Global Change Biology (2011), doi: 10.1111/j.1365-2486.2011.02562.x

Terrestrial biosphere models need better representation of vegetation phenology: results from the North American Carbon Program Site Synthesis

ANDREW D. RICHARDSON*, RYAN S. ANDERSON†, M. ALTAF ARAIN‡, ALAN G. BARR§, GIL BOHRER¶, GUANGSHENG CHEN**, JING M. CHEN††, PHILIPPE CIAIS‡‡, KENNETH J. DAVIS§§, ANKUR R. DESAI¶¶, MICHAEL C. DIETZE***, DANILO DRAGONI†††, MUSTAPHA EL MAAYAR‡‡‡, STEVEN GARRITY§§§, CHRISTOPHER M. GOUGH¶¶¶, ROBERT GRANT****, DAVID Y. HOLLINGER††††, HANK A. MARGOLIS‡‡‡‡, HARRY MCCAUGHEY§§§§, MIRCO MIGLIAVACCA¶¶¶¶, RUSSELL K. MONSON*****, J. WILLIAM MUNGER†††††, BENJAMIN POULTER‡‡‡‡‡, BRETT M. RACZKA§§, DANIEL M. RICCIUTO§§§§§, ALOK K. SAHOO¶¶¶¶¶, KEVIN SCHAEFER******, HANQIN TIAN††††††, RODRIGO VARGAS‡‡‡‡‡‡, HANS VERBEECK§§§§§§, JINGFENG XIAO¶¶¶¶¶¶ and YONGKANG XUE*******

*Department of Organismic and Evolutionary Biology, Harvard University, HUH, 22 Divinity Ave., Cambridge, MA 02138, USA, †Numerical Terradynamic Simulation Group, University of Montana, Missoula, MT 59812, USA, ‡School of Geography and Earth Sciences, McMaster University, Hamilton, Ontario L8S 4K1, Canada, §Science and Technology Branch, Environment Canada, Saskatoon, Saskatchewan S7N 3H5, Canada, ¶Department of Civil & Environmental Engineering & Geodetic Science, Ohio State University, Columbus, OH 43210, USA, **School of Forestry and Wildlife Sciences, Auburn University, Auburn, AL 36849, USA, ††Department of Geography, University of Toronto, Toronto, Ontario M5S 3G3, Canada, ‡‡Laboratoire des Sciences du Climatet de l'Environnement (LSCE), CEA CNRS UVSQ, Orme des Merisiers, 91190, Gif-sur-Yvette France, §§Department of Meteorology, The Pennsylvania State University, University Park, PA 16802, USA, ¶¶Department of Atmospheric & Oceanic Sciences, University of Wisconsin–Madison, Madison, WI 53706, USA, ***Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA, †††Department of Geography, Indiana University, Bloomington, IN 47405, USA, ‡‡‡Energy, Environment and Water Research Center, The Cyprus Institute, CY-1645, Nicosia, Cyprus, §§§Department of Civil & Environmental Engineering, Ohio State University, Columbus, OH 43210, USA, ¶¶¶Department of Biology, Virginia Commonwealth University, Richmond, VA 23284, USA, ****Department of Renewable Resources, University of Alberta, Edmonton, Alberta T6G 2E3, Canada, ††††Northern Research Station, USDA Forest Service, Durham, NH 03824, USA, ‡‡‡‡Faculty of Forestry, Geography and Geomatics, Laval University, Quebec City, Quebec G1V 0A6, Canada, §§§§Department of Geography, Queen's University, Kingston, Ontario K7L 3N6, Canada, ¶¶¶¶European Commission – DG Joint Research Centre, Institute for Environment and Sustainability, Climate Change and Air Quality Unit, I-21027, Ispra (VA), Italy, *****Department of Ecology & Evolutionary Biology and Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO 80309, USA, †††††School of Engineering and Applied Sciences and Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, USA, ‡‡‡‡‡Swiss Federal Research Institute WSL, 8903, Birmensdorf, Switzerland, §§§§§Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA, ¶¶¶¶¶Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ 08544, USA, ******National Snow and Ice Data Center, University of Colorado, Boulder, CO 80309, USA, ††††††International Center for Climate and Global Change Research, School of Forestry and Wildlife Sciences, Auburn University, Auburn, AL 36849, USA, ‡‡‡‡‡‡Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), Ensenada, BC Mexico, §§§§§§Laboratory of Plant Ecology, Faculty of Bioscience Engineering, Ghent University, Belgium,, ¶¶¶¶¶¶Complex Systems Research Center, University of New Hampshire, Durham, NH 03824, USA, *******Department of Geography, University of California–Los Angeles, Los Angeles, CA 90095, USA

Abstract

Phenology, by controlling the seasonal activity of vegetation on the land surface, plays a fundamental role in regulating photosynthesis and other ecosystem processes, as well as competitive interactions and feedbacks to the climate system. We conducted an analysis to evaluate the representation of phenology, and the associated seasonality of ecosystem-scale CO_2 exchange, in 14 models participating in the North American Carbon Program Site Synthesis. Model predictions were evaluated using long-term measurements (emphasizing the period 2000–2006) from 10 forested sites within the AmeriFlux and Fluxnet-Canada networks. In deciduous forests, almost all models consistently predicted

Correspondence: Andrew D. Richardson, tel. + xxxxxx, fax + xxxxxx, e-mail: arichardson@oeb.harvard.edu

that the growing season started earlier, and ended later, than was actually observed; biases of 2 weeks or more were typical. For these sites, most models were also unable to explain more than a small fraction of the observed interannual variability in phenological transition dates. Finally, for deciduous forests, misrepresentation of the seasonal cycle resulted in over-prediction of gross ecosystem photosynthesis by +160 ± 145 g C m^{-2} yr^{-1} during the spring transition period and +75 ± 130 g C m^{-2} yr^{-1} during the autumn transition period (13% and 8% annual productivity, respectively) compensating for the tendency of most models to under-predict the magnitude of peak summertime photosynthetic rates. Models did a better job of predicting the seasonality of CO_2 exchange for evergreen forests. These results highlight the need for improved understanding of the environmental controls on vegetation phenology and incorporation of this knowledge into better phenological models. Existing models are unlikely to predict future responses of phenology to climate change accurately and therefore will misrepresent the seasonality and interannual variability of key biosphere–atmosphere feedbacks and interactions in coupled global climate models.

Keywords: autumn senescence, carbon cycle, land surface model (LSM), leaf area index (LAI), model error, North American Carbon Program (NACP), phenology, seasonal dynamics, spring onset

Received 10 June 2011; revised version received 13 September 2011 and accepted 20 September 2011

Introduction

Phenological transitions drive the seasonal progression of vegetation through stages of dormancy, active growth, and senescence. Although phenology has traditionally been concerned with physical changes in structure (e.g., leaf development and abscission), the inherent seasonality of mass and energy exchange between terrestrial ecosystems and the atmosphere can, more generally, be viewed as phenological in nature (Gu *et al*., 2003). In deciduous forests, the relationships between the phenology of canopy structure and function are obvious. In evergreen forests, physiological changes within existing foliage (and not the production of new foliage) regulate the annual rhythms of photosynthesis and transpiration (e.g., Monson *et al*., 2005; Richardson *et al*., 2009b). In both forest types, phenological switches, rather than fast responses to high-frequency variation in environmental drivers, are controlling the seasonal patterns.

Phenology is thus a key regulator of ecosystem processes and biosphere feedbacks to the climate system (Peñuelas *et al*., 2009). Phenology influences both spatial and temporal (at seasonal-to-interannual time scales) variability in ecosystem productivity (Baldocchi *et al*., 2001; Churkina *et al*., 2005; Richardson *et al*., 2009a, 2010; Dragoni *et al*., 2011), and it is of fundamental importance for ecosystem carbon cycling, terrestrial carbon sequestration, and mitigation of anthropogenic CO_2 emissions. Furthermore, phenology affects the following: *hydrology* (Hogg *et al*., 2000), as leaf-out is accompanied by an increase in evapotranspiration and reduced throughfall; *nutrient cycling processes* (Cooke & Weih, 2005), as senescence results in fresh litter inputs to the soil; and *atmospheric and climate system feedbacks* (Schwartz, 1992), as the amount and condition of foliage present affects albedo, surface energy balance, and surface roughness (Moore *et al*., 1996; Sakai *et al*., 1997; Peñuelas *et al*., 2009).

It is, therefore, essential that terrestrial biosphere models simulating the temporal dynamics of biological processes on the land surface have an accurate representation of phenology. This is true whether the model is simple or complex (in terms of the number of biogeochemical processes it features, and the degree to which processes are coupled or interact with each other) and whether the model is being run for a single site or the entire globe. Indeed, Levis & Bonan (2004) highlight the importance of accurate prognostic modeling of phenology, and the associated seasonal patterns of canopy leaf area index (LAI), for climate model runs that couple a land surface scheme to an atmospheric general circulation model.

A number of previous studies have evaluated the phenology submodels included in state-of-the-art land surface schemes and ecosystem models and concluded that these routines tend to be overly simplistic and result in biased predictions (Kucharik *et al*., 2006; Ryu *et al*., 2008). Randerson *et al.* (2009) included phenological metrics as part of a systematic framework, the Carbon-LAnd Model intercomparison Project (C-LAMP), to assess the biogeochemical component of coupled climate–carbon models. They concluded that model bias toward under-predicting temperate and boreal forest uptake of CO_2 could be attributed to a 1–3 month delay in predicting the timing of maximum LAI in these ecosystems, compared to estimates derived from MODIS data. Randerson *et al.* (2009) also noted that the two models they evaluated tended to predict a longer growing season than was actually observed in temperate ecosystems, with photosynthetic uptake occurring too early in the spring and too late in the autumn, compared with ground observations. Errors in LAI would likely propagate to errors in partitioning the available energy to latent and sensible heat fluxes, and errors in the timing of photosynthetic uptake would also affect the seasonality of modeled atmospheric CO_2 concentrations, emphasizing the importance of accurate representation of phenologically mediated processes.

In this study, we describe an analysis of the representation of phenology, in terms of the seasonality of LAI, gross ecosystem photosynthesis (GEP), and net ecosystem exchange (NEE), in 14 terrestrial biosphere models that contributed model runs to the North American Carbon Program (NACP) Site Synthesis. (We do not intend for this study to be considered a comprehensive analysis of all aspects of model performance; complementary NACP efforts include the earlier work by Schwalm *et al.* (2010), and work in preparation by Schaeffer *et al.*, Dietze *et al.*, Stoy *et al.*, and Raczka *et al.*.) The five deciduous broadleaf forest (DBF) and five evergreen needleleaf forest (ENF) sites selected for the analysis are all members of either the AmeriFlux or Fluxnet-Canada networks. Our analysis draws on the continuous eddy covariance measurements of forest-atmosphere CO_2 fluxes that have been made at each site for the last decade or more. At the deciduous forest sites, the flux measurements are complemented by above- and below-canopy measurements of photosynthetically active radiation, with which the seasonal trajectory of LAI can be estimated (e.g., Turner *et al.*, 2003).

The objectives of our analysis are as follows: (1) to assess the accuracy with which spring and autumn phenological transitions are predicted by different models; (2) to evaluate how these patterns vary between deciduous and evergreen forest types; and (3) to quantify how much of the total bias in modeling annual GEP can be attributed to errors in modeling the spring and autumn phenological transitions.

Data and method

Field measurements

The present analysis uses field measurements and model runs contributed to the NACP Site Synthesis project (http://nacp.ornl.gov/mast-dc/int_synthesis.shtml). We restrict our analysis to temperate and boreal deciduous broadleaf and evergreen needleleaf sites (five DBF and five ENF sites), with summer active/winter dormant seasonality, selected from the NACP 'Priority 1' list (Table 1).

Eddy covariance measurements of net ecosystem exchange of CO_2 (NEE; μmol CO_2 m^{-2} s^{-1}) supplied by site investigators were gap filled and partitioned to GEP and ecosystem respiration according to Barr *et al.* (2004). The partitioning algorithm was compared with

Table 1 Locations of deciduous broadleaf (DBF) and evergreen needleleaf (ENF) sites used in this analysis. Mean (±1 SD) spring and autumn transition dates (day of year), as extracted from measured data, are given for leaf area index(LAI) (20% seasonal amplitude of LAI; only reported for deciduous sites) and gross ecosystem photosynthesis (GEP) (20% GEP_{max}). Further details are provided in text

Site	Name	Country	Lat. °N	Long. °W	Years	IGBP	20% LAI		20% GEP		Reference
							Spring	Autumn	Spring	Autumn	
Ca-Oas	SSA Old Aspen	Canada	53.63	106.20	1997–2006	DBF	142 ± 9	279 ± 9	140 ± 8	266 ± 3	Barr *et al.* (2007)
US-Ha1	Harvard Forest (EMS Tower)	USA	42.54	72.17	1992–2006	DBF	137 ± 6	308 ± 2	131 ± 6	287 ± 3	Urbanski *et al.* (2007)
US-MMS	Morgan Monroe State Forest	USA	39.32	86.41	1999–2005	DBF	116 ± 3	304 ± 5	113 ± 2	290 ± 3	Schmid *et al.* (2000)
US-UMB	Univ. of Mich. Biological Station	USA	45.56	84.71	1999–2006	DBF	142 ± 5	294 ± 8	135 ± 9	283 ± 5	Gough *et al.* (2008)
US-WCr	Willow Creek	USA	45.81	90.08	2000–2005	DBF	138 ± 7	283 ± 5	138 ± 7	270 ± 7	Cook *et al.* (2004)
Ca-Obs	SSA Old Black Spruce	Canada	53.99	105.12	2000–2006	ENF			110 ± 11	288 ± 3	Krishnan *et al.* (2008)
Ca-Ojp	SSA Old Jack Pine	Canada	53.92	104.69	2000–2006	ENF			112 ± 10	289 ± 3	Kljun *et al.* (2006)
Ca-Qfo	Chibougamau (Mature Forest)	Canada	49.69	74.34	2004–2006	ENF			119 ± 12	297 ± 4	Bergeron *et al.* (2007)
US-Ho1	Howland Forest (Main Tower)	USA	45.20	68.74	1996–2004	ENF			98 ± 7	308 ± 10	Hollinger *et al.* (2004)
US-NR1	Niwot Ridge	USA	40.03	105.55	1999–2007	ENF			120 ± 8	298 ± 6	Monson *et al.* (2002)

a variety of other approaches by Desai *et al.* (2008). Most partitioning methods were found to yield similar seasonal cycles, and estimates of annual GEP that were within 10% of each other. This gives us confidence in the partitioning at daily-to-annual time scales. Furthermore, we argue that the narrow moving window used by the Barr *et al.* algorithm makes it well suited to the seasonality analyses we conduct. Filtering of nocturnal NEE measurements during periods of inadequate turbulence was conducted on a site-by-site basis using a u* change-point detection algorithm, described and evaluated by Barr *et al.* (2009, 2011).

For deciduous forests (the necessary data were generally not available for evergreen conifer sites), the seasonal trajectory of canopy leaf area index (LAI, $\mathrm{m}^2\ \mathrm{m}^{-2}$) was estimated as follows. First, we calculated the gap fraction, P, as $P = Q_\mathrm{t}/Q_\mathrm{o}$, where Q_o is incident solar photosynthetic photon flux density (PPFD) measured above the canopy and Q_t is the PPFD measured below the canopy. We used measurements of P when the solar zenith angle was closest to 57° (one sample in the morning and one sample in the afternoon) and then calculated LAI for each sample as $\mathrm{LAI} = -\log(P)/K$ where $K = G(57)/\cos(57)$. We restricted our analyses to zenith angles nearest 57 degrees because at this point, all leaf inclination distribution functions (G) converge to 0.5. Then, we obtained daily LAI by averaging the two LAI values to consider foliar clumping effects (Ryu *et al.*, 2010). Noise in the resulting time series, which we attribute mostly to cloud effects (variability in direct beam and diffuse PPFD), was smoothed with a spline function. We then re-scaled the seasonal trajectory of LAI so that the seasonal peak LAI derived in this manner matched with measured LAI as reported on the AmeriFlux web page (http://ameriflux.ornl.gov/) or in published manuscripts for each site (see references in Table 1), and the seasonal minimum LAI was zero.

With this approach, we obtained essentially continuous estimates of changes in LAI over time for each of the deciduous forest sites. However, the heterogeneous nature of the below-canopy light environment raises questions about the degree to which these estimates may be representative of leaf area dynamics across the larger tower footprint. At four of the five deciduous sites, only a single below-canopy quantum sensor was used to measure Q_t. At US-MMS, there were four below-canopy sensors, which no doubt provided better sampling of spatial variability in Q_t. At none of the sites were field campaigns to measure LAI (e.g., plant canopy analyzer or hemispherical photography) conducted at a sufficiently high temporal resolution (e.g., weekly) to permit accurate estimation of phenological transition dates. However, where such data are available at a lower temporal frequency (e.g., monthly), they provide a context for evaluating our LAI estimates. As shown in Fig. 1, the mean seasonal course (over multiple years) of LAI estimated from $P = Q_\mathrm{t}/Q_\mathrm{o}$ (solid black lines) is in good agreement with that obtained by the LAI-2000 (Li-Cor Biosciences, Lincoln, NE, USA) plant canopy analyzer instrument (open circles) across a network of plots at each site (no LAI-2000 data for US-WCr). For US-MMS and US-UMB, the timing of spring and autumn transitions was consistent between the two methods. For US-Ha1, there is a clear divergence in early autumn, with the LAI-2000 data indicating an earlier decline in leaf area. However, this could simply be an artifact of sampling in different years: although Q_t data were available for 9 years and LAI-2000 for 4 years, there was only 1 year of overlap between the two data sets at US-Ha1. For Ca-Oas, LAI-2000 measurements were made in years with substantial variability in the timing of canopy development, and this variability (as shown in Fig. 1) masks the otherwise good concordance between LAI measured with the LAI-2000 and estimated from $P = Q_\mathrm{t}/Q_\mathrm{o}$. Both methods were in agreement, for example, on the exceptionally early springs in 1998 and 2001 and the late onset of leaf development in 1997 and 2004. The above patterns, and the overall strong correlation between LAI from the two methods at each site ($r > 0.95$ for Ca-Oas, US-MMS, and US-UMB; $r = 0.85$ for US-Ha1), give us confidence in our retrievals.

We conclude by noting that although satellite data (e.g., vegetation indices as well as more targeted products related to LAI and phenology) offer the promise of global coverage, they suffer from tradeoffs between spatial and temporal resolution and have their own substantial uncertainties (e.g., Zhang *et al.*, 2006; Garrigues *et al.*, 2008; White *et al.*, 2009), which in the context of the present analysis make them less suitable benchmarks for model evaluation.

Model runs

Participation in the NACP site synthesis was on a volunteer basis, and an open invitation was sent to the terrestrial biosphere modeling community. Modeling teams were free to choose as simple or as complex a model as they desired. Models were run on a site-by-site basis, using measured environmental drivers (gap filled as necessary using a standardized method, Ricciuto *et al.* in preparation), and site-specific initial conditions (as judged necessary), following a standard protocol (http://nacp.ornl.gov/mast-dc/docs/Site_

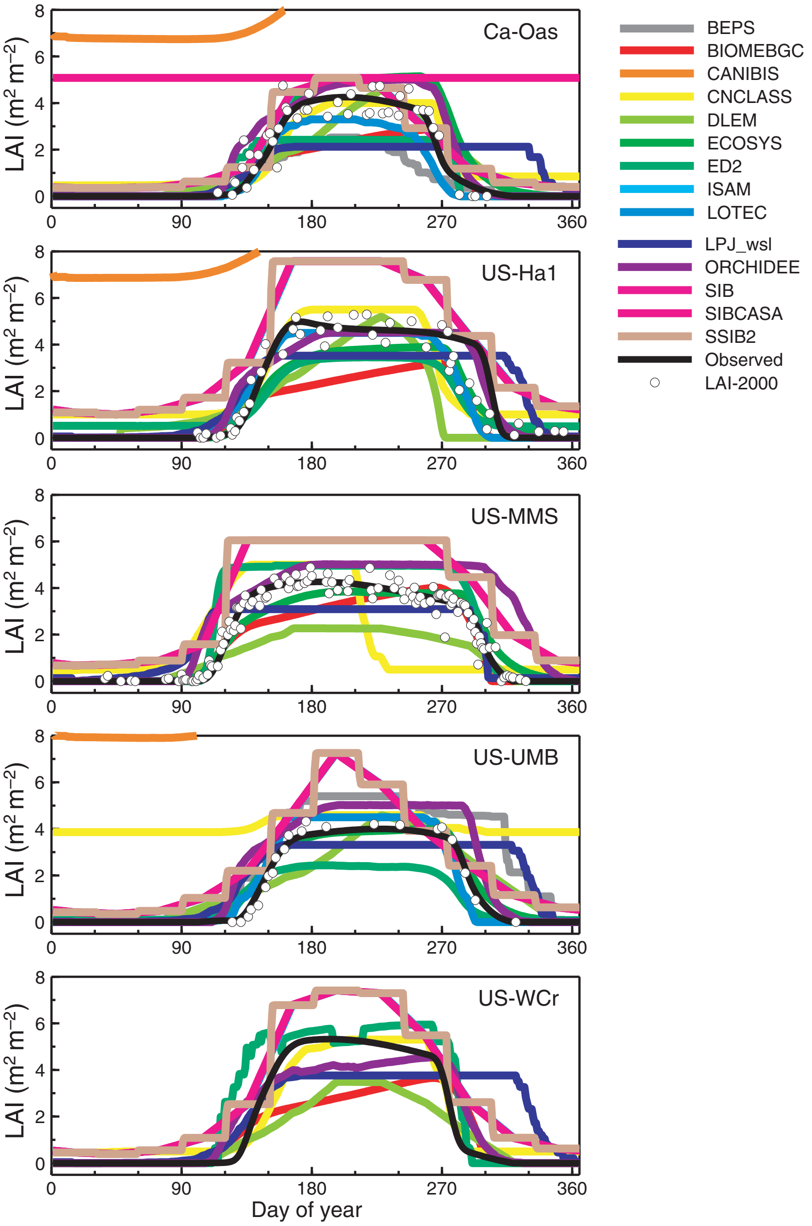


Fig. 1 The mean seasonal trajectory of observed and simulated leaf area index (LAI), for five deciduous broadleaf sites. The continuous observed LAI is derived from gap fraction estimates based on above- and below-canopy measurements of photosynthetically active radiation. Open circles show periodic LAI measurements made using an LAI-2000 plant canopy analyzer.

Synthesis_Protocol_v7.pdf). The protocol specified spin-up of carbon pools to steady state, but the way in which this was implemented varied among modeling teams.

Of the variety of models for which output was submitted to the NACP database, we included only those that appeared to at least superficially capture the seasonal trajectory of ecosystem activity. For example, models that did not predict winter dormancy (i.e., if they instead predicted significant wintertime photosynthetic uptake) were not included in this analysis. The 14 models that were included in our analysis are listed in Table 2. Note that the LoTEC model was run in a data assimilation mode, and model parameters were optimized, on a site-by-site basis, conditional on the flux measurements. This no doubt contributes to the better

Table 2 Summary of models used in this analysis and their representation of phenology and seasonality of leaf area index (LAI). For models with 'prognostic' phenology, the seasonality of LAI is predicted based on climatic drivers; for those with 'prescribed' phenology, an average seasonal LAI cycle, as derived on a site-by-site basis from satellite (AVHRR) data, was used. Models with semi-prescribed and semi-prognostic phenology represent a hybrid of these approaches. GDD is growing degree days; T is temperature; C is carbon; PFT is plant functional type

Model name	Resolution	Leaf onset	Control on LAI	Leaf loss	Source
BEPS	Daily	Satellite	Satellite	Satellite	Ju *et al.* (2006)
Biome-BGC	Daily	GDD and radiation sum	Dynamic C allocation	Daylength and low temperature	Thornton *et al.* (2002)
Can-IBIS	Half-hourly	T threshold	GDD and dynamic C	Prescribed	El Maayar *et al.* (2002)
CN-CLASS	Half-hourly	C balance	C balance	Daylength and low temperature	Arain *et al.* (2006)
DLEM	Daily	$T_{7\text{-day}}$ > threshold	GDD to PFT limit	Daylength and low temperature	Tian *et al.* (2010)
Ecosys	Hourly	Hours above T threshold	Dynamic C allocation	Hours below T threshold	Grant *et al.* (2009)
ED2	Half-hourly	Semi-prescribed	Dynamic C allocation	GDD and leaf turnover	Medvigy *et al.* (2009)
ISAM	Half-hourly	Prescribed	Prescribed	Prescribed	Jain & Yang (2005)
LoTEC	Half-hourly	GDD	GDD	T-dependent turnover	Hanson *et al.* (2004)
LPJ_wsl	Daily	GDD	GDD	Leaf longevity (prescribed)	Sitch *et al.* (2003)
ORCHIDEE	Half-hourly	GDD and chilling	Dynamic C allocation	Decreasing T and T threshold	Krinner *et al.* (2005)
SiB3	Half-hourly	Prescribed	Prescribed	Prescribed	Baker *et al.* (2008)
SiBCASA	10 min	Prescribed	Prescribed	Prescribed	Schaefer *et al.* (2008)
SSiB2	Half-hourly	Prescribed	Prescribed	Prescribed	Zhan *et al.* (2003)

performance of LoTEC compared with some of the other models in this analysis.

Model approaches to phenological variation

Critical phenological events influencing carbon uptake in deciduous forests relate to the timing of leaf appearance, leaf expansion (increase of LAI), and leaf loss. For evergreen forests, a similar classification of foliage activity may be made from winter dormant to fully active and then dormant again. The models analyzed here use a variety of methods to determine the amount of foliage present and active in a canopy (Table 2). The simplest approach is to prescribe a fixed seasonal course of LAI. This approach encompasses the onset and development of foliage and also the dynamics of leaf loss. The original version of SiB (Sellers *et al*., 1986) used monthly LAI values that were specific for each plant functional type. A slightly more complicated approach to phenology is to prescribe the presence and amount of foliage based on remote sensing data. SiB2 [8] (Sellers *et al*., 1995) used AVHRR data to determine seasonal NDVI and then f_{PAR} and LAI. In the present analysis, BEPS relies on a global LAI dataset (Deng *et al*., 2006) derived from SPOT4 VEGETATION images and corrected for clumping via multi-angle POLDER observations (Chen *et al*., 2005). ED2 is designed to operate with the MODIS LAI product or other phenological drivers (Medvigy *et al*., 2009). Because satellite data are sometimes not available (e.g., for prognostic runs), models that use remotely sensed phenological observations may use multi-year average LAI and often maintain the flexibility of using other sources. The results presented here for the SiB class of models (as well as ISAM), for example, use a single average seasonal [9] course of LAI determined for each site.

Foliage onset and development in plants have long been related to temperature thresholds and cumulative heat sums (see, e.g., Shelford, 1930). This kind of approach is used by a number of the models (Table 2). The Ecosys model requires soil surfaces temperatures to exceed specified thresholds depending upon latitude for a certain number of hours. Can-IBIS operates in a similar fashion based on air temperatures while for deciduous sites DLEM uses a 7-day moving average that must exceed a threshold (for evergreen sites, DLEM uses prescribed phenology). The heat sum or growing degree day (GDD) approach is used by several models (including Biome-BGC, ED2, LoTEC, LPJ_wsl, and ORCHIDEE) for deciduous broadleaf trees but details of the implementations vary widely. Biome-BGC combines GDDs with a radiation sum (White *et al*., 1997). ORCHIDEE requires a chilling sum (Botta *et al*., 2000) prior to leaf initiation; other models implicitly account for chilling by summing degree days after a particular date such as January 1. LoTEC is distinct from the other models in that it optimizes parameter values to best fit the observed data. Included in the LoTEC optimizations are a leaf initiation GDD threshold and a full canopy GDD sum. These values are determined individually for each site and across the data record (the phenological parameters do not vary by year). Some of the models (Biome-BGC, ORCHIDEE) use a GDD approach for deciduous trees only; there is no explicit phenology (removal of dormancy) for evergreens.

A final approach, exhibited by CN-CLASS, is to prognostically calculate foliage carbon balance. In this method, leaf onset starts when daily photosynthesis of virtual leaves exceeds daily respiration for seven consecutive days (Arora & Boer, 2005). Virtual photosynthesis and respiration are both functions of temperature, and photosynthesis is also a function of soil moisture. Once leaf onset occurs, a number of models (Table 2) calculate LAI based on carbon allocation principles. These include seasonal allocation rules for different tissues and net foliage C gain.

Several different schemes are used for determining foliage inactivation or shedding. These include prescribed LAI (the SiB class models, BEPS, Can-IBIS, and ISAM), prognostic leaf longevity that varies according to GDD (LPJ_wsl), and various low temperature thresholds. Several of the models (Biome-BGC, CN-CLASS, DLEM) combine daylength and temperature thresholds based on the results of White *et al.* (1997). Ecosys requires a set number of hours below plant functional type-specific thresholds and shortening photoperiods to initiate litterfall in deciduous species and foliage inactivation in evergreens. In some models (e.g., ED2, ORCHIDEE), in addition to changes in LAI through the season, there are decreases in photosynthetic capacity driven by leaf aging. Leaf loss in LoTEC is determined by a low temperature parameter that is optimized for each site.

Data processing and extraction of phenological transition dates

For the deciduous sites, we extracted phenological transition dates from measured and modeled LAI trajectories. In addition, for both deciduous and evergreen sites, we included a series of diagnostic phenological metrics extracted from the measured and modeled time series of NEE and GEP. Transition dates and thresholds were estimated from smoothing splines fit to measured and modeled data at the daily time step, as illustrated in Richardson *et al.* (2010). The phenological transition dates we estimated from the data were as follows:

1. The first spring and last autumn dates at which measured and modeled LAI = 20%, 50%, and 80% of the seasonal LAI amplitude (deciduous sites only);
2. The first spring and last autumn dates at which estimated and modeled daily GEP = 20%, 50%, and 80% of the seasonal maximum GEP; and
3. The first spring and last autumn dates of NEE source/sink transition (restricted to source/sink transitions that began or concluded periods of 14 continuous days of net CO_2 uptake).

The relative thresholds (20%, 50%, and 80%) were selected to correspond to a range of developmental stages, so as to evaluate the ability of models to predict not only 'start' and 'end' of the growing season but also the overall seasonal pattern. We used relative rather than absolute (e.g., LAI = 1.0 m^2 m^{-2} or GEP = 2 g C m^{-2} day^{-1}) measures to account for differences in magnitude of both leaf area and CO_2 fluxes across sites (see also Richardson *et al*., 2010 for a similar approach; alternative methods have been proposed elsewhere, e.g., Gu *et al*., 2003).

We calculated transition date anomalies on a site-by-site basis, i.e., as interannual departures from the site mean.

For transition dates, model bias was calculated as the difference, in days, between the modeled and the observed transition date. Thus, a negative bias indicates that the model predicted the transition date too early in the year, whereas a positive bias indicates the model predicted the transition date too late in the year.

To evaluate the impact of errors in modeling phenological transitions on annual ecosystem productivity estimates, we calculated several metrics, beginning with the total model bias in annual GEP:

$$\text{Total model bias} = (\text{annual model GEP}) - (\text{annual tower GEP}). \quad (1)$$

Three main sources of model error (which may reflect some combination of errors in model structure and errors in model parameterization) that contribute to the total model bias in GEP are (1) errors in the overall magnitude of modeled GEP; (2) errors in the seasonality of modeled GEP; and (3) errors in the sensitivity of modeled GEP to high-frequency variability in environmental drivers (e.g., incorrect representation of GEP sensitivity to vapor pressure deficit, soil water stress, or air temperature, among other factors). Errors in GEP magnitude (1), are likely due to incorrect specification of photosynthetic parameters such as $A_{\max}$ or V_{cmax} and can occur independently of errors in (2) or (3). To correct for (1) and emphasize instead a focus on errors in seasonality, we re-scaled modeled GEP (daily values) so that the seasonal maxima (as determined by the 95th percentile value) of measured and observed GEP were the same. Thus,

$$\text{model GEP}_{\text{re-scaled}} = \text{model GEP} \times \frac{\text{tower GEP}_{\max}}{\text{model GEP}_{\max}}. \tag{2}$$

We then calculated the total model bias in annual GEP, corrected for differences in $\mathrm{GEP}_{\max}$, as follows:

$$\begin{aligned}&\text{Total model bias (scaled for GEP}_{\max})\\&\quad = (\text{annual model GEP}_{\text{re-scaled}})\\&\qquad - (\text{annual tower GEP}).\end{aligned} \tag{3}$$

To quantify how much of the model bias could be attributed to errors in seasonality, we defined ‘spring’ as the period between the first date (for each site-year of data, for each model run) when either model GEP or tower GEP rose to 20% of $\mathrm{GEP}_{\max}$, and the last date when either model GEP or tower GEP rose to 80% of $\mathrm{GEP}_{\max}$. ‘Autumn’ was similarly defined as the period between the first date when either model GEP or tower GEP dropped below 80% of $\mathrm{GEP}_{\max}$, and the last date when either model GEP or tower GEP dropped below 20% of $\mathrm{GEP}_{\max}$. Put differently, ‘spring’ and ‘autumn’ model biases were calculated from model $\mathrm{GEP}_{\text{re-scaled}}$ and tower GEP, with GEP integrals calculated for the period during which either model GEP or tower GEP is in the ‘increasing GEP’ phase ($0.2 \leq$ daily GEP/$\mathrm{GEP}_{\max} \leq 0.8$) in spring or ‘decreasing GEP’ phase in autumn ($0.8 \geq$ daily GEP/$\mathrm{GEP}_{\max} \geq 0.2$). (Thus, the first and last dates of ‘spring’ and ‘autumn’ varied from year-to-year and across sites, but in each instance, the same dates were used to define the period of integration for both model GEP and tower GEP. Note also that the periods of integration could thus vary among models.)

Results

Leaf area dynamics

The modeled seasonal trajectory of deciduous forest LAI suffered from errors in the timing of both spring increases and autumn decreases in LAI, as well as errors in the amplitude of the seasonal cycle. No single model characterized LAI dynamics well at all five sites (Fig. 1). Reasonable performance at one site did not guarantee good performance at other sites: compare, for example, CN-CLASS at Ca-Oas and US-WCr with the same model at US-MMS (autumn decrease in LAI predicted approximately 3 months early) or US-UMB (virtually no seasonality to LAI, perhaps indicating that the site may have been mistyped as evergreen, rather than deciduous, forest). For some models, LAI dynamics were poor at most or all sites (e.g., Can-IBIS, DLEM, LPJ_wsl, SSiB2). For Can-IBIS, deciduous forest LAI was too high (≥ 5 m^2 m^{-2}) during the winter dormant season and did not decline in autumn. Biome-BGC captured the beginning and end of the season but LAI increased too slowly over the season. LoTEC, which optimized leaf onset, development, and abscission at each site, reproduced leaf onset and LAI values well but was surprisingly poor at capturing autumn senescence.

Models generally initiated greening up of the canopy too early in spring (Fig. 2): across all years, sites, and models, the mean (±1 SD) bias in the date at which model LAI reached 20% of maximum LAI was −10 ± 12 days, with a range from −53 to +28 days. There was, however, considerable variation among models in the magnitude of this bias. For example, across all sites, bias in the date at which the 20% LAI threshold was reached were largest for LPJ_wsl and BEPS (−19 ± 6 days) and smallest for Ecosys (−5 ± 6 days) and CN-CLASS (+3 ± 13 days). Can-IBIS was the only model commonly biased toward late predictions (+7 ± 10 days) of the onset of spring increases in LAI (recall however that this model retained substantial leaf area through the dormant season).

By comparison, across all years, sites, and models, the mean bias in the autumn date at which model LAI dropped to 20% of maximum LAI was +1 ± 32 days, with a range from −98 to +84 days. For this indicator, model predictions were generally biased early for ED2 (−5 ± 10 days), LoTEC (−13 ± 9 days), and CN-CLASS (−30 ± 32 days), but late for all the other models (particularly LPJ_wsl, +38 ± 28 days). Predictions were consistently biased late for US-WCr, US-UMB, and Ca-Oas. At the other two deciduous sites, model performance was mixed (Fig. 2).

In spite of these errors in modeling the overall seasonal trajectory of LAI, most models were able to predict some of the interannual variability in phenology (that is, year-to-year phenological anomalies) correctly in spring, but not in autumn (Table 3). Thus, although model predictions of spring onset dates were biased overall, the models did correctly represent a significant fraction of the interannual variability in canopy development associated with ‘early’ vs. ‘late’ spring onset. However, even in the best cases, with correlation coefficients typically in the range of $r \approx 0.5$–0.8, between 30% and 75% of the observed interannual variation

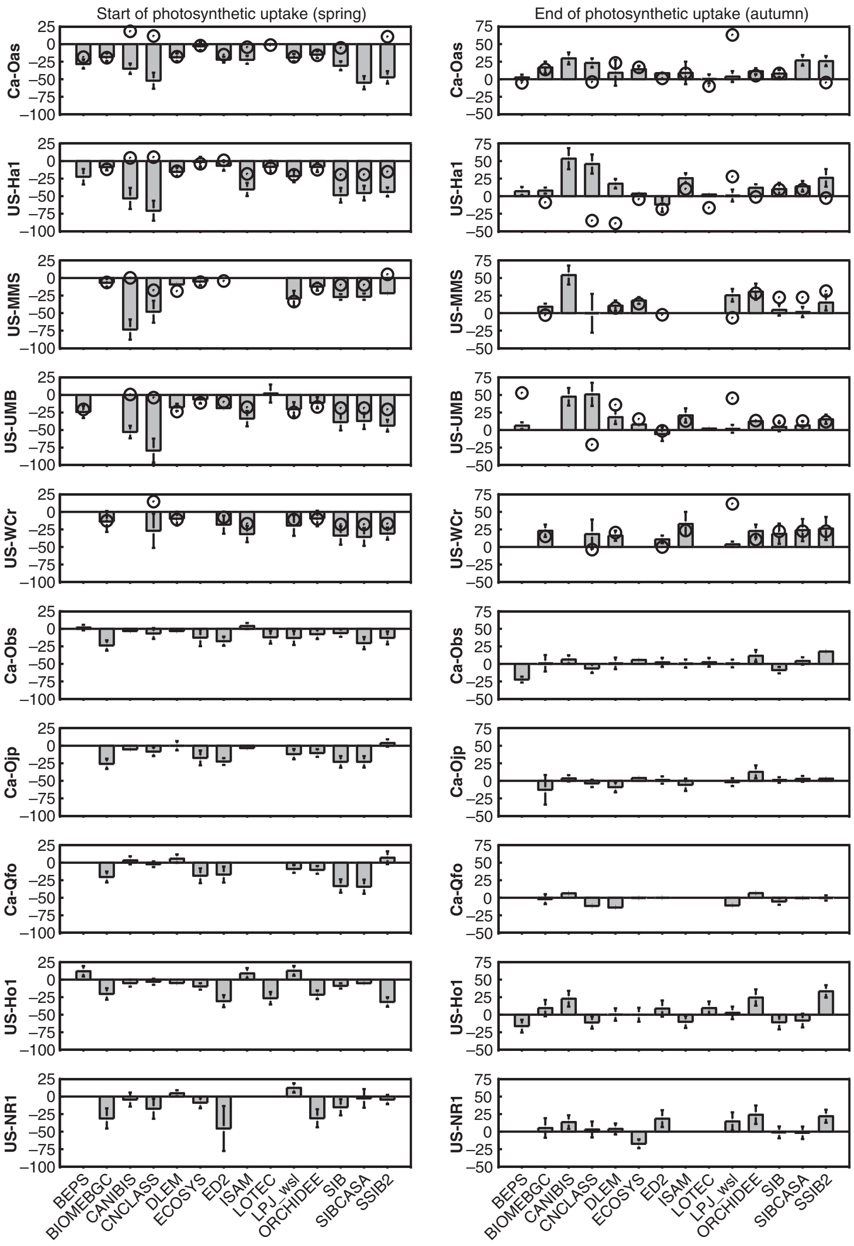


Fig. 2 Difference in number of days (*y* axis) between observed and modeled start (left panels) and end (right panels) of photosynthetically active period for five deciduous broadleaf and five evergreen needleleaf sites. The start and end of photosynthetic uptake are defined as the first and last dates at which daily gross ecosystem photosynthesis (GEP) = 20% of maximum daily GEP. For each of 14 models (*x* axis), bars indicate the mean bias, with error bars indicating the standard deviation across multiple years. Circles indicate bias in the spring and autumn dates at which 20% of seasonal amplitude of leaf area index (LAI) was reached (deciduous sites only). Negative values indicate that the modeled transition occurred *prior* to the observed transition.

remained unexplained by the models. For eight models, the correlation between anomalies in observed and modeled dates at which LAI reached 20% of the seasonal maximum was highly significant in spring (all $P < 0.001$; Table 3). However, for only one model, Ecosys ($r = 0.42$, $P < 0.05$), was there a significant correlation between the observed and modeled anomalies in the autumn date at which this same threshold was reached. Furthermore, although models could predict some of the variation in the spring dates at which LAI reached 20% and 50% of the seasonal maximum, they were much less successful at predicting spring dates at which LAI reached 80% of the seasonal maximum. This indicates deficiencies in model representation of rates

Table 3 Correlation coefficient between observations and model predictions of leaf area index (LAI) transition date anomalies (i.e., years with 'earlier' vs. 'later' spring). Anomalies were calculated on a site-by-site basis, across all years of data for each site (deciduous sites only). LAI transition dates were estimated based on dates at which specific relative thresholds of seasonal development, i.e., 20%, 50%, 80% of seasonal LAI amplitude, were reached

	Spring LAI thresholds			Autumn LAI thresholds		
Model	20%	50%	80%	80%	50%	20%
BEPS	0.18	0.48	0.12	−0.25	−0.20	0.01
Biome-BGC	0.64***	0.64***	0.24	0.23	0.02	−0.29
Can-IBIS	0.57***	0.46*	−0.19			
CN-CLASS	0.69***	0.61***	0.40*	−0.10	−0.14	0.05
DLEM	−0.41*	−0.03	−0.06	−0.39*	−0.11	0.02
Ecosys	0.71***	0.71***	0.66***	0.36	0.47**	0.42*
ED2	0.86***	0.75***	0.41*	0.48**	0.59***	0.26
LoTEC	0.82***	0.92***	0.77***	0.02	0.09	−0.06
LPJ_wsl	0.64***	0.79***	0.59***	0.19	0.00	0.20
ORCHIDEE	0.62***	0.58***	0.37*	0.00	−0.02	−0.12

Note: SiB3, SiBCASA, SSiB2, and ISAM were excluded from this analysis because prescribed phenology did not vary from year-to-year. No autumn correlations are reported for Can-IBIS because predicted LAI did not decrease in autumn.
Asterisks denote statistical significance:
*$P < 0.05$;
**$P < 0.01$;
***$P < 0.001$.

of leaf growth and the sensitivity of leaf growth to interannual climate variability. Finally, the amount of interannual variability predicted by the models was highly variable; CN-CLASS and LoTEC generally predicted too much variability in both spring and autumn developmental threshold dates, whereas in the models with prescribed phenology, there was of course no interannual variability in the dates when different thresholds were reached.

Overall, then, models were generally inadequate in their representation of the timing, and interannual variability in the timing, of both spring green-up and autumn senescence of deciduous forest sites. Better input data could rectify this problem for the subset of models using prescribed LAI (Table 2), but in those with prognostic LAI routines, either model structure or model parameters need to be improved.

Start and end of photosynthetic activity

For deciduous sites, virtually every one of the 14 models included in this analysis predicted an earlier onset of photosynthetic activity (defined as the first date at which daily GEP = 20% of maximum daily GEP) than was indicated by the eddy covariance measurements (Fig. 2). Across all models, sites, and years, the mean bias in photosynthetic onset date was −28 ± 21 days, with a range from −108 to +19 days. Relatively small biases were observed for some models (Ecosys, −3 ± 6 days; LoTEC, −4 ± 9 days), but large biases, of more than 6 weeks, were typical for other models (Can-IBIS, −52 ± 17 days; CN-CLASS, −60 ± 22 days). For evergreen sites, the same pattern was apparent – with three minor exceptions (BEPS, +7 ± 8 days, ISAM, +4 ± 7 days, and LPJ_wsl, +1 ± 14 days); predicted dates of the onset of photosynthesis were earlier than the observed dates – but the biases tended to be somewhat smaller than for deciduous forests (mean bias, all models, sites, and years, −11 ± 15 days). Larger errors were observed for ORCHIDEE (−18 ± 12 days), Biome-BGC (−25 ± 10 days), and ED2 (−29 ± 20 days).

In autumn, most models predicted that the photosynthetic activity of deciduous sites (here judged as the last date at which daily GEP = 20% of maximum daily GEP) persisted later than was actually observed (mean error, all models, sites, and years, +15 ± 17 days) (Fig. 2). Biases were again largest for Can-IBIS (+45 ± 16 days) and CN-CLASS (+32 ± 23 days); the smallest errors were observed for LoTEC (+2 ± 5 days) and ED2 (−2 ± 13 days). By comparison, for evergreen sites there was not a strong bias one way or the other (mean bias, all models, sites, and years, +3 ± 14 days) with respect to predicting the end of photosynthetic uptake. On average, some models were too early, and some models were too late, but for a given model and a given site, even the sign of the bias could vary from year-to-year. However, ORCHIDEE (+18 ± 12 days) and SSiB2 (+18 ± 14 days) were notable exceptions to this general pattern, because both models predicted that photosynthetic uptake would continue, on average, for two and a half weeks longer than was actually observed at the evergreen sites.

For deciduous sites, only about half of the models could predict interannual variation in the timing of the spring onset of photosynthesis with any degree of success, and none of the models consistently captured observed variation in the end of photosynthetic uptake. Six models showed statistically significant correlation with observed anomalies in spring dates at which GEP = 20% of maximum GEP (Table 4, top). However, even for the two best models (ED2, r = 0.66; LoTEC, r = 0.67), these correlation coefficients, although statistically significant, indicate that models capture no more than 50% of the observed interannual variation in the onset of photosynthesis. For the 50% and 80% spring thresholds of maximum GEP, the number of correlations significant at $P < 0.001$ was even lower (three and two models, respectively). For evergreen sites, models (with the exception of ED2) were generally more successful at predicting interannual variation in the timing of the spring onset of photosynthesis (Table 4, bottom). Once more, however, the 80% threshold of maximum GEP was predicted less well than the 20% threshold. At the end of the growing season, most models, whether simulating deciduous or evergreen sites, were unable to explain more than a very small proportion of the interannual variation in the timing of autumn declines in GEP (Table 4).

Table 4 Correlation between observed and modeled gross ecosystem photosynthesis (GEP) transition date anomalies (e.g., years with 'earlier' vs. 'later' transition). Transition dates were determined as the first (spring) and last (autumn) day at which daily GEP = 20%, 50%, or 80% of maximum daily GEP. Anomalies were calculated on a site-by-site basis, across all years of data for each site

	Spring GEP thresholds			Autumn GEP thresholds		
	20%	50%	80%	80%	50%	20%
Deciduous forests						
BEPS	0.45^{**}	0.51^{**}	0.63^{***}	−0.10	0.19	0.11
Biome-BGC	0.56^{***}	0.36^{*}	0.43^{*}	0.19	0.26	−0.16
Can-IBIS	0.20	0.56^{***}	0.32	0.18	0.05	−0.28
CN-CLASS	0.08	0.14	0.38^{*}	0.21	−0.15	−0.20
DLEM	0.53^{***}	0.27	0.23	0.47^{**}	0.37^{*}	−0.21
Ecosys	0.61^{***}	0.72^{***}	0.53^{***}	0.31	0.15	0.45^{**}
ED2	0.66^{***}	0.40^{*}	0.10	0.22	−0.02	0.27
ISAM	0.23	0.06	0.42^{*}	0.13	−0.02	−0.23
LoTEC	0.67^{***}	0.77^{***}	0.50^{**}	0.00	−0.09	0.28
LPJ_wsl	0.41^{**}	0.28	−0.01	−0.06	0.03	0.32^{*}
ORCHIDEE	0.51^{***}	0.30	0.11	0.45^{**}	0.37^{*}	0.03
SiB	−0.03	0.16	0.42^{**}	0.33^{*}	0.27	-0.34^{*}
SiBCASA	−0.07	0.34^{*}	0.13	0.33^{*}	0.31^{*}	−0.29
SSiB2	0.12	−0.18	0.36^{*}	0.17	0.31^{*}	−0.29
Evergreen forests						
BEPS	0.71^{**}	0.37	0.55^{*}	0.34	0.48	0.32
Biome-BGC	0.47^{**}	0.67^{***}	0.57^{***}	0.32	0.38^{*}	0.24
Can-IBIS	0.75^{***}	0.63^{***}	0.26	0.59^{***}	0.09	0.42^{*}
CN-CLASS	0.65^{***}	0.71^{***}	0.02	0.37^{*}	0.29	0.38^{*}
DLEM	0.88^{***}	0.56^{***}	0.05	0.18	0.14	0.38^{*}
Ecosys	0.51^{**}	0.58^{***}	0.49^{**}	0.12	0.27	0.41^{*}
ED2	0.21	0.49^{**}	0.12	0.36^{*}	−0.26	0.18
ISAM	0.80^{***}	0.56^{**}	−0.05	0.45^{*}	0.21	0.27
LoTEC	0.67^{**}	0.69^{**}	0.49	−0.04	0.35	0.42
LPJ_wsl	0.64^{***}	0.66^{***}	0.43^{**}	0.20	0.01	0.31
ORCHIDEE	0.59^{***}	0.46^{**}	0.65^{***}	0.45^{**}	0.40^{*}	0.18
SIB	0.57^{***}	0.53^{***}	0.37^{*}	0.47^{**}	0.30	0.29
SiBCASA	0.45^{**}	0.61^{***}	0.16	0.41^{*}	0.32	0.34^{*}
SSiB2	0.65^{***}	0.64^{***}	−0.09	−0.06	0.24	0.37^{*}

Asterisks denote statistical significance: $^{*}P < 0.05$;
$^{**}P < 0.01$;
$^{***}P < 0.001$.

For deciduous sites, errors in modeling the seasonal dynamics of GEP could largely be attributed to errors in modeling seasonal dynamics of LAI. For both spring (most models) and autumn (some models), biases in modeled dates at which 20% thresholds of LAI and GEP were reached were strongly correlated (for each model, across all sites, and years) with each other (Fig. 2; Table 5). Thus, errors in modeling the beginning of canopy development, and the end of canopy senescence, typically translated directly to corresponding errors in modeling the timing of seasonal dynamics of GEP. Errors in modeling the dates at which 80% thresholds of LAI and GEP were reached were less strongly correlated, probably reflecting a decoupling between photosynthesis and leaf area in models once the canopy is more than half-full. There were, however, some obvious exceptions to these patterns (Fig. 2). For example, for Can-IBIS, the deciduous sites were modeled with a large LAI in winter, and no autumn decline in LAI. CN-CLASS retained varying amounts of leaf area through winter, and for some sites, GEP could increase in spring before any new foliage was formed. For LPJ_wsl, leaf area was retained in autumn much longer than photosynthesis was sustained, resulting in large biases for LAI, but not GEP, threshold dates.

Source/sink transition dates

As was the case with GEP, virtually every model predicted an earlier spring source/sink transition than was actually observed for the deciduous sites (Fig. 3; mean across all models, sites, and years, −32 ± 36 days). Biases of more than 6 weeks were typical for some models (Can-IBIS, −54 ± 42 days; SSiB2, −71 ± 27 days; CN-CLASS, −90 ± 18 days), but biases were negligible for other models (e.g., LoTEC, −4 ± 7 days and Ecosys, −5 ± 9 days). Among deciduous sites, the largest biases (30 days or more) in modeling the spring source/sink transition data were observed for Ca-Oas, US-UMB, and US-Ha1.

Model bias in predicting spring source/sink transition dates was smaller for evergreen sites (mean across all models, sites, and years, −8 ± 33 days). Several models (ED2, −36 ± 28 days; SSiB2, −71 ± 25 days) predicted the transition more than a month early. However, other models (most notably Can-IBIS, +17 ± 32) were generally late in predicting the spring source/sink transition date for evergreen sites.

Across all models, sites, and years, autumn sink/source transition dates were modeled somewhat better than spring source/sink transition dates for deciduous (mean bias, −1 ± 37 days) sites, but considerably worse for evergreen (+42 ± 58 days) sites. However, within each forest type, large variation was observed. For example, for deciduous sites, some models predicted excessively early autumn source/sink transition dates (SSiB2, −68 ± 35 days) while others were late (Can-IBIS, +45 ± 24 days). For evergreen sites, mean model bias in predicting autumn sink/source transition dates was relatively small (although somewhat variable) for some sites (Ca-Qfo, +5 ± 23 days) but very large for other sites (Ca-Ojp, +80 ± 37 days; US-NR1, +80 ± 52

Table 5 Correlation between errors in modeled dates at which gross ecosystem photosynthesis (GEP) and leaf area index (LAI) thresholds (20%, 50%, 80% of seasonal maximum) were reached, across all deciduous sites

	Errors in spring threshold			Errors in autumn thresholds		
Model	20%	50%	80%	80%	50%	20%
BEPS	0.31	−0.06	0.34	0.65^{*}	0.22	0.82^{***}
Biome-BGC	0.87^{***}	0.38	0.23	0.00	0.24	0.50^{*}
Can-IBIS	0.67^{***}	0.36	−0.05			
CN-CLASS	0.38^{*}	0.55^{**}	0.50^{**}	0.38^{*}	0.25	−0.25
DLEM	0.54^{**}	0.66^{***}	0.57^{**}	0.05	0.33	0.44^{*}
Ecosys	0.84^{***}	0.87^{***}	0.31	0.23	0.43^{*}	0.66^{***}
ED2	0.77^{***}	0.50^{**}	0.51^{**}	0.23	0.36	−0.02
ISAM	0.27	0.39^{*}	−0.04	−0.02	−0.06	0.14
LoTEC	0.94^{***}	0.82^{***}	0.03	0.59^{**}	0.00	0.74^{***}
LPJ_wsl	0.56^{**}	0.31	0.42^{*}	0.00	−0.08	0.10
ORCHIDEE	0.85^{***}	0.55^{**}	0.46^{**}	0.22	0.25	0.69^{***}
SiB3	0.49^{**}	0.37^{*}	0.07	0.25	0.50^{**}	0.50^{**}
SiBCASA	0.41	0.18	−0.25	0.08	0.57^{**}	0.53^{*}
SSiB2	0.73^{***}	0.35	0.06	0.26	0.55^{**}	0.53^{**}

Asterisks denote statistical significance: *$P < 0.05$;
**$P < 0.01$;
***$P < 0.001$.

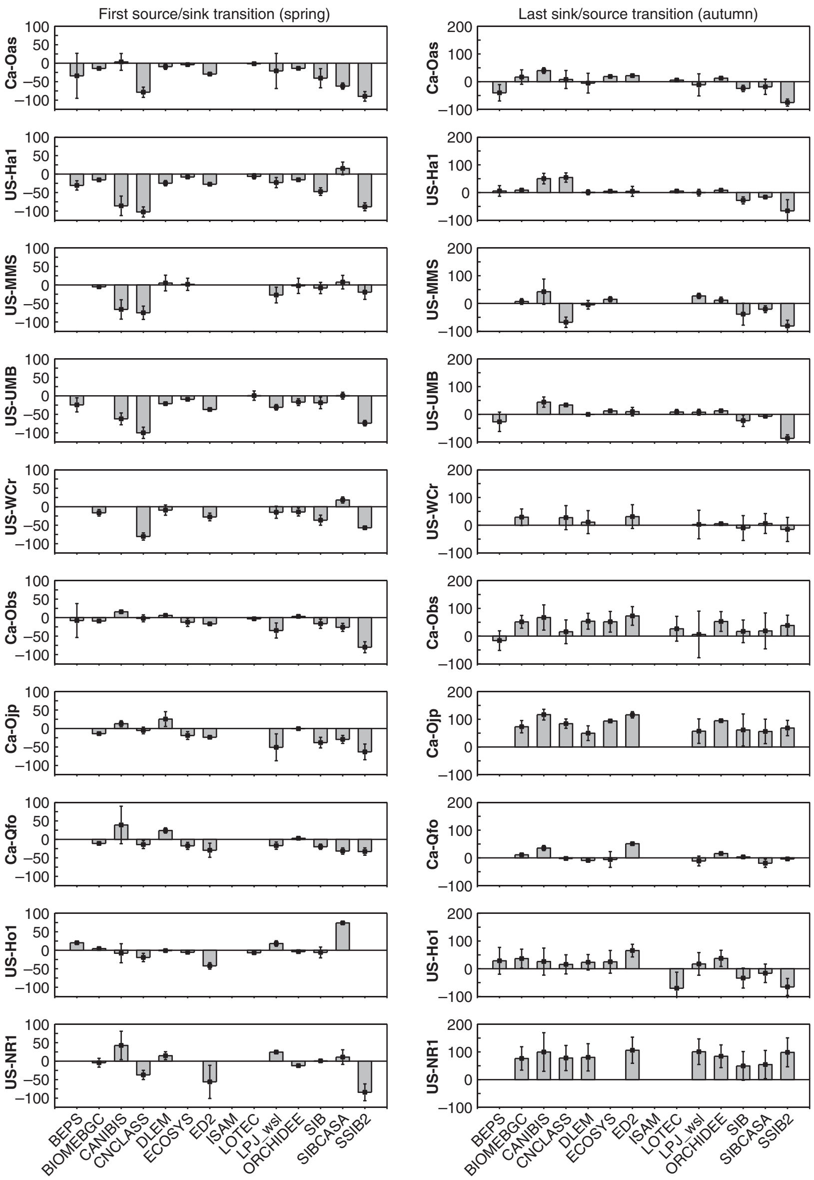


Fig. 3 Difference in number of days (*y* axis) between observed and modeled start (left panels) and end (right panels) of carbon uptake period for five deciduous broadleaf and five evergreen needleleaf forest sites. The start and end of the carbon uptake period are defined as the first spring date and last autumn date, respectively, on which daily net ecosystem exchange of CO_2 (NEE) crossed from a source to a sink, or vice versa. For 13 of 14 models (*x* axis), bars indicate the mean bias, with error bars indicating the standard deviation across multiple years. Negative values indicate that the modeled transition occurred *prior* to the observed transition. No results shown for ISAM, which did not provide NEE output.

days). Both ED2 (+87 ± 38 days) and Can-IBIS (+72 ± 59 days) models predicted autumn source/sink transition dates for evergreen sites that were more than 2 months later than actually observed.

Although many models predicted a statistically significant proportion of the interannual variation in the first date at which daily GEP = 20% of maximum daily GEP (Table 4), only a handful of models could predict the interannual variation in spring source/sink transition dates (Table 6). Of those, only Biome-BGC and LoTEC could explain at least half of the observed variability in spring source/sink transition dates for both deciduous and evergreen sites. Model skill in predicting autumn sink/source transition date anomalies was consistently poor (Table 6).

Just as errors in modeling the seasonal cycle of LAI explained much of the error in modeling the seasonal cycle of GEP, errors in modeling the seasonal cycle of

Table 6 (Left columns) Correlation between observed and modeled NEE source/sink transition date anomalies (e.g., years with 'earlier' vs. 'later' transition). Transition dates were determined as the first (spring) and last (autumn) day at which NEE source/sink transition occurred. Anomalies were calculated on a site-by-site basis, across all years of data for each site. (Right columns) Correlation between errors in modeled dates at which daily GEP = 20% maximum GEP and NEE source/sink transition dates, in spring and autumn. NEE is net ecosystem exchange of CO_2, and GEP is gross ecosystem photosynthesis

	NEE source/sink transition		Errors in GEP and NEE transitions	
	Spring	Autumn	Spring	Autumn
Deciduous forests				
BEPS	−0.27	−0.06	0.20	−0.16
Biome-BGC	0.70***	0.34	0.59**	0.83***
Can-IBIS	0.11	0.42**	0.46**	0.60***
CN-CLASS	0.24	−0.07	0.58***	0.24
DLEM	−0.12	0.04	0.29	0.67***
Ecosys	0.53***	0.38*	0.79***	0.64***
ED2	0.77***	0.00	0.76**	0.34
LoTEC	0.71***	0.08	0.83***	0.77***
LPJ_wsl	0.24	0.06	0.24	0.28
ORCHIDEE	0.25	0.90***	0.73***	0.47**
SiB3	−0.01	−0.19	0.33*	0.47**
SiBCASA	0.14	0.19	0.38*	0.36*
SSiB2	−0.16	−0.01	0.57***	0.46**
Evergreen forests				
BEPS	0.38	−0.07	0.03	0.31
Biome-BGC	0.79***	0.37	−0.10	−0.11
Can-IBIS	−0.03	−0.28	0.51**	−0.01
CN-CLASS	0.60***	0.16	−0.10	0.08
DLEM	0.35	0.09	0.19	−0.26
Ecosys	0.45*	−0.24	0.95***	0.52**
ED2	0.07	0.28	0.88***	−0.05
LoTEC	0.76***	0.15	0.46	0.76***
LPJ_wsl	−0.36	0.04	0.07	−0.06
ORCHIDEE	0.85***	0.40*	0.54**	−0.10
SiB3	0.23	0.03	0.22	0.26
SiBCASA	0.21	−0.16	0.26	0.25
SSiB2	−0.26	0.22	0.46	−0.18

Note: ISAM was excluded from this analysis because model NEE was not provided.
Asterisks denote statistical significance:
*$P < 0.05$;
**$P < 0.01$;

GEP explained, for deciduous sites in both spring and autumn, a sizable fraction of the error in modeling the observed NEE source/sink and sink/source transitions (Table 6). Only for BEPS and LPJ_wsl were errors in predicting NEE transition dates not significantly correlated with errors in predicting GEP transition dates at deciduous sites. But, for evergreen sites, errors in modeling GEP transition dates generally did not explain the errors in modeling NEE transition dates, presumably indicating that interannual variability in the seasonality of ecosystem respiration may contribute substantially to variability in NEE source/sink and sink/source transition dates in this forest type.

Errors in GEP integrals from incorrect representation of seasonality

The total bias in modeled annual GEP was +35 ± 365 g C m^{-2} yr^{-1} for deciduous forests and +70 ± 335 g C m^{-2} yr^{-1} for evergreen forests (mean ± 1 SD across all sites, models, and years; for reference, mean annual GEP was 1250 ± 200 g C m^{-2} yr^{-1} in deciduous forests and 950 ± 375 g C m^{-2} yr^{-1} in evergreen forests). By comparison, the total bias in modeled annual GEP, after correcting for model bias in GEP$_{max}$, was +260 ± 250 for deciduous forests and +55 ± 130 g C m^{-2} yr^{-1} for evergreen forests. Thus, for deciduous sites, re-scaling GEP generally increased the total model bias in annual GEP, indicating that biases in GEP_{max} were effectively compensating for other model deficiencies. As an example, ED2 consistently under-estimated annual GEP in deciduous sites because model GEP_{max} was much smaller than the observed GEP_{max}. However, when modeled daily GEP was re-scaled to account for differences in GEP_{max}, the model typically over-estimated annual GEP because the model predicted a longer growing season than was actually observed (Figs 2 and 4).

Biases in annual simulated GEP were driven nearly equally by misrepresenting the timing of spring and fall transitions and by biases in simulated GEP_{max}. We showed (Fig. 2) that for most deciduous sites, models generally predicted an earlier spring onset (first date at which GEP = 20% of GEP_{max}) of GEP than was actually observed, and a later autumn termination (last date at which GEP = 20% of GEP_{max}) of GEP than was actually observed. Thus, not surprisingly, models over-estimated GEP during both the spring and autumn transition periods (Fig. 4). What is perhaps surprising is the magnitude of this bias; across all sites, models, and years, the mean (±1 SD) total model bias (scaled for GEP_{max}) in deciduous forests was +160 ± 145 g C m^{-2} yr^{-1} in spring and +75 ± 130 g C m^{-2} yr^{-1} in autumn (Fig. 4). Indeed, together (+235 ± 230 g C m^{-2} yr^{-1}) these two biases essentially offset the model error that could be attributed to differences in modeled vs. observed GEP_{max} (−225 ± 440 g C m^{-2} yr^{-1}), and accounted for virtually all of the model bias that remained after correcting for differences in GEP_{max} (+260 ± 250 g C m^{-2} yr^{-1}) for deciduous sites.

Biases in model predictions of spring and autumn GEP transition dates were generally much smaller for evergreen sites than for deciduous sites (Fig. 3). As a result, model biases in GEP during seasonal transition periods (after re-scaling to account for differences in GEP_{max}) tended to be smaller than those for deciduous sites (Fig. 4): +40 ± 80 and −5 ± 65 g C m^{-2} yr^{-1} in spring and autumn, respectively. And, whereas for deciduous sites, spring (≈20% of all sites, models, and years) and autumn (≈10% of all sites, models, and years) GEP errors of ≥ 250 g C m^{-2} yr^{-1} were common (Fig. 4), they were rare to nonexistent (≈1% and 0%, respectively) for evergreen sites.

Discussion

Overview and relevance to modeling climate system feedbacks

The above analysis of predictions from 14 different terrestrial biosphere models has identified four key weaknesses in the representation of phenology and phenologically mediated processes:

1. Most models are biased toward predicting a growing season that is substantially longer than the observed growing season for deciduous forest sites, but biases are smaller and less consistent for evergreen forest sites.
2. Models are typically unable to capture more than a small fraction (albeit, in more than a few cases, a statistically significant fraction) of the observed interannual variability in phenological transition dates. This problem is more pronounced for deciduous forest sites than evergreen forest sites and more pronounced in autumn than in spring.
3. For deciduous sites, errors in modeling the seasonality of LAI often appear to propagate to errors in modeling the seasonality of GEP. This, in turn, leads to errors in modeling the seasonality of NEE.
4. Accumulated biases in GEP during spring and autumn transition periods, attributed to misrepresentation of the seasonality of GEP, are large and highly variable for deciduous sites: +160 ± 145 and +75 ± 130 g C m^{-2} yr^{-1} (≈13% and ≈8% of total annual GEP), respectively. These tend to offset errors associated with under-estimation of the magnitude of the seasonal peak GEP in deciduous sites. Thus, compensating errors may lead to erroneous conclusions about model performance at the annual time step.

From the perspective of global change science, the results presented here are important because (1) phenology is sensitive to climate change and variability; and (2) phenology controls many vegetation feedbacks to the climate system (Morisette *et al.*, 2009). Analyses of diverse data sets provide compelling evidence for phenological shifts toward earlier spring onset and delayed autumn senescence over the last four decades (Peñuelas *et al.*, 2002; Badeck *et al.*, 2004; Schwartz *et al.*, 2006; Parmesan, 2007; Parry *et al.*, 2007). These patterns have largely been attributed to climate change, particularly recent warming trends. However, this analysis suggests that current models are unable to portray adequately the seasonality of either LAI or processes related to ecosystem carbon cycling under present climate scenarios. The analysis by Desai (2010) showed that accurate representation of interannual variability in phenology is important if the corresponding variability in net uptake of CO_2 is to be predicted correctly. We expect that most models (especially those in which phenology is prescribed) will not accurately predict the associated phenological responses to future climate change and variability either, which limits the usefulness of these models for prognostic studies. As will be discussed below, this is

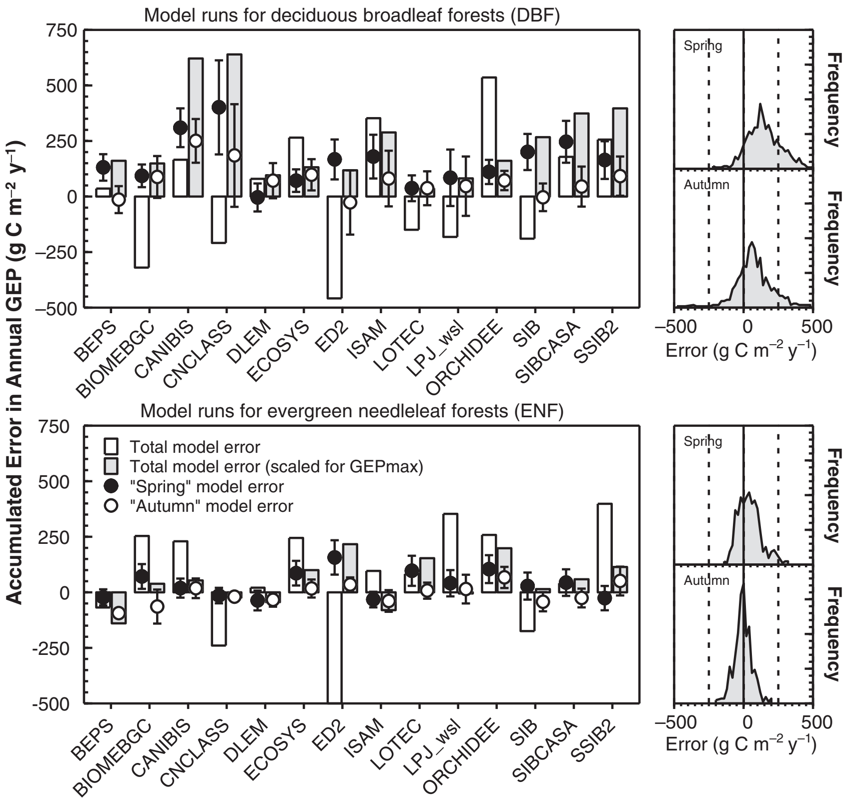


Fig. 4 Errors in modeled gross ecosystem photosynthesis (GEP) for deciduous broadleaf (top) and evergreen needleleaf (bottom) forests. Left panels show error, by model (means and standard deviations across multiple years of data for $n = 5$ sites), as follows: 'Total model bias' is (annual model GEP) − (annual tower GEP); 'Total model bias (scaled for GEP_{max})' re-scales the modeled GEP to account for order-of-magnitude differences between model and tower GEP_{max} and then calculates the bias as (annual model $GEP_{re\text{-}scaled}$) − (annual tower GEP); 'spring' and 'autumn' model biases are calculated from model $GEP_{re\text{-}scaled}$ and tower GEP, with sums calculated for the period during which either model GEP or tower GEP is in the 'increasing GEP' phase ($0.2 \leq$ daily $GEP/GEP_{max} \leq 0.8$) in spring or 'decreasing GEP' phase in autumn ($0.8 \geq$ daily $GEP/GEP_{max} \geq 0.2$). The right panels show the frequency distribution of these spring and autumn biases in re-scaled model GEP, across all models, sites, and years of data, for each forest type: models show a strong and consistent bias in both spring and autumn in deciduous, but not evergreen, forests. The sign convention is that positive bias means that modeled GEP > tower GEP.

an outstanding challenge for phenological modeling in general.

In terms of the second point, our analysis was limited to the seasonality of LAI and ecosystem-atmosphere fluxes of CO_2. In all likelihood, however, these and similar models would also misrepresent other key feedbacks of terrestrial vegetation to the climate system during spring and autumn transition periods, e.g., through changing albedo, surface energy balance adjustment, and the changing partitioning of available energy to latent and sensible heat fluxes. This is of great importance because in addition to an influence on microclimate (e.g., ambient surface temperature, humidity, and radiative transfer through the canopy), phenology has effects on the planetary boundary layer, regional-to-global circulation patterns, and thus continental-scale climatic patterns (Hayden, 1998; Pielke *et al.*, 1998; Chapin *et al.*, 2000; Hogg *et al.*, 2000; Fitzjarrald *et al.*, 2001). Failure to represent phenology accurately in models that couple the land surface to the atmosphere could lead to large errors in the seasonal evolution of regional weather patterns, for example. The study by Levis & Bonan (2004) demonstrated on a regional scale that when phenology was prescribed, model runs using the Community Land Model coupled to the Community Atmosphere Model could not replicate observations that document a reduction in the rate of increase in surface air temperature that occurs coincident with spring leaf emergence and associated increases in transpiration. By comparison, when a prognostic phenology scheme was implemented, the important coupling between biological processes on the land surface and feedbacks to the atmosphere was restored, thereby improving model performance for this diagnostic. Thus, accurate model representation of phenology is critical because of the multitude of climate system feedbacks that are mediated by phenology.

Improving models

What steps are needed to improve phenological submodels in terrestrial biosphere models? For deciduous sites, large biases in predicting the start and end of the growing season need to be resolved, but models also need to do a better job of reproducing the interannual variability in phenology as well. Undoubtedly, progress requires better understanding of the controls on vegetation phenology, and the phenology of ecosystem processes, in different biomes and across plant functional types. For example, although the phenology of temperate, deciduous forests is well studied, there is remarkably little agreement regarding the degree to which photoperiod, cold temperatures, and warm temperatures combine to regulate spring budburst in these ecosystems (Chuine *et al*., 2010; Körner & Basler, 2010). Consequently, numerous models to predict budburst have been described in the literature (e.g., Chuine, 2000; Hänninen & Kramer, 2007; Richardson & O'Keefe, 2009), but there is no consensus on which model works best across species or across geographically distinct populations of a given species.

Forecasts of budburst dates for future climate scenarios are highly uncertain because the predicted response to warming depends strongly on the underlying model structure. Failure to incorporate photoperiodic control and chilling requirements, as is the case for most of the phenology submodels in the 14 models analyzed here, will likely result in over-estimation of the response to future warming (Körner & Basler, 2010). Even ignoring the impact that such errors would have on other aspects of biogeochemical cycling and climate system feedbacks, it is obvious that such biases would only further exaggerate the patterns for LAI and GEP reported here.

In general, neither prescribed nor prognostic schemes did well in simulating site-level phenology and its impact on LAI and GEP. However, this general pattern obscures some notable differences in model structure and performance, especially as related to the representation of environmental controls on photosynthesis and why the models tend to do better at evergreen sites than deciduous sites. Importantly, the LAI at evergreen sites stays nearly constant with time whereas the deciduous sites lose (nearly) all their leaves each winter. The timing of spring uptake at evergreen sites is controlled primarily by temperature, whereas at deciduous sites spring uptake is controlled by both temperature and the production of new foliage. Correspondingly, there are a variety of prognostic schemes for each of several aspects of phenology; (1) the start of leaf onset for deciduous species and removal of dormancy for evergreens, (2) the progress to full LAI in deciduous species and full photosynthetic capacity in evergreens, and (3) the onset and schedule of leaf loss in deciduous species and entry into dormancy for cold winter evergreens.

Models using GDD sums which explicitly or implicitly included a chilling requirement did relatively well in capturing the onset of LAI and GEP for deciduous and evergreen forests. Models that used the GDD approach for deciduous but not evergreen forests (Biome-BGC, ORCHIDEE) consequently did relatively better for the deciduous type. LoTEC demonstrated that optimization of GDD parameters to multi-year mean values improved the model's ability to capture interannual variability in spring LAI. However, optimized phenological parameters were site specific, giving this approach limited power in long-term climate change simulations.

For autumn leaf loss in deciduous species and photosynthetic deactivation in evergreens, temperature thresholds combined with a shorter photoperiod have some predictive utility although there was a range of success in the models employing this approach (Ecosys, Biome-BGC, CN-CLASS, DLEM). In general, most of the variance in site-level autumn decreases in LAI, GEP, and NEE were unaccounted for by the models. Historically, there has been much less of an emphasis on developing models for autumn phenology, but the analysis presented here illustrates the need for new efforts in this direction.

Data needs

Increasing recognition of the importance of phenology should motivate progress toward disentangling the mechanisms and environmental drivers of annual phenological cycles. As a starting point, Stöckli *et al.* (2008) and Randerson *et al.* (2009) have emphasized the 12 importance of developing better data sets with which to test and evaluate model predictions. Comparative analyses of different phenological models have typically used data from only a single site (e.g., Richardson & O'Keefe, 2009), which hinders the development and parameterization of generalized models. Only in a very few studies (e.g., Schaber & Badeck, 2003) have attempts been made to constrain phenological models using data across a wide geographic range. Thus, results from most analyses reflect over-fitting of models to individuals from a particular population, when in reality there may be genetic variation across the native range of a species with respect to the phenological sensitivity to climatic drivers (Chuine *et al*., 1998). An additional drawback is that single-site, short-term observational data often do not span sufficiently wide ranges of environmental or climatic conditions to falsify model predictions and thus distinguish among competing model structures (Hänninen, 1995). Satellite data offer the promise of global coverage, but are hindered by issues of both spatial and temporal resolution. Long-term, spatially extensive ground observations (ideally characterizing the entire seasonal trajectory of canopy development and senescence) are therefore urgently needed to elucidate environmental controls on phenology and improve phenological models. Ecosystem-scale modeling provides an additional challenge, in that it requires methods for scaling up from the phenology of individual species to correctly represent the aggregate phenology of mixed-species stands.

More than a decade ago, Baldocchi *et al.* (1996) recognized the value of eddy covariance time series of CO_2 and H_2O exchanges for evaluating and improving model representation of seasonal vegetation dynamics – i.e., phenology, and its role in regulating ecosystem processes related to carbon and water cycling. Indeed, this has motivated the present analysis. Beyond the NACP Site Synthesis, there are opportunities for related analyses that are even broader in scope. With close to a thousand site-years of measurements, from ecosystems spanning much of the globe's climate and vegetation space, the FLUXNET 'La Thuile' database (http://www.fluxdata.org) is a virtual goldmine for the earth system modeling community (e.g., Williams *et al.*, 2009). Also offering promise for improving phenological models are ground observations from continental-scale monitoring networks. For example, citizen science efforts, such as the USA National Phenology Network (http://www.usanpn.org) or webcam-based efforts such as PhenoCam (http://phenocam.sr.unh.edu; see Richardson *et al.*, 2007, 2009b), could potentially yield spatially extensive data on the phenology of key plant functional types. Phenological observations from multifactor, manipulative global change experiments (e.g., Cleland *et al.*, 2006) would be valuable for constraining model predictions under novel climatic or environmental conditions. Combining these diverse observations within a data model fusion framework (as described in Williams *et al.*, 2009), and in conjunction with objective model selection criteria (as previously applied to phenology models by Richardson & O'Keefe, 2009), it should be possible to develop and parameterize new phenological models and make substantial progress toward reducing biases and uncertainties of the type that have been documented here.

Conclusion

This analysis has shown that errors in simulating phenology and the seasonality of GEP result in large biases in modeling the productivity of DBF, whereas model performance was better for evergreen forests. Improving deciduous forest phenological models, particularly the controls on the seasonality of LAI and relationships among LAI, canopy photosynthesis and environmental drivers, should therefore be seen as a priority for the terrestrial biosphere modeling community. This is a prerequisite to better forecasts of vegetation responses to climate change and variability and is also essential for reducing errors in model representation of many biosphere–atmosphere interactions and climate system feedbacks.

Acknowledgements

We thank the NACP Site Synthesis, the modeling teams, and the AmeriFlux and Fluxnet-Canada Research Network/Canadian Carbon Program PIs who provided the data on which this analysis is based. We also thank the funding agencies that have supported model development and long-term flux measurements. A. D. R. thank Mark Friedl and Steve Running for feedback on a draft manuscript, and Youngryel Ryu for assistance with leaf area index estimates. A. D. R. and D. Y. H. acknowledge support from Office of Science (BER), US Department of Energy, through the Terrestrial Carbon Program under Interagency Agreement No. DE-AI02-07ER64355 and through the Northeastern Regional Center of the National Institute for Climatic Change Research. A. D. R. acknowledges additional support from the National Science Foundation through the Macrosystems Biology program, award EF-1065029. A. R. D. and K. J. D. acknowledge support from the Midwestern Regional Center of the National Institute for Global Environmental Change under Cooperative Agreement No. DE-FC03-90ER61010, USDA Northern Research Station Joint Venture agreement 09-JV-11242306-105, and the Wisconsin Focus on Energy. D. D. acknowledges support from the Office of Science (BER), US Department of Energy, through agreement DE-FG02-07ER64371. G. B. and C. M. G. acknowledge support from the Midwestern Regional Center of the National Institute for Global Environmental Change under Cooperative Agreement No. DE-FC02-06ER64158. K. S. acknowledges support from NOAA Award NA07OAR431011.

References

Arain MA, Yuan FM, Black TA (2006) Soil-plant nitrogen cycling modulated carbon exchanges in a western temperate conifer forest in Canada. *Agricultural and Forest Meteorology*, **140**, 171–192.

Arora VK, Boer GJ (2005) A parameterization of leaf phenology for the terrestrial ecosystem component of climate models. *Global Change Biology*, **11**, 39–59.

Badeck F-W, Bondeau A, Böttcher K, Doktor D, Lucht W, Schaber J, Sitch S (2004) Responses of spring phenology to climate change. *New Phytologist*, **162**, 295–309.

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-Biogeosciences*, **113**, ????–????. 13

Baldocchi D, Valentini R, Running S, Oechel W, Dahlman R (1996) Strategies for measuring and modelling carbon dioxide and water vapour fluxes over terrestrial ecosystems. *Global Change Biology*, **2**, 159–168.

Baldocchi D, Falge E, Gu LH *et al.* (2001) FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. *Bulletin of the American Meteorological Society*, **82**, 2415–2434.

Barr AG, Black TA, Hogg EH, Kljun N, Morgenstern K, Nesic Z (2004) Inter-annual variability in the leaf area index of a boreal aspen-hazelnut forest in relation to net ecosystem production. *Agricultural and Forest Meteorology*, **126**, 237–255.

Barr AG, Black TA, Hogg EH *et al.* (2007) Climatic controls on the carbon and water balances of a boreal aspen forest, 1994–2003. *Global Change Biology*, **13**, 561–576.

Barr A, Hollinger D, Richardson AD (2009) CO_2 flux measurement uncertainty estimates for NACP. *Eos Trans. AGU*, **90**, ???–??? (Fall Meeting Supplement), Abstract B54A-04. [14]

Barr AG, Richardson AD, Hollinger DY (2011) Use of change-point detection for u*-threshold evaluation for the North American Carbon Program Interim Synthesis. In preparation. [15]

Bergeron O, Margolis HA, Black TA, Coursolle C, Dunn AL, Barr AG, Wofsy SC (2007) Comparison of carbon dioxide fluxes over three boreal black spruce forests in Canada. *Global Change Biology*, **13**, 89–107.

Botta A, Viovy N, Ciais P, Friedlingstein P, Monfray P (2000) A global prognostic scheme of leaf onset using satellite data. *Global Change Biology*, **6**, 709–725.

Chapin FS, McGuire AD, Randerson J *et al.* (2000) Arctic and boreal ecosystems of western North America as components of the climate system. *Global Change Biology*, **6**, 211–223.

Chen JM, Menges CH, Leblanc SG (2005) Global mapping of foliage clumping index using multi-angular satellite data. *Remote Sensing of Environment*, **97**, 447–457.

Chuine I (2000) A unified model for budburst of trees. *Journal of Theoretical Biology*, **207**, 337–347.

Chuine I, Cour P, Rousseau DD (1998) Fitting models predicting dates of flowering of temperate-zone trees using simulated annealing. *Plant Cell and Environment*, **21**, 455–466.

Chuine I, Morin X, Bugmann H (2010) Warming, photoperiods, and tree phenology. *Science*, **329**, 277–278.

Churkina G, Schimel D, Braswell BH, Xiao XM (2005) Spatial analysis of growing season length control over net ecosystem exchange. *Global Change Biology*, **11**, 1777–1787.

Cleland EE, Chiariello NR, Loarie SR, Mooney HA, Field CB (2006) Diverse responses of phenology to global changes in a grassland ecosystem. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 13740–13744.

Cook BD, Davis KJ, Wang WG *et al.* (2004) Carbon exchange and venting anomalies in an upland deciduous forest in northern Wisconsin, USA. *Agricultural and Forest Meteorology*, **126**, 271–295.

Cooke JEK, Weih M (2005) Nitrogen storage and seasonal nitrogen cycling in Populus: bridging molecular physiology and ecophysiology. *New Phytologist*, **167**, 19–30.

Deng F, Chen JM, Plummer S, Chen MZ, Pisek J (2006) Algorithm for global leaf area index retrieval using satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, **44**, 2219–2229.

Desai AR (2010) Climatic and phenological controls on coherent regional interannual variability of carbon dioxide flux in a heterogeneous landscape. *Journal of Geophysical Research-Biogeosciences*, **115**, ???–????. [16]

Desai AR, Richardson AD, Moffat AM *et al.* (2008) Cross-site evaluation of eddy covariance GPP and RE decomposition techniques. *Agricultural and Forest Meteorology*, **148**, 821–838.

Dragoni D, Schmid HP, Wayson CA, Potter H, Grimmond CSB, Randolph JC (2011) Evidence of increased net ecosystem productivity associated with a longer vegetated season in a deciduous forest in south-central Indiana, USA. *Global Change Biology*, **17**, 886–897.

El Maayar M, Price DT, Black TA, Humphreys ER, Jork EM (2002) Sensitivity tests of the integrated biosphere simulator to soil and vegetation characteristics in a Pacific coastal coniferous forest. *Atmosphere-Ocean*, **40**, 313–332.

Fitzjarrald DR, Acevedo OC, Moore KE (2001) Climatic consequences of leaf presence in the eastern United States. *Journal of Climate*, **14**, 598–614.

Garrigues S, Lacaze R, Baret F *et al.* (2008) Validation and intercomparison of global leaf area index products derived from remote sensing data. *Journal of Geophysical Research-Biogeosciences*, **113**, ???–???. [17]

Gough CM, Vogel CS, Schmid HP, Su HB, Curtis PS (2008) Multi-year convergence of biometric and meteorological estimates of forest carbon storage. *Agricultural and Forest Meteorology*, **148**, 158–170.

Grant RF, Barr AG, Black TA *et al.* (2009) Interannual variation in net ecosystem productivity of Canadian forests as affected by regional weather patterns - a fluxnet-Canada synthesis. *Agricultural and Forest Meteorology*, **149**, 2022–2039.

Gu L, Post WM, Baldocchi D, Black TA, Verma SB, Vesala T, Wofsy SC (2003) Phenology of vegetation photosynthesis. In: *Phenology: An Integrative Environmental Science* (ed. Schwartz MD), pp. 467–485. Kluwer, Dordrecht.

Hänninen H (1995) Effects of climatic change on trees from cool and temperate regions: an ecophysiological approach to modeling of bud burst phenology. *Canadian Journal of Botany*, **73**, 183–199.

Hänninen H, Kramer K (2007) A framework for modelling the annual cycle of trees in boreal and temperate regions. *Silva Fennica*, **41**, 167–205.

Hanson PJ, Amthor JS, Wullschleger SD *et al.* (2004) Oak forest carbon and water simulations: model intercomparisons and evaluations against independent data. *Ecological Monographs*, **74**, 443–489.

Hayden BP (1998) Ecosystem feedbacks on climate at the landscape scale. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, **353**, 5–18.

Hogg EH, Price DT, Black TA (2000) Postulated feedbacks of deciduous forest phenology on seasonal climate patterns in the western Canadian interior. *Journal of Climate*, **13**, 4229–4243.

Hollinger DY, Aber J, Dail B *et al.* (2004) Spatial and temporal variability in forest-atmosphere CO_2 exchange. *Global Change Biology*, **10**, 1689–1706.

Jain AK, Yang XJ (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**, ???–???. [18]

Ju WM, Chen JM, Black TA, Barr AG, Liu J, Chen BZ (2006) Modelling multi-year coupled carbon and water fluxes in a boreal aspen forest. *Agricultural and Forest Meteorology*, **140**, 136–151.

Kljun N, Black TA, Griffis TJ *et al.* (2006) Response of net ecosystem productivity of three boreal forest stands to drought. *Ecosystems*, **9**, 1128–1144.

Körner C, Basler D (2010) Phenology under global warming. *Science*, **327**, 1461–1462.

Krinner G, Viovy N, de Noblet-Ducoudre N *et al.* (2005) A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, **19**, ???–???. [19]

Krishnan P, Black TA, Barr AG, Grant NJ, Gaumont-Guay D, Nesic Z (2008) Factors controlling the interannual variability in the carbon balance of a southern boreal black spruce forest. *Journal of Geophysical Research-Atmospheres*, **113**, ???–???. [20]

Kucharik CJ, Barford CC, El Maayar M, Wofsy SC, Monson RK, Baldocchi DD (2006) A multiyear evaluation of a Dynamic Global Vegetation Model at three AmeriFlux forest sites: vegetation structure, phenology, soil temperature, and CO_2 and H_2O vapor exchange. *Ecological Modelling*, **196**, 1–31.

Levis S, Bonan GB (2004) Simulating springtime temperature patterns in the community atmosphere model coupled to the community land model using prognostic leaf area. *Journal of Climate*, **17**, 4531–4540.

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-Biogeosciences*, **114**, ???–???. [21]

Moffat AM, Papale D, Reichstein M *et al.* (2007) Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. *Agricultural and Forest Meteorology*, **147**, 209–232. [22]

Monson RK, Turnipseed AA, Sparks JP, Harley PC, Scott-Denton LE, Sparks K, Huxman TE (2002) Carbon sequestration in a high-elevation, subalpine forest. *Global Change Biology*, **8**, 459–478.

Monson RK, Sparks JP, Rosenstiel TN *et al.* (2005) Climatic influences on net ecosystem CO_2 exchange during the transition from wintertime carbon source to springtime carbon sink in a high-elevation, subalpine forest. *Oecologia*, **146**, 130–147.

Moore KE, Fitzjarrald DR, Sakai RK, Goulden ML, Munger JW, Wofsy SC (1996) Seasonal variation in radiative and turbulent exchange at a deciduous forest in central Massachusetts. *Journal of Applied Meteorology*, **35**, 122–134.

Morisette JT, Richardson AD, Knapp AK *et al.* (2009) Tracking the rhythm of the seasons in the face of global change: phenological research in the 21st century. *Frontiers in Ecology and the Environment*, **7**, 253–260.

Parmesan C (2007) Influences of species, latitudes and methodologies on estimates of phenological response to global warming. *Global Change Biology*, **13**, 1860–1872.

Parry ML, Canziani OF, Palutikof JP, van der Linden PJ, Hanson CE (eds) (2007) *Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge UP, Cambridge.

Peñuelas J, Filella I, Comas P (2002) Changed plant and animal life cycles from 1952 to 2000 in the Mediterranean region. *Global Change Biology*, **8**, 531–544.

Peñuelas J, Rutishauser T, Filella I (2009) Phenology feedbacks on climate change. *Science*, **324**, 887–888.

Pielke RA, Avissar R, Raupach M, Dolman AJ, Zeng XB, Denning AS (1998) Interactions between the atmosphere and terrestrial ecosystems: influence on weather and climate. *Global Change Biology*, **4**, 461–475.

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

Richardson AD, O'Keefe J (2009) Phenological differences between understory and overstory: a case study using the long-term Harvard Forest records. In: *Phenology of Ecosystem Processes* (ed. Noormets A), pp. 87–117. Springer, New York.

Richardson AD, Jenkins JP, Braswell BH, Hollinger DY, Ollinger SV, Smith ML (2007) Use of digital webcam images to track spring green-up in a deciduous broadleaf forest. *Oecologia*, **152**, 323–334.

Richardson AD, Hollinger DY, Dail DB, Lee JT, Munger JW, O'Keefe J (2009a) Influence of spring phenology on seasonal and annual carbon balance in two contrasting New England forests. *Tree Physiology*, **29**, 321–331.

Richardson AD, Braswell BH, Hollinger DY, Jenkins JP, Ollinger SV (2009b) Near-surface remote sensing of spatial and temporal variation in canopy phenology. *Ecological Applications*, **19**, 1417–1428.

Richardson AD, Black TA, Ciais P *et al.* (2010) Influence of spring and autumn phenological transitions on forest ecosystem productivity. *Philosophical Transactions of the Royal Society, Series B*, **???**, ???–???. 23

Ryu SR, Chen J, Noormets A, Bresee MK, Ollinger SV (2008) Comparisons between PnET-Day and eddy covariance based gross ecosystem production in two Northern Wisconsin forests. *Agricultural and Forest Meteorology*, **148**, 247–256.

Ryu Y, Nilson T, Kobayashi H, Sonnentag O, Law BE, Baldocchi DD (2010) On the correct estimation of effective leaf area index: Does it reveal information on clumping effects? *Agricultural and Forest Meteorology*, **150**, 463–472.

Sakai RK, Fitzjarrald DR, Moore KE (1997) Detecting leaf area and surface resistance during transition seasons. *Agricultural and Forest Meteorology*, **84**, 273–284.

Schaber J, Badeck FW (2003) Physiology-based phenology models for forest tree species in Germany. *International Journal of Biometeorology*, **47**, 193–201.

Schaefer K, Collatz GJ, Tans P *et al.* (2008) Combined Simple Biosphere/Carnegie-Ames-Stanford Approach terrestrial carbon cycle model. *Journal of Geophysical Research-Biogeosciences*, **113**, ???–???. 24

Schmid HP, Grimmond CSB, Cropley F, Offerle B, Su HB (2000) Measurements of CO_2 and energy fluxes over a mixed hardwood forest in the mid-western United States. *Agricultural and Forest Meteorology*, **103**, 357–374.

Schwalm CR, Williams CA, Schaefer K *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**, ???–???. 25

Schwartz MD (1992) Phenology and springtime surface-layer change. *Monthly Weather Review*, **120**, 2570–2578.

Schwartz MD, Ahas R, Aasa A (2006) Onset of spring starting earlier across the Northern Hemisphere. *Global Change Biology*, **12**, 343–351.

Sellers PJ, Mintz Y, Sud YC, Dalcher A (1986) A Simple Biosphere Model (SiB) for use within general-circulation models. *Journal of the Atmospheric Sciences*, **43**, 505–531.

Sellers PJ, Heiser MD, Hall FG *et al.* (1995) Effects of spatial variability in topography, vegetation cover and soil moisture on area-averaged surface fluxes: a case study using the FIFE 1989 data. *Journal of Geophysical Research-Atmospheres*, **100**, 25607–25629.

Shelford VE (1930) Phenology and one of its modern descendants. *Quarterly Review of Biology*, **5**, 207–216.

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

Stöckli R, Rutishauser T, Dragoni D *et al.* (2008) Remote sensing data assimilation for a prognostic phenology model. *Journal of Geophysical Research-Biogeosciences*, **113**, ???–???. 26

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

Tian HQ, Chen GS, Liu ML *et al.* (2010) Model estimates of net primary productivity, evapotranspiration, and water use efficiency in the terrestrial ecosystems of the southern United States during 1895-2007. *Forest Ecology and Management*, **259**, 1311–1327.

Turner DP, Urbanski S, Bremer D, Wofsy SC, Meyers T, Gower ST, Gregory M (2003) A cross-biome comparison of daily light use efficiency for gross primary production. *Global Change Biology*, **9**, 383–395.

Urbanski S, Barford C, Wofsy S *et al.* (2007) Factors controlling CO_2 exchange on timescales from hourly to decadal at Harvard Forest. *Journal of Geophysical Research-Biogeosciences*, **112**, ???–???. 27

White MA, Thornton PE, Running SW (1997) A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochemical Cycles*, **11**, 217–234.

White MA, de Beurs KM, Didan K *et al.* (2009) Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006. *Global Change Biology*, **15**, 2335–2359.

Williams M, Richardson AD, Reichstein M *et al.* (2009) Improving land surface models with FLUXNET data. *Biogeosciences*, **6**, 1341–1359.

Zhan XW, Xue YK, Collatz GJ (2003) An analytical approach for estimating CO_2 and heat fluxes over the Amazonian region. *Ecological Modelling*, **162**, 97–117.

Zhang XY, Friedl MA, Schaaf CB (2006) Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS): evaluation of global patterns and comparison with in situ measurements. *Journal of Geophysical Research-Biogeosciences*, **111**, ???–???. 28

Author Query Form

Journal: GCB
Article: 2562

Dear Author,
During the copy-editing of your paper, the following queries arose. Please respond to these by marking up your proofs with the necessary changes/additions. Please write your answers on the query sheet if there is insufficient space on the page proofs. Please write clearly and follow the conventions shown on the attached corrections sheet. If returning the proof by fax do not write too close to the paper's edge. Please remember that illegible mark-ups may delay publication.
Many thanks for your assistance.

Query reference	**Query**	**Remarks**
1	AUTHOR: A running head short title was not supplied; please check if this one is suitable and, if not, please supply a short title that can be used instead.	
2	AUTHOR: Please check that authors and their affiliations are correct.	
3	AUTHOR: Please provide the full name of the authors M. Altaf Arain and J. William Munger.	
4	AUTHOR: Please provide telephone and fax numbers for corresponding author.	
5	AUTHOR: Schwalm et al. 2011 has been changed to Schwalm et al., 2010 so that this citation matches the Reference List. Please confirm that this is correct.	
6	AUTHOR: Please provide all the author names with initials for this “work in preparation by Schaeffer et al., Dietze et al., Stoy et al., and Raczka et al.”	
7	AUTHOR: Please provide all the author names with initials for this “Ricciuto et al. in preparation.“	
8	AUTHOR: Sellers et al. 1996 has been changed to Sellers et al., 1995 so that this citation matches the Reference List. Please confirm that this is correct.	
9	AUTHOR: Please define ISAM.	
10	AUTHOR: Pielke 1998 has been changed to Pielke et al., 1998 so that this citation matches the Reference List. Please confirm that this is correct.	

11	AUTHOR: Fitzjarrald et al. 2000 has been changed to Fitzjarrald et al., 2001 so that this citation matches the Reference List. Please confirm that this is correct.	
12	AUTHOR: Randerson (2009) has been changed to Randerson et al. (2009) so that this citation matches the Reference List. Please confirm that this is correct.	
13	AUTHOR: Please provide the page range for reference Baker et al. (2008).	
14	AUTHOR: Please expand the journal name and provide the page range for reference Barr et al. (2009).	
15	AUTHOR: Papers that have not yet been accepted for publication should not be included in the Reference List; they should be cited in the text as 'X. XX, unpublished data'. Please update Reference Barr et al. (2011) if it has now been published, use 'in press' if it has been accepted for publication.	
16	AUTHOR: Please provide the page range for reference Desai (2010).	
17	AUTHOR: Please provide the page range for reference Garrigues et al. (2008).	
18	AUTHOR: Please provide the page range for reference Jain and Yang (2005).	
19	AUTHOR: Please provide the page range for reference Krinner et al. (2005).	
20	AUTHOR: Please provide the page range for reference Krishnan et al. (2008).	
21	AUTHOR: Please provide the page range for reference Medvigy et al. (2009).	
22	AUTHOR: Moffat et al. (2007) has not been cited in the text. Please indicate where it should be cited; or delete from the Reference List.	
23	AUTHOR: Please provide the volume number, page range for reference Richardson et al. (2010).	
24	AUTHOR: Please provide the page range for reference Schaefer et al. (2008).	
25	AUTHOR: Please provide the page range for reference Schwalm et al. (2010).	

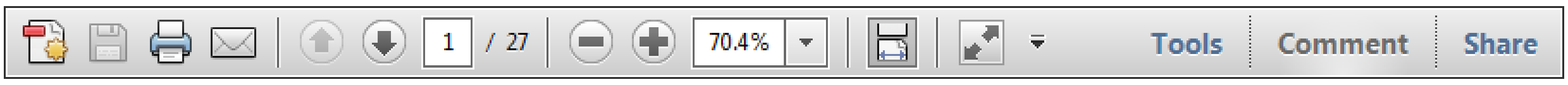
26	AUTHOR: Please provide the page range for reference Stöckli et al. (2008).	
27	AUTHOR: Please provide the page range for reference Urbanski et al. (2007).	
28	AUTHOR: Please provide the page range for reference Zhang et al. (2006).	

WILEY-BLACKWELL

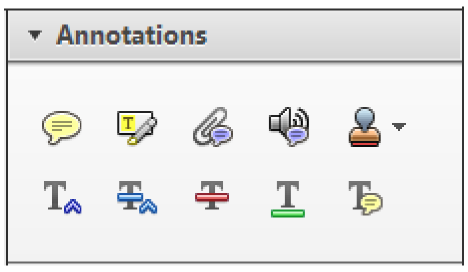
USING e-ANNOTATION TOOLS FOR ELECTRONIC PROOF CORRECTION

Required software to e-Annotate PDFs: Adobe Acrobat Professional or Adobe Reader (version 8.0 or above). (Note that this document uses screenshots from Adobe Reader X)
The latest version of Acrobat Reader can be downloaded for free at: http://get.adobe.com/reader/

Once you have Acrobat Reader open on your computer, click on the Comment tab at the right of the toolbar:



This will open up a panel down the right side of the document. The majority of tools you will use for annotating your proof will be in the Annotations section, pictured opposite. We've picked out some of these tools below:

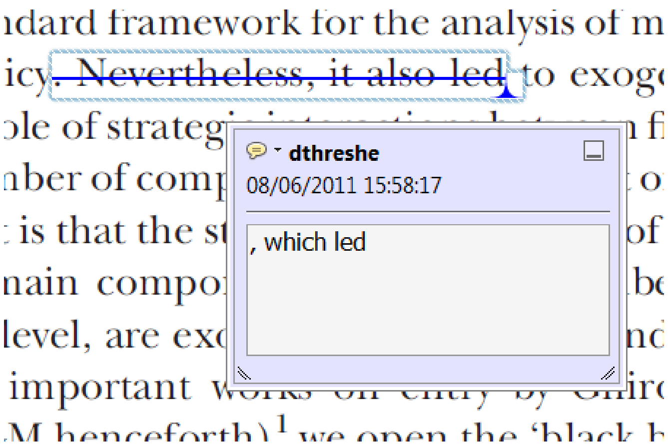


1. Replace (Ins) Tool – for replacing text.

Strikes a line through text and opens up a text box where replacement text can be entered.

How to use it

- Highlight a word or sentence.
- Click on the Replace (Ins) icon in the Annotations section.
- Type the replacement text into the blue box that appears.



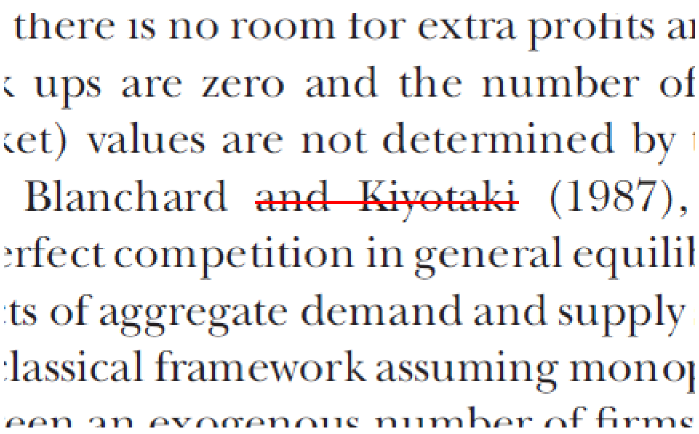
dthreshe
08/06/2011 15:58:17
, which led

2. Strikethrough (Del) Tool – for deleting text.

Strikes a red line through text that is to be deleted.

How to use it

- Highlight a word or sentence.
- Click on the Strikethrough (Del) icon in the Annotations section.

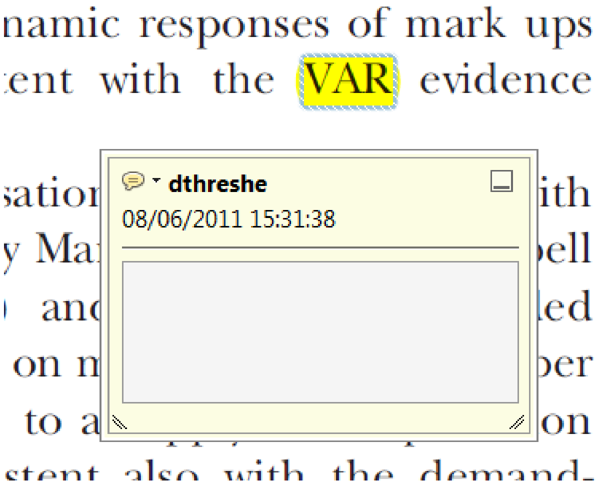


3. Add note to text Tool – for highlighting a section to be changed to bold or italic.

Highlights text in yellow and opens up a text box where comments can be entered.

How to use it

- Highlight the relevant section of text.
- Click on the Add note to text icon in the Annotations section.
- Type instruction on what should be changed regarding the text into the yellow box that appears.



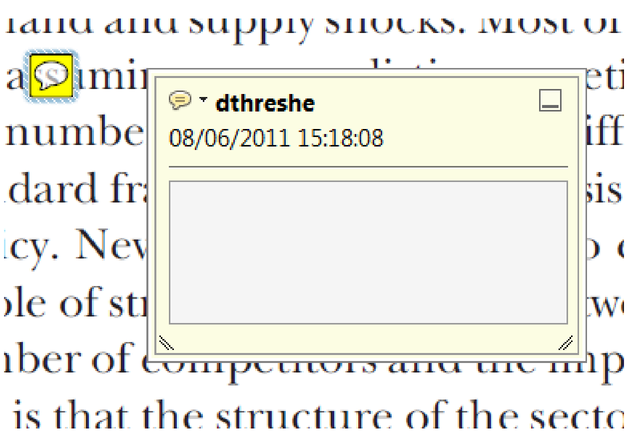
dthreshe
08/06/2011 15:31:38

4. Add sticky note Tool – for making notes at specific points in the text.

Marks a point in the proof where a comment needs to be highlighted.

How to use it

- Click on the Add sticky note icon in the Annotations section.
- Click at the point in the proof where the comment should be inserted.
- Type the comment into the yellow box that appears.



dthreshe
08/06/2011 15:18:08

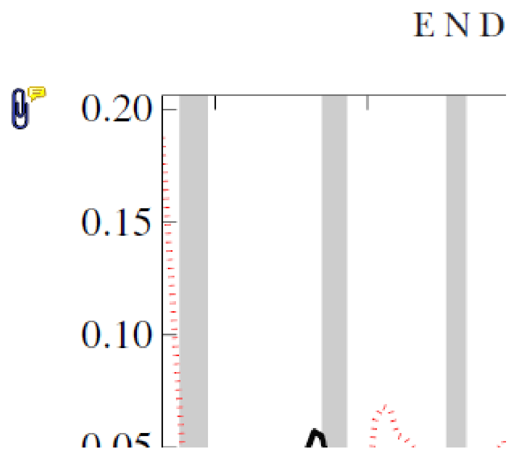
USING e-ANNOTATION TOOLS FOR ELECTRONIC PROOF CORRECTION

5. Attach File Tool – for inserting large amounts of text or replacement figures.

Inserts an icon linking to the attached file in the appropriate pace in the text.

How to use it

- Click on the Attach File icon in the Annotations section.
- Click on the proof to where you'd like the attached file to be linked.
- Select the file to be attached from your computer or network.
- Select the colour and type of icon that will appear in the proof. Click OK.

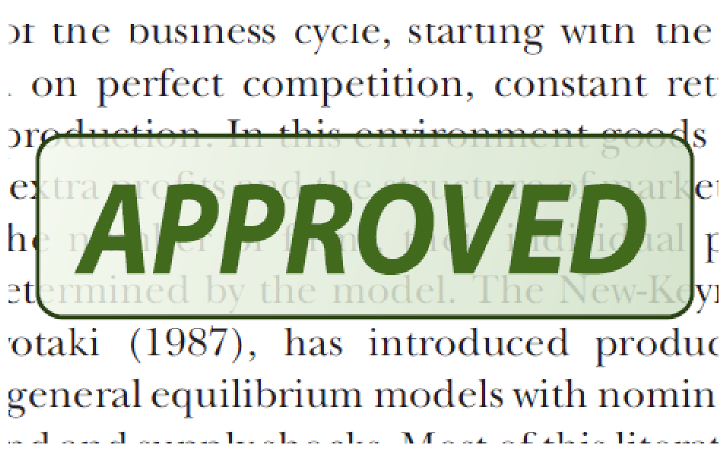


6. Add stamp Tool – for approving a proof if no corrections are required.

Inserts a selected stamp onto an appropriate place in the proof.

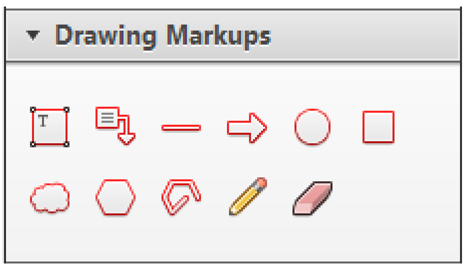
How to use it

- Click on the Add stamp icon in the Annotations section.
- Select the stamp you want to use. (The Approved stamp is usually available directly in the menu that appears).
- Click on the proof where you'd like the stamp to appear. (Where a proof is to be approved as it is, this would normally be on the first page).



the business cycle, starting with the
on perfect competition, constant ret
production. In this environment goods
extra profits and the structure of marke
he number of firms the individual p
etermined by the model. The New-Key
otaki (1987), has introduced produc
general equilibrium models with nomin

APPROVED

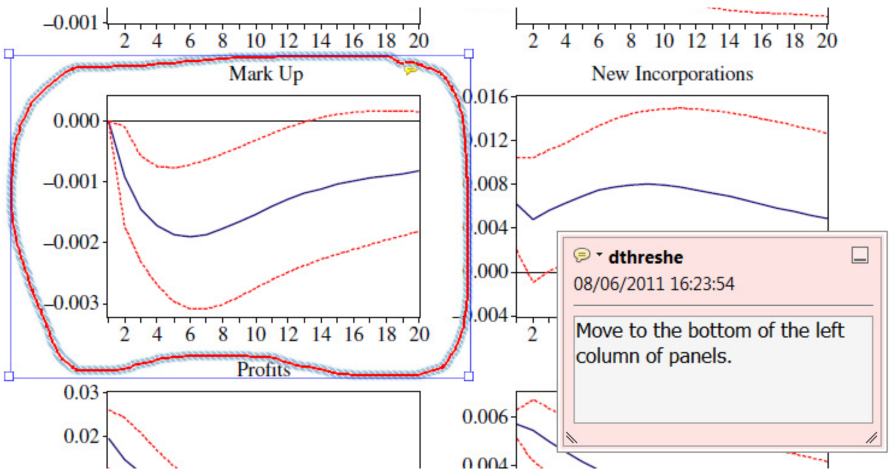


7. Drawing Markups Tools – for drawing shapes, lines and freeform annotations on proofs and commenting on these marks.

Allows shapes, lines and freeform annotations to be drawn on proofs and for comment to be made on these marks..

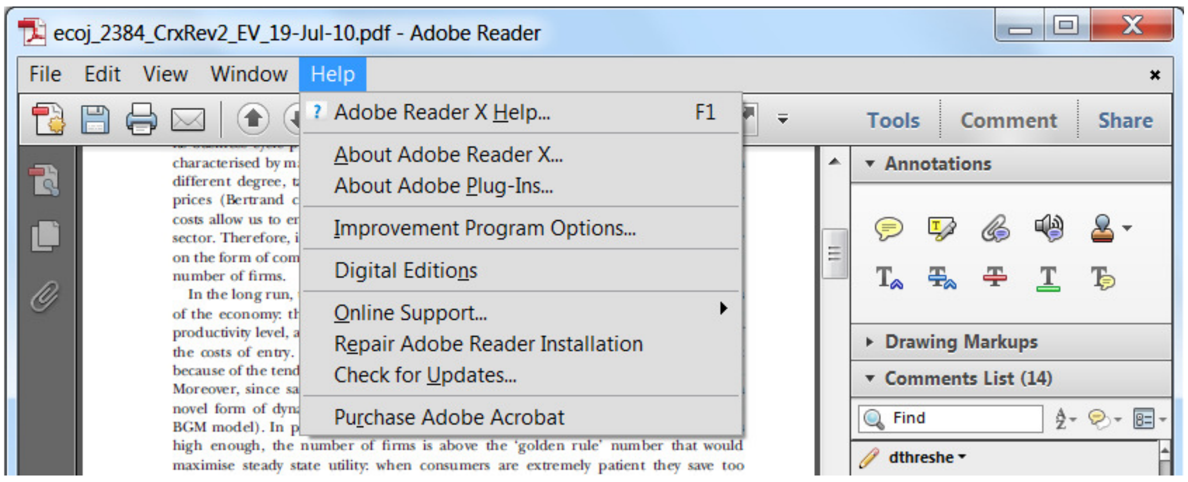
How to use it

- Click on one of the shapes in the Drawing Markups section.
- Click on the proof at the relevant point and draw the selected shape with the cursor.
- To add a comment to the drawn shape, move the cursor over the shape until an arrowhead appears.
- Double click on the shape and type any text in the red box that appears.



dthreshe
08/06/2011 16:23:54
Move to the bottom of the left column of panels.

For further information on how to annotate proofs, click on the Help menu to reveal a list of further options:



ecoj_2384_CrxRev2_EV_19-Jul-10.pdf - Adobe Reader