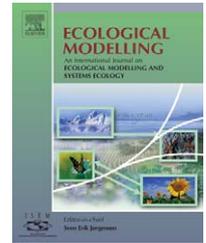


available at www.sciencedirect.comjournal homepage: www.elsevier.com/locate/ecolmodel

A multiyear evaluation of a Dynamic Global Vegetation Model at three AmeriFlux forest sites: Vegetation structure, phenology, soil temperature, and CO₂ and H₂O vapor exchange

Christopher J. Kucharik^{a,*}, Carol C. Barford^a, Mustapha El Maayar^b, Steven C. Wofsy^c, Russell K. Monson^d, Dennis D. Baldocchi^e

^a Center for Sustainability and the Global Environment (SAGE), The Nelson Institute for Environmental Studies, 1710 University Avenue, University of Wisconsin-Madison, Madison, WI 53726, USA

^b Department of Geography, Program in Planning, 100 St. George Street, University of Toronto, Toronto, Ont., Canada M5S3G3

^c Division of Engineering and Applied Sciences, Department of Earth and Planetary Sciences, 29 Oxford Street, Pierce Hall, Harvard University, Cambridge, MA 02138, USA

^d Department of Environmental, Population and Organismic Biology, University of Colorado, Boulder, CO 80309, USA

^e Ecosystem Science Division, Department of Environmental Science, Policy, and Management, 345 Hilgard Hall, University of California-Berkeley, Berkeley, CA 94720, USA

ARTICLE INFO

Article history:

Received 25 March 2005

Received in revised form 9

November 2005

Accepted 15 November 2005

Published on line 27 December 2005

Keywords:

DGVM

IBIS

Carbon balance

Evapotranspiration

Phenology

ABSTRACT

We utilized eddy-covariance observations of carbon dioxide (CO₂) and water vapor exchange at three AmeriFlux mid-latitude forest stands to evaluate IBIS, a Dynamic Global Vegetation Model (DGVM). Measurements of leaf area index (LAI), soil moisture and temperature, runoff, soil carbon (C), and soil respiration (R) were also compared with model output. An experimental approach was designed to help attribute model errors to the vegetation dynamics and phenology formulations versus simulated biological processes. Continental scale phenology sub-models poorly represented the timing of budburst and evolution of canopy LAI in deciduous forests. Biases of vegetation green-up of 6 weeks and delayed senescence were noted. Simulated soil temperatures were overestimated (underestimated) during the summer (winter) on average by 2–5 °C. Ecosystem R was overestimated during the growing season, on average, by 20–60 g C m⁻² month⁻¹, and underestimated during the winter by 10–20 g C m⁻² month⁻¹ at all sites. Simulated soil R failed to capture observed mid-summer peak rates and was generally lower than observed in winter. The overall comparison of simulated net ecosystem production (NEP) to observations showed a significant underestimate of growing season NEP of 25–100 g C m⁻² month⁻¹, and an overall positive bias of 10–40 g C m⁻² month⁻¹ during the winter. Excellent agreement between annual average NEP observations and IBIS simulations in “fixed vegetation” mode resulted from offsetting seasonal model biases. The magnitude of simulated variation in seasonal and inter-annual C

* Corresponding author. Tel.: +1 608 2631859; fax: +1 608 2654113.

E-mail address: kucharik@wisc.edu (C.J. Kucharik).

exchange was generally dampened with respect to observations. The parameterization, and in some cases the formulations (e.g., ecosystem R and phenology) limited model capacity to capture the seasonal fluctuations of C and water exchange. Model parameterizations and formulations were originally constrained and generalized for application to a wide range of global climate and soil conditions and plant functional types (PFTs), likely contributing to model biases. This problem potentially applies to other DGVMs and biosphere models, and will likely become increasingly relevant if investigators apply their models at higher spatial resolution. We suggest that revisions to DGVMs should focus on advancing the capabilities of current phenology formulations to account for photoperiod, soil moisture and frost in addition to temperature. Model representations of PFTs and formulations of ecosystem R need to be rethought, particularly with respect to use of Q_{10} temperature functions as modifiers. Surface energy balance, C allocation, soil R, and plant response to nutrient stress deserve attention as well.

© 2005 Elsevier B.V. All rights reserved.

1. Introduction

Humans have significantly altered the Earth's atmosphere, oceans, and terrestrial ecosystems via greenhouse gas emissions and land-use change (Goolsby et al., 2000; Houghton and Hackler, 2001; Pielke et al., 2002). In order to help guide important societal choices, we must better understand how global systems have been perturbed, and how they may change in the future. In particular, we need basic understanding of exchanges of heat, moisture and trace gases within the planetary boundary layer (PBL). These fluxes are dynamic and responsive to each other, and they combine with land and ocean surface characteristics to influence local microclimates, regional and global atmospheric circulation, and therefore, large-scale climate patterns (Charney, 1975; Hogg et al., 2000). In turn, the Earth's climate system affects vegetation structure, function, and global distribution over long timescales (e.g., decades to millennia). To improve understanding of these climate/biosphere feedbacks, the study of environmental change has gradually adopted an integrated approach, in which global systems are often studied in unison rather than separately, typically with numerical models.

This integrated approach is exemplified by Dynamic Global Vegetation Models (DGVMs), a new class of models that has emerged in the last decade (Foley et al., 1996, 1998; Friend et al., 1997; Woodward et al., 1998; Kucharik et al., 2000; Cramer et al., 2001; Bachelet et al., 2003; Bonan et al., 2003; Sitch et al., 2003; Hickler et al., 2004). These models integrate biogeography, soil biogeochemistry, and soil-vegetation-atmosphere transfer (SVAT) components into the same framework (e.g., Foley et al., 1996; Friend et al., 1997; Kucharik et al., 2000), allowing for vegetation characteristics (e.g., leaf area index, height, biomass and type, root biomass, and albedo), soil moisture, and nutrient availability to respond to atmospheric forcing (climate and carbon dioxide [CO_2]) and land management change. Individual tree, grass, and shrub species are typically grouped into generic plant functional types (PFTs; such as "temperate deciduous tree"), whose spatial distribution is determined by bioclimatic rules and by competition for light and water resources. However, a few approaches simulate competition between individual plants that occurs at finer scales (Smith et al., 2001). These dynamics give rise to simulated "biome" types, e.g., savanna, temperate deciduous forest, boreal evergreen coniferous forest. In addition,

DGVMs have been coupled to Global Climate Models (GCMs) (Foley et al., 1998; Levis et al., 1999; Delire et al., 2003, 2004), where in "coupled mode" these models can explicitly represent the bi-directional feedbacks between vegetation and climate. Although some formulations in DGVMs are simplified relative to ecosystem-scale models (e.g., vegetation dynamics, cf. Hickler et al., 2004; Moorcroft, 2003), and the spatial resolution of the "coupled mode" is relatively coarse, the explicit links between ecological processes and the atmosphere in DGVMs represent a considerable step forward (Arora, 2002; Hickler et al., 2004), and DGVMs are currently indispensable for the study of biome distribution, ecosystem function, and climate feedbacks in the context of both global climate change and land use change (Foley et al., 2000).

Thorough evaluations of DGVMs across varied spatial and temporal scales are needed in order to refine and improve model performance. To date, there have been several validation studies performed with DGVMs at continental to global scales (e.g., Kucharik et al., 2000; Lucht et al., 2002; Bachelet et al., 2003; Bonan et al., 2003; Sitch et al., 2003; Gordon et al., 2004; Gerten et al., 2004; Hickler et al., 2004), where simulated land cover, hydrology, and global carbon (C) balance have been evaluated with land observations, satellite data, or other model output (Cramer et al., 2001). A small number of published studies have evaluated DGVM water, carbon, and energy balance at the individual field scale (Delire and Foley, 1999; El Maayar et al., 2001, 2002; Sitch et al., 2003). To our knowledge, comparisons of DGVM simulated plant phenology with field observations are rarely made, but have recently been performed using satellite information (Lüdeke et al., 1996; McCloy and Lucht, 2004; Kim and Wang, 2005). One downfall of previous ecosystem modeling has been that vegetation structure and/or plant phenology (such as leaf area index [LAI]) have been prescribed in many cases (Arora, 2002), making it difficult to project ecosystem response to future changes in climate or other global change drivers. In the case of DGVMs that simulate changing ecosystem structure (e.g., biomass and C allocation, and plant species distributions) and plant phenology, it is important to examine seasonal and annual carbon, water, and energy exchanges in conjunction with changing vegetation dynamics. Until recently, this type of DGVM validation exercise was not feasible due to the lack of high-quality, longer term observations of vegetation structure and function, together with necessary site descriptions.

The FLUXNET network (and its sub-network of USA sites, called AmeriFlux) was initiated to help address the need for longer term comprehensive data for ecosystem model evaluation (Baldocchi et al., 2001). A key objective of the network is to provide physiologists and modelers with long-term continuous measurements of soil–plant–atmosphere CO₂, H₂O, and energy exchanges within several distinct ecosystem types (e.g., tropical, temperate, and boreal), along with other supplementary ecosystem data (e.g., LAI and phenology, soil C, root biomass, and volumetric soil moisture). At some FLUXNET sites, observations now span over a decade, enabling quantification of mean ecosystem response to climate forcing and identification of anomalies on hourly to inter-annual timescales (Urbanski et al., in preparation). Although these observations are still not long enough to fully validate the simulated vegetation response to climate (i.e., shifting biome distribution), the data records at FLUXNET sites are approaching the needed length, and already show the non-equilibrium species and biomass composition of these ecosystems (Barford et al., 2001; Baker et al., 2003). Nonetheless, modelers now have access to longer term observations of plant phenology in conjunction with observations of short timescale carbon, water, and energy exchange between the biosphere and the atmosphere, and can use these data to assess the ability of their models to reproduce land surface responses to climatic variability (Baldocchi et al., 2001). In some cases, secondary relationships derived from site-specific data can lend additional means to test model formulations. For example, Baldocchi et al. (2005) found a significant relationship between the initiation of net CO₂ uptake of deciduous forests and the time at which springtime soil temperature reached the mean annual air temperature.

The key objective of this study is to evaluate the ability of the Integrated Biosphere Simulator (IBIS) DGVM to simulate seasonal and annual surface-atmosphere exchanges of C and H₂O within three mid-latitude, AmeriFlux forest study sites (Walker Branch, Harvard Forest, and Niwot Ridge), while allowing the simulated vegetation structure to interact dynamically with changing atmospheric and soil conditions. This approach to DGVM validation has yet to be fully explored at the individual site level. We used measurements of vegetation phenology (LAI development and senescence), biomass, canopy height, soil C density, and surface runoff at these sites to further assess model formulations. Specific questions we chose to address are: (a) Does the model satisfactorily simulate forest stand characteristics, phenology (or other stand structure in conifers), and the associated seasonal and annual carbon and water fluxes? (b) How applicable are global scale, generalized model parameterizations at a local scale when simulations of forest stand characteristics and quantification of seasonal and annual carbon and water fluxes are desired? (c) How well do certain commonly used process formulations in the model (e.g., Q₁₀ type equations for simulating respiration) behave with respect to observations, at hourly to inter-annual timescales? Thus, we are interested in testing for model failures due to simulation of the wrong ecosystem characteristics (a), and for model failures due to inappropriate description of ecosystem processes (b or c). The structure and formulations of IBIS are described below. In general, we are interested in quantifying the accuracy and level of error associated with

the modeling approach, and where model improvement is most needed in the future. The results presented here should be of significant interest to many other ecosystem modelers because some of the key parameterizations and formulations used in the IBIS DGVM (such as generic PFTs, the phenology model, and Q₁₀ type functions) are common approaches used in other biosphere models.

2. Methods

2.1. Study site characteristics

We evaluated IBIS at three mid-latitude forest sites in the AmeriFlux network, chosen for their contrasting climate (and elevation) and leaf habit, and their longer term continuity of both eddy-covariance measurements and micrometeorological data. Here, we briefly describe the Walker Branch, Harvard Forest, and Niwot Ridge sites that we used for model validation (Tables 1 and A.1). The Walker Branch (hereafter referred to as WB) watershed field site is located near Oak Ridge, Tennessee (35.97°N latitude, 84.28°W longitude, elevation 365 m). The vegetation at this 60-year-old broad-leaved temperate deciduous forest site regenerated on abandoned pasture (Wilson et al., 2001), and now consists of oak (*Quercus alba* L. and *Q. prinus* L.), hickory (*Carya ovata* (Mill.) K. Koch), maple (*Acer rubrum* L.), and tulip poplar (*Liriodendron tulipifera* L.), and averages 26.0 m in height (Wilson and Baldocchi, 2001; Table 1). The forest is situated on infertile cherty silt loam soils. The annual average air temperature and precipitation (1961–1990) are 13.8 °C and 1355 mm, respectively (Table A.1).

The Harvard Forest (HF) field site is in the central uplands of New England, near Petersham, Massachusetts. The experimental site (42.53°N latitude, 72.17°W longitude, elevation 340 m) has developed from previous agricultural land-use and succession following extensive damage from a hurricane in 1938 (Foster et al., 1992). The 50–70-year-old mixed-hardwood forest (25 m in height) is situated on acidic sandy loam soils (Typic Distrochrepts) and is currently comprised primarily of red oak (*Quercus rubra* L.), black oak and red maple, with interspersed hemlock (*Tsuga canadensis* L.), white pine (*Pinus strobus* L.), and red pine (*Pinus resinosa* Aiton) (Foster et al., 1992; Savage and Davidson, 2001; Table 1). The annual average air temperature and precipitation (1961–1990) are 7.8 °C and 1066 mm, respectively (Table A.1).

The Niwot Ridge (NR) AmeriFlux site is located near Nederland, Colorado (40.03°N latitude, 105.55°W longitude, elevation of 3050 m). The experimental site is surrounded by a 97-year-old sub-alpine coniferous forest (recovering from logging in the very early 1900s; Monson et al., 2002), approximately 11.4 m in height (Table 1), and dominated by sub-alpine fir (*Abies lasiocarpa*), Engelmann spruce (*Picea engelmannii*), and lodgepole pine (*Pinus contorta*) (Monson et al., 2002; Scott-Denton et al., 2003). The soil composition is mainly coarse-textured sand, with a 10 cm organic horizon, and is covered with a sparse amount of lichens and some moss (Scott-Denton et al., 2003). The annual average air temperature and precipitation (1961–1990) are 4 °C and 800 mm, respectively (Monson et al., 2002; Table A.1). Thus, among the study sites chosen, Niwot Ridge is the coolest and driest, while Walker Branch is

Table 1 – Summary of biophysical characteristics at each experimental site

	Walker Branch	Harvard Forest	Niwot Ridge
Location	Tennessee (USA)	Massachusetts (USA)	Colorado (USA)
Latitude (N)	35.97°	42.53°	40.03°
Longitude (W)	84.28°	72.17°	105.55°
Elevation (m)	365	340	3050
Forest type	Temperate broad-leaved deciduous	Temperate broad-leaved deciduous	Sub-alpine conifer evergreen
Stand age (year)	~70	~60	~90
Growing season length (days)	~200	~160	~180
Canopy height (m)	25.0	26.0	11.4
Maximum leaf area index (m ² m ⁻²)	6.0	5.5	2.8–4.2
V _{max} at 15 °C (μmol m ⁻² s ⁻¹)	40	55	20
Soil carbon density (kg C m ⁻²)	7.9	8.8	5.0
Soil type	Silty loam	Sandy loam	Sandy loam

the warmest and wettest, and experiences the longest annual average growing season of the three sites (~200 days).

2.2. AmeriFlux data and ecological and soil measurements

Half-hourly average measurements of net ecosystem exchange (NEE) of CO₂ and water vapor (evapotranspiration; ET) used for model validation were collected from 1995 to 1998 at WB, from 1992 to 1999 at HF, and from 1999 to 2001 at NR, using eddy-covariance. Actual quantities were calculated following Goulde et al. (1996) and Baldocchi et al. (2001). In brief, the eddy-covariance measurement technique employs three-dimensional sonic anemometers to quantify wind velocity and temperature fluctuations within and above vegetative canopies, and infrared absorption gas analyzers are used to quantify corresponding CO₂ concentration fluctuations (Baldocchi et al., 2001; Wilson and Baldocchi, 2001). A detailed description of the instrumentation and methodology used to collect environmental data at AmeriFlux sites is reported by Baldocchi and Wilson (2001) for WB, Goulde et al. (1996) for HF, and Monson et al. (2002) for NR. Empirically derived relationships between CO₂ fluxes and climate forcing, as described in Falge et al. (2001), were generally used to fill-in data gaps caused by missing (e.g., instrumentation failure, power outages, maintenance) or rejected data (e.g., weak vertical mixing) after quality control (Wilson and Baldocchi, 2001). For example, at WB, missing or rejected NEE data were replaced by values derived from empirical relationships between CO₂ fluxes, photosynthetically active radiation (PAR), and air temperature, averaged over 15-day periods (Wilson and Baldocchi, 2001). Daytime ecosystem respiration (R) at WB and NR was estimated, following the approach of Wofsy et al. (1993), using empirical relationships derived from measurements of nighttime NEE (respiration) and soil temperature at 2 and 5 cm depths, respectively. At HF, daytime R was calculated as a function of air temperature, extrapolated from the relationship between observations of nighttime R and air temperature (Goulde et al., 1996). At all sites, a significant portion (on the order of half) of the hourly fluxes were filled using the above methods (Falge et al., 2001), and thus published flux observations are actually mixtures of observed and modeled quantities. However, critical analysis of the WB, HF, and NR flux products as they relate to IBIS

validation is beyond the scope of this paper, and therefore all were treated as equivalent observations. In addition, values of gross ecosystem production (GEP) were subsequently calculated as the summation of net ecosystem production (NEP = -NEE) and R. Inferred values of GEP and R (sum of heterotrophic and autotrophic components) that are derived from measured soil or air temperature and nighttime NEE of carbon are also hereafter referred to as *observed* quantities from the AmeriFlux sites.

Measurements of soil temperature and moisture, runoff, soil carbon density, soil surface CO₂ efflux, leaf area index and leaf onset date, vegetation biomass, and canopy height, when available at the AmeriFlux sites, were compared with simulated values. All sites reported vegetation biomass, average canopy height, and soil carbon density data. At WB, daily average soil temperature at 4 cm, and periodic (weekly to bi-weekly) measurements of LAI (inferred from the transmission of solar radiation through the canopy; Wilson and Baldocchi, 2001; Baldocchi and Wilson, 2001) and volumetric soil water content (0–15 cm; measured by gravimetric and time-domain reflectivity [TDR] techniques; Wilson and Baldocchi, 2001) were available. Estimates of annual soil respiration were produced by a biophysical model (CANOAK; Wilson and Baldocchi, 2001), which was parameterized with data collected from a closed gas-exchange soil chamber system (Hanson et al., 1993). Watershed runoff (summation of surface runoff and drainage) data used in this study are described in Luxmoore and Huff (1989) and Wilson et al. (2001). At HF, continuous 5 cm soil temperature data from 1992 to 1999, periodic LAI measurements from 1998 to 1999, and leaf onset information (<http://harvardforest.fas.harvard.edu/research/lter3/online.html>) from 1990 onward are available. Soil respiration measurements (made with 25 cm PVC collars and an infrared gas analyzer), coupled with 10 cm soil temperature and 0–15 cm soil moisture data (TDR), were collected from June 1995 to 1999 (Savage and Davidson, 2001; available at <http://harvardforest.fas.harvard.edu/data/p00/hf006/hf006.html>). We chose to compare our simulations with data obtained from measurement locations that were well-drained (designated SWF) to the southwest of the HF eddy-covariance tower because of the soil type (fine sandy loam) and dominant vegetation cover (mixed hardwoods) located in that transect (Savage and Davidson, 2001). At NR, continuous measurements of 5 cm soil temperature were available from 1999 to 2001. Because NR is

an evergreen conifer site, only very periodic measurements of LAI are available. Investigators have previously reported a maximum average value of $4.2 \text{ m}^2 \text{ m}^{-2}$ but additional data suggests that the LAI at the site ranges between 2.8 and 4.2. Soil moisture, soil CO_2 respiration data, and additional soil temperature data were provided by Laura Scott-Denton at the University of Colorado (Scott-Denton et al., 2003).

2.3. IBIS model description

Version 2.6 of IBIS (Kucharik et al., 2000) was used in this investigation. The model is designed with a hierarchical conceptual framework (i.e., Fig. 1 in Kucharik et al., 2000), and includes several sub-models that are organized with respect to their characteristic temporal scale:

- *Land surface processes*—the model simulates the energy, water, carbon, and momentum balance of the soil–plant–atmosphere system at a half-hourly time step using the LSX land surface scheme of Pollard and Thompson (1995). IBIS includes two vegetation layers with eight potential forest plant functional types in the upper canopy, and two grasses (cool and warm season) and two shrub PFTs in the lower canopy. This number of PFTs is generally higher than several other DGVMs (e.g., HYBRID, TRIFFID, VECODE, and SDGVM), and is on the same level of differentiation as the LPJ model (Cramer et al., 2001). The model state description includes six soil layers of varying thicknesses to a 4 m depth (0–10, 10–25, 25–50, 50–100, 100–200, and 200–400 cm), which are parameterized with biome-specific root biomass distributions of Jackson et al. (1996), and varied soil texture and corresponding physical attributes (Campbell and Norman, 1998). Physiologically based formulations of leaf-level photosynthesis (Farquhar et al., 1980), stomatal conductance (Ball, 1988; Collatz et al., 1991, 1992), and respiration (Ryan, 1991) control canopy exchange processes, and parameters vary according to generalized vegetation categories (e.g., trees, shrubs, or C3 and C4 grasses). This approach to leaf-level physiology is shared with other current DGVMs (e.g., LPJ and SDGVM; Cramer et al., 2001). Leaf-level photosynthesis is scaled to the canopy level by assuming that photosynthesis is proportional to the absorbed photosynthetically active radiation (APAR) within the canopy. Stem and root respiration are dependent on the magnitude of the C pool, a constant maintenance respiration value, and for stems, the sapwood fraction of the total stem biomass (Kucharik et al., 2000). An Arrhenius temperature function modifies the autotrophic respiration rate (Lloyd and Taylor, 1994), where the reference temperature and temperature sensitivity value are consistent for all biome types.
- *Vegetation phenology*—IBIS uses a simplified relationship between accumulated growing degree-days and budburst. Simple temperature thresholds are used to initiate litter-fall for deciduous vegetation; changing photoperiod is not accounted for at the end of the growing season for leaf senescence. In general, soil moisture, soil frost depth, and the availability of soil nutrients and stored carbohydrates are not accounted for in the generalized global-scale phenology formulation.

- *Vegetation dynamics*—IBIS simulates changes in vegetation structure on an annual time step through PFT competition for light (e.g., shading differences) and water (e.g., root profile differences) from common resource pools. The competition between PFTs is driven by differences in resource availability (light and water), carbon allocation, phenology (evergreen, deciduous), leaf-form (needleleaf, broadleaf), and photosynthetic pathway (C3 versus C4) (Foley et al., 1996; Kucharik et al., 2000). Carbon allocation to leaf, stem, and root C pools and the residence times in each are constant values that only vary according to PFT (Kucharik et al., 2000). This approach is similar to several other published DGVMs that do not simulate competition between individual plants (e.g., HYBRID of Friend et al., 1997; SDGVM of Woodward et al., 1998). The vegetation distribution results are area-averaged, and the horizontal and vertical representations of canopy structure are homogeneous. There are currently very few DGVMs that explicitly simulate individual plants or cohorts and competition for light and water between them (e.g., LPJ-GUESS; Smith et al., 2001).
- *Soil biogeochemistry*—IBIS accounts for daily flows of C and nitrogen through vegetation, detritus, and soil organic matter similarly to the CENTURY model (Parton et al., 1987) and the biogeochemistry model of Verberne et al. (1990), which is a strategy adopted by other DGVMs (Cramer et al., 2001). The rate of litter and soil C pool decomposition is modified by an Arrhenius temperature function (Lloyd and Taylor, 1994) and an empirical relationship between soil water-filled pore space and the relative rate of microbial activity (Linn and Doran, 1984). The parameterizations of these functions are consistent regardless of litter and soil C pools and biome type.

The current version of the model does not account for leaf nitrogen effects on photosynthesis, nor the effects of herbivory or disease on LAI and accumulated biomass. Furthermore, there is no explicit representation of fire disturbance such as is found in the LPJ model (Sitch et al., 2003).

2.4. Model parameterizations

Thirty-minute micrometeorological observations from each AmeriFlux site were used to drive IBIS simulations. These observations included air temperature, precipitation, downward shortwave and longwave radiation, wind speed, and relative humidity. Annual weather data summaries for each site are provided in Table A.1. Precipitation data were only available at a daily time step at HF; thus, half-hourly data were derived by dividing the daily values by 48. Downward longwave radiation at HF was calculated using Brutsaert's formulae (Brutsaert, 1975). To parameterize leaf-level photosynthesis equations, the maximum rate of carboxylation (V_{cmax}) was set to $40 \mu\text{mol m}^{-2} \text{ s}^{-1}$ for WB (at 15°C ; Wilson et al., 2000a), $55 \mu\text{mol m}^{-2} \text{ s}^{-1}$ at HF (at 15°C ; Williams et al., 1996), and $20 \mu\text{mol m}^{-2} \text{ s}^{-1}$ at NW (at 15°C ; Wullschlegel, 1993) (Table 1). Observed soil texture information (sand, silt, and clay fractions) was used for model simulations at WB. For simulations of the HF and NR sites, specific soil textural observations were not available as a function of depth; therefore, dominant soil textural information from the STATSGO dataset (Miller and

White, 1998) were substituted. The atmospheric CO₂ concentration was set constant at 360 ppm for all simulations.

As with other biophysical models, IBIS produces output (characterized by the simulation of a specific, single “point”) that is an averaged representation of the mean ecosystem response to changing environmental conditions above (e.g., cloudiness, precipitation, temperature, wind, and humidity), below (soil water and temperature), and within (e.g., CO₂, temperature, light, water, and wind) each vegetative canopy. We performed two separate simulations for each forest ecosystem in order to understand whether simulated processes, parameterizations, or a combination of both affected model performance.

2.4.1. Fixed vegetation runs

Fixed vegetation simulations are used specifically to evaluate the functioning of the SVAT components of the model separately from the vegetation dynamics routines. Thus, it helps discriminate between model errors originating from the land surface parameterizations and processes (e.g., such as photosynthesis and respiration), and errors resulting from simulations of the development of each forest stand species structure, with accompanying errors in vegetation phenology. In fixed vegetation (FV) simulations, the vegetation dynamics routines in IBIS were not used. Therefore, competition for light (shading differences) and water uptake between individual PFTs was not simulated, nor were the phenology algorithms implemented. The WB and HF simulations were parameterized with a temperate broadleaf deciduous tree PFT, and NR was parameterized with a temperate needleleaf evergreen tree PFT, based on the current vegetation description (Table 1). Observations of canopy height and soil carbon density were also used to parameterize the model state description for each site, while available observations of LAI were used to prescribe the vegetation phenology. The partitioning of total soil C into active, intermediate, and stabilized pools (Kucharik et al., 2000) was based upon soil C accumulations resulting from an additional “spin-up” simulation that allowed the aboveground vegetation and soil to develop from an initial state of bare ground and zero soil C storage (i.e., a dynamic vegetation run). In these simulations, forest re-growth was allowed to respond to varying weather and soil environmental conditions, using representative micrometeorological data from 1994, 1992, and 1999 for WB, HF, and NR, respectively.

2.4.2. Dynamic vegetation runs

In simulations where the dynamic vegetation (DV) routines were employed (e.g., continuous competition for light and water between upper and lower canopy PFTs and annual changes in plant phenology), only site-specific half-hourly micrometeorological data and soil textural information were used as model drivers. Each simulation is used to portray the development of a forest stand through time (e.g., the simulated time is the approximate age of each site; 58 years for WB, 60 years for HF, and 90 years for NR; Table 1), with particular attention to the mature vegetation structure (aboveground biomass in the upper and lower canopies, canopy height, and LAI) and below-ground soil C and nitrogen storage (Table 1), for comparison with observations. Dynamic vegetation simulations were initialized with bare ground (no vegetation), zero

snow cover and soil carbon, and uniform soil temperature (5 °C) and soil moisture (50% of pore space) at time = 0. We used two different phenology approaches to test the simulation of plant phenology and LAI, total biomass, canopy height, and soil carbon.

Plant phenology is an extremely important component of dynamic vegetation models because it controls the seasonal evolution of LAI and therefore variations in carbon and water exchange between the land surface and the atmosphere (White et al., 1999; White and Nemani, 2003; Baldocchi et al., 2005). The global-scale version of IBIS (v. 2.6) used in this study adopts simplified leaf onset and senescence algorithms (hereafter referred to as IBIS-Simple). For cold deciduous plants, budburst is initiated when the accumulated growing degree-days (GDD) on a 0 °C basis exceeds 100 for tree PFTs. For grass and shrub PFTs, budburst is initiated when accumulated GDD on a –5 °C basis exceeds 150. The accumulation of GDD begins when the 10-day average air temperature exceeds 0 °C for trees, and –5 °C for grasses and shrubs. Leaf senescence for grasses and shrubs is initiated when the 10-day average air temperature is less than or equal to 0 °C. For tree PFTs, leaf senescence is initiated when the 10-day average air temperature is less than 0 °C, or is less than the mean temperature of the coldest month of the year (based on 30 years climatology) plus 5 °C. The daily LAI is the product of the peak or maximum LAI and the leaf display function that is calculated as a function of GDD, which fluctuates between 0 and 1. The peak LAI is calculated annually in IBIS for each plant functional type by dividing the simulated leaf carbon, during the previous year, by the specific leaf area (m² leaf area per kg dry matter). IBIS assumes a 15-day transition period to increase LAI from the minimum to maximum value and also to decrease LAI during senescence for cold deciduous PFTs.

Because all of our study sites are located in the United States (USA), we also examined the performance of another phenology approach that was calibrated using satellite observations limited to the conterminous USA (White et al., 1997). This approach (hereafter referred to as IBIS-White) incorporates an air temperature function, and other empirical functions of precipitation, soil temperature, and day length. We performed the same suite of DV simulations using these calibrated algorithms to examine if simulated phenology could be improved over the global-scale phenology parameterizations. White et al. (1997) reported that a 10-day error in satellite-derived observations of greenness onset and offset dates is not unreasonable. All carbon and water balance output that is compared with observations use the simplified global-scale phenology algorithms.

2.5. Model testing and validation approach

At each site, observations of monthly NEE (g C m⁻²) and latent heat flux densities (MJ m⁻²) were used for model comparisons. The monthly latent heat flux values were converted to values of evapotranspiration (mm month⁻¹). Daily average soil temperature at the flux tower sites and instantaneous soil moisture and temperature readings, some of which were collected concurrently with soil surface CO₂ efflux measurements at other plots surrounding the NR and HF flux towers, were compared with model output. Annual estimates of soil respiration

at WB were compared with those provided by a biophysical model parameterized for the study site (CANOAK; Wilson and Baldocchi, 2001).

Model performance was quantified in several ways. The correlation between observed and simulated monthly NEE, ET, soil respiration, soil temperature, and moisture was used to calculate the coefficient of determination (r^2), which was used as a relative index of model performance, and was based on the following equation:

$$r^2 = \left[\frac{\sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P})}{\left(\sum_{i=1}^n (O_i - \bar{O})^2 \right)^{1/2} \left(\sum_{i=1}^n (P_i - \bar{P})^2 \right)^{1/2}} \right]^2$$

in which O_i and P_i are the individual observed and model simulated values, respectively, and \bar{O} and \bar{P} are the mean of the observed and simulated values, respectively. Correlation between measured and simulated annual totals of NEE, ET, soil respiration, and leaf onset at HF was used to determine if the model could satisfactorily capture the measured longer term inter-annual variability of these quantities. We also calculated mean bias errors (MBE) and mean absolute bias (MAB) for monthly and annual totals using the following equations:

$$MBE = \frac{\sum_{i=1}^n (P_i - O_i)}{n}$$

$$MAB = \frac{\sum_{i=1}^n |P_i - O_i|}{n}$$

The MBE calculations provide an estimate of whether the model has tendencies to over-predict (i.e., positive bias) or under-predict (i.e., negative bias) quantities with respect to observations, whereas the MAB value is an absolute measure of the amount of simulated mean deviation from observations (Hanson et al., 2004). Ratios of belowground to total NPP (Gower et al., 1999), NPP to GPP (Waring et al., 1998), root to total soil respiration, and plant to total ecosystem respiration (Sanderman et al., 2003) on an annual basis were also computed for comparison with published observations (Baldocchi and Wilson, 2001).

The model performance was evaluated to allow assessment of: (1) the realism of simulated dynamic vegetation processes (e.g., plant competition) that lead to present-day vegetation structure and phenology; (2) the ability of the model to simulate the observed inter-annual and seasonal variability of coupled water-carbon fluxes; (3) the validity of global-scale PFT parameterizations for simulating carbon and water exchanges at smaller spatial scales; (4) the validity of certain process formulations in the model (e.g., phenology and respiration) across a range of timescales.

3. Results and discussion

3.1. Walker Branch: 1995–1998

3.1.1. Leaf phenology and canopy structure

The DV simulations for Walker Branch yielded significant differences between observed and simulated mature canopy structure in terms of plant biomass, canopy height, and total

organic soil C for both phenology approaches (Table 2). Vegetation height, which is crudely approximated by IBIS as a simple linear function of biomass density, influences other simulated processes through calculations of surface roughness length. Both phenology schemes (Simple and White, respectively) underestimated: (1) mature stand total biomass (by 12.3 and 26.5%), (2) canopy height (by 25.6 and 37.6%), and (3) total soil C (by 43.2 and 23.5%; Table 2). The global phenology scheme produced better agreement with observations than the approach calibrated with USA data (White), largely because the two phenology approaches led to the eventual development of different PFT compositions. The 4-year average of weekly LAI values (Fig. 1) suggested: (1) leaf onset and green-up occurred approximately 7 and 4 weeks early in the IBIS-Simple and IBIS-White simulations, respectively; (2) the maximum average LAI was 15–20% higher (5.9 and 6.4 for IBIS-Simple and IBIS-White, respectively) than observations; (3) leaf senescence was approximated satisfactorily by the IBIS-White scheme, but was 2 weeks later than observations using the IBIS-Simple scheme (Fig. 1); (4) average wintertime LAI was 2.6 in IBIS-White (versus ~0.5 in IBIS-Simple, excluding the 1997–1998 El Niño episode), indicating a significant presence of the temperate evergreen coniferous PFT.

Because temperate deciduous tree PFT leaf onset occurred 3 weeks later using the White et al. (1997) approach (compared to IBIS-Simple), the temperate evergreen coniferous PFT was able to capture additional shortwave radiation early in the growing season, and thereby competed more efficiently for light with the temperate deciduous tree PFT during the early years of simulated re-growth. Model parameterizations of specific leaf area (SLA; 12.5 m² kg⁻¹ for temperate evergreen conifers and 25.0 m² kg⁻¹ for temperate deciduous), V_{max} (40 μmol m⁻² s⁻¹ for temperate deciduous and 30 μmol m⁻² s⁻¹ for temperate evergreen conifers), and C allocation (30 and 40% to stems and roots, respectively, for temperate evergreen conifers; 50 and 20% to stem and roots, respectively, for temperate deciduous trees) contributed to the differences in accumulated biomass, canopy height, and soil C between the IBIS-White and IBIS-Simple simulations. The higher fraction of C allocated belowground for evergreen conifers contributed to larger accumulation of soil C in the IBIS-White simulation (Table 2). Because of the higher V_{max} , higher SLA, and larger aboveground woody allocation parameterizations, the deciduous tree PFT has the potential for higher LAI, C assimilation, and biomass accumulation than evergreen coniferous trees. During the middle of the growing season, the peak LAI for the evergreen coniferous tree PFT was 0.5 m² m⁻² in the IBIS-Simple simulation (5.4 m² m⁻² for temperate deciduous PFT) and 2.6 m² m⁻² in the IBIS-White simulation (3.8 m² m⁻² for the temperate deciduous tree PFT). The observed forest canopy at WB is nearly all broad-leaved deciduous, with maximum LAI of 5.5–6.0 (Wilson and Baldocchi, 2001).

3.1.2. Soil temperature

Simulated daily average 0–10 cm soil temperature was compared with measured daily average 4 cm temperatures. The correlation between measured and computed values for the FV and DV runs yielded a coefficient of determination (r^2) of 0.96 and 0.98, respectively. Although IBIS captured the over-

Table 2 – Measured and modeled forest biomass density (Mg C ha^{-1}), soil carbon density (Mg C ha^{-1}), and canopy height (m)

Site and model simulation	Total biomass	Canopy height	Soil carbon
Walker Branch			
Measured	110.3	25.0	79.0
Dynamic-run/Simple	96.7	18.6	44.9
Dynamic-run/White	81.1	15.6	60.4
Harvard Forest			
Measured	128.6 ^a	26.0	88.0
Dynamic-run/Simple	111.2	21.5	68.8
Dynamic-run/White	63.8	12.3	69.4
Niwot Ridge			
Measured	61.0 ^b	11.4	50.0
Dynamic-run/Simple	57.7	10.7	91.4
Dynamic-run/White	34.6	6.7	123.6

The terms Simple and White refer to the simulations performed using simple global-scale phenology algorithms of version 2 of IBIS (Kucharik et al., 2000) and White et al. (1997), respectively. Simulated values are averages over 1995–1998, 1992–1999, and 1999–2001 periods for Walker Branch, Harvard, and Niwot Ridge forests, respectively.

^a Data on biomass of leaves are not available and hence are not accounted for.

^b Data taken from the NR web-site (<http://spot.colorado.edu/~monsonr/sitesdes.html>) and assuming that a 1 kg dry mass contains 0.5 kg C. Root biomass is not included (not available).

all seasonal changes of soil temperature (>95% of observed daily variability), a warm surface soil temperature bias existed during the growing season of about 2.5°C , and a wintertime cold bias of approximately -1.5°C in both FV and DV model runs (Fig. A.1). Because IBIS produced a substantial bias in both simulation types, we cannot attribute the errors to the improper simulation of canopy structure and phenology. In fact, we expected soil temperature in the DV simulation to be less than observations during the growing season

because LAI was 20% higher than measurements, which would have decreased shortwave radiation reaching the soil surface. However, two other IBIS validation studies (Delire and Foley, 1999; El Maayar et al., 2001) reported similar soil temperature biases and overestimation of soil heat fluxes, a problem that is inherent within many land surface models (El Maayar et al., 2001). The reasons for such biases include improper simulation of wintertime snow dynamics and the absence of organic soil properties and a surface litter (thatch) layer in IBIS,

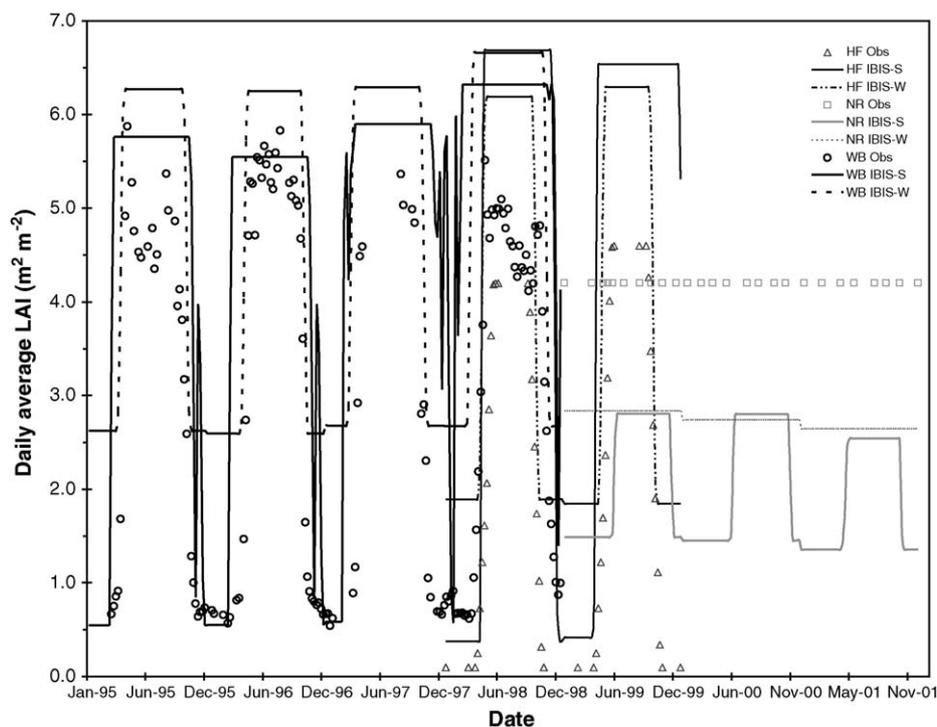


Fig. 1 – Comparisons of LAI observations (obs) with IBIS DV simulations for two phenology approaches (IBIS-Simple [S] and IBIS-White [W]) at the Harvard Forest (HF), Niwot Ridge (NR), and Walker Branch (WB) AmeriFlux sites.

which would act to insulate the soil surface (Delire and Foley, 1999).

3.1.3. Soil moisture and surface runoff

We compared simulated daily average soil volumetric water content (0–10 cm) with periodic, instantaneous measurements at WB (0–15 cm) obtained using gravimetric techniques in 1995, and time-domain reflectometry (TDR) in 1996–1998. The correlation between observed and the FV and DV simulations yielded r^2 of 0.29 and 0.24, respectively. The average 0–15 cm volumetric water content measured over the 4 years was $0.23 \text{ m}^3 \text{ m}^{-3}$, whereas averages for the FV and DV simulations on the same measurement days yielded 0.33 and $0.29 \text{ m}^3 \text{ m}^{-3}$, equivalent to MBEs of 47 and 27%, respectively (Fig. A.2). Thus, the model had a significant wet bias in the top 10 cm, but simulations captured the general timing of seasonal fluctuations of soil drying and wetting (Fig. A.2).

When comparing IBIS runoff with observations, the quantity refers to the sum of soil surface and sub-surface (drainage) components. The two streams that drain the east and west catchments of the 97.5 ha WB watershed have been monitored for discharge, and runoff is calculated by dividing the discharge by the total area of the catchment. The observed annual average runoff during 1995–1998 was 900 mm, while the FV and DV simulations yielded 815 and 895 mm (Table A.2;

Fig. 2a), respectively (MBE of -9.5 and -0.6% underestimates for FV and DV). The FV and DV simulations captured 53 and 57%, respectively, of the inter-annual variability in observed runoff. When the simulated monthly totals ($n=48$) of runoff were compared with observations, the DV run captured 72% of the monthly variability, while the FV run yielded an r^2 of 0.49 (Fig. 2b). The monthly MAB values were 31.0 and $22.7 \text{ mm month}^{-1}$ for the FV and DV simulations, respectively. We also compared the observed monthly average runoff ($n=12$) with simulations across all years using correlation. We found that both FV and DV runs underestimated the peak seasonal runoff in March and April by approximately 20% or 25 mm (Fig. 2c). The DV and FV simulations captured 87 and 58%, respectively, of the observed runoff seasonality.

3.1.4. Evapotranspiration

Evapotranspiration is simulated in IBIS as the sum of three components: soil evaporation, plant transpiration, and water that is intercepted by canopy foliage and evaporated. The annual average ET measured by eddy-covariance (latent heat flux densities) was 557 mm, compared to 743 and 626 mm for FV and DV simulations, respectively, which is equivalent to MBEs of 33 and 12% (Fig. 3a). Hanson et al. (2004) reported ET simulated by 12 ecosystem models run at WB; the mean value was 650 mm, with a range of 596–706 mm from 1993 to

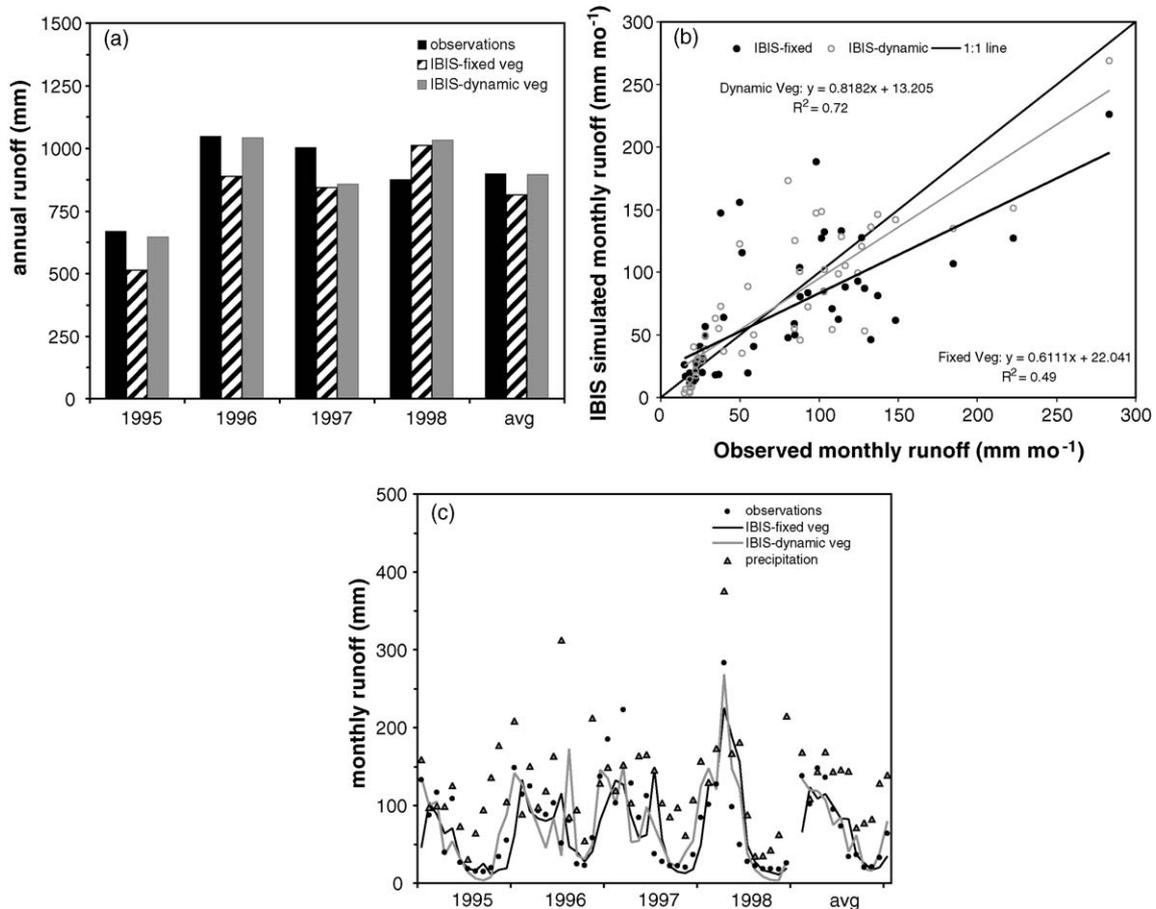


Fig. 2 – (a) Annual total runoff simulated by IBIS fixed vegetation and dynamic vegetation runs compared to observations for Walker Branch; (b) IBIS simulated monthly runoff plotted against observations; (c) comparison of simulated and observed seasonal runoff changes.

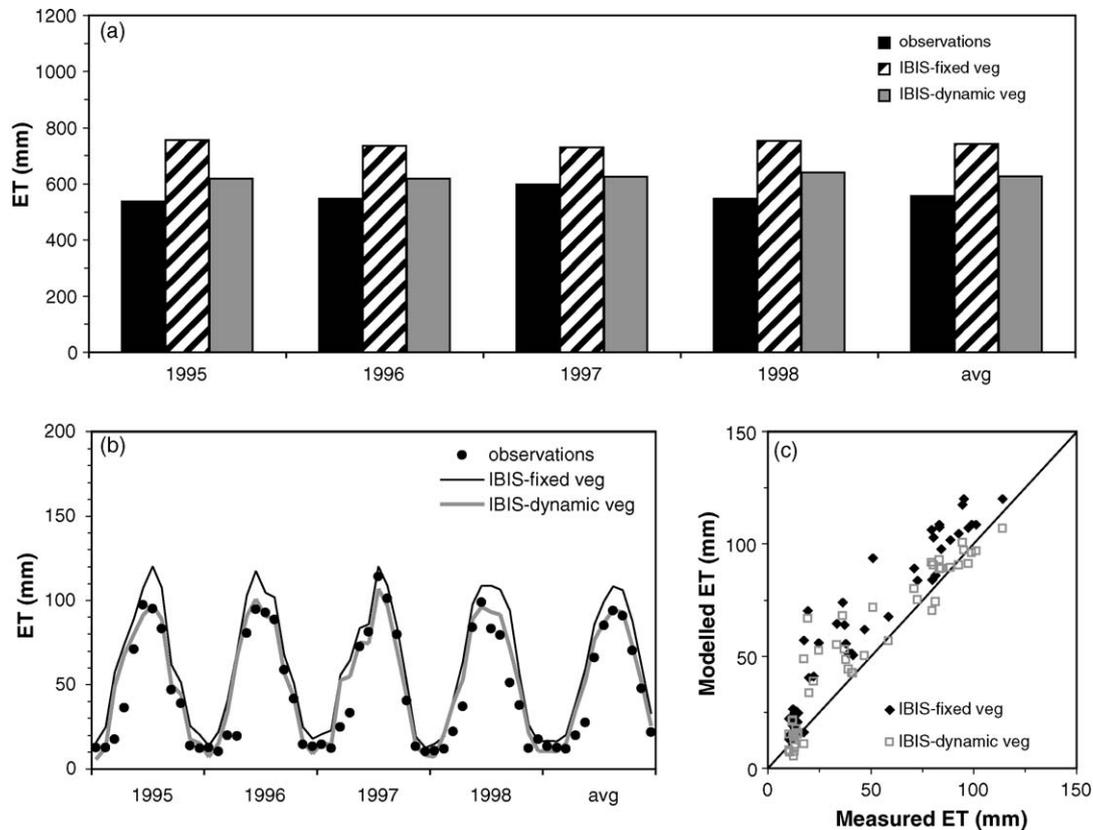


Fig. 3 – Simulated and observed (a) annual total ET at Walker Branch watershed; (b) monthly ET; (c) correlation of simulated and observed monthly ET.

2000. Thus, it appears that many numerical models overestimate the observed ET at WB. The FV and DV simulations also failed to capture inter-annual ET variability, with an r^2 of 0.57 and 0.0, respectively. The FV simulation captured 90% of the overall monthly variability ($n=48$), while the DV run yielded an r^2 of 0.88 (Fig. 3b and c). However, monthly MBEs of 15.5 and 8.7 mm month⁻¹ for the FV and DV simulations, respectively, still existed. We compared the observed average ET for each month ($n=12$) with simulations for the 1995–1998 period. The FV and DV simulations captured 94 and 92% of the observed seasonality, respectively. However, monthly average ET during the growing season was overestimated by ~15–30% (MBE 10–30 mm month⁻¹) in the FV scenario, with the maximum error occurring in March and April (MBE of 130 and 116%, respectively; Table A.2). The DV simulation overestimated ET by 3–5 mm month⁻¹ from May to October, but had the largest error in March and April (MBE 106 and 92%, respectively).

Because LAI was higher in the DV run, the fraction of total ET that was re-evaporated (intercepted) water increased from 6% (FV run) to 10%. Transpiration comprised 81% of the total DV ET (compared with 65% in FV run), and soil evaporation subsequently decreased from 29 to 9% of total ET. For a simple comparison with other ecosystem models, Hanson et al. (2004) reported that a 12-model ensemble average of transpiration at WB was 444 mm, which comprised 68% of the total simulated ET. The higher DV LAI also contributed to cooler soil temperatures, which resulted in higher fractional snow cover

and duration from December to March. As a result, wintertime runoff increased in DV runs.

Based on the consensus of numerous ecosystem model results for WB (e.g., Hanson et al., 2004), and the known difficulty of achieving energy budget closure at WB, we do not necessarily conclude that the overestimation of observed ET by IBIS in both FV and DV simulations is due to model process formulation errors. Because ET depends primarily on the available energy partitioning at the surface, we verified that simulated net radiation (R_n) was in agreement with observations. While the model adequately characterized monthly R_n within 3% of measurements (data not shown), the measured annual energy budget components at the WB site exhibited an annual imbalance of 654 MJ m⁻² year⁻¹ or 22% of the R_n (Table 3), which is equivalent to 261 mm of evaporated water. Based on the observed energy partitioning of annual R_n at WB (Table 3; Wilson et al., 2000b), approximately 50% of the imbalance could be attributed to latent heat flux. If we adjust the observations accordingly, the average annual measured ET would increase to 687 mm, which is 8% less than IBIS simulated ET. The inability of the eddy-covariance technique to achieve energy budget closure at several FLUXNET sites was recently reported by Wilson et al. (2002). This suggests that IBIS estimates of annual ET at WB may be more representative of the true conditions (Hanson et al., 2004), and biosphere models like IBIS could be used to diagnose errors in eddy-covariance data. Unfortunately, if simulated plant phenology is not correct, then the simulated R_n will also be in error.

Table 3 – Measured energy balance components at each of the experimental sites

Site	R_n	H	LE	G	Imbalance	Imbalance (%)
Walker Branch	2990	940	1399	–4	654	22
Harvard Forest	2367	1049	1116	na	202	9
Niwot Ridge	2732	1328	1641	–3	–234	–9

Shown are net radiation (R_n), sensible heat flux (H), latent heat flux (LE), and soil heat flux (G). Energy imbalance was calculated as ($R_n - H - LE - G$). Fluxes are in $\text{MJ m}^{-2} \text{ year}^{-1}$. na: not available. Values are averages over 1995–1998, 1992–1999, and 1999–2001 periods for Walker Branch, Harvard, and Niwot Ridge, respectively.

3.1.5. Carbon fluxes

3.1.5.1. *Gross ecosystem production.* Fig. 4a–c and Table A.2 show simulated annual and monthly GEP compared with observations. The observed annual average GEP was 1594 gC m^{-2} , and FV and DV simulations produced annual average values of 1579 gC m^{-2} (–1.0% MBE) and 1734 gC m^{-2} (8.8% MBE), respectively. Our simulated values were comparable to two other independent annual average GEP estimates of 1473 and 1693 gC m^{-2} reported by Falge et al. (2002) and Hanson et al. (2004), respectively, that were made over longer time periods. The DV run captured 98% of the observed inter-annual variability, but had a positive MBE, and exhibited a higher range of annual GEP (i.e., difference between minimum and maximum values was 220 gC m^{-2} compared to 157 gC m^{-2} for observations), both presumably due to overestimating seasonal LAI changes, and simulating evergreen coniferous trees ($\text{LAI} = 0.5 \text{ m}^2 \text{ m}^{-2}$). The FV simulation had an r^2 of 0.01, demonstrating little sensitivity to inter-annual variability of environmental conditions. The range of annual GEP for the FV run was comparable to observations, with a value of 125 gC m^{-2} .

Fig. 4b and c suggests that IBIS produced consistent seasonal biases for both FV and DV simulations. The FV run reproduced the observed monthly GEP variability ($r^2 = 0.94$) slightly better than the DV simulation ($r^2 = 0.80$), particularly during the growing season. Simulated monthly MBE (Table A.2) was between –4 and –15% from May to September in the FV run and –11 to –26% in the DV run. The most significant error was present from late fall through late spring in both model simulations, when observed GEP was below 50 gC m^{-2} . The large errors evident in the DV run from November to February were likely due to C assimilated by temperate evergreen conifer trees while soil temperatures are above 0.0°C . The DV simulated errors in March and April were attributed to the model bias of early leaf onset.

3.1.5.2. *Ecosystem respiration.* Fig. 4a, d, and e, and Table A.2 illustrate simulated annual and monthly R compared with observations. The annual average observed R was 1029 gC m^{-2} , and FV and DV simulations yielded annual average values of 1233 gC m^{-2} (19.7% MBE) and 1448 gC m^{-2} (40.6% MBE), respectively. The FV and DV simulations captured 61 and 32%, respectively, of the observed inter-annual R variability, and produced a much higher range of annual R between the minimum and maximum values (131 and 146 gC m^{-2} , respectively) compared to observations (28 gC m^{-2}). The FV scenario had a significant overall positive bias when measured monthly R was $>50 \text{ gC m}^{-2}$. However, the correlation between the monthly IBIS output and observations suggested that the

FV ($r^2 = 0.92$) and DV ($r^2 = 0.89$) simulations demonstrated similar behavior in capturing the observed changes in seasonal R. The apparent model errors in simulating R may be attributed to the following: (1) a prolonged period of C assimilation resulting from poor characterization of the site phenology and the peak LAI value; (2) a generalized model formulation of stem and root respiration and their relationship to temperature; (3) simulated heterotrophic respiration that is modified by generalized temperature (Arrhenius; Lloyd and Taylor, 1994) and soil moisture (water-filled pore space; Linn and Doran, 1984) functions; (4) difficulty in measuring nighttime ecosystem R (i.e., nighttime NEE) using gap-filling techniques (Curtis et al., 2002) at the WB site. In support of the last point above, Hanson et al. (2004) found that the mean annual simulated R at WB made by 10 independent ecosystem models was 1306 gC m^{-2} , which was 27% higher than observations. Thus, there is corroborating evidence that a portion of the simulated IBIS error may be attributed to underestimates of observed nighttime R, given the known problem of weak vertical canopy mixing at WB.

3.1.5.3. *Net ecosystem production.* The simulated and observed annual total and seasonal patterns of NEP are presented in Fig. 4f and g, and Table A.2. The observed annual average eddy-covariance NEP was 564 gC m^{-2} , compared to 345 gC m^{-2} (–39% MBE) and 286 gC m^{-2} (–49% MBE) for the IBIS FV and DV simulations, respectively. A study by Curtis et al. (2002) also estimated average annual NEP using two methods: (1) combining NPP and soil respiration (R_s), assuming heterotrophic respiration (R_H) was 50% of total R_s and (2) from measurements of C pool size changes (e.g., stocks). Curtis et al. (2002) reported annual NEP of 252 and 264 gC m^{-2} using the $\text{NPP} + R_s$ and carbon stocks methods, respectively. In comparison, the Hanson et al. (2004) model evaluation study reported high and low biometric NEP values of 253 and 191 gC m^{-2} , respectively, and the average from nine ecosystem models was 258 gC m^{-2} . All of these were based on just the 1995–1998 period. These annual net carbon storage estimates agreed well with our simulated average results. Conversely, other studies by Falge et al. (2002) and Sanderman et al. (2003) have reported annual average NEP of 757 and 500 gC m^{-2} , respectively, which are two to three times greater than other observations and model simulations.

At the inter-annual timescale, the DV run captured 80% (r^2) of the observed variability, and exhibited an absolute range (minimum–maximum) of annual NEP (130.6 gC m^{-2}) that was comparable to eddy-covariance measurements (148.6 gC m^{-2}), but the relative variability was almost a factor of two greater ($\text{CV} = 19.2\%$ compared to 11.6%). The correlation between the annual observed NEP and the FV simulation

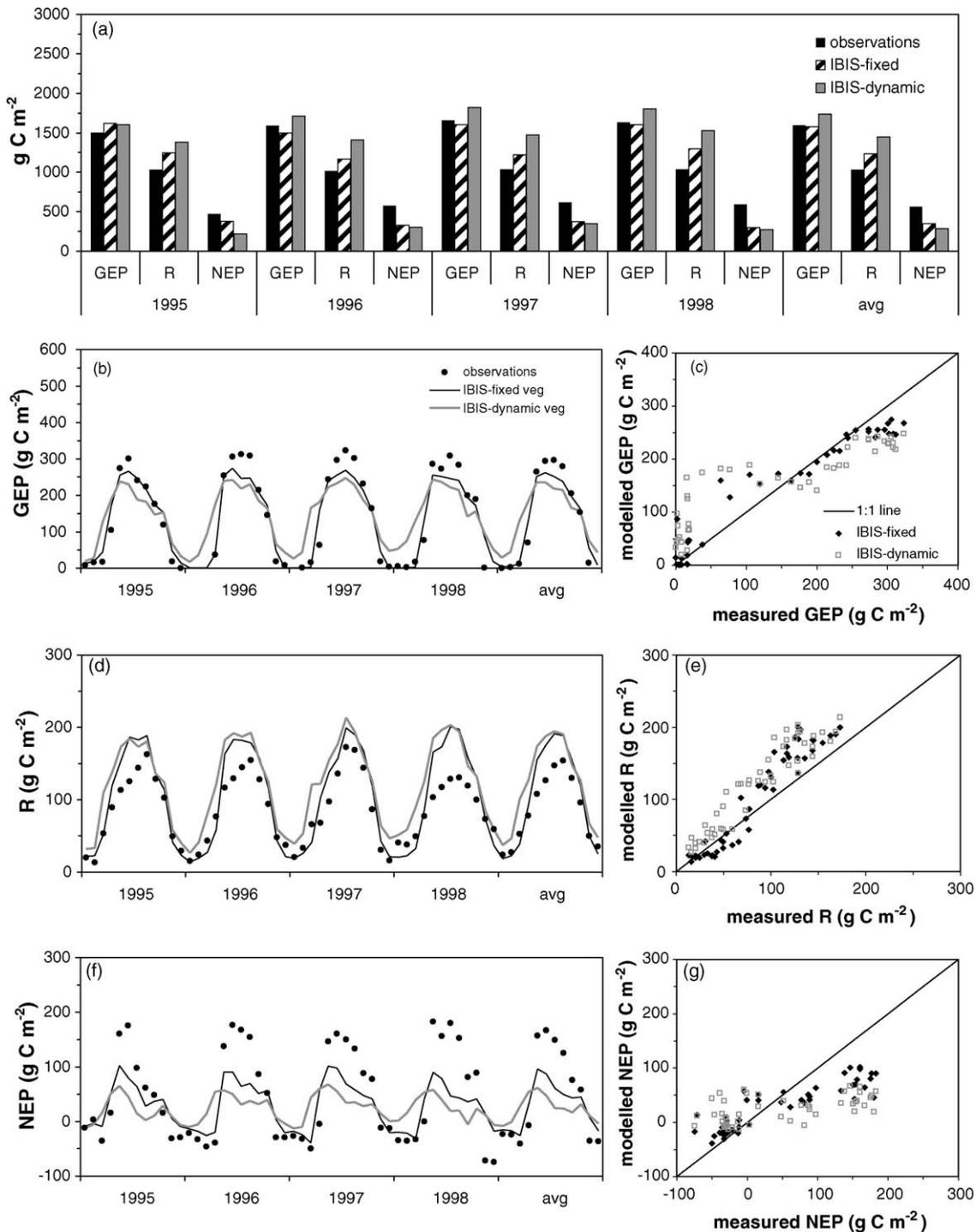


Fig. 4 – Walker Branch watershed carbon cycling. (a) Simulated and observed annual total GEP, R, and NEP; (b) monthly simulated and observed GEP; (c) correlation of simulated and observed monthly GEP; (d) monthly simulated and observed ecosystem respiration; (e) correlation of simulated and observed monthly ecosystem respiration; (f) monthly simulated and observed NEP; (g) correlation of simulated and observed monthly NEP.

yielded an r^2 of only 0.15. The range of annual NEP for the FV run (75 g C m^{-2}) was only about 50% of eddy-covariance measurements.

At a monthly time step, a significantly damped seasonal signal of NEP (Fig. 5), compared to measurements, is apparent in both FV and DV model simulations (Fig. 4f and g). When

compared to the eddy-covariance observations, the DV simulation represented only 30% of the month-to-month variation in NEP, and had a very minimal seasonal signal (slope 0.158; intercept 16.4). This is a result of large growing season R offsetting GEP, suggesting that formulations of R are too sensitive to temperature fluctuations at the site. From May to October,

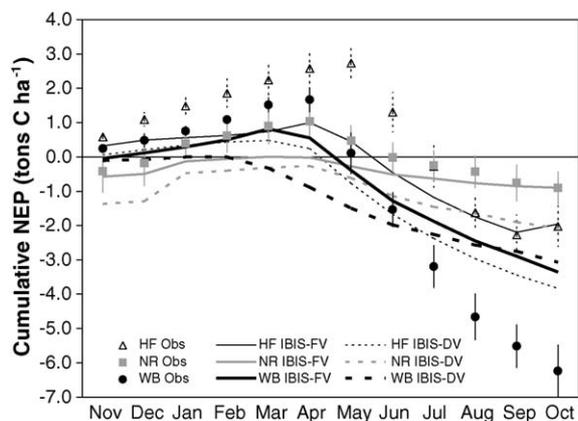


Fig. 5 – Cumulative average monthly NEP for IBIS fixed vegetation (FV) and dynamic vegetation (DV) simulations compared with observations at Harvard Forest (HF) from 1992 to 1998, Niwot Ridge (NR) from 1999 to 2001, and Walker Branch (WB) watershed from 1995 to 1998. Error bars represent ± 1 standard deviation.

the DV run underestimated observed monthly NEP values by 45–85% (Table A.2). Large overestimates in March and April, when the DV run simulated WB as a net C sink, compared to the observed net C source, are attributed to biases in leaf onset, warmer than observed soil temperatures, and generalized PFT physiology and C allocation parameters. The FV run captured the observed monthly variability significantly better than the DV simulation ($r^2 = 0.79$), but still showed a weak amplitude of the net carbon exchange seasonal cycle compared to observations (Fig. 4f and g). Average simulated monthly MBEs for the FV run were between -24 and -60% from May to October (Table A.2). The simulated positive soil temperature bias in spring and model parameterization errors may contribute to error in the simulated seasonal timing of net CO_2 uptake. For example, FV simulated April NEP is a 32.7 g C m^{-2} sink compared to a -7.0 g C m^{-2} source suggested by eddy-covariance measurements (Table A.2).

Fig. 5 depicts several potential model limitations. First, the magnitude of net ecosystem carbon loss during the winter season (November–March) is underestimated by IBIS. In fact, the DV simulation suggested that the deciduous forest stand at WB was not a cumulative net source of C to the atmosphere at any point during the winter and late spring. Second, the onset of cumulative positive carbon uptake by the forest was approximately 2 weeks early in the FV simulation (around May 1) and 3 months early in the DV run. Finally, the magnitude of the average seasonal cycle of ecosystem carbon release and uptake is quite minimal with respect to eddy-covariance measurements.

3.1.5.4. Component carbon flux differences between model and observations. We analyzed the ratios among simulated and observed component C fluxes at WB to further investigate model functioning and performance. We were particularly interested in the partitioning of autotrophic respiration (R_A) and R_H as parts of total soil CO_2 respiration, and their total contribution to R. We compared our simulated values with those

produced by an independent biophysical model (CANOAK) that was parameterized using data previously collected at WB (Wilson and Baldocchi, 2001), and also with other ecosystem model results (Hanson et al., 2004). The annual average IBIS simulated soil CO_2 efflux was 618 g C m^{-2} (S.D. 35.1) and 763 g C m^{-2} (S.D. 38.7) for FV and DV runs, respectively, compared to an average of 743 g C m^{-2} (S.D. 3.7) produced by the CANOAK model (Wilson and Baldocchi, 2001). Hanson et al. (2004) reported that the average soil CO_2 respiration value for ecosystem models in their study was 785 g C m^{-2} for 1993–2000.

Sanderman et al. (2003) and Hanson et al. (2000) reported that in forest ecosystems, the root contribution to total soil CO_2 efflux is roughly 0.48–0.50, while Landsberg and Gower (1997) reported an average of 0.45, with a large range from 0.33 to 0.62. Sanderman et al. (2003) also reported that the R_H and R_A contributions to total ecosystem respiration average 33 and 67%, respectively, and that belowground respiration is approximately two-thirds of ecosystem R. Our simulations produced root respiration that only comprised 27 and 29% of the total soil respiration in FV and DV runs, respectively. The R_H (R_A) contribution to ecosystem R was $\sim 37\%$ (63%) in both model simulations, and the ratio of soil respiration to total ecosystem R was 0.50 and 0.53 in the FV and DV simulations, respectively. Hanson et al. (2004) showed that several ecosystem models simulated a 30% average contribution of R_H to R at WB, and that soil respiration comprised 60% of R. The annual average NPP:GPP was 0.49 and 0.51 for FV and DV model runs, respectively. The ratio of NPP to GPP has been suggested to be a relatively constant fraction (~ 0.47 ; Waring et al., 1998; Gifford, 2003).

Based on these comparisons, we concluded that IBIS produced reasonable values of R_A at WB; the 63% contribution to total ecosystem R and NPP:GPP = 0.5 are in agreement with other reported values (Sanderman et al., 2003; Gifford, 2003; Hanson et al., 2004). There is some concern that the partitioning of soil respiration to microbial and root components was not adequately captured by IBIS. The temperature functions (Lloyd and Taylor, 1994) that modify root respiration are generalized in the model to be applicable across all biome types, and are dependent on properly simulated root growth dynamics (e.g., turnover) and C allocation to fine roots. IBIS assumes that all fine roots turnover once each year, with no differentiation between PFTs. The microbial respiration component is also likely to be in error if soil C and nitrogen dynamics are not satisfactorily captured. Unfortunately, it is extremely difficult to validate model simulations of three distinct soil C pools and their vertical distribution in the soil profile.

3.2. Harvard Forest: 1992–1999

3.2.1. Leaf phenology and canopy structure: LAI, leaf onset records, total biomass

We used a combination of LAI observations in 1998–1999 and a portion of the long-term phenology record at HF to evaluate further the varied phenology models (Fig. 1). The IBIS-Simple and IBIS-White approaches both overestimated the observed maximum green LAI by 70–80% (~ 6.3 – 6.7 versus 3.7). Leaf onset was 5 and 14 days later in the IBIS-White simulation in 1998 and 1999, respectively, and was 10 days later and 6

days earlier for the IBIS-Simple simulation in 1998 and 1999, respectively. More significant errors were noted when comparing the dates at which maximum leaf area was reached; in the IBIS-Simple simulations, maximum LAI was reached 36 days (1998) and 45 days (1999) earlier than observations, and 29 days (1998) and 24 days (1999) earlier using the White et al. (1997) algorithms. Leaf offset was triggered 8–10 weeks later using the simple global-scale algorithms, but was only 1–2 weeks later using the White et al. (1997) approach. The transition between the observed minimum and maximum LAI took approximately 7–8 weeks, which is considerably longer than the 2-week period imposed by the model (Fig. 1). It is evident that the rather simplistic phenology approaches, which fail to account for soil moisture, frost, and photoperiod at the end of the growing season are unable to adequately characterize LAI dynamics.

We used phenology data collected for oak and maple species at HF from 1992 to 1999 (available online at <http://harvardforest.fas.harvard.edu/data/p00/hf003/hf003.html>), specifically the reported day of 50% bud break threshold (i.e., 50% of tree buds have broken revealing leaves) to compare with the simulated average day that 50% of the maximum LAI was reached in each simulation. This very conservative comparison also indicated a significant bias of earlier than observed plant development, in which the IBIS-Simple simulation reached 50% maximum LAI on average 18 days earlier than the average 50% bud break threshold for oak and maple species at the site, and the IBIS-White simulation had a 7.2 days early bias.

Table 2 suggests that significant differences existed between observed and simulated canopy structure in terms of total plant biomass and canopy height, and also for total organic soil carbon, regardless of the phenology approach used. Both of the prognostic phenology schemes (Simple and White, respectively) underestimated the mature stand total biomass (MBEs = –13.5 and –50.3%), underestimated the canopy height (MBEs = –17.3 and –52.6%), and underestimated total soil C by 21% (Table 2). The simulated seasonal course of LAI using the White et al. (1997) approach was better suited to capture the observed plant phenology. However, the subtle differences in phenology and species dynamics contributed to the existence and development of the temperate evergreen coniferous PFT that was not dominant in the IBIS-Simple simulation. The prevalence of evergreen conifer trees in the IBIS-White simulation is reflected by a winter season LAI of 1.8 (Fig. 1), which was 30% of the total LAI. This is a factor of two greater than the measured green winter LAI of 0.7 at HF, where 19% of the total LAI was contributed by evergreen conifers.

3.2.2. Soil temperature

We compared the simulated daily average 0–10 cm soil temperature with the observed daily average 5 cm soil temperatures for the 1992–1999 period (Fig. A.1). The correlation between measured and computed values for the FV and DV runs yielded a coefficient of determination of 0.92 and 0.95, respectively. While these statistics and the average seasonal soil temperature changes shown in Fig. A.1 suggest that IBIS captured seasonal soil temperature fluctuations at HF very well, the model produced a warm surface soil temperature bias during

the growing season of about 3.5 °C, and a maximum cold bias of approximately –3.5 °C during the winter (Fig. A.1). The lack of explicit simulation of a litter layer in IBIS has been shown to influence seasonal temperature biases, particularly in summer due to changes in soil evaporation, and problems with the numerical representation of seasonal snow dynamics may partially explain wintertime cold biases (Vano et al., submitted for publication).

3.2.3. Soil surface CO₂ efflux

A 5-year record (1995–1999) of weekly to bi-weekly soil surface CO₂ flux measurements and corresponding 0–15 cm volumetric water content and 0–10 cm soil temperature observations (Savage and Davidson, 2001) was used to further evaluate the IBIS model at HF (available at <http://harvardforest.fas.harvard.edu/data/p00/hf006/hf006.html>). This record contains measurements from a total of six locations near the HF eddy flux tower, which vary in soil drainage characteristics and position with respect to the tower. We chose to compare our simulations with data collected in mixed hardwood vegetation southwest from the tower (SWF; 225 m from tower), with well-drained soil characteristics.

The correlation between observed and FV run volumetric water content values yielded an r^2 of 0.27. The correlation for the DV simulation produced an r^2 of 0.23. These correlations produced a similar model bias to the comparison made using the 5 cm soil temperature measured at the flux tower. Simulations of soil moisture showed that most values were concentrated in the 0.1–0.2 m³ m⁻³ range, whereas observations were more variable between 0.1 and 0.4 m³ m⁻³ (Fig. A.2). Figs. A.1 and A.2 show the seasonal changes in soil temperature and volumetric water content measured along the SWF transect. These results suggest the model had significant difficulty in capturing a highly variable (temporal) quantity such as soil moisture, and that several of the maximum simulated water content values greater than 0.35 occur during winter.

The correlation between observed and simulated daily soil respiration suggested that in both FV and DV simulations, there was a tendency to underestimate mid summer maximum observed respiration values by about 45% (Fig. 6), but the general seasonal cycle appeared to be captured. The correlation between measured and computed daily mean soil respiration (g C m⁻² day⁻¹) for the FV and DV runs yielded r^2 of 0.54 and 0.68, respectively. Savage and Davidson (2001) reported that the average wintertime (December–April) soil fluxes were 0.0147 g C m⁻² day⁻¹, which compared well with the FV average of 0.0136 g C m⁻² day⁻¹, but was 35% lower than the DV run value of 0.023 g C m⁻² day⁻¹.

The general underestimate of growing season soil respiration led to an annual average MBEs of –20.3 and –6.2% for FV and DV runs, respectively. The annual average observed soil respiration was 748 g C m⁻² year⁻¹, and was 582 and 684 g C m⁻² year⁻¹ for FV and DV runs, respectively. The most glaring weakness that we detected was the inability of the model to capture the observed inter-annual variability of total annual soil respiration. First, the model failed to satisfactorily produce the magnitude of the inter-annual variability; the CV was only 5.7 and 4.5% for FV and DV simulations (range of ~50 g C m⁻² year⁻¹ between high and low values), respectively, while the observed CV from 1995 to 1999 was 22.4% (a

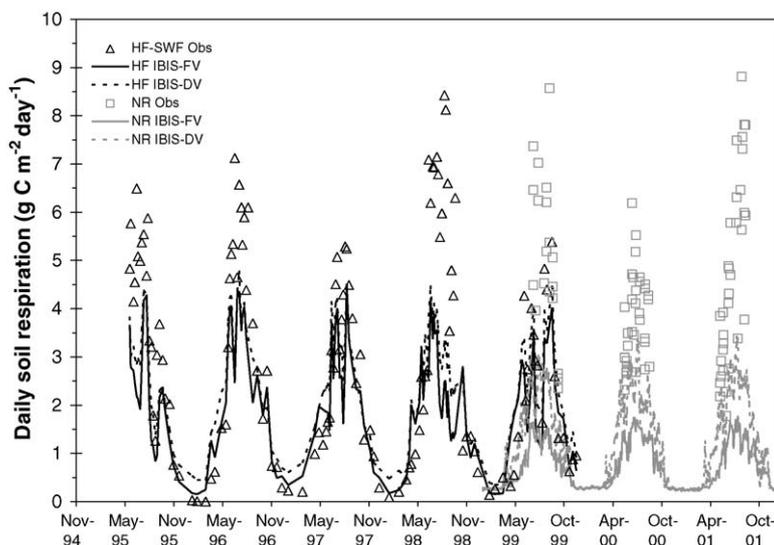


Fig. 6 – Seasonal cycle of IBIS simulated daily soil respiration for fixed vegetation (FV) and dynamic vegetation (DV) scenarios, plotted with periodic soil respiration measurements at Harvard Forest (HF) along the SWF transect for 1995–1998 (Savage and Davidson, 2001) and Niwot Ridge (NR) for 1999–2001 (Scott-Denton et al., 2003).

range of $335 \text{ g C m}^{-2} \text{ year}^{-1}$). The FV simulation only captured 27% of the variability in observations, whereas the DV simulation captured 51% of the annual variation. The simplified temperature and water-filled pore space functions that modify decomposition in IBIS would be possible reasons for simulated error, particularly in their interaction and representation of actual processes (Davidson et al., 1998).

3.2.4. Carbon fluxes: GEP, R, NEP, and soil carbon

3.2.4.1. Gross ecosystem production. We compared observations of annual GEP with the HF FV and DV simulations (Fig. 7a and Table A.3). The “observed” GEP is again determined from a combination of NEP values and ecosystem respiration estimates. The annual average observed GEP was 1286 g C m^{-2} (Barford et al., 2001), and FV and DV simulations yielded annual average values of 1265 (MBE = -1.7%) and 1701 g C m^{-2} (MBE = 32.2%), respectively. The significant positive bias in the DV run is most likely attributed to higher than observed LAI. The FV simulated annual GEP value was only slightly higher than the secondary estimates of 1122 and $1180 \text{ g C m}^{-2} \text{ year}^{-1}$ reported by Falge et al. (2002) and Sanderman et al. (2003), respectively, for HF. Neither of the model simulations was able to replicate the observed inter-annual variability of GEP ($r^2 < 0.01$), and both exhibited a significantly lower range in annual GEP values during the 1992–1999 period of study. The range between the highest and lowest annual GEP was $2.5 \text{ Mg C ha}^{-1} \text{ year}^{-1}$ for observations compared to $1.9 \text{ Mg C ha}^{-1} \text{ year}^{-1}$ for the FV run and $1.4 \text{ Mg C ha}^{-1} \text{ year}^{-1}$ for the DV run (Barford et al., 2001).

Fig. 7b and c clearly suggest that the model produced consistent seasonal biases for FV and DV simulations. The FV run reproduced monthly variability ($r^2 = 0.91$) with slightly greater success than the DV simulation ($r^2 = 0.78$). The average simulated MBE (Table A.3) was between -2 and -13% from June to September in the FV run and comparable (MBE = -6 to 14%) in the DV run. The most significant error was present from

late fall through late spring in both model simulations, when monthly measured GEP was less than $10 \text{ g C m}^{-2} \text{ month}^{-1}$. As evident in model runs for the WB site, the extremely large errors in the DV run from October to May are most likely due to C assimilated by the evergreen conifer PFT when simulated soil temperatures were above 0.0°C .

3.2.4.2. Ecosystem respiration. The annual average observed R was 1087 g C m^{-2} , while FV and DV simulations yielded average annual values of 1080 g C m^{-2} (-0.6% error) and 1324 g C m^{-2} (21.8% error), respectively (Fig. 7a). The two model runs had difficulty capturing the observed inter-annual variability of R. The correlation between annual FV run values and observations produced an r^2 of 0.16, while the DV simulation yielded an r^2 of only 0.04. The range between the highest and lowest observed yearly R was $2.1 \text{ Mg C ha}^{-1} \text{ year}^{-1}$ compared to $1.7 \text{ Mg C ha}^{-1} \text{ year}^{-1}$ for the FV run and $1.3 \text{ Mg C ha}^{-1} \text{ year}^{-1}$ for the DV run. These results suggested that when IBIS was parameterized with site-specific vegetation information (in the FV run), the model simulated the annual average amount of respiration that was observed at HF with great success, although offsetting seasonal biases led to the apparent annual agreement. Year-to-year variations were difficult to replicate. The larger model error in simulating R in the DV run was confounded by a significantly prolonged growing season with higher than observed LAI, an overestimate of the peak LAI by about $2.5 \text{ m}^2 \text{ m}^{-2}$, and the simulated prevalence of the temperate evergreen conifer PFT.

Fig. 7d and e shows the comparison between monthly observations of R and the two model simulations for the 1992–1999 time period. The correlation between the model output and observations suggested that the FV ($r^2 = 0.85$) and DV ($r^2 = 0.86$) simulations captured the observed seasonal variations, but model biases existed in the winter and summer seasons. Total monthly R was overestimated (Table A.3) by 10 – 36% from May to September in the FV run and by 33 – 49%

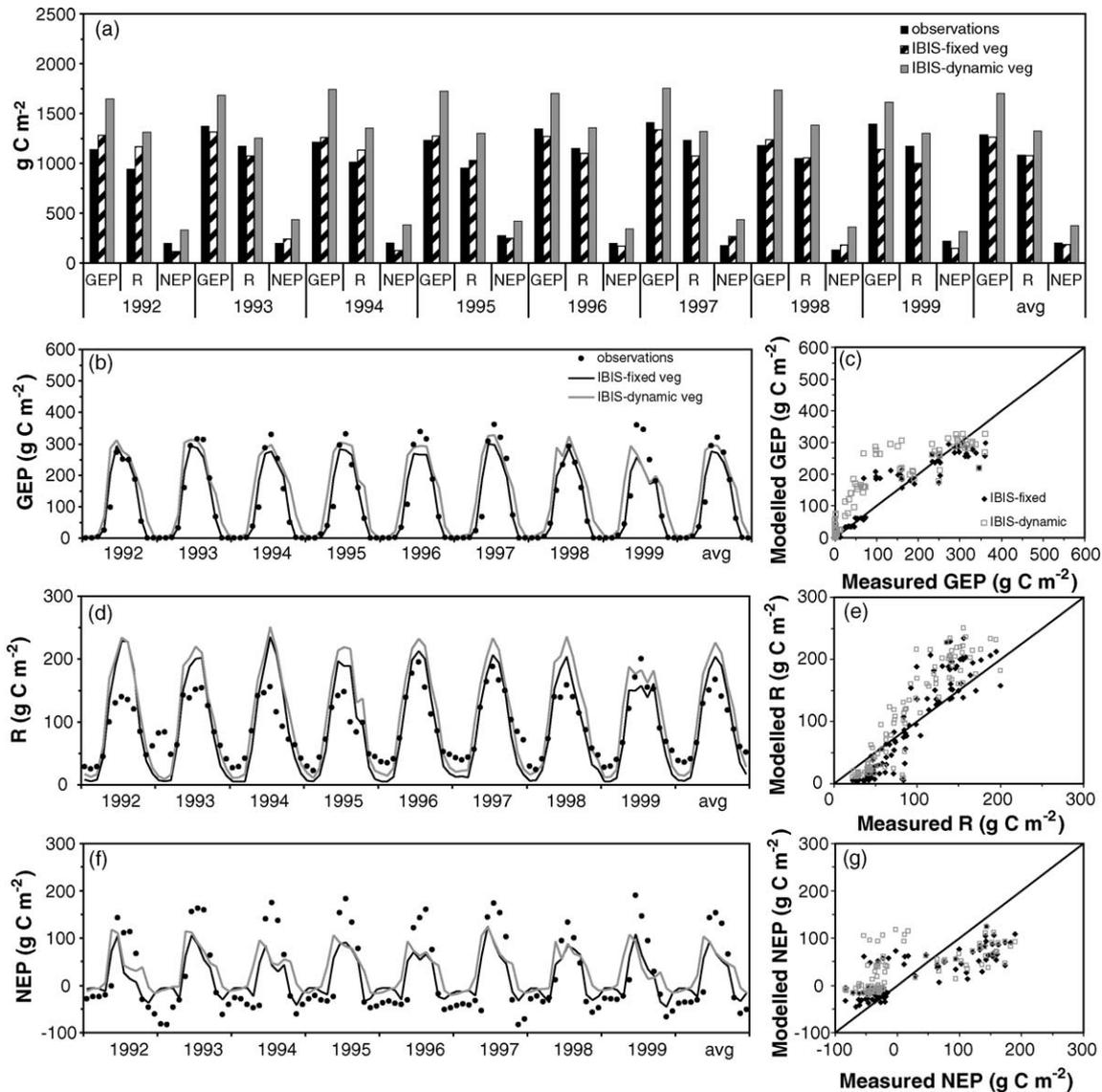


Fig. 7 – Harvard Forest carbon cycling components. (a) Simulated and observed annual total GEP, R, and NEP; (b) monthly simulated and observed GEP; (c) correlation of simulated and observed monthly GEP; (d) monthly simulated and observed ecosystem respiration; (e) correlation of simulated and observed monthly ecosystem respiration; (f) monthly simulated and observed NEP; (g) correlation of simulated and observed monthly NEP.

during the same period in the DV run. This is due to the fact that model formulations for respiration do not account for soil moisture effects and seasonal ecosystem changes (e.g., allocation), which are at least as important as temperature in driving respiration. Both model runs produced significant underestimates (i.e., MBEs = -40 to -80%) of monthly average ecosystem respiration between November and March, which may be attributed to simulation of colder than observed (-3 to 4°C) soil temperatures and because the parameterized respiration-temperature dependence is too sensitive for HF.

3.2.4.3. Net ecosystem production. The simulated and observed annual total and seasonal patterns of NEP at HF are compared in Fig. 7a, f, and g, and Table A.3. The observed annual average NEP was 199 g C m^{-2} (Barford et al., 2001) compared with totals of 184.4 g C m^{-2} (MBE = -7.6%) and

376.7 g C m^{-2} (MBE = 89%) for the IBIS FV and DV simulations, respectively. Curtis et al. (2002) reported that the annual net carbon storage was 165 g C m^{-2} using available component NPP measurements combined with R_H estimates to calculate NEP, and a value of 175 g C m^{-2} was derived from carbon stock changes. Thus, we have good confidence that the results for the FV run are acceptable given the range of reported NEP values for HF.

While observations suggested that a significant amount of year-to-year variation in NEP occurred at HF from 1992 to 1999 (e.g., range between minimum and maximum values was 146.8 g C m^{-2}), the two model runs were only satisfactory in capturing the magnitude of inter-annual variations over the 8-year period and not the specific year-to-year variation. The correlation through the annual observed NEP and the FV simulation yielded an r^2 of only 0.02. The range of annual NEP for

the FV run from 1992 to 1999 (S.D. 58.6, range of 152.6 g C m^{-2}) was comparable to eddy-covariance measurements. The DV run only captured 3% (r^2) of the observed inter-annual variability, and exhibited a comparable absolute range of annual NEP (S.D. 48.0, range of 121.3 g C m^{-2}). However, the average annual NEP value was significantly overestimated due to the problems discussed in Sections 3.2.4.1 and 3.2.4.2.

A significantly dampened seasonal (monthly) signal of NEP, compared to measurements, is apparent in both IBIS model runs at HF (Fig. 7f and g). The DV simulation captured 44% of the month-to-month variation in NEP, exhibiting a weaker than observed seasonal signal, which resulted from the exaggerated simulated seasonal cycle of ecosystem respiration (Fig. 7d). While the FV simulation fared slightly better ($r^2 = 0.70$), the carbon uptake was still underestimated during the growing season. The observed maximum average monthly NEP occurred in July (153.3 g C m^{-2}), while the average minimum NEP occurred in November (-59.8 g C m^{-2}). In contrast, the peak monthly NEP in the FV run occurred 1 month earlier in June (92.6 g C m^{-2}), but did capture the timing of the observed minimum monthly NEP in November (-34.3 g C m^{-2}).

From June to September, the DV run underestimated observed monthly NEP values by 30–58% (Table A.3). Large overestimates of NEP in the DV run from March to May are attributed to the model bias of the faster than observed vegetation green-up, warmer than observed soil temperatures, and generalized PFT physiology and carbon allocation. Average simulated monthly MBEs for the FV run were between –35 and –61% from June to September (Table A.3). The likelihood of the forest being a carbon sink in May, suggested by the FV simulated May NEP (57.1 g C m^{-2}) is not supported by eddy-covariance measurements (-14.4 g C m^{-2} ; Table A.3). This bias cannot be attributed to errors in plant phenology. Conversely, the errors can be attributed to the failure to capture seasonal changes in growth respiration and the availability of leaf litter substrate from the previous fall feeding soil respiration.

It is evident that while the FV simulation produces an average NEP value that is only 7.6% less than the observations, offsetting seasonal biases contribute to the apparently good result (Fig. 5). The time series plotted in Fig. 5 illustrates three key model shortcomings: (1) ecosystem respiration during the dormant season is underestimated by IBIS, (2) the onset of positive carbon uptake by the forest is 1 month early in the FV simulation and 2 months early in the DV run, and (3) the magnitude of the seasonal cycle of ecosystem carbon release and uptake is about 32% less in the FV simulation than eddy-covariance observations.

3.2.4.4. Component carbon flux differences between model and observations. Our simulations at HF produced root respiration that was 15 and 27% of the total soil respiration in FV and DV runs, respectively. The R_H (R_A) contribution to ecosystem R was 47% (53%) in the FV simulation and 38% (62%) in the DV simulation. The ratio of belowground respiration to total ecosystem R was 0.56 and 0.51 in the FV and DV simulations, respectively. The annual average ratio of NPP to GPP was 0.55 and 0.52 for FV and DV model runs, respectively. Sanderman et al. (2003) reported a NPP:GPP value of 0.45 for HF.

3.2.5. Water budget terms: evapotranspiration, and soil moisture and runoff feedbacks

The annual average ET observed at HF was 446 mm, whereas simulated values were 537 and 608 mm for FV and DV simulations (Fig. 8a), respectively, which constitutes 20 and 36% model overestimates. Both of the FV and DV simulations suggested that the model poorly represented the observed year-to-year fluctuations in total ET, as the r^2 s were 0.18 and 0.36, respectively.

When the simulated monthly totals ($n = 96$) of ET were correlated with observations, we found that the FV simulation captured 90% of the monthly variability (Fig. 8b and c), while the DV run yielded an r^2 of 0.89. The general bias of overestimated annual ET was equivalent to monthly MABs of 10.4 and 14.3 mm month⁻¹ for the FV and DV simulations, respectively. We compared the observed average ET for each month ($n = 12$) with simulations for the 8-year measurement period. The FV run captured, on average, 96% of the observed seasonality, and the DV run captured 95% of the average month-to-month variability. These results suggest excellent agreement between the predicted and measured ET in terms of the magnitude of typical seasonal changes (Table A.3).

We found that in the FV simulation, monthly average ET during the growing season was overestimated by 10–60% (MBE = 10–20 mm month⁻¹), with the maximum error occurring in April and May (MBE = 58 and 43%, respectively; Table A.3). The DV simulation overestimated ET in every month, with the greatest deviation from observations during March–May (MBE = 50–80%; Table A.3), and a secondary error peak in October and November. The timing and magnitude of monthly biases in simulated ET is similar to the simulated biases in soil temperature for the FV simulation; namely, ET was overestimated when soil temperatures were simulated higher than observed, and slightly underestimated during winter when non-existent frozen soils were simulated. The simulated overestimate of growing season ET (April–August) was likely due to a combination of higher than observed LAI and surface soil temperatures. We note that as the difference between the observed and simulated LAI decreased from April to August, the error decreased accordingly from 85 to 12% (Table A.3).

Model response for HF showed significant changes in water balance between the FV and DV run, triggered primarily by changes in plant phenology, LAI, and species existence. The higher LAI in the DV run caused the fraction of total ET that was re-evaporated as intercepted water to increase from 24% (FV run) to 39%. Transpiration comprised 67% of the total DV annual ET (compared with 51% in FV run), and soil evaporation was only 8% of annual ET in the DV run and 25% in the FV simulation. The most important outcome was that the model tended to overestimate total ET in both FV and DV simulations.

Measured values exhibited an average energy imbalance of $202 \text{ MJ m}^{-2} \text{ year}^{-1}$ (Table 3), which is equivalent to 81 mm of evaporated water each year. Based on observations, about 47% of the net radiation measured at HF was partitioned into latent heat energy. Thus, by assuming that 47% of the energy imbalance can be attributed to latent heat (38.1 mm) to close the measured energy budget, the difference between annual observed and simulated ET would be adjusted to 52.9 mm in the FV run (10.9% error) and 124 mm in the DV simulation

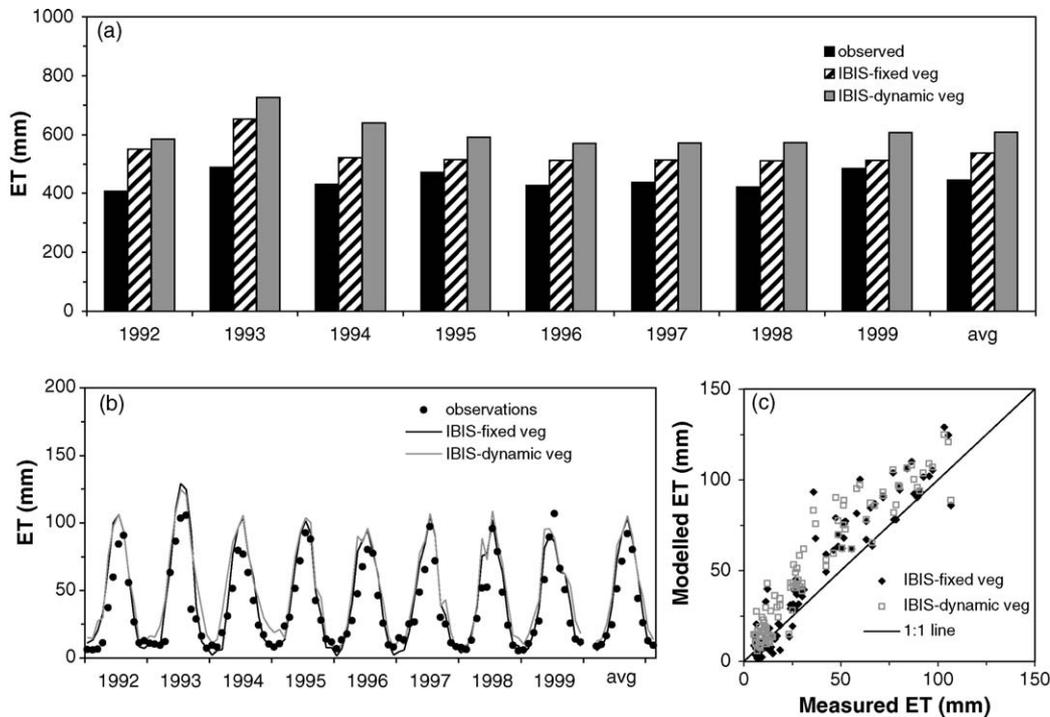


Fig. 8 – Simulated and observed (a) annual total ET; (b) monthly ET; (c) the correlation of simulated and observed monthly ET at Harvard Forest.

(25.6% error). Another possible source of the simulated bias at HF could be our approximation of the seasonal variation of LAI between 1992 and 1997, used in the FV simulation.

3.3. Niwot Ridge: 1999–2001

3.3.1. Canopy structure: LAI and total biomass

For the NR forest site, the DV simulation using the simplified global phenology scheme generated a mixed-forest stand for the upper canopy (needleleaf evergreen conifers and broadleaf deciduous trees), and developed a substantial lower canopy consisting of a deciduous shrub PFT. The NR site is currently dominated by temperate needleleaf evergreen trees (e.g., sub-alpine fir, Engelmann spruce and lodge pole pine; Monson et al., 2002). The DV simulation that implemented the White et al. (1997) phenology scheme led to the satisfactory development of a coniferous forest stand dominated by boreal evergreen conifer trees, but a lower canopy of shrubs also existed. Table 2 shows that the prognostic phenology schemes (Simple and White, respectively) underestimated the mature stand total biomass by 5.4 and 43.3%, underestimated the upper canopy height by 6.1 and 41.2%, and overestimated the observed total soil carbon by 83 and 147%, respectively (Table 3).

While the IBIS-White simulation led to the development of an upper forest canopy that closely resembles the current evergreen species at NR, the LAI was 35% lower than published observations (~2.7 versus 4.2), and the simulated lower canopy vegetation had an LAI of approximately $3.0 \text{ m}^2 \text{ m}^{-2}$. The lower canopy at NR is only described as being sparse (Monson et al., 2002). The simplified global phenology scheme coupled with the IBIS vegetation dynamics sub-model produced a signif-

icant amount of broadleaf deciduous tree cover as LAI and therefore high seasonal variability in LAI, which does not occur at the site. The upper canopy LAI in this dynamic run varied from 1.5 to 2.8 during the growing season (Fig. 1). The LAI of evergreen conifer PFTs was approximately 1.5, and the LAI of broadleaf deciduous trees varied from 0.2 to 1.5.

3.3.2. Soil temperature

The FV and DV (IBIS-Simple) simulated daily average 0–10 cm soil layer temperature was compared with daily average 5 cm soil temperatures (Fig. A.1). The correlation between measured and computed values for the FV run yielded a coefficient of determination of 0.89. The correlation performed for the DV run produced an r^2 of 0.87. There appears to be a discontinuity in simulated soil temperatures around 0.0°C , potentially due to phase changes between soil water and ice. As was noted in both of the WB and HF simulations, the overall seasonality of soil temperature changes are well simulated by IBIS, but the model again demonstrated a warmer than observed soil surface layer temperature bias during the growing season, on average, of between 2 and 5°C , and a cold bias of approximately -1 to -3°C during the winter in both FV and DV model runs (Fig. A.1).

3.3.3. Soil surface CO_2 efflux

A 3-year record (1999–2001) of soil surface CO_2 flux measurements and corresponding 8 cm depth volumetric water content and 5 cm depth soil temperature observations (Scott-Denton et al., 2003) were used to examine the simulation of total soil respiration. A maximum of 5 transects consisting of 12 individual collars each were measured weekly to bi-weekly

between July and September 1999, June and October 2000, and May and August 2001. We compared the instantaneous transect averages of soil respiration with IBIS.

Significant differences between chamber measurements and simulated soil respiration rates were noted during the majority of each growing season from 1999 to 2001 (Fig. 6). Because nighttime NEE values reported by Monson et al. (2002) and soil chamber measurements (Scott-Denton et al., 2003) were in approximate agreement (e.g., average rates were between 3 and 6 $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$ during summer in each year), we are fairly confident that IBIS had a significant bias of underestimating the peak rates. In contrast to the two independent estimates of soil respiration, the peak simulated soil respiration rates were only 1.4 $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$ in the FV run and 3.5 $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$ in the DV simulation. Overall, the FV and DV simulations underestimated growing season respiration values, on average, by about 65 and 40%, respectively, and underestimated wintertime respiration by 69%. Observed nighttime NEE (i.e., a surrogate for soil respiration) averaged 0.64 and 0.76 $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$ for January and March, respectively, from 1999 to 2001, while the model simulated an average of 0.2 $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$. The correlation between measured and simulated soil respiration for the FV run yielded an r^2 of 0.08, and the DV simulation produced an r^2 of 0.28. The annual average soil respiration in the FV and DV simulations were 272.4 and 445.4 $\text{g C m}^{-2} \text{ year}^{-1}$, respectively, and the coefficient of variation was only 2.3 and 1.8%, respectively. The studies of Monson et al. (2002) and Scott-Denton et al. (2003) suggested that the annual average ecosystem respiration was 169–444 g C m^{-2} , with a significant contribution from soil.

Fig. A.2 shows the relationship between soil moisture observations (instantaneous 8 cm) of Scott-Denton et al. (2003) and simulations (0–10 cm daily average) for volumetric water content. The correlation between measured and computed volumetric water content values for the FV run yielded an r^2 of 0.09 and 0.17 for the DV simulation. Fig. A.2 shows the observed seasonal changes in soil volumetric water content, and suggest that the DV run fared better during 2000 and 2001 than the FV run, but measurements were consistently higher than the DV model run in 1999.

We attributed the higher simulated soil respiration in the DV run (compared to FV run) to the existence of more organic matter and litterfall available for rapid decomposition. The simulated soil carbon density of 91.4 Mg C ha^{-1} was 82% higher than observations (which were used to parameterize the IBIS soil biogeochemistry sub-model in the FV run), and is attributed to a higher GEP and more annual litter production associated with the existence of lower and upper canopy deciduous trees and shrubs. While the leaf carbon allocation constants are similar for all tree PFT's in IBIS, the leaf turnover time constant for the boreal evergreen conifer PFT in IBIS is 2.5 years. The root respiration was lower in the DV simulation because the prescribed allocation of C to roots is 20% in boreal and temperate deciduous PFT trees, compared to 40% for boreal and temperate evergreen conifer PFTs. Overall, soil respiration contributed to 60 and 55% of the total ecosystem respiration in the FV and DV simulations, respectively, and heterotrophic respiration comprised 48 and 70% of soil respiration for FV and DV simulations, respectively. We ruled out as causal factors the differences in soil temperature and mois-

ture that occurred between the two model runs because they were not significantly different.

3.3.4. Carbon budget: GEP, R, and NEP

3.3.4.1. *Gross ecosystem production.* We compared observations of annual GEP with the NR FV and DV IBIS simulations (Fig. 9a). The annual average observed GEP was 446 g C m^{-2} , while FV and DV simulations yielded annual average values of 519 g C m^{-2} (MBE = 16.3%) and 964 g C m^{-2} (MBE = 116.1%), respectively. The larger error associated with the DV run was due to: (1) simulated LAI that was higher than the observed value, due to the simulation of an upper and lower canopy and (2) the existence of deciduous tree PFT's that had higher photosynthetic capacity than the boreal, sub-alpine evergreen conifers. Neither of the model simulations captured the range of observed inter-annual variability of GEP. The CV of observations over the 3 years of observations was 11.4% compared with 2.7 and 5.6% for the FV and DV simulations, respectively. Each model run poorly replicated the observed increasing trend in GEP in consecutive years. In particular, the FV simulation GEP values decreased from the highest simulated value of 535 $\text{g C m}^{-2} \text{ year}^{-1}$ in 1999 to 508 $\text{g C m}^{-2} \text{ year}^{-1}$ in 2001, whereas observations increased from 387 to 477 $\text{g C m}^{-2} \text{ year}^{-1}$.

Fig. 9b and c shows how the model produced a consistent bias for the DV simulation, due to the aforementioned problems with development of a significant presence of deciduous tree PFTs. The FV run correlated with monthly variability exceedingly well ($r^2 = 0.92$), although there was still simulated GEP during the observed dormant season (November–April) when GEP of 0.0 (Table A.4) was a function of R and NEP. The simulated winter monthly average GEP > 0.0 could be an artifact of the model parameterization of leaf physiological response to temperature and water stress, which allowed leaf photosynthesis to occur when temperatures were less than 0.0°C. However, there is still large uncertainty as to how much carbon evergreen conifer forests can sequester during the non-growing season, particularly because most leaf-level gas-exchange measurements are made during the typical growing season (Gough et al., 2004). While the DV simulation successfully represented the seasonal patterns of observed GEP changes ($r^2 = 0.93$), the monthly magnitudes were in poor agreement with the actual observations. The average simulated monthly MBE (Table A.4) was between –8 and 21% from May to October in the FV run and exceedingly large (82–126%) in the DV run. The slight difference in monthly average simulated GEP between the FV to DV runs during the dormant season is due to a large fraction of vegetation that was simulated as deciduous forest in the DV run, causing the “green” LAI capable of photosynthesizing in the DV to be less than that in FV simulations during the winter season.

3.3.4.2. *Ecosystem respiration.* The annual average observed R was 369 g C m^{-2} , while the FV and DV simulations yielded values of 453 g C m^{-2} (MBE = 22.6%) and 810 g C m^{-2} (MBE = 119.4%), respectively (Fig. 9a). Clearly, the existence of the lower canopy vegetation, while contributing to an elevated GEP in the DV run, also led to an exaggerated simulated value of R. The two model runs poorly represented the observed inter-annual variability of R. The correlation between annual

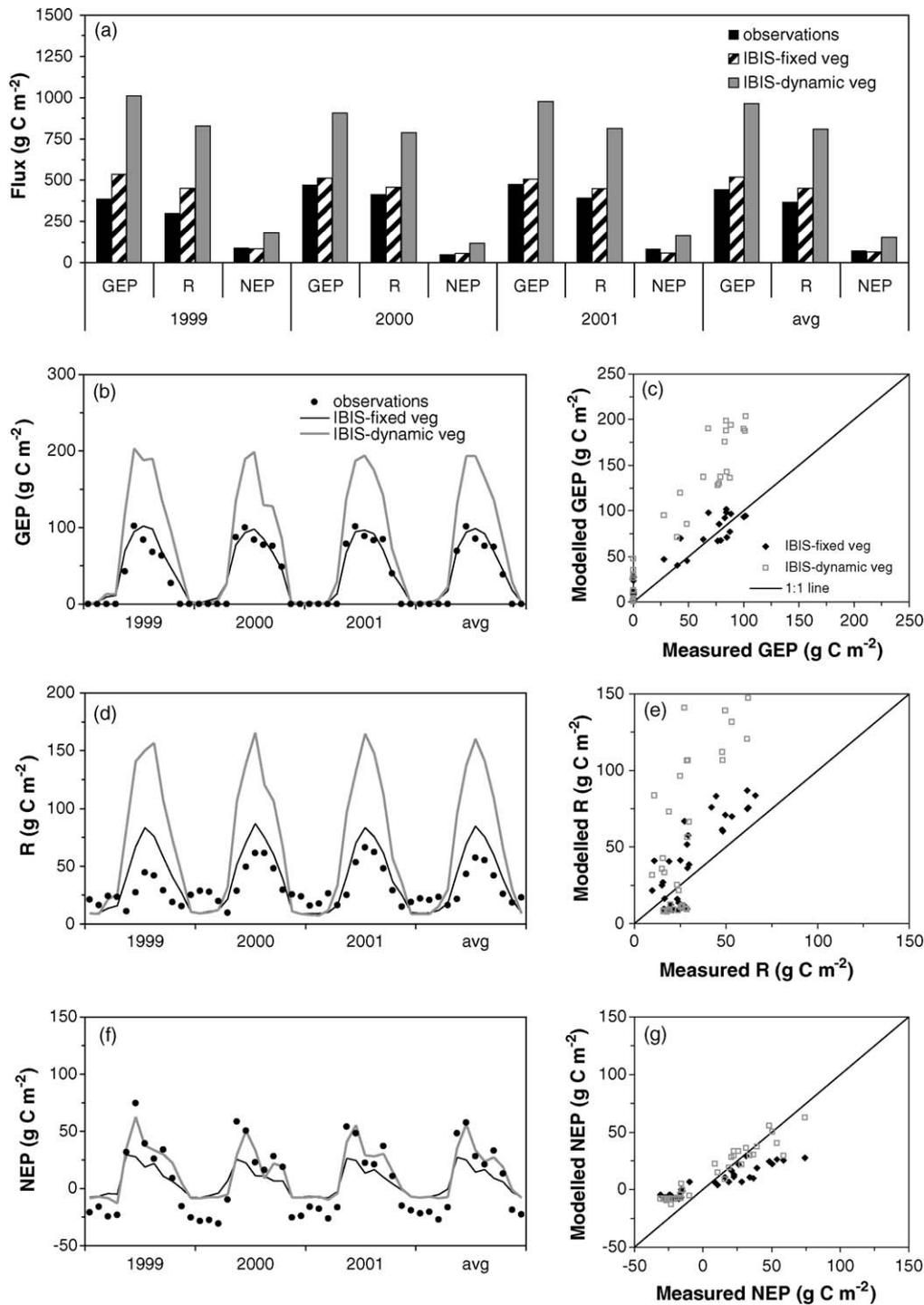


Fig. 9 – Niwot Ridge carbon cycling components from 1999 to 2001. (a) Simulated and observed annual total GEP, R, and NEP; (b) monthly simulated and observed GEP; (c) correlation of simulated and observed monthly GEP; (d) monthly simulated and observed ecosystem respiration; (e) correlation of simulated and observed monthly ecosystem respiration; (f) monthly simulated and observed NEP; (g) correlation of simulated and observed monthly NEP.

FV values and observations produced an r^2 of 0.17, while the DV simulation yielded an r^2 of 0.77, but had an inverse relationship with observations. The range between the highest and lowest observed annual R was $116 \text{ g C m}^{-2} \text{ year}^{-1}$ compared to only $10 \text{ g C m}^{-2} \text{ year}^{-1}$ for the FV run and $42 \text{ g C m}^{-2} \text{ year}^{-1}$ for the DV run. These results suggested that when the model was

parameterized with site-specific vegetation information in the FV run, IBIS was able to simulate the average observed respiration at NR, although the year-to-year variations were more difficult to replicate.

Fig. 9d and e shows the seasonal changes and correlation between monthly calculations of R at NR and the

two model simulations. The general pattern in the FV run was of an amplified seasonal cycle; growing-season peak observed monthly R values were overestimated by the model by 40–60% (from May to October), and non-growing season R (December–March) was underestimated by 50–70% (Fig. 9d and Table A.4), due to low soil respiration values compared to observed nighttime NEE. The correlation between FV modeled values and observations produced an $r^2 = 0.67$ (Fig. 9e). This behavior is consistent with the simulations of the different forest ecosystems at WB and HF, suggesting that a significant functional problem exists with simulating the emergent property of ecosystem R.

The seasonality and amplitude of R in the DV run was greatly exaggerated because of the improper simulation of vegetation structure and function at the site. While the DV did reproduce seasonal changes in R with some success ($r^2 = 0.61$; Fig. 9d), there was a factor of two to three difference in peak summertime respiration values. The average simulated MBE in the DV run was between 150 and 342% from May to October. Interestingly, the level of simulated error was of similar magnitude for both model simulations during winter months. The consistent underestimate of R in both the RV and DV model runs during winter suggests that the soil respiration was too low; this points to a potential problem with temperature and soil moisture-respiration parameterizations for soil decomposition, and that the temperature function that controls root respiration was too sensitive to temperature.

3.3.4.3. Net ecosystem production. Observed and simulated monthly and annual NEP at NR are shown in Fig. 9a, f, and g, and Table A.4. The observed annual average NEP was 73.2 g C m^{-2} , compared with totals of 66.4 g C m^{-2} (MBE = -9.2%) in the FV simulation and 153.9 g C m^{-2} (MBE = 110%) for the DV simulation. The CV of annual NEP observed at NR from 1999 to 2001 was 30%, compared with 23% for the FV run and 21% for the DV run. The correlation of simulated and observed annual NEP values suggested that 51 and 98% of the inter-annual variability was represented by the FV and DV simulations, respectively, and the range of annual NEP for the FV run (28.5 g C m^{-2}) was very comparable to the range captured with eddy-covariance measurements (40.4 g C m^{-2}). However, offsetting seasonal biases of simulated R led to the close agreement between observations and the FV simulated values. For this coniferous forest, the simulated NPP:GPP in the FV simulation (0.38) was lower than the average value (0.47 ± 0.04) reported by Waring et al. (1998) and Gifford (2003), but was still within the observed range (0.3–0.7) reported by Landsberg and Gower (1997) for several coniferous forests. The simulated average NPP:GPP for the DV simulation was higher (0.48).

It is evident that while the FV simulation produces an average NEP value that was only underestimated by 9.2%, the offsetting seasonal biases contributed to the apparent excellent agreement. The observed versus simulated values shown in Fig. 5 suggested that ecosystem carbon loss via plant and soil respiration from January to April is underestimated by IBIS, and the net carbon uptake is underestimated during the growing season, particularly in the FV simulation. The magnitude of the seasonal cycle of ecosystem carbon release and uptake is about $0.9 \text{ t C ha}^{-1} \text{ year}^{-1}$ in the FV simulation,

$1.8 \text{ t C ha}^{-1} \text{ year}^{-1}$ in the DV run, and $1.93 \text{ t C ha}^{-1} \text{ year}^{-1}$ from eddy-covariance observations. The timing of onset of net positive carbon uptake by the model agreed with observations, showing that the ecosystem changed from a net source to sink of carbon from April to May. The model runs also captured the months of January and May, respectively, as being the periods when the most significant month-to-month changes (e.g., ecosystem response to environmental conditions) in carbon exchange took place (Fig. 5). The DV simulation captured a subtle observed feature of increase in the rate of net carbon uptake by the forest in September, with the average observed and simulated rates being less during the months of July, August, and October. We speculate that this secondary peak late in the growing season is due to decreased soil respiration rates with the change in soil temperature and moisture conditions, while vegetation is still capable of sequestering significant amounts of carbon (Monson et al., 2002).

The significantly dampened seasonal (monthly) signal of NEP, compared to measurements, is apparent in the FV simulation (Fig. 9f and g). The FV simulation captured 88% of the month-to-month variation in NEP, but exhibited a weaker than observed seasonal signal. While the DV simulation performed slightly better ($r^2 = 0.92$) it still tended to underestimate the carbon release during the cold weather months. The observed average monthly NEP was between -20 and -27 g C m^{-2} from November to April (Table A.4), while the average simulated monthly minimum NEP values during the winter were typically less than 10 g C m^{-2} . This resulted in an underestimate of net carbon release of about 60–80% during this time (Table A.4).

3.3.5. Evapotranspiration

The annual average observed ET at NR was 643.6 mm, and the FV and DV simulated values were 513.6 mm (MBE = -20.2%) and 600.9 mm (MBE = -6.6%) (Fig. 10a). Linear correlation between observed and simulated values showed that IBIS captured about 47% of the inter-annual variability in the DV run, and 58% in the FV run. When the simulated monthly totals ($n = 36$) of ET were compared with observations using correlation, we found that the FV simulation captured 84% of the monthly variability and seasonality (Fig. 10b and c), while the DV run yielded an r^2 of 0.86. These results suggest high agreement between the predicted and measured ET in terms of the magnitude of typical seasonal changes (Table A.4).

In the FV simulation, the monthly average ET during the growing season was underestimated by 18–24% (MBE = 10–20 mm month⁻¹), with the maximum error occurring during December and January (MBE = -43 and -39%, or $\sim 15 \text{ mm month}^{-1}$, respectively; Table A.4). The DV simulation also slightly underestimated ET during the growing season by less than 10%, with the greatest deviation from observations also occurring during the winter months of December and January (-30%; Table A.4). Similar to the correlations found at HF and WB, the timing and magnitude of monthly biases in simulated ET appear to be weakly related to the simulated cold biases in soil temperature for both model simulations in December and January, but the warmer than observed simulated soil temperatures during the growing season did not lead to a positive bias in ET. Interestingly, the energy balance measurements made at this site show that net radiation was less than the sum of energy budget components (Table 3), suggest-

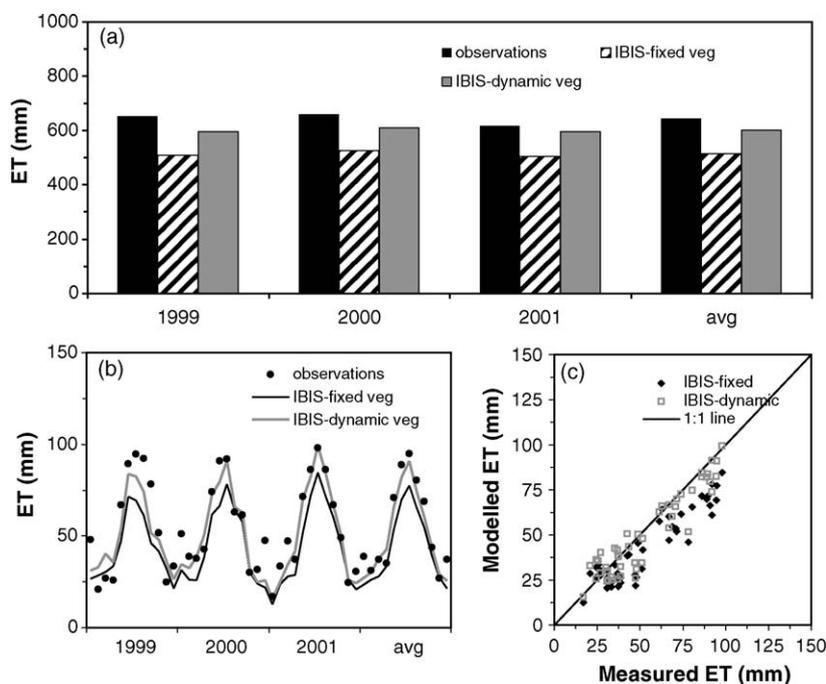


Fig. 10 – Simulated and observed (a) annual total ET; (b) monthly ET; (c) the correlation of simulated and observed monthly ET at Niwot Ridge.

ing that either measurements underestimated net radiation or overestimated turbulent energy fluxes (Wilson et al., 2002). This may also explain the negative bias of simulated ET, in contrast to the positive bias obtained at WB and HF, where measured net radiation were more than the sum of measured energy turbulent fluxes.

4. Conclusions

The growing database of multiyear, short timescale (seconds–minutes) observations of coupled water–carbon fluxes has given the ecosystem modeling community a wealth of information to use in model testing. However, we argue that model testing protocols should be extended to include other ecosystem variables such as LAI (phenology), soil respiration, and soil temperature and moisture. Baldocchi and Wilson (2001) suggest that confidence in long-term simulations of coupled carbon–water fluxes will be gained by the integration of leaf phenology models into biophysical and biogeochemical modeling frameworks.

In this study, a dynamic vegetation and biosphere model (IBIS), was used to simulate the seasonal and annual changes in carbon and water exchange in three mid-latitude forest locations in the AmeriFlux observational network. The approach was designed to identify model limitations and overall sources of error at the individual-site level. To do so, IBIS was applied in two ways: (1) a “fixed” vegetation mode, where observations of canopy structure, species, and soil properties were used as input, in addition to the site-specific micrometeorological data and (2) a “dynamic” vegetation mode, in which only soil textural information and weather data were used,

thereby allowing the “dynamic vegetation” component of the model to simulate the long-term (decadal) development of the current vegetation (species or PFT) distribution and allow for LAI to change (phenology) seasonally as part of simulations. The hypothesis was that the application of the model in these two modes would allow for separation of model error due to parameterizations and formulations of plant physiology, canopy and soil physics, soil biogeochemistry, ecosystem respiration, and global-scale generalizations of PFT communities; versus errors due to the vegetation dynamics portion of the model.

In a general sense, the model reproduced mean annual water and carbon exchange with the atmosphere fairly well, a conclusion that has been noted in previous work (Kucharik et al., 2000). Fig. 11a–c illustrates that IBIS captured annual GEP, R, and NEP differences between biomes. The DV runs clearly had a positive bias for GEP and R. Fig. 11d suggested that inter-annual variability in simulated ET was decreased compared to observations, but the key biome differences were replicated. On shorter timescales, IBIS was able to capture greater than 80% of the observed month-to-month variability in carbon and water exchange (Table A.5), and was relatively consistent in simulations of soil temperature and net radiation. However, the magnitude of seasonal changes in ecosystem respiration were greatly exaggerated, thus the annual average agreement between simulated values and observations was due to compensating errors. While the model was able to capture the significant differences in carbon and water vapor exchange across the varied forest types in this study, largely based on the generalized parameterizations of plant functional groupings, it was apparent that capturing the inter-annual variability of coupled carbon–water exchanges at each individual site

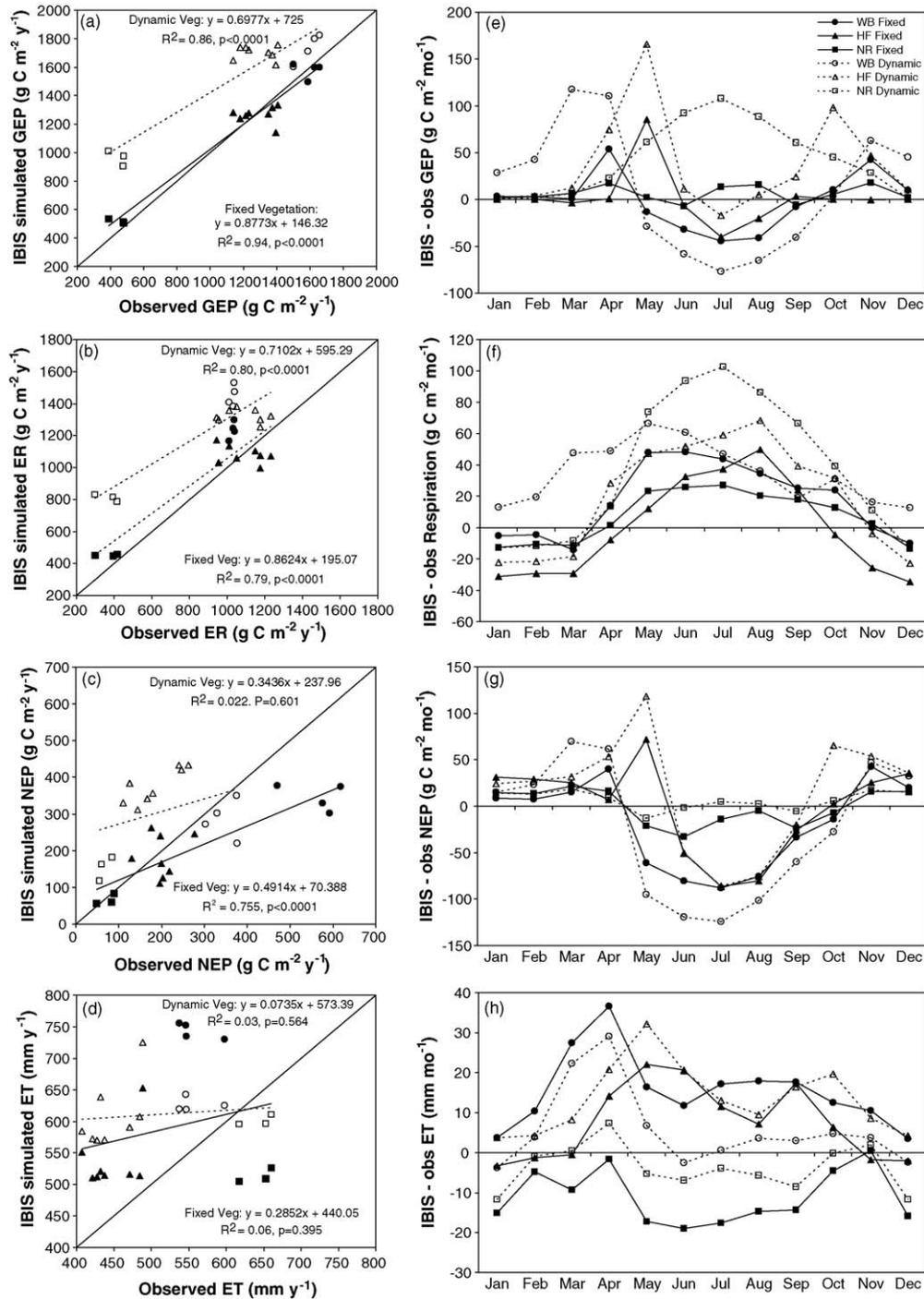


Fig. 11 – Monthly average differences between simulated (fixed vegetation and dynamic vegetation runs) values and observations for (a) GEP, (b) ecosystem respiration, (c) NEP, and (d) ET. Closed symbols are for fixed vegetation runs and open symbols are for dynamic vegetation runs. Circles are for WB, triangles for HF, and squares for NR. Correlations of annual values for simulated and observed (e) GEP, (f) ecosystem respiration, (g) NEP, and (h) ET across all sites and all years of comparison.

was a more formidable task (Table A.5). Thus, we expect that the model may simulate successfully the variation in average biome NEP and ET across larger regions or the global scale (e.g., Kucharik et al., 2000), but simulations of seasonal changes and year-to-year fluctuations would likely be poor. Below we highlight the important conclusions about model performance.

- The global to continental scale phenology models appear to have significant difficulty capturing the timing of budburst and evolution of canopy leaf out in deciduous forests at the individual site level. The problem is also present in leaf offset at the end of the growing season. Biases of early season green-up of 6 weeks were not uncommon. Consid-

- ering the importance of changes in carbon exchange during the spring to early summer, the simulated error leads to significant problems in simulating the correct response of vegetation growth and carbon and water exchange. Baldocchi et al. (2005) found that 64% of the variance in the timing of net CO₂ uptake by several deciduous forests was explained by the point at which soil temperatures reached the annual average air temperature. Application of these prognostic phenology approaches across regional to global scales would likely lead to substantial errors in the simulated seasonality of carbon exchange, and erroneous results if applied in climate change studies. Furthermore, when running the model in dynamic mode, slight changes in plant phenology can lead to widely varying PFT development due to differences in shading, leading to incorrect vegetation trajectories over time. The simplified vegetation dynamics routine sometimes led to the development of a species mixture at the sites that was not characteristic of the observed current vegetation type. Smith et al. (2001) suggested that an individual plant based approach to modeling vegetation dynamics may have advantages at regional to continental scales when trying to capture ecosystem structure and function.
- Simulated gross ecosystem production (Fig. 11e) shows a pattern of early growing season positive bias and a general midseason deficit. Dynamic vegetation simulations tend to produce larger error than fixed vegetation runs because of the inability to represent accurately the site-level phenology.
 - Ecosystem respiration (Fig. 11f) is generally overestimated during the growing season, on average, by 20–60 gC m⁻² month⁻¹, and underestimated during the winter by 10–20 gC m⁻² month⁻¹ across all sites. Fundamental model formulation errors with respect to root, stem, and soil respiration are apparent; the sensitivity of respiration to temperature is too pronounced for stem, leaf, and root. This suggests that in a climate-warming scenario, increased rates of simulated decomposition and respiration could lead to net loss of biomass in an ecosystem. The results here are indicative of what could be produced by a larger genre of biosphere models commonly used in integrated assessments of climate change impacts to ecosystems and the carbon cycle. The error generally increases in dynamic vegetation runs (compared to fixed vegetation simulations) because simulated plant phenology usually leads to a bias of earlier budburst, significantly accelerated leaf-out to maximum LAI, overestimated maximum LAI, and the existence of PFTs that are generally either not observed at the site or are located in a simulated lower canopy. Soil respiration appears to be a difficult process to capture in terms of the magnitude of exchange (summer maximum) and during wintertime. This could either be related to the simplistic functions that are applied to modify the rate of microbial activity and organic matter decomposition (Arrhenius for temperature (Lloyd and Taylor, 1994) and water-filled pore space for soil moisture (Linn and Doran, 1984)). It is also conceivable that individual carbon pools were poorly characterized in the model. It appears that a major weakness is the relatively low contribution of root respiration to total soil respiration.
 - The overall comparison of simulated NEP (Fig. 11g) to observations showed that in the majority of model runs across all sites, there was a significant underestimate of growing season NEP (May–September) of 25–100 gC m⁻² month⁻¹, and a overall positive bias of 10–40 gC m⁻² month⁻¹ during the winter season, related to improper characterization of R. At the two sites with deciduous vegetation, there was a early season peak in overestimated NEP during April and May, presumably due to incorrect simulation of budburst (early bias) in dynamic runs, and potentially because of warmer than observed soil temperatures near the surface.
 - Simulated soil temperatures are overestimated during the summer, and underestimated during the winter. This was a common theme across all sites and it most likely related to the fact that the model does not account for a litter or thatch layer on the soil surface—leading to a more pronounced seasonal cycle of soil temperature changes. This may have contributed to the simulated early-season bias of enhanced vegetation green-up and carbon uptake. Point measurements of soil moisture at the three sites proved to be a difficult test for IBIS. The model could not replicate the substantial variation of this quantity (e.g., all correlations between observations and simulations yielded $r^2 < 0.35$).
 - The comparison of simulated and observed ET (Fig. 11h) showed biases were related to the dominant vegetation type at each AmeriFlux site (evergreen conifer vs. deciduous) and whether the measured energy balance (latent + sensible + ground heat) resulted in a net deficit or surplus when compared to the measured net radiation. The deciduous forest site (WB and HF) energy balance measurements showed that measured net radiation was 22 and 9% greater than the summation of the individual partitioned terms. Correspondingly, IBIS simulations produced an annual average 33 and 22% overestimate of ET (latent heat fluxes) at the WB and HF sites, respectively. At the NR conifer site, observations showed that annual average net radiation was 9% smaller than the summation of the individual energy balance terms. In this case, IBIS simulated an annual average –20% error in ET. Thus, it appears that the model could be used as a tool in closing the energy budget at these sites.
 - The magnitude of simulated variation in seasonal and inter-annual carbon exchange is generally dampened with respect to observations. These discrepancies show up particularly well for the cumulative monthly NEP and for soil respiration. It is clear that in some instances IBIS is able to simulate annual average NEP exceedingly well, but that these simulations are artifacts of offsetting errors in seasonal carbon exchange. The parameterization of global-scale biosphere or ecosystem models may be such that they are constrained to produce acceptable output across a wide range of climate and soil conditions and biome types across the globe. It is also possible that important site-level processes are not represented by the model at a fine scale. If this is the case, then we have to use these general models cautiously at the fine scale and for climate change studies. Furthermore, when using the models at a larger scale (or coarser resolution) we must then operate with the assumption that site-level variability is much greater than observed variability across a larger scale (e.g., 0.5° resolution).

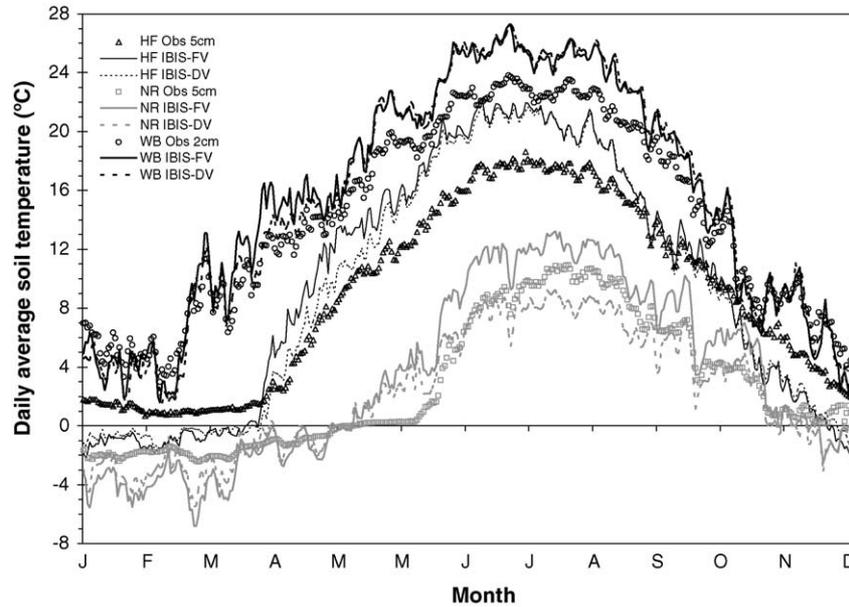


Fig. A.1 – Average daily soil temperature for IBIS fixed vegetation (FV) and dynamic vegetation (DV) simulations compared with daily average observations (obs) for the 1992–1999 period at Harvard Forest (HF), the 1999–2001 period at Niwot Ridge (NR), and the 1995–1998 period at Walker Branch (WB).

We expected to capture the seasonal variation in carbon and water exchange better when given such important quantities as LAI as model input, rather than relying upon the simulation of long-term (decadal to century) species competition (re-growth) and short-term (daily) phenology. However, the enhanced model errors in dynamic vegetation runs were not entirely due to poor representation of plant phenology. While this obviously contributed errors in the seasonal timing of the net carbon uptake, the growth of tree and shrub species not

representative of each site was a more troubling error. At NR, for instance, a substantial lower canopy of shrubs developed in the DV simulation. In fairness to the model, IBIS was not originally designed to simulate the precise vegetation species composition at single sites. The IBIS DGVM was originally developed for global-scale simulations—to look specifically at the impact of changing climate and atmospheric CO₂ on vegetation distribution and carbon and water exchange. The scale at which the model had been previously used was in the

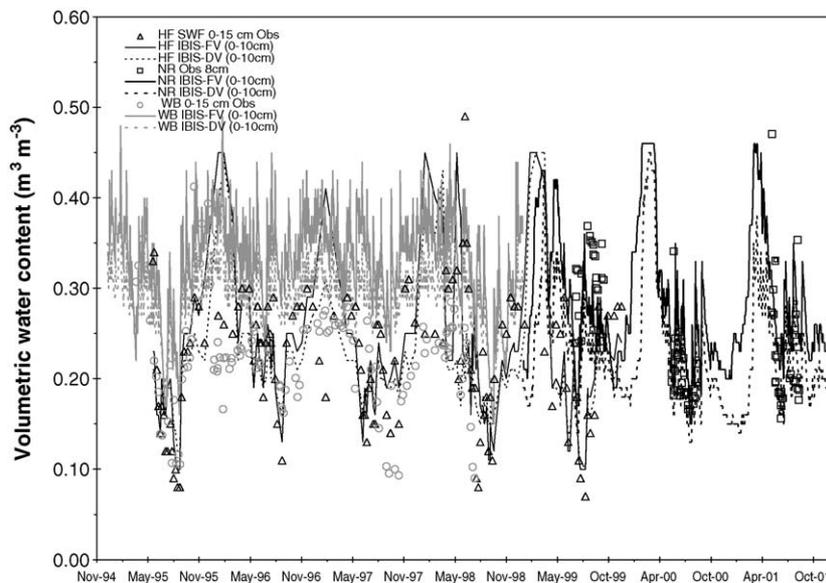


Fig. A.2 – Average daily volumetric water content for IBIS fixed vegetation (FV) and dynamic vegetation (DV) simulations (0–10 cm) compared with periodic observations (obs) for the 1992–1999 period at Harvard Forest (HF SWF transect, 0–15 cm), the 1999–2001 period at Niwot Ridge (NR, 8 cm), and the 1995–1998 period at Walker Branch (WB, 0–15 cm).

Table A.1 – Climate summary at each of the experimental sites, over the period of field measurements

Site	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	Normals 1961–1990
Walker Branch											
Mean temperature				14.5	13.7	14.2	16.1				13.8
Precipitation	–	–	–	1253	1705	1450	1663	–	–	–	1352
Solar radiation				5444	5429	5417	5276				–
Harvard Forest											
Mean temperature	8.2	7.5	7.5	7.7	7.4	6.9	8.9	9.0			7.8
Precipitation	1095	1262	1332	1083	1556	924	949	1049	–	–	1066
Solar radiation	3849	4057	4028	4109	4086	4224	4102	4331			–
Niwot Ridge											
Mean temperature								2.3	2.7	2.9	4.0
Precipitation	–	–	–	–	–	–	–	766	662	660	800
Solar radiation								6382	6044	5881	–

Data are observed mean annual air temperature (°C), total annual precipitation (mm), and total incident solar radiation ($\text{MJ m}^{-2} \text{year}^{-1}$). Normal values were taken from the AmeriFlux website: <http://public.ornl.gov/fluxnet/dataindex.cfm>.

0.5–4° grid cell (latitude by longitude) size range. In order for such simulations to be realistic, the existence rules and plant physiological parameterizations for the 12 PFTs are broad. There is no guarantee that the parameterizations, which are based on various site-specific studies, will apply to all indi-

vidual sites and there is no reason to believe that the resulting errors will average out to a value of zero across a larger landscape.

The results of this model testing suggest several ways to improve the modeling approach. Given the fact that the

Table A.2 – Walker Branch, TN monthly averages from 1995 to 1998 (top); GEP, R, and NEP values are in units of $\text{g C m}^{-2} \text{month}^{-1}$; ET and runoff in mm month^{-1} ; (bottom) average simulated monthly MBE (%)

Month	GEP			Respiration			NEP			ET			Runoff		
	Obs	FV	DV	Obs	FV	DV	Obs	FV	DV	Obs	FV	DV	Obs	FV	DV
January	0.6	4.0	29.0	24.2	18.9	37.3	–23.6	–14.9	–8.3	12.6	16.3	8.8	137.7	66.0	134.4
February	2.9	5.7	45.5	27.2	22.6	46.7	–24.2	–16.9	–1.3	11.8	22.3	15.6	101.7	123.9	119.6
March	11.6	12.6	128.9	52.8	38.7	100.5	–41.1	–26.1	28.4	21.1	48.6	43.4	147.8	108.8	118.9
April	70.8	124.5	181.2	77.8	91.8	126.6	–7.0	32.7	54.7	31.6	68.2	60.7	136.3	115.0	107.6
May	264.6	251.5	236.1	107.7	155.5	174.1	156.9	96.0	61.9	77.0	93.5	83.8	94.9	99.4	75.3
June	293.9	262.0	235.4	126.8	175.1	187.7	167.1	86.9	47.7	93.1	104.9	90.6	73.0	83.3	84.2
July	296.3	252.0	219.1	147.5	191.2	194.4	148.8	60.9	24.7	96.3	113.4	96.9	34.0	82.7	40.1
August	279.4	238.3	214.4	154.2	188.6	190.4	125.2	49.7	24.0	88.1	106.1	91.8	36.6	36.9	61.4
September	205.8	197.8	165.1	129.8	155.0	148.9	76.0	42.7	16.2	59.2	76.9	62.2	20.2	26.1	18.5
October	154.3	164.5	157.9	95.8	119.9	126.7	58.5	44.6	31.2	39.8	52.3	44.5	20.8	17.2	16.2
November	14.4	56.9	77.4	50.3	50.1	66.6	–35.9	6.8	10.8	13.5	23.9	17.2	33.0	20.6	39.0
December	–0.9	8.9	44.3	35.7	25.6	48.3	–36.6	–16.7	–4.0	13.3	16.8	10.9	63.8	34.5	79.6
Annual	1593.8	1578.7	1734.3	1029.7	1233.1	1448.2	564.0	345.6	286.1	557.3	743.1	626.4	899.9	814.5	894.8
Month	GEP		Respiration		NEP		ET		Runoff						
	FV	DV	FV	DV	FV	DV	FV	DV	FV	DV					
January	535.9	4525.3	–21.9	53.8	36.7	64.9	29.8	–29.6	–52.0	–2.4					
February	94.9	1442.3	–16.6	72.0	30.2	94.8	88.4	32.2	21.8	17.6					
March	8.0	1008.2	–26.8	90.3	36.6	169.1	130.4	105.7	–26.4	–19.6					
April	75.9	156.0	18.1	62.8	570.3	885.4	115.7	91.9	–15.6	–21.1					
May	–4.9	–10.8	44.4	61.7	–38.8	–60.5	21.4	8.8	4.7	–20.7					
June	–10.9	–19.9	38.1	48.0	–48.0	–71.4	12.7	–2.7	14.0	15.3					
July	–14.9	–26.0	29.6	31.8	–59.1	–83.4	17.8	0.6	142.9	17.7					
August	–14.7	–23.3	22.3	23.5	–60.3	–80.8	20.4	4.1	0.9	67.8					
September	–3.9	–19.8	19.4	14.7	–43.8	–78.7	29.9	5.2	29.5	–8.3					
October	6.6	2.4	25.1	32.3	–23.7	–46.7	31.5	11.9	–17.5	–22.3					
November	295.8	438.0	–0.3	32.4	119.0	130.1	77.7	27.6	–37.7	18.2					
December	1091.3	5051.1	–28.3	35.5	54.3	89.0	26.4	–18.3	–45.9	24.8					
Annual	–0.95	8.82	19.74	40.64	–38.72	–49.28	33.34	12.39	–9.49	–0.57					

model has been largely successful when energy, water and carbon balance have been validated in “fixed mode” (Delire and Foley, 1999; Lenters et al., 2000; Kucharik et al., 2001; El Maayar et al., 2001, 2002), the model holds great promise for future scaling work from the site level to regional or global scale. However, in order to fully understand how future ecosystems may exchange carbon and water with the atmosphere, vegetation characteristics (such as LAI) cannot be prescribed in our biosphere models. This would defeat the purpose of developing such modeling tools. Thus, we need to develop better simulations of the dynamic aspects of vegetation, in the context of changing atmospheric conditions—at least at a daily timescale, if not shorter. However, another fundamental outcome of this study was the overall inability of IBIS to capture the inter-annual variability of flux measurements—even when the model was in fixed vegetation mode. This suggests that there are also problems in either leaf-level physiological response to water or nutrient stress, and/or in scaling algorithms that take average leaf photosynthesis and convert it to a canopy value. The lack of simulations of nutrient stress, pests and their impact on plant productivity may be part of the reason for low simulated inter-annual variability.

We suggest that revisions to DGVMs should focus on developing suitable phenology schemes that account for pho-

toperiod, soil moisture and frost in addition to temperature. Model formulations of ecosystem respiration need to be rethought, particularly with respect to use of Q_{10} or Arrhenius temperature functions as modifiers. Surface energy balance (litter layer), carbon allocation, soil respiration, and plant response to varied stresses will have to be improved as well. Furthermore, if DGVMs are intended to simulate biosphere–atmosphere carbon and water vapor exchange and, in particular, the distribution of vegetation types in response to changes in climate and atmospheric CO_2 , we need to rethink how PFTs are represented in these models, and decide whether other processes, such as species migration (Neilson et al., 2005), need to be included.

Acknowledgements

The authors wish to thank the numerous AmeriFlux investigators who contributed to the collection of essential field data at Niwot Ridge, Walker Branch, and Harvard Forest. The senior author personally sends gratitude to Kell Wilson, Laura Scott-Denton, Kathleen Savage, Eric Davidson, and Paul Hanson for providing an array of useful model validation data. Runoff and the hourly climate data were

Table A.3 – Harvard Forest, MA monthly averages from 1992 to 1999 (top); GEP, R, and NEP values are in units of $g\ C\ m^{-2}\ month^{-1}$; ET in $mm\ month^{-1}$; (bottom) average simulated monthly MBE (%)

Month	GEP			Respiration			NEP			ET		
	Obs	FV	DV	Obs	FV	DV	Obs	FV	DV	Obs	FV	DV
January	0.2	0.1	2.6	38.0	7.0	15.7	-37.9	-6.9	-13.5	8.8	5.5	12.5
February	0.1	0.1	3.9	35.8	6.5	14.2	-35.7	-6.3	-8.5	9.8	8.5	14.0
March	4.2	0.7	16.7	40.9	11.8	22.7	-36.3	-11.0	-4.6	16.5	15.9	24.8
April	33.3	34.1	108.2	66.9	59.0	95.4	-31.9	-24.7	21.3	24.3	38.5	45.1
May	112.4	197.5	278.2	128.7	140.9	176.1	-14.4	57.1	103.7	51.1	73.1	83.3
June	289.0	282.0	300.6	150.8	183.4	202.7	142.5	92.6	92.0	71.6	92.2	92.0
July	312.6	272.7	295.5	167.0	204.2	226.0	153.3	66.9	67.6	92.1	103.6	105.1
August	272.7	252.8	277.6	140.2	190.0	208.8	131.3	51.1	55.6	80.1	87.2	89.7
September	176.2	179.6	200.5	118.7	142.5	158.0	65.7	39.6	45.7	44.0	61.3	60.4
October	59.8	60.1	157.8	88.1	83.5	119.8	-26.3	-23.6	38.8	26.0	32.4	45.6
November	0.8	0.7	47.6	60.7	35.0	56.6	-59.8	-34.3	-5.8	12.8	11.0	21.3
December	0.1	0.3	10.8	51.2	16.5	28.4	-51.1	-16.2	-15.6	9.4	7.4	13.6
Annual	1261.6	1280.8	1699.9	1086.9	1080.4	1324.3	199.5	184.4	376.7	446.3	536.6	607.6
Month	GEP		Respiration		NEP		ET					
	FV	DV	FV	DV	FV	DV	FV	DV				
January	-41.6	1172.8	-81.7	-58.8	81.9	64.5	-37.5	42.6				
February	1.4	2617.4	-81.7	-60.5	82.3	76.1	-13.2	43.2				
March	-83.0	292.9	-71.2	-44.5	69.6	87.3	-3.5	50.3				
April	2.5	225.1	-11.9	42.5	22.6	166.7	58.3	85.6				
May	75.7	147.5	9.5	36.9	498.1	822.5	43.1	63.1				
June	-2.4	4.0	21.7	34.4	-35.1	-35.4	28.9	28.6				
July	-12.7	-5.5	22.3	35.3	-56.4	-55.9	12.5	14.2				
August	-7.3	1.8	35.5	48.9	-61.1	-57.6	8.9	11.9				
September	1.9	13.8	20.1	33.2	-39.7	-30.4	39.5	37.4				
October	0.3	163.7	-5.2	36.0	10.3	247.6	24.4	75.5				
November	-18.6	5635.5	-42.4	-6.7	42.7	90.3	-13.6	67.3				
December	137.7	10108	-67.7	-44.6	68.2	69.6	-21.6	45.1				
Annual	1.52	34.75	-0.60	21.83	-7.61	88.81	20.22	36.13				

Table A.4 – Niwot Ridge, CO monthly averages from 1999 to 2001 (top); GEP, R, and NEP values are in units of $\text{g C m}^{-2} \text{ month}^{-1}$; ET in mm month^{-1} ; (bottom) average simulated monthly MBE (%)

Month	GEP			Respiration			NEP			ET		
	Obs	FV	DV	Obs	FV	DV	Obs	FV	DV	Obs	FV	DV
January	0.0	1.7	0.7	21.9	9.2	8.9	-21.9	-7.5	-8.1	38.7	23.5	27.0
February	0.0	2.8	1.4	20.5	9.9	8.7	-20.5	-7.0	-7.3	31.0	26.2	30.1
March	0.0	6.6	6.5	23.5	12.1	15.2	-27.3	-5.5	-8.7	37.3	28.0	37.8
April	0.0	17.6	22.7	16.5	18.0	30.1	-16.5	-0.4	-7.4	35.2	33.6	42.6
May	69.6	71.7	130.8	21.6	44.7	95.5	48.0	27.0	35.3	70.8	53.6	65.5
June	101.2	94.2	193.4	43.4	69.3	137.2	57.8	24.8	56.2	88.9	70.0	81.9
July	85.6	98.9	193.4	57.4	84.6	160.1	28.2	14.3	33.3	94.9	77.4	91.0
August	76.2	91.9	165.1	55.2	75.5	141.5	21.0	16.4	23.6	80.4	65.7	74.7
September	74.8	68.9	135.8	41.8	59.8	108.4	33.0	9.1	27.4	68.9	54.6	60.3
October	38.7	44.3	84.0	25.9	38.6	65.2	12.8	5.7	18.8	43.6	39.1	43.5
November	0.0	18.0	28.5	18.8	21.1	29.8	-18.8	-3.1	-1.3	26.9	27.5	28.9
December	0.0	2.4	1.5	22.8	9.8	9.4	-22.8	-7.4	-7.9	37.2	21.3	25.6
Annual	446.2	519.1	964.0	369.3	452.6	810.1	73.2	66.4	153.9	653.8	520.5	609.0

Month	GEP		Respiration		NEP		ET	
	FV	DV	FV	DV	FV	DV	FV	DV
January			-58.1	-59.6	65.9	63.0	-39.1	-30.1
February	-	-	-51.9	-57.4	65.8	64.3	-15.5	-2.9
March	-	-	-48.7	-35.3	79.7	68.1	-24.9	1.4
April	-	-	9.2	83.1	97.6	55.0	-4.6	21.1
May	3.1	87.9	107.2	342.9	-43.7	-26.5	-24.4	-7.5
June	-6.9	91.2	59.9	216.5	-57.1	-2.8	-21.3	-7.9
July	15.6	125.9	47.4	179.0	-49.2	18.0	-18.5	-4.1
August	20.6	116.5	36.7	156.1	-21.9	12.4	-18.3	-7.0
September	-7.9	81.5	42.9	159.3	-72.3	-17.0	-20.7	-12.4
October	14.4	117.0	49.3	151.8	-55.9	46.8	-10.3	-0.3
November	-	-	12.4	59.1	83.4	93.1	2.0	7.2
December	-	-	-56.9	-59.0	67.4	65.5	-42.7	-31.3
Annual	16.34	116.06	22.57	119.37	-9.20	110.30	-20.40	-6.86

obtained from the Walker Branch Watershed Long-term Data Archive funded by the Program for Ecosystem Research, Environmental Sciences Division, Office of Health and Environmental Research, U.S. Department of Energy. This research

was supported by a grant through the Office of Science, Biological and Environmental Research Program (BER), U.S. Department of Energy, through the South Central Regional Center of the National Institute for Global Environmental

Table A.5 – Coefficient of determination (r^2) between simulated and observed values for components of carbon, water, and energy cycles

Site	GEP		R		NEP		ET		R_{net}		T_{soil}		Soil M		Soil CO ₂		Runoff	
	FV	DV	FV	DV	FV	DV	FV	DV	FV	DV	FV	DV	FV	DV	FV	DV	FV	DV
Seasonal ^{a,b}																		
Harvard Forest	0.91	0.78	0.85	0.86	0.70	0.44	0.90	0.89	0.90	0.93	0.92	0.95	0.27	0.23	0.54	0.68	NA	NA
Walker Branch	0.94	0.80	0.92	0.89	0.79	0.30	0.90	0.88	0.97	0.98	0.96	0.98	0.29	0.25	NA	NA	0.49	0.72
Niwot Ridge	0.92	0.93	0.67	0.61	0.88	0.92	0.84	0.86	0.97	0.97	0.89	0.87	0.09*	0.17	0.08*	0.28	NA	NA
Annual ^c																		
Harvard Forest	0.00	0.01	0.16	0.04	0.02	0.03	0.18	0.36	0.79*	0.82*	NA	NA	NA	NA	0.21	0.52	NA	NA
Walker Branch	0.01	0.98*	0.62	0.33	0.15	0.80	0.55	0.00	0.63	0.84	NA	NA	NA	NA	NA	NA	0.53	0.57
Niwot Ridge	0.96	0.54	0.15	0.76	0.51	0.98	0.02	0.06	0.14	0.11	NA	NA	NA	NA	NA	NA	NA	NA

^a All seasonal quantities are monthly values with the exception of soil moisture (Soil M) and soil CO₂ efflux which were measured periodically.

^b All seasonal correlations were significant at the $P < 0.0001$ level, with the exception of r^2 values followed by asterisk (*), which were significant at $P < 0.05$.

^c All annual correlations were not statistically significant, with the exception of those r^2 values followed by asterisk (*), which were significant at $P < 0.05$.

Change (NIGEC) under Cooperative Agreement No. DE-FC03-90ER61010.

Appendix A

See Figs. A.1 and A.2 and Tables A.1–A.5.

REFERENCES

- Arora, V., 2002. Modeling vegetation as a dynamic component in soil–vegetation–atmosphere transfer schemes and hydrological models. *Rev. Geophys.* 40 (2), Art. No. 1006.
- Bachelet, D., Neilson, R.P., Hickler, T., Drapek, R.J., Lenihan, J.M., Sykes, M.T., Smith, B., Sitch, S., Thonicke, K., 2003. Simulating past and future dynamics of natural ecosystems in the United States. *Global Biogeochem. Cycles* 17 (2), 1045, doi:10.1029/2001GB001508.
- Baker, I., Denning, A.S., Hanan, N., Prihodko, L., Uliasz, M., Vidale, P.L., Davis, K., Bakwin, P., 2003. Simulated and observed fluxes of sensible and latent heat and CO₂ at the WLEF-TV tower using SiB2.5. *Global Change Biol.* 9, 1262–1277.
- Baldocchi, D.D., Wilson, K.B., 2001. Modelling CO₂ and water vapor exchange of a temperate broadleaved forest on daily to decadal time scales. *Ecol. Modell.* 142, 155–184.
- Baldocchi, D.D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, Ch., Davis, K., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, J.W., Oechel, W., Pilegaard, K., Schmid, H.P., Valentini, R., Verma, S., Vesala, T., Wilson, K., Wofsy, S., 2001. FLUXNET: a new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor and energy flux densities. *Bull. Am. Meteorol. Soc.* 82, 2415–2435.
- Baldocchi, D.D., Black, T.A., Curtis, P.S., Falge, E., Fuentes, J.D., Granier, A., Gu, L., Knohl, A., Pilegaard, K., Schmid, H.P., Valentini, R., Wilson, K., Wofsy, S., Xu, L., Yamamoto, S., 2005. Predicting the onset of net carbon uptake by deciduous forests with soil temperature and climate data: a synthesis of FLUXNET data. *Int. J. Biometeorol.* 49, 377–387.
- Ball, J.T., 1988. The ci/cs ratio: a basis for predicting stomatal control of photosynthesis. Ph.D. Thesis. Stanford University.
- Barford, C.C., Wofsy, S.C., Goulden, M.L., Munger, J.W., Pyle, E.H., Urbanski, S.P., Hutrya, L., Saleska, S.R., Fitzjarrald, D., Moore, K., 2001. Factors controlling long- and short-term sequestration of atmospheric CO₂ in a mid-latitude forest. *Science* 294, 1688–1691.
- Bonan, G.B., Levis, S., Sitch, S., Vertenstein, M., Oleson, K.W., 2003. A dynamic global vegetation model for use with climate models: concepts and description of simulated vegetation dynamics. *Global Change Biol.* 9, 1543–1566.
- Brutsaert, W.H., 1975. On a derivable formula for long-wave radiation from clear skies. *Water Res.* 11, 742–744.
- Campbell, G.S., Norman, J.M., 1998. *An Introduction to Environmental Biophysics*. Springer, New York, 286 pp.
- Charney, J.G., 1975. Dynamics of deserts and drought in the Sahel. *Q. J. R. Met. Soc.* 101, 193–202.
- Collatz, G.J., Ball, J.T., Grivet, C., Berry, J.A., 1991. Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer. *Agric. Forest Meteorol.* 53, 107–136.
- Collatz, G.J., Ribas-Carbo, M., Berry, J.A., 1992. Coupled photosynthesis-stomatal conductance model for leaves of C4 plants. *Aust. J. Plant Physiol.* 19, 519–538.
- Cramer, W., Bondeau, A., Woodward, F.I., Prentice, I.C., Betts, R.A., Brovkin, V., Cox, P.M., Fisher, V., Foley, J., Friend, A.D., Kucharik, C., Lomas, M.R., Ramankutty, N., Sitch, S., Smith, B., White, A., Young-Molling, C., 2001. Global response of terrestrial ecosystem structure and function to CO₂ and climate change: results from six dynamic global vegetation models. *Global Change Biol.* 7, 347–373.
- Curtis, P.S., Hanson, P.J., Bolstad, P., Barford, C., Randolph, J.C., Schmid, H.P., Wilson, K.B., 2002. Biometric and eddy-covariance based estimates of annual carbon storage in five eastern North American deciduous forests. *Agric. Forest Meteorol.* 113, 3–19.
- Davidson, E.A., Belk, E., Boone, R.D., 1998. Soil water content and temperature as independent or confounded factors controlling soil respiration in a temperate mixed hardwood forest. *Global Change Biol.* 4, 217–227.
- Delire, C., Foley, J.A., 1999. Evaluating the performance of a land surface/ecosystem model with biophysical measurements from contrasting environments. *J. Geophys. Res. (Atmos.)* 104 (D14), 16,895–16,909.
- Delire, C., Foley, J.A., Thompson, S., 2003. Evaluating the carbon cycle of a coupled atmosphere-biosphere model. *Global Biogeochem. Cycles* 17 (1), doi:10.1029/2002GB001870.
- Delire, C., Foley, J.A., Thompson, S., 2004. Long-term variability in a coupled atmosphere-biosphere model. *J. Clim.* 20 (October), 3947–3959.
- El Maayar, M., Price, D.T., Delire, C., Foley, J.A., Back, A.T., Bessemoulin, P., 2001. Validation of the Integrated Biosphere Simulator over Canadian deciduous and coniferous boreal forest stands. *J. Geophys. Res.* 106 (D13), 14,339–14,355.
- El Maayar, M., Price, D.T., Black, T.A., Humphreys, E.R., Jork, E.M., 2002. Sensitivity tests of the Integrated Biosphere Simulator to soil and vegetation characteristics in a Pacific Coastal Coniferous Forest. *Atmos. Ocean* 40 (3), 313–332.
- Falge, E., Baldocchi, D., Olson, R., et al., 2001. Gap filling strategies for defensible annual sums of net ecosystem exchange. *Agric. Forest Meteorol.* 107 (1), 43–69.
- Falge, E., Baldocchi, D., Tenhunen, J., et al., 2002. Seasonality of ecosystem respiration and gross primary production as derived from FLUXNET measurements. *Agric. Forest Meteorol.* 113, 53–74.
- Farquhar, G.D., von Caemmerer, S., Berry, J.A., 1980. A biogeochemical model of photosynthetic CO₂ assimilation in leaves of C3 species. *Planta* 149, 78–90.
- Foley, J.A., Prentice, I.C., Ramankutty, N., Levis, S., Pollard, D., Sitch, S., Haxeltine, A., 1996. An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. *Global Biogeochem. Cycles* 10 (4), 603–628.
- Foley, J.A., Levis, S.I.C., Prentice, I.C.D., Pollard, D., Thompson, S.L., 1998. Coupling dynamic models of climate and vegetation. *Global Change Biol.* 4, 561–579.
- Foley, J.A., Levis, S., Costa, M.H., Cramer, W., Pollard, D., 2000. Incorporating dynamic vegetation cover within global climate models. *Ecol. Appl.* 10 (6), 1620–1632.
- Foster, D.R., Zebryk, T., Schoonmaker, P., Lezberg, A., 1992. Post-settlement history of human land-use and vegetation dynamics of a hemlock woodlot in central New England. *J. Ecol.* 80, 773–786.
- Friend, A.D., Stevens, A.K., Knox, R.G., Cannell, M.G.R., 1997. A process-based terrestrial biosphere model of ecosystem dynamics (Hybrid v3.0). *Ecol. Modell.* 95, 249–287.
- Gerten, D., Schaphoff, S., Haberlandt, U., Lucht, W., Sitch, S., 2004. Terrestrial vegetation and water balance-hydrological evaluation of a dynamic global vegetation model. *J. Hydrol.* 286, 249–270.
- Gifford, R.M., 2003. Plant respiration in productivity models: conceptualization, representation and issues for global

- terrestrial carbon-cycle research. *Funct. Plant Biol.* 30, 171–186.
- Goolsby, D.A., Battaglin, W.A., Aulenbach, B.T., Hooper, R.P., 2000. Nitrogen flux and sources in the Mississippi River Basin. *Sci. Total Environ.* 248 (2–3), 75–86.
- Gordon, W.S., Famiglietti, J.S., Fowler, N.L., Kittel, T.G.F., Hibbard, K.A., 2004. Validation of simulated runoff from six terrestrial ecosystem models: results from VEMAP. *Ecol. Appl.* 14 (2), 527–545.
- Gough, C.M., Seiler, J.R., Johnsen, K.H., Sampson, D.A., 2004. Seasonal photosynthesis in fertilized and nonfertilized loblolly pine. *Forest Sci.* 50, 1–9.
- Goulden, M.L., Munger, J.S., Fan, S.M., et al., 1996. Measurements of carbon sequestration by long-term eddy covariance: methods and a critical evaluation of accuracy. *Global Change Biol.* 2 (3), 169–182.
- Gower, S.T., Kucharik, C.J., Norman, J.M., 1999. Direct and indirect estimate of leaf area index f_{apar} and net primary production of terrestrial ecosystems. *Remote Sens. Environ.* 70 (1), 29–51.
- Hanson, P.J., Wullschlegel, S.D., Bohlman, S.A., Todd, D.E., 1993. Seasonal and topographic patterns of forest floor CO₂ efflux from an upland oak forest. *Tree Physiol.* 13, 1–15.
- Hanson, P.J., Edwards, N.T., Garten, C.T., Andrews, J.A., 2000. Separating root and soil microbial contributions to soil respiration: a review of methods and observations. *Biogeochemistry* 48, 115–146.
- Hanson, P.J., Amthor, J.S., Wullschlegel, S.D., Wilson, K.B., Grant, R.F., Hartley, A., Hui, D., Hunt Jr., E.R., Johnson, D.W., Kimball, J.S., King, A.W., Luo, Y., McNulty, S.G., Sun, G., Thornton, P.E., Wang, S., Williams, M., Baldocchi, D.D., Cushman, R.M., 2004. Oak forest carbon and water simulations: model intercomparisons and evaluations against independent data. *Ecol. Monogr.* 74, 443–489.
- Hickler, T., Smith, B., Sykes, M.T., Davis, M.B., Sugita, S., Walker, K., 2004. Using a generalized vegetation model to simulate vegetation dynamics in northeastern USA. *Ecology* 85 (2), 519–530.
- Hogg, H.E., Price, D.T., Black, T.A., 2000. Postulated feedbacks of deciduous forest phenology on seasonal climate patterns in the western Canadian interior. *J. Clim.* 13, 4229–4243.
- Houghton, R.A., Hackler, J.L., 2001. Carbon Flux to the Atmosphere from Land-Use Changes: 1850 to 1990. NDP-050/R1, Carbon Dioxide Information Analysis Center. Oak Ridge National Laboratory, Oak Ridge.
- Jackson, R.B., Canadell, J., Ehleringer, J.R., 1996. A global analysis of root distributions for terrestrial biomes. *Oecologia* 108, 389–411.
- Kim, Y., Wang, G., 2005. Modeling seasonal vegetation variation and its validation against Moderate Resolution Imaging Spectroradiometer (MODIS) observations over North America. *J. Geophys. Res.* 110, D04106, doi:10.1029/2004JD005436.
- Kucharik, C.J., Brye, K.R., Norman, J.M., Foley, J.A., Gower, S.T., Bundy, L.G., 2001. Measurements and modeling of carbon and nitrogen cycling in agroecosystems of southern Wisconsin: Potential for SOC sequestration during the next 50 years. *Ecosystems* 4, 237–258.
- Kucharik, C.J., Foley, J.A., Delire, C., Fisher, V.A., Coe, M.T., Gower, S.T., Lenters, J., Molling, C., Norman, J.M., Ramankutty, N., 2000. Testing the performance of a dynamic global ecosystem model: water balance, carbon balance, and vegetation structure. *Global Biogeochem. Cycles* 14 (3), 795–825.
- Landsberg, J.J., Gower, S.T., 1997. Applications of Physiological Ecology to Forest Management. Academic Press, San Diego, CA, 354 pp.
- Lenters, J.D., Coe, M.T., Foley, J.A., 2000. Surface water balance of the continental United States, 1963–1995: regional evaluation of a terrestrial biosphere model and the NCEP/NCAR reanalysis. *J. Geophys. Res. (Atmos.)* 105 (D17), 22,393–22,425.
- Levis, S., Foley, J.A., Brovkin, V., Pollard, D., 1999. On the stability of the high-latitude climate–vegetation system in a coupled atmosphere–biosphere model. *Global Ecol. Biogeogr.* 8 (6), 489–500.
- Linn, D.M., Doran, J.W., 1984. Effect of water-filled pore space on carbon dioxide and nitrous oxide production in tilled and non tilled soils. *Soil Sci. Soc. Am. J.* 48, 1267–1272.
- Lloyd, J., Taylor, J.A., 1994. On the temperature dependence of soil respiration. *Funct. Ecol.* 8, 315–323.
- Lucht, W., Prentice, I.C., Myneni, R.B., Stith, S., Friedlingstein, P., Cramer, W., Bousquet, P., Buermann, W., Smith, B., 2002. Climatic control of high-latitude vegetation greening trend and Pinatubo effect. *Science* 296, 1687–1689.
- Lüdeke, M.K.B., Ramge, P.H., Kohlmaier, G.H., 1996. The use of satellite NDVI data for the validation of global vegetation phenology models: application to the Frankfurt Biosphere Model. *Ecol. Modell.* 91, 255–270.
- Luxmoore, R.J., Huff, D.D., 1989. Water. In: Johnson, D.W., Van Hook, R.I. (Eds.), *Analysis of Biogeochemical Cycling Processes in Walker Branch Watershed*. Springer, New York, pp. 164–196.
- McCloy, K.R., Lucht, W., 2004. Comparative evaluation of seasonal patterns in long time series of satellite image data and simulations of a global vegetation model. *IEEE Trans. Geosci. Remote Sens.* 42 (1), 140–153.
- Miller, D.A. White, R.A., 1998. A conterminous United States multilayer soil characteristics dataset for regional climate and hydrology modelling. *Earth Interactions* 2, paper 2.
- Monson, R.K., Turnipseed, A.A., Sparks, J.P., Harley, P.C., Scott-Denton, L.E., Sparks, K., Huxman, T.E., 2002. Carbon sequestration in a high-elevation, subalpine forest. *Global Change Biol.* 8, 459–478.
- Moorcroft, P.R., 2003. Recent advances in ecosystem–atmosphere interactions: an ecological perspective. *Proc. R. Soc. Lond. B* 270, 1215–1227.
- Neilson, R.P., Pitelka, L.F., Solomon, A.M., Nathan, R., Midgley, G.F., Fragoso, J.M.F., Lischke, H., Thompson, K., 2005. Forecasting regional to global plant migration in response to climate change. *Bioscience* 55, 749–759.
- Parton, W.J., Schimel, D.S., Cole, C.V., 1987. Analysis of factors controlling soil organic matter levels in Great Plains grassland. *Soil Sci. Soc. Am. J.* 51, 1173–1179.
- Pielke, R.A., Marland, G., Betts, R.A., Chase, T.N., Eastman, J.L., Niles, J.O., Niyogi, D.D.S., Running, S.W., 2002. The influence of land-use change and landscape dynamics on the climate system: relevance to climate-change policy beyond the radiative effect of greenhouse gases. *Philos. Trans. R. Soc. Lond. Ser. A: Math. Phys. Eng. Sci.* 360 (1797), 1705–1719.
- Pollard, D., Thompson, S.L., 1995. Use of a land-surface-transfer scheme (LSX) in a global climate model: the response to doubling stomatal resistance. *Global Planet. Change* 10, 129–161.
- Ryan, M.G., 1991. A simple method for estimating gross carbon budgets for vegetation in forest ecosystems. *Tree Physiol.* 9, 255–266.
- Sanderman, J., Amundson, R.G., Baldocchi, D.D., 2003. Application of eddy covariance measurements to the temperature dependence of soil organic matter mean residence time. *Global Biogeochem. Cycles* 17 (2), 1061, doi:10.1029/2001GB001833.
- Savage, K.E., Davidson, E.A., 2001. Interannual variation of soil respiration in two New England forests. *Global Biogeochem. Cycles* 15 (2), 337–350.

- Scott-Denton, L.E., Sparks, K.L., Monson, R.K., 2003. Spatial and temporal controls of soil respiration rate in a high-elevation, subalpine forest. *Soil Biol. Biochem.* 35, 525–534.
- Sitch, S., Smith, B., Prentice, I.C., Arneeth, A., Bondeau, A., Cramer, W., Kaplan, J.O., Levis, S., Lucht, W., Sykes, M.T., Thonicke, K., Venevsky, S., 2003. Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biol.* 9, 161–185.
- Smith, B., Prentice, I.C., Sykes, M.T., 2001. Representation of vegetation dynamics in the modeling of terrestrial ecosystems: comparing two contrasting approaches within European climate space. *Global Ecol. Biogeogr.* 10, 621–637.
- Vano, J.A., Foley, J.A., Kucharik, C.J., Coe, M.T. Evaluating the seasonal and interannual variations in water balance in northern Wisconsin, USA, using a land surface model. *J. Geophys. Res.: Biogeosci.*, submitted for publication.
- Verberne, E.L., Hassink, J., De Willigen, P., Groot, J.J.R., Van Veen, J.A., 1990. Modeling organic matter dynamics in different soils. *Netherlands J. Agric. Sci.* 38, 221–238.
- Waring, R.H., Landsberg, J.J., Williams, M., 1998. Net primary production of forests: a constant fraction of gross primary production? *Tree Physiol.* 18, 129–134.
- White, M.A., Thornton, P.E., Running, S.W., 1997. A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochem. Cycles* 11 (2), 217–234.
- White, M.A., Running, S.W., Thornton, P.E., 1999. The impact of growing-season length variability on carbon assimilation and evapotranspiration over 88 years in the eastern US deciduous forest. *Int. J. Biometeorol.* 42, 139–145.
- White, M.A., Nemani, R.R., 2003. Canopy duration has little influence on annual carbon storage in the deciduous broad leaf forest. *Global Change Biol.* 9, 967–972.
- Williams, M., Rastetter, E.B., Fernandes, D., Goulden, M.L., Wofsy, S.C., Shaver, G.R., Melillo, J.M., Munger, J.W., Fan, S.-M., Nadelhoffer, K.J., 1996. Modeling the soil–plant atmosphere continuum in a *Quercus-Acer* stand at Harvard Forest: the regulation of stomatal conductance by light nitrogen and soil/plant hydraulic properties. *Plant Cell Environ.* 19, 911–927.
- Wilson, K.B., Baldocchi, D.D., Hanson, P.J., 2000a. Spatial and seasonal variability of photosynthetic parameters and their relationship to leaf nitrogen in a deciduous forest. *Tree Physiol.* 20, 565–578.
- Wilson, K.B., Hanson, P.J., Baldocchi, D.D., 2000b. Factors controlling evaporation and energy partitioning beneath a deciduous forest over an annual cycle. *Agric. Forest Meteorol.* 102, 83–103.
- Wilson, K.B., Baldocchi, D.D., 2001. Comparing independent estimates of carbon dioxide exchange over five years at a deciduous forest in the southern United States. *J. Geophys. Res.* 106, 34167.
- Wilson, K.B., Hanson, P.J., Mulholland, P., Baldocchi, D.D., Wullschleger, S., 2001. A comparison of methods for determining forest evapotranspiration and its components: sap flow, soil water budget, eddy covariance and catchment water balance. *Agric. Forest Meteorol.* 106, 153–168.
- Wilson, K.B., Goldstein, A.H., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P., Bernhofer, Ch., Ceulemans, R., Dolman, H., Field, C., Grelle, A., Law, B., Meyers, T., Moncrieff, J., Monson, R., Oechel, W., Tenhunen, J., Valentini, R., Verma, S., 2002. Energy balance closure at FLUXNET sites. *Agric. Forest Meteorol.* 113, 223–243.
- Wofsy, S.C., Goulden, M.L., Munger, J.W., Fan, S.-M., Bakwin, P.S., Daube, B.C., Bassow, S.L., Bazzaz, F.A., 1993. Net exchange of CO₂ in a mid-latitude forest. *Science* 260, 1314–1317.
- Woodward, F.I., Lomas, M.R., Betts, R.A., 1998. Vegetation-climate feedbacks in a greenhouse world. *Philos. Trans. R. Soc. Lond. Ser. B* 353, 29–38.
- Wullschleger, S.D., 1993. Biochemical limitations to carbon assimilation in C(3) plants—a retrospective analysis of the A/CI curves from 109 species. *J. Exp. Bot.* 44, 907–920.