation and Cloudiness of a Global Terrestrial Grid. International Institute for Applied Systems Analysis (IIASA), RR-91-18.

Lenihan, J.M., Bachelet, D., Neilson, R.P., Drapek, R., 2008. Simulated response of conterminous United States ecosystems to climate change at different levels of fire suppression, CO_2 emission rate, and growth response to CO_2. Global and Planetary Change 64 (1–2), 16–25.

Loveland, T.R., Belward, A.S., 1997. The International Geosphere Biosphere Programme Data and Information System global land cover data set (DISCover). Acta Astronautica 41 (4–10), 681–689.

Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, J., Yang, L., Merchant, J.W., 2000. Development of a Global Land Cover Characteristics Database and IGBP DISCover from 1-km AVHRR Data. International Journal of Remote Sensing 21 (6/7), 1303–1330.

McGuire, A.D., Hayes, D.J., Kicklighter, D.W., Manizza, M., Zhuang, Q., Chen, M., et al., 2010. An analysis of the carbon balance of the Arctic Basin from 1997 to 2006. Tellus 62B, 455–474, doi:10.1111/j.1600-0889.2010.00497.x.

Medvigy, D.S., Wofsy, C., Munger, J.W., Moorcroft, P.R., 2010. Responses of terrestrial ecosystems and carbon budgets to current and future environmental variability. Proceedings of the National Academy of Science of the United States of America 201 (18), 8275–8280, doi:10.1073/pnas.0912032107.

Melillo, J.M., McGuire, A.D., Kicklighter, D.W., Moore, Vorosmarty, C.J., Schloss, A.L., 1993. Global climate-change and terrestrial net primary production. Nature 363 (6426), 234–240.

Melillo, J.M., Borchers, J., Chaney, J., Fisher, H., Fox, S., Haxeltine, A., et al., 1995. Vegetation ecosystem modeling and analysis project – 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 (4), 407–437.

NRCan and USGS, 2003. Land Cover Database of North America 2000.

Oades, J.M., 1988. The retention of organic-matter in soils. Biogeochemistry 5 (1), 35–70.

Pacala, S.W., et al., 2001. Consistent land- and atmosphere-based US carbon sink estimates. Science 292 (5525), 2316–2320.

Pan, Y.D., Melillo, J.M., McGuire, A.D., Kicklighter, D.W., Pitelka, L.F., Hibbard, K., Pierce, et al., 1998. Modeled responses of terrestrial ecosystems to elevated atmospheric CO_2: a comparison of simulations by the biogeochemistry models of the vegetation/ecosystem modeling and analysis project (VEMAP). Oecologia 114, 389–404.

Parton, W.J., Schimel, D.S., Cole, C.V., Ojima, D.S., 1987. Analysis of factors controlling soil organic-matter levels in great-plains grasslands. Soil Science Society of America Journal 51 (5), 1173–1179.

Potter, C., Klooster, S., Huete, A., Genovese, V., 2007. Terrestrial carbon sinks for the United States predicted from MODIS satellite data and ecosystem modeling. Earth Interactions 11.

Poulter, B.D., Frank, C., Hodson, E.L., Zimmerman, N.E., 2011. Impacts of land cover and climate data selection on understanding terrestrial carbon dynamics and the CO_2 airborne fraction. Biogeosciences Discuss 8, 1617–1642, doi:10.5194/bgd-9-1617-2011.

Prentice, I.C., 2001. The Carbon Cycle and Atmospheric Carbon Dioxide Rep. 3rd Assessment. Intergovernmental Panel on Climate Change.

Randerson, J.T., Thompson, M.V., Conway, T.J., Fung, I.Y., Field, C.B., 1997. The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide. Global Biogeochemical Cycles 11 (4), 535–560.

Randerson, J.T., Hoffman, F.M., Thornton, P.E., Mahowald, N.M., Lindsay, K., Lee, Y.H., et al., 2009. Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models. Global Change Biology 15, 2462–2484, doi:10.1111/j.1365-2486.2009.01912.x.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Bebigier, P., et al., 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Global Change Biology 11, 1424–1439.

Reichstein, M., Beer, C., 2008. Soil respiration across scales: the importance of a model–data integration framework for data interpretation. Journal of Plant Nutrition and Soil Science 171, 2344–2354.

Running, S.W., Nemani, R.R., Heinsch, F.A., Zhao, M.S., Reeves, M., Hashimoto, H., 2004. A continuous satellite-derived measure of global terrestrial primary production. Bioscience 54 (6), 547–560.

Schimel, D.S., Enting, I.G., Heimann, M., WIgley, Raynaud, D., Alves, D., Siegenthaler, U., 2000. CO_2 and the carbon cycle. In: Wigley, T.M.L. (Ed.), The Carbon Cycle. Cambridge University Press.

Schwalm, C.R., Williams, C.A., Schaefer, K., Anderson, R., Arain, M.A., Baker, I., et al., 2010. A model-data intercomparison of CO_2 exchange across North America: results from the North American Carbon Program site synthesis. Journal of Geophysical Research-Biogeosciences 115, doi:10.1029/2009JG001229.

Sitch, S., et al., 2003. Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. Global Change Biology 9 (2), 161–185.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. U.S. General Soil Map (STATSGO2) for (State). Available online at http://soildatamart.nrcs.usda.gov accessed.

Thornton, P.E., Lamarque, J.F., Rosenbloom, N.A., Mahowald, N.M., 2007. Influence of carbon–nitrogen cycle coupling on land model response to CO_2 fertilization and climate variability. Global Biogeochemical Cycles 21, GB4018, doi:10.1029/2006GB002868.

Thornton, P.E., Doney, S.C., Lindsay, K., Moore, J.K., Mahowald, N., Randerson, J.T., et al., 2009. Carbon–nitrogen interactions regulate climate-carbon cycle feedbacks: results from an atmosphere–ocean general circulation model. Biogeosciences 6 (10), 2099–2120.

Tian, H.Q., Chen, G., Liu, M., Zhang, C., Sun, G., Lu, C., et al., 2010. Model estimates of ecosystem net primary productivity, evapotranspiration, and water use efficiency in the Southern United States during 1895–2007. Forest Ecology and Management 259, 1311–1327.

van der Werf, G.R., Randerson, J.T., Collatz, G.J., Giglio, L., Kasibhatla, P.S., Arellano, A.F., et al., 2004. Continental-scale partitioning of fire emissions during the 1997 to 2001 El Nino/La Nina period. Science 303 (5654), 73–76.

van der Werf, G.R., Randerson, J.T., Giglio, L., Collatz, G.J., Kasibhatla, P.S., Arellano, A.F., 2006. Interannual variability in global biomass burning emissions from 1997 to 2004. Atmospheric Chemistry and Physics 6, 3423–3441.

Viovy, N., Francois, C., Bondeau, A., Krinner, G., Polcher, J., Kergoat, L., et al., 2000. Assimilation of Remote Sensing Measurements into the ORCHIDEE/STOMATE DGVM Biosphere Model.

Wang, Z., Grant, R.F., Arain, M.A., Chen, B.N., Coops, N., Hember, R., Kurz, W.A., Price, D.T., Stinson, G., Trofymow, J.A., Yeluripati, J., Chen, Z., 2011. Evaluating weather effects on interannual variation in net ecosystem productivity of a coastal temperate forest landscape: a model intercomparison. Ecological Modelling 222, 3236–3249.

Waring, R.H., Running, S.W., 2007. Forest Ecosystems: Analysis at Multiple Scales, third ed. Elsevier Academic Press, Burlington, MA.

Wofsy, S.C., Harriss, R.C., 2002. The North American Carbon Program (NACP). Report of the NACP Committee of the U.S. Interagency Carbon Cycle Science Program, Rep. U.S. Global Change Research Program, Washington, DC, 56 pp.

Xiao, J.F., et al., 2008. Estimation of net ecosystem carbon exchange for the conterminous United States by combining MODIS and AmeriFlux data. Agricultural and Forest Meteorology 148 (11), 1827–1847.

Xiao, J.F., et al., 2010. A continuous measure of gross primary productivity for the conterminous U.S. derived from MODIS and AmeriFlux data. Remote Sensing of Environment 114, 576–591, doi:10.1016/j.rse.2009.10.013.

Xiao, J.F., et al., 2011. Assessing net ecosystem carbon exchange of U.S. terrestrial ecosystems by integrating eddy covariance flux measurements and satellite observations. Agricultural and Forest Meteorology 151, 60–69, doi:10.1016/j.agrformet.2010.09.002.

Yang, X.J., Wittig, V., Jain, A.K., Post, W., 2009. Integration of nitrogen cycle dynamics into the Integrated Science Assessment Model for the study of terrestrial ecosystem responses to global change. Global Biogeochemical Cycles, 23.

Yuan, W.P., et al., 2007. Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes. Agricultural and Forest Meteorology 143 (3–4), 189–207.

Zeng, N., 2003. Glacial–interglacial atmospheric CO_2 change – the glacial burial hypothesis. Advances in Atmospheric Sciences 20 (5), 677–693.

Zeng, N., Qian, H.F., Munoz, E., Iacono, R., 2004. How strong is carbon cycle-climate feedback under global warming? Geophysical Research Letters 31 (20).

Zeng, N., Mariotti, A., Wetzel, P., 2005. Terrestrial mechanisms of interannual CO_2 variability. Global Biogeochemical Cycles 19 (1).

Zhao, M.S., Heinsch, F.A., Nemani, R.R., Running, S.W., 2005. Improvements of the MODIS terrestrial gross and net primary production global data set. Remote Sensing of Environment 95 (2), 164–176.

Zhao, Y., Ciais, P., Peylin, P., Viovy, N., Longdoz, B., Bonnefond, J.M., et al., 2011. How errors on meteorological variables impact simulated ecosystem fluxes: a case study for six French sites. Biogeosciences Discuss 8, 2522–5467, doi:10.5194/bgd-8-2467-2011.

Zobler, L., 1986. A World Soil File for Global Climate Modeling. Technical Memorandum 87802. NASA Goddard Institute for Space Studies (GISS), New York, NY.