OAK FOREST CARBON AND WATER SIMULATIONS: MODEL INTERCOMPARISONS AND EVALUATIONS AGAINST INDEPENDENT DATA

P. J. Hanson,[1,15] J. S. Amthor,[1,14] S. D. Wullschleger,[1] K. B. Wilson,[2] R. F. Grant,[3] A. Hartley,[4]
D. Hui,[5] E. R. Hunt, Jr.,[6] D. W. Johnson,[7] J. S. Kimball,[8] A. W. King,[1] Y. Luo,[5] S. G. McNulty,[9]
G. Sun,[9] P. E. Thornton,[10] S. Wang,[11] M. Williams,[12] D. D. Baldocchi,[13] and R. M. Cushman[1]

[1]*Environmental Sciences Division, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, Tennessee 37831-6422 USA*
[2]*Atmospheric Turbulence Diffusion Division, National Oceanic Atmospheric Administration (NOAA),*
Oak Ridge, Tennessee 37840 USA
[3]*University of Alberta, Edmonton, Alberta, Canada T6G 2E3*
[4]*Florida International University, Department of Environmental Studies, Miami, Florida 33199 USA*
[5]*Department of Botany and Microbiology, University of Oklahoma, Norman, Oklahoma 73019 USA*
[6]*USDA Agricultural Research Service, Hydrology and Remote Sensing Laboratory, Beltsville, Maryland 20705 USA*
[7]*Environmental and Resource Sciences, University of Nevada, Reno, Nevada 89512-0013 USA*
[8]*Flathead Lake Biological Station and School of Forestry, University of Montana, Polson, Montana 59860 USA*
[9]*Southern Global Change Program, USDA Forest Service, Raleigh, North Carolina 27606 USA*
[10]*National Center for Atmospheric Research, Boulder, Colorado 80503 USA*
[11]*Canada Centre for Remote Sensing, Ottawa, Ontario, Canada K1A 0Y7*
[12]*School of GeoSciences, University of Edinburgh, Edinburgh EH9 3JU, UK*
[13]*Ecosystem Science Division, Department of Environmental Science, Policy Management, University of California,*
Berkeley, Berkeley, California 94720 USA

Abstract. Models represent our primary method for integration of small-scale, process-level phenomena into a comprehensive description of forest-stand or ecosystem function. They also represent a key method for testing hypotheses about the response of forest ecosystems to multiple changing environmental conditions. This paper describes the evaluation of 13 stand-level models varying in their spatial, mechanistic, and temporal complexity for their ability to capture intra- and interannual components of the water and carbon cycle for an upland, oak-dominated forest of eastern Tennessee. Comparisons between model simulations and observations were conducted for hourly, daily, and annual time steps. Data for the comparisons were obtained from a wide range of methods including: eddy covariance, sapflow, chamber-based soil respiration, biometric estimates of stand-level net primary production and growth, and soil water content by time or frequency domain reflectometry. Response surfaces of carbon and water flux as a function of environmental drivers, and a variety of goodness-of-fit statistics (bias, absolute bias, and model efficiency) were used to judge model performance.

A single model did not consistently perform the best at all time steps or for all variables considered. Intermodel comparisons showed good agreement for water cycle fluxes, but considerable disagreement among models for predicted carbon fluxes. The mean of all model outputs, however, was nearly always the best fit to the observations. Not surprisingly, models missing key forest components or processes, such as roots or modeled soil water content, were unable to provide accurate predictions of ecosystem responses to short-term drought phenomenon. Nevertheless, an inability to correctly capture short-term physiological processes under drought was not necessarily an indicator of poor annual water and carbon budget simulations. This is possible because droughts in the subject ecosystem were of short duration and therefore had a small cumulative impact. Models using hourly time steps and detailed mechanistic processes, and having a realistic spatial representation of the forest ecosystem provided the best predictions of observed data. Predictive ability of all models deteriorated under drought conditions, suggesting that further work is needed to evaluate and improve ecosystem model performance under unusual conditions, such as drought, that are a common focus of environmental change discussions.

Key words: autotrophic respiration; carbon budget; computer models; evaporation; evapotranspiration; NEE; NPP; transpiration; water budget.

INTRODUCTION

The Intergovernmental Panel on Climate Change (IPCC) concluded that continued increases in greenhouse gas concentrations will result in temperature increases from 1.4° to 5.8°C during the 21st century which are expected to modify the global hydrologic cycle, resulting in increased winter precipitation at high latitudes; more hot days and fewer cold days; and changes in the frequency of droughts and floods (Houghton et al. 2001). Such simulated changes raise concerns about terrestrial ecosystem productivity, biogeochemical cycling, and the availability of water resources (Kirschbaum and Fischlin 1996, Melillo et al. 1990). Furthermore, such concerns are amplified because ecological change is expected to occur much faster than species adaptation or migration. Models represent our primary method for integration of small-scale, process-level phenomena into a cohesive description of forest-stand or ecosystem function. They also represent the only viable means for testing hypotheses about regional forest ecosystem response to changing environmental conditions.

Although models must be used to assess the impact of climatic change on forest ecosystem processes (Hunt et al. 1996, McNulty et al. 1998, Jenkins et al. 2000, Aber et al. 2001, Bachelet et al. 2001, Malcolm et al. 2002), few have been rigorously tested against the same comprehensive and long-term measurements. Aber (1997) emphasized that prior to the application of a given model for the purpose of prediction, appropriate documentation of the model structure, parameterization process, and testing of model output against independent data (i.e., validation) must be accomplished. A number of model tests and/or intercomparisons have been conducted (VEMAP Members 1995, Ryan et al. 1996a, b, Landau et al. 1998, Weinstein et al. 1998, Homann et al. 2000). This paper represents an intercomparison for upland oak forests following the framework established by Amthor et al. (2001).

Oak forests comprise a large fraction of the forested land area of the eastern deciduous forest of the United States (Powell et al. 1992). In addition to their extensive land cover, these forests are among the most highly productive ecosystems (Whittaker and Likens 1975), they store substantial amounts of carbon (Fan et al. 1998, Canadell et al. 2000), they are a key source of clean water (Satterlund 1972, Gleick and Adams 2000), and they support a high degree of biodiversity (Whittaker 1956, Iverson and Prasad 2001). The importance of temperate deciduous forests to society is increased by their proximity to large population centers of eastern North America. Given the anticipated rate of climatic change and the importance of the upland-oak component of the eastern deciduous forest, good models must be available for evaluating the potential impacts of future environmental change.

This paper evaluates the efficacy of 13 stand-level forest ecosystem models for predicting the carbon and water budgets of an upland-oak forest in eastern Tennessee under current climate conditions. The models vary substantially in their spatial, mechanistic, and temporal complexity. Questions being addressed in the current intercomparison include:

1) How well do models compare to one another and independent measurements?

2) Are the results from general models different from those of site-calibrated models?

3) Can model projections set logical bounds on ecosystem responses to environmental change appropriate to policy questions?

Independent field data used to judge the adequacy of predictions were derived from the Walker Branch Throughfall Displacement Experiment (TDE; Hanson et al. 2001a, b, 2003a) and AmeriFlux site (Wilson and Baldocchi 2001) for the period from 1993 through 2000. This 8-year period includes a dynamic range of intra- and interannual precipitation patterns and hydrologic conditions that provide a stringent test of the versatility of ecosystem models.

DESCRIPTION OF MODELS

A total of 13 models were used in this study covering a range of temporal scales, spatial complexity, and mechanistic detail (Fig. 1). The participating models represented modeling groups able and willing to respond to an open invitation. Eight models used an hourly time step, four models used a daily time step, and only one model used a monthly time step (i.e., PnET-II). Most models provided estimates for both stand carbon and water cycles, but others were limited to either the carbon (MAESTRA) or water (LINKAGES and NuCM) cycle. The models with greatest mechanistic complexity (*ecosys*, EALCO, LaRS, and CANOAK) also used a complete energy balance. A brief description of each of the models is provided below, and details of the structural and physiological characteristics of each model are summarized in Tables 1 and 2, respectively. Model descriptions along with archived environmental data, site characterization information, and original model results can be found online at the TDE model intercomparison web site.[16]

BIOME-BGC (John S. Kimball and Peter E. Thornton)

BIOME-BGC (v 4.1.1 used in this study) is a general terrestrial ecosystem model designed to simulate the coupling of carbon, nitrogen, and water cycles in the plant–litter–soil system for both woody and herbaceous vegetation (Thornton et al. 2002). BIOME-BGC is a daily time-step model with a single homogeneous canopy layer divided into sun and shade leaves, but mixed-type simulations are possible (e.g., Law et al. 2001b).

---

[16] ⟨http://cdiac.ornl.gov/epubs/tdemodel/tdemodel.html⟩
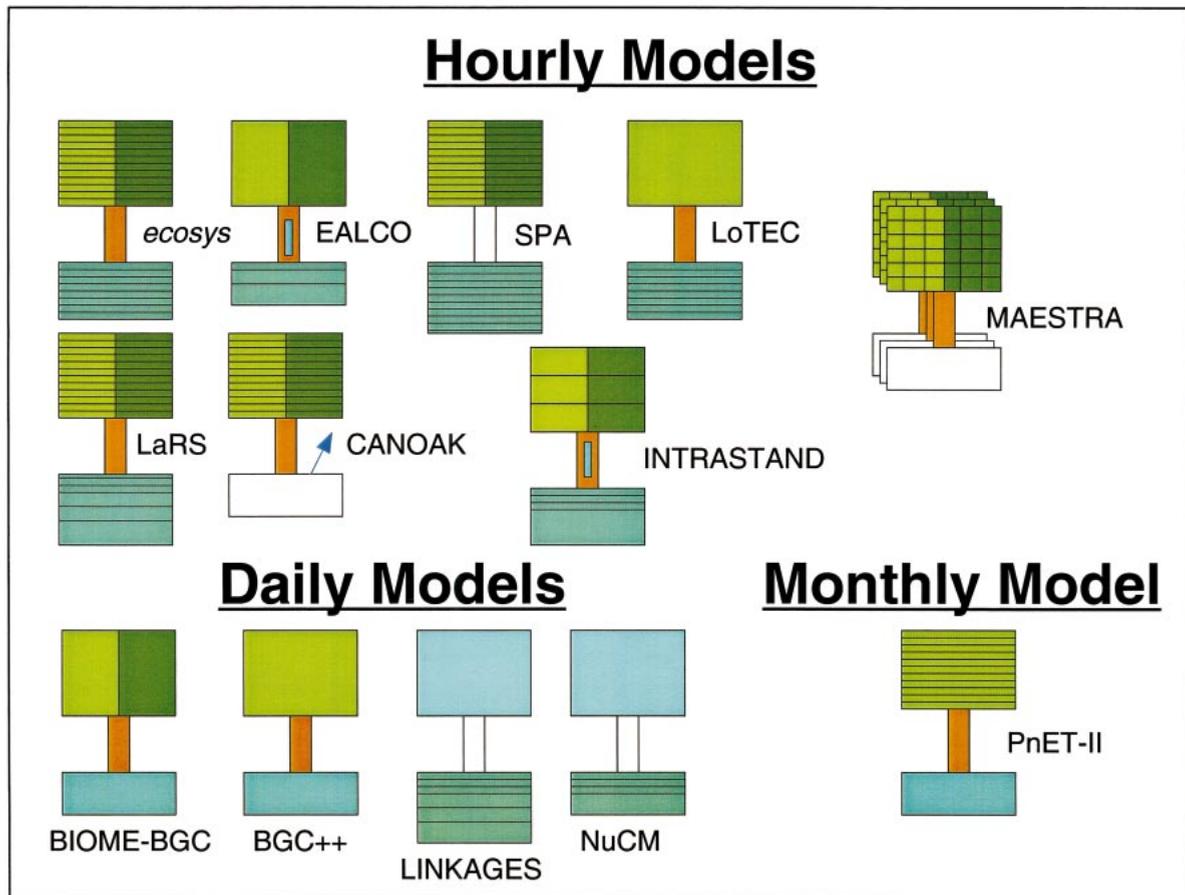
Fig. 1. Schematic diagram of the 13 models used, including eight hourly models, four daily models, and one monthly time-step model. The shaded portions of each diagram indicate corresponding functions in the respective models. Light and dark shading of the canopy indicates that sun and shade leaves are simulated. Horizontal lines in the canopy or belowground portions show the level of vertical detail. All models except one (MAESTRA) are one dimensional.

The model has a single soil water layer. The model is driven by daily surface weather (temperature, precipitation, radiation, and humidity), and requires a set (34) of constants describing each plant functional type (White et al. 2000). All plant, litter, and soil carbon, nitrogen, and water pools and fluxes are entirely prognostic. This is in contrast to models such as Forest-BGC (Running and Coughlan 1988) and CANOAK (Baldocchi and Harley 1995) with prescribed canopy and other biomass pools. Canopy carbon and water dynamics are treated separately for sunlit and shaded fractions, with the two fractions having the same mass-based leaf nitrogen concentration and different specific-leaf-area, a realistic assumption for these forests (Wilson et al. 2000b). Allocation to new growth depends on the availability of both assimilated carbon and soil mineral nitrogen (augmented by an internal pool of retranslocated foliar nitrogen). Plants compete with soil heterotrophs for a single pool of soil mineral nitrogen, with down-regulation of both carbon assimilation and nitrogen immobilizing steps in the trophic processing of litter and soil organic matter when nitrogen is limiting. Individual plant tissues have static C:N ratios, but whole-plant C:N changes as the relative amounts of different tissues change over time (e.g., as wood accumulates relative to foliage during forest stand development). Leaf area phenology for temperate deciduous systems is determined by the model of White et al. (1997).

The model is designed with special attention to the long-term controls on net ecosystem carbon exchange (NEE), and an important aspect of this design is the use of spin-up and perturbation simulations, the purpose of which are to bring model state variables to a configuration that is both internally consistent and consistent with known patterns of land use and disturbance history (Law et al. 2001b, Thornton et al. 2002). The spin-up simulation begins with minimal soil organic matter (SOM) content and a nascent canopy, and proceeds until the SOM and plant pools have reached a steady state with respect to a repeated sequence of surface weather drivers. A second step moves this steady state from preindustrial to current conditions for atmospheric CO_2 concentration and mineral nitrogen de-

TABLE 1.   Structural characteristics of the 13 models.

Structural parameters	BGC++	BIOME-BGC	CANOAK	EALCO	ecosys	INTRASTAND
Time step	d/yr	d	h	0.5 h	h	h
Carbon cycle	yes	yes	yes	yes	yes	yes
Water cycle	yes	yes	yes	yes	yes	yes
Energy cycle	partial	partial	yes	yes	yes	no
Nutrient cycle(s)	nitrogen	yes	no	yes	yes	no
Ozone effects	no	no	no	no	no	yes
NMHC	no	no	no	no	no	yes
Turbulent transport	no	no	yes	yes	yes	no
Initial conditions	obs.	spun up	obs.	obs.	spun up	obs.
Biomass pools	4	9	NA	5	8	3
Canopy layers	1	1	30	1	15	3
Sun vs. shade	no	yes	yes	yes	yes	yes
Soil layers						
Water	1	1	0	3	15	4
Energy	partial	1	5–10	3	15	NA
Leaf phenology	obs.	pred.	obs.	pred.	pred.	obs.
Leaf growth	pred.	pred.	obs.	pred.	pred.	obs.
Stem growth	pred.	pred.	obs.	pred.	pred.	obs.
Root growth	pred.	pred.	obs.	pred.	pred.	obs.
Soil temperature	pred.	pred.	pred.	pred.	pred.	obs.
Litter C pools	1	3	no	3	4	1

Notes: GPP = gross primary production; NA = not applicable; NMHC = nonmethane hydrocarbons; obs. = observed patterns; pred. = predicted patterns; spun up = models run for a period of years to set initial conditions; h = hour; d/yr = day/year; mo = month.

position. Overlaid on these simulations are instantaneous modifications to the prognostic model state variables to represent the known historical sequence of land use and/or disturbance history (e.g., removal of mass from plant carbon and nitrogen pools to simulate different levels of harvest or fire). This protocol avoids the inevitable transient responses in NEE that result from forcing a simulation to start with observed plant and soil state variables, since these observed states would never be entirely consistent with the internal dynamics of any particular model.

BGC++ (E. Raymond Hunt, Jr.)

The BGC++ model (Hunt et al. 1999) is a daily time-step model with a homogeneous big-leaf canopy and a single soil water layer. The model calculates fluxes and pools for the carbon, water, nutrient, and a partial energy cycle for different terrestrial ecosystems. It was derived from BIOME-BGC (for BioGeochemical Cycles; Running and Hunt 1993, Hunt et al. 1996), for the purpose of simulating allocation and growth over long time periods under different climatic conditions. Similar to BIOME-BGC and its predecessor, the FOR-EST-BGC model (Running and Coughlan 1988, Running and Gower 1991), BGC++ uses a dual daily and annual time step, where the hydrologic, photosynthesis, and respiration processes are simulated daily and carbon and nitrogen allocation are simulated annually. However, BGC++ differs from BIOME-BGC in three ways: (1) BGC++ simulates the fluxes of dominant and subdominant life forms separately, then combines the fluxes assuming a simple mixture of leaf area index (LAI), (2) allocation of carbon and nitrogen to the leaves, stems, coarse roots, and fine roots is based on

the ratio of demand to availability of nitrogen; and (3) BGC++ uses a new algorithm for estimating incident solar radiation (Winslow et al. 2001).

The key variable for the daily time step is LAI (Hunt et al. 1996), which was set from the data. The three climatic variables used to drive the model are daily maximum and minimum temperature, and daily precipitation. All other meteorological variables (e.g., solar radiation and VPD) were calculated from these simple inputs. Transpiration is based on the Penman-Monteith equation using absorbed solar radiation. Soil evaporation is based on the transmitted solar radiation. Intercepted water evaporated from the canopy is linearly related to LAI. Using the model of Zheng et al. (1993), LAI strongly controls soil temperature over the year, which in turn controls the rates of root maintenance respiration, litter decomposition, and turnover of soil organic matter. For predictions regarding the TDE, phenology of the leaves and fine roots was handled with constants derived from site observations (Hanson et al. 2003c, Joslin and Wolfe 2003).

During the annual time step, plant-available nitrogen is calculated from litter decomposition, turnover of soil organic matter, internal nitrogen from retranslocation, and nitrogen losses with water runoff (Hunt et al. 1999). Plant available carbon is the net primary production, and is allocated to new leaves and fine roots based on a soil water index and a soil nitrogen index, such that with high levels of soil moisture and nitrogen, there is preferential allocation to the leaves. Excess plant carbon and nitrogen are allocated to stems and coarse roots based on life form, 85% and 15% are allocated to tree stems and grass life forms, respectively. Nitrogen is allocated to leaves independently of carbon,

TABLE 1. Extended.

LaRS	LINKAGES	LoTEC	MAESTRA	NuCM	PnET-II	SPA
h	d	h	h	d	mo	0.5 h
yes	no	yes	yes	no	yes	GPP only
yes	yes	yes	partial	yes	yes	yes
yes	no	no	partial	no	no	yes
no	yes	no	no	yes	no	no
no	no	no	no	no	no	no
yes	no	no	no	no	no	no
yes	no	no	no	no	no	no
spun up	spun up	obs.	obs.	obs.	spun up	obs.
4	2	4	2	3	3	NA
20	1	1	10	NA	50	10
yes	no	no	yes	no	no	yes
15	12	14	NA	10	1	20
15	NA	NA	NA	NA	NA	20
pred.	obs.	obs.	obs.	obs.	pred.	obs.
pred.	pred.	obs.	obs.	obs.	pred.	obs.
pred.	pred.	pred.	NA	obs.	pred.	obs.
pred.	NA	pred.	NA	obs.	pred.	obs.
pred.	no	obs.	NA	obs.	NA	pred.
0	1	2	NA	NA	0	NA

so the leaf-nitrogen to carbon ratio is a model predic-
tion, and is used to determine the V_{max} of Rubisco and
J_{max} of the RuBP regeneration for photosynthesis.

CANOAK (Kell B. Wilson)

CANOAK (Version 1 modified) is a one-dimension-
al, multilayer biosphere–atmosphere model that com-
putes water vapor, CO_2 and sensible heat flux densities
at an hourly time step. The model has been described
and tested for growing season conditions (Baldocchi
and Harley 1995, Baldocchi 1997), applied to a 20-
year climate record (Baldocchi and Wilson 2001), and
to a discussion of environmental and structural impacts
on canopy carbon, water, and energy flux (Baldocchi
et al. 2002). The model consists of coupled micro-
meteorological and ecophysiological modules. The mi-
crometeorological modules compute leaf (sunlit and
shaded) and soil energy exchange, turbulent (Lagrang-
ian) diffusion, scalar concentration profiles, and radi-
ative transfer through the canopy using observed me-
teorological conditions above the canopy. The physi-
ological modules are driven by physiological param-
eters that are obtained directly from extensive chamber
measurements performed in the field. The predicted
micrometeorology drives leaf photosynthesis and res-
piration, stomatal conductance, and transpiration at 40
canopy layers. Canopy leaf area profiles were assumed
to follow a beta distribution, with a heavier concen-
tration of leaves in the upper canopy (Hutchinson and
Baldocchi 1989). NEE in CANOAK is obtained by
summing each component of the carbon flux: bole and
soil/root respiration, and leaf photosynthesis and res-
piration.

CANOAK does not explicitly simulate soil water
content dynamics. As a result, drought impacts on mod-
eled physiological processes are only affected by
changing atmospheric vapor pressure deficit.

EALCO (Shusen Wang)

The EALCO model was developed to simulate the
ecosystem radiation, water, carbon, and nitrogen pro-
cesses of various land cover types using remote sensing
and GIS data inputs. EALCO is an hourly time-step
model with a single canopy layer divided into sun and
shade leaves and a multilayered soil. The model em-
phasizes the interactions between the ecosystem water
and carbon dynamics by including a water balance
equation (Wang et al. 2002a) and the photosynthesis-
based canopy conductance algorithm (Ball et al. 1987).
The plant carbon and nitrogen calculations are modified
from Wang et al. (2001). It includes one substrate car-
bon/nitrogen pool and three structural carbon/nitrogen
pools representing the three physiologically active
parts of the plant, i.e., foliage; sapwood of branch,
stem, and coarse root; and fine roots. The soil carbon
and nitrogen calculations include the transformations
of both litterfall and soil organic matter in/on the soil
(Wang et al. 2002b). Plant litterfall and soil organic
matter are separated into different pools based on their
biochemical resistance to microbial decomposition: lit-
terfall (water-extractable, cellulose and hemicellulose,
and lignin) and soil organic matter (active, slow, and
humus). A separate microbial biomass pool and a min-
eral N pool are also represented in each soil layer.
While litterfall obtained in the plant simulations pro-
vides the carbon and nitrogen input to the soil, the
organic matter turnover in/on the soil controls the min-
eral nitrogen release rates and its availability for plant
uptake, which determines the plant nitrogen conditions

TABLE 2. Physiological characteristics of the 13 models (NA = not applicable or available).

Physiological parameters	BGC++	BIOME-BGC	CANOAK	EALCO
Photosynthesis†	Farquhar	Farquhar	Farquhar	Farquhar
Leaf conductance (g)	Multiplier × g_{max}‡	Multiplier × g_{max}‡	Ball-Berry§	Ball-Berry§
Leaf maintenance respiration‖	$f(M, T)$	$f(M, N, T)$	$f(M, N, T)$	$f(N, T)$
Stem maintenance respiration¶	$f(T, W)$	$f(T, N, W)$	$f(T)$	$f(T, N)$
Growth cost (g C cost/g C built)				
Leaves	0.3	0.3	NA	0.29
Stems	0.2	0.3	NA	0.29
Roots	0.25	0.3	NA	0.29
Coarse roots	0.2	0.3	NA	0.29
Carbohydrate feedbacks	no	yes	no	no
Soil N/plant C feedbacks	yes	yes	no	yes
ET approach	Pen/Mon††	Pen/Mon††	energy balance	energy balance
Canopy interception	yes	yes	no	yes
Litter evaporation	no	no	no	no
Soil evaporation	yes	yes	no	yes
Hydraulic lift	no	no	no	off
Stem capacitance	no	no	no	yes
Soil organic pools	1	4	NA	3
Litter pools	1	3	NA	3
Soil/heterotrophic respiration#	$f(T, SW, L/N, C)$	$f(T, SW, C)$	$f(T, SWP)$	$f(T, SWP, C)$
Rate dynamics	first order	first order	first order	first order
Coarse wood decomposition	yes	yes	no	no

† Farquhar et al. (1980) and Farquhar and von Caemerer (1982).

‡ Maximum stomatal conductance in modified by a series of multipliers from 0 to 1. BGC++ = $f(T, VPD, PAR, CO_2,$ SWP, minimum night temp.); BIOME-BGC = $f(T, VPD, PAR, CO_2, SWP)$; LaRS = $f(T, VPD, PAR, SWP)$; LINKAGES = $f(T, VPD, PAR,$ ext. soil water). T = temperature.

§ Ball et al. (1987).

‖ Leaf respiration is a function of the following: M = leaf mass, T = temperature, N = nitrogen, abs PAR = absorbed photosynthetically active radiation.

¶ Stem respiration is a function of the following: T = temperature, W = wood or sapwood mass. For BGC++, stem respiration per cubic meter assumes a wood density of 400 kg dry matter/m³.

Soil respiration is a function of the following: T = temperature, SW = soil water content, SWP = soil water potential, L/N = lignin/nitrogen ratio, C = carbon, O = oxygen.

†† Penman-Monteith formula (e.g., Jarvis and McNaughton 1986).

and affects the plant CO_2 fixation and thereafter the ecosystem water balance.

ecosys (Robert F. Grant)

The *ecosys* model is an hourly time-step model with multiple canopy layers separated into sun and shade-leaf components and with a multilayered soil. Carbon, water, nutrient, and energy cycles are simulated. The *ecosys* model is designed to represent terrestrial ecosystems subject to a range of management practices (e.g., fertilization, tillage, irrigation, planting, harvesting, thinning) and environmental changes (e.g., atmospheric CO_2, temperature, precipitation, water quality) at patch (one-dimensional) and landscape (two- or three-dimensional) scales. This model simulates CO_2 fixation, N and P uptake, plant growth, autotrophic respiration, litter production, microbial growth, heterotrophic respiration, and soil C, N, and P transformations. All reactants and products undergo convective–dispersive transport in gaseous and aqueous phases driven by soil water and energy transfers coupled to first-order closure schemes for energy exchange at plant, residue, soil, and snow surfaces. The *ecosys* model represents a user-selected number of canopy and soil layers (a maximum of 15 for each) that provides a framework

for a user-selected number of plant species from functional types within groupings of trees, crops, forages, rangeland grasses, and mosses). These species compete for light, water, and nutrients (N and P) based on vertical distributions of leaves and roots. For this study, a temperate, deciduous overstory was selected with 10 canopy layers, 9 soil layers, and a one-dimensional formulation. The *ecosys* model used site data for soil and weather attributes, but all other state variables were generated internally during a 60-year spin-up, thereby minimizing use of prescribed site data. A parameter used to set timing of leaf-out was adjusted to site conditions. Further details are in Grant (2001).

INTRASTAND (Paul J. Hanson)

INTRASTAND is an hourly time-step model designed for use in the interpolation of measured physiological data over time for the calculation of daily and intraannual forest stand carbon and water budgets. The model structure contains three canopy foliage layers, branch and bole stem components, four soil layers, and stem capacitance. By design, the structural detail is limited so as not to exceed the availability of measured input data. Carbon uptake is based on the coupled Farquhar/Ball-Berry photosynthetic and stomatal conduc-

TABLE 2. Extended.

ecosys	INTRASTAND	LaRS	LINKAGES	LoTEC
Farquhar	Approximate Farquhar	Biophysics and Biochem.	NA	Farquhar
Ci:Ca ratio	Ball-Berry	Multiplier $\times\ g_{max}$‡	Multiplier $\times\ g_{max}$‡	Ball-Berry
$f(N, T)$	$f(M, N, T)$	$f(M, N, T,$ abs PAR$)$	NA	$f(M, N, T)$
$f(N, T)$	$f(T, W)$	$f(T, W)$	NA	$f(T, N, W)$
0.49	0.28	$f(N)$	NA	0.22
0.32	0.22	$f(N)$	NA	0.28
0.49	0.23	$f(N)$	NA	0.22
NA	NA	$f(N)$	NA	0.28
yes	no	yes	no	yes
yes	no	no	yes	no
energy balance	leaf/litter	leaf/soil	Pen/Mon††	big leaf
yes	yes	yes	yes	yes
yes	yes	no	no	no
yes	no	yes	no	no
yes	no	yes	no	no
yes	yes	no	no	no
3	NA	NA	NA	two active; one passive
4	1	0	NA	2
$f(T, SW, C, O)$	$f(T, SWP)$	$f(T, SW)$	$f(T, SW)$	$f(T, SWD, C)$
microbial kinetics	NA	NA	first order	first order
no	no	yes	NA	NA

tance model as described by Harley et al. (1992) and parameterized according to data in Wilson et al. (2001a), a modified version of the water budget model PROSPER (Huff et al. 1977) with the inclusion of a stem capacitance, a model of stem respiration (Edwards and Hanson 1996), and a model of forest floor CO_2 efflux (Hanson et al. 1993, 2003b). The canopy is divided into three layers of equal LAI, and diffuse and direct light penetration is calculated as suggested by Norman (1982). The modeled canopy net foliar assimilation rates were calibrated to yield maximum assimilation in early June following full leaf expansion in accordance with reported values for the upland oak forest on Walker Branch watershed (Verma et al. 1986, Harley and Baldocchi 1995, Baldocchi and Vogel 1996, Baldocchi 1997). Isoprene emission was also estimated for this stand (a small component of carbon flux) based on the observations and models of Harley et al. (1997). The model is coded using "Stella" modeling software (High Performance Systems, Hanover, New Hampshire, USA). The combined mechanistic model was developed for application to intra-annual carbon and water budgeting but has been used here over a multiyear period by transferring soil water, litter mass, and stored carbohydrate data to the initial conditions of model runs for subsequent years. Growth of stems, roots, and leaves, the timing of leaf-out and leaf senescence, climate variables, and required physiological variables are inputs from direct measurements on the study site. Respiratory costs of growing tissues were calculated by a modified Penning de Vries approach as outlined by Amthor (1996). Published results for the water budget and carbon flux components can be found in Edwards and Hanson (1996), Hanson et al. (1998, 2001a, 2003b,

2003d) and Johnson et al. (2002). Results presented in this manuscript represent the first comprehensive test of its utility against a multiyear data set and other models.

LaRS (Jeffrey S. Amthor)

LaRS is an hourly time-step model designed to simulate tree physiology and growth within the context of a forest ecosystem. LaRS simulates a multilayered canopy divided into sun and shade leaves and includes a multilayered soil. It calculates pools and fluxes for a carbon, water, and energy cycle. Solar radiation–forest interactions are simulated with a ray-tracing procedure that accounts for diffuse solar (sky) radiation arising from 10 elevation bands and for direct-beam radiation (based on solar elevation). Solar radiation is divided into PAR and NIR wave bands, with differing optical properties of canopy elements specified for each wave band. The canopy is divided into multiple horizontal layers, and radiation from each source (moving toward the ground or, after reflection, toward the sky) is either absorbed by, reflected by, or transmitted through leaves, branches, and boles in each canopy layer. The vertical distribution of leaf and branch/bole "clumping" is accounted for. The forest floor also absorbs and reflects solar radiation. During rain, water is stored on leaf and branch/bole surfaces based on rain amount and the surface areas of leaves and branches/boles. That "intercepted" water is then evaporated during energy balance calculations in subsequent hours.

Leaf growth is initiated in the spring based on "temperature sums" and day length. Leaf growth rate during each hour (for the period from leaf bud burst to leaf physiological maturity) is based on temperature and

TABLE 2. Extended.

Physiological parameters	MAESTRA	NuCM	PnET-II	SPA
Photosynthesis†	Farquhar	NA	empirical	Farquhar
Leaf Conductance (g)	Ball-Berry	NA	NA	$f(SWP)$
Leaf maintenance respiration‖	$f(M, T)$	NA	$f(M, T)$	$f(M, N, T)$
Stem maintenance respiration¶	$f(M, T)$	NA	$f(T, W)$	NA
Growth cost (g C cost/g C built)				
Leaves	NA	NA	dynamic	NA
Stems	NA	NA	dynamic	NA
Roots	NA	NA	dynamic	NA
Coarse roots	NA	NA	NA	NA
Carbohydrate feedbacks	no	NA	yes	NA
Soil N/plant C feedbacks	no	NA	yes	NA
ET approach	Pen/Mon††	calibrated	big leaf	Pen/Mon††
Canopy interception	no	NA	yes	yes
Litter evaporation	no	NA	no	yes
Soil evaporation	no	NA	no	yes
Hydraulic lift	no	NA	no	no
Stem capacitance	no	no	no	no
Soil organic pools	NA	NA	0	NA
Litter pools	NA	NA	0	NA
Soil/heterotrophic respiration#	NA	NA	$f(T, SW)$	NA
Rate dynamics	NA	NA	exponential	NA
Coarse wood decomposition	NA	NA	no	NA

availability of substrate (i.e., stored and current photosynthate). Potential LAI and structural leaf mass per unit leaf area for a site are inputs to the model. Once leaves reach maturity (based on a spring/summer temperature sum), they mobilize nitrogen continuously, with the rate of mobilization based on temperature and soil water content. Autumnal leaf senescence is based on temperature and day length.

Leaf temperature (from an energy balance), photosynthesis, maintenance respiration, and phloem translocation are simulated for sunlit and shaded leaves in each canopy layer. Transpiration is an output of the energy balance calculations. Leaf photosynthetic rate is calculated from a light-response curve defined by the quantum efficiency and photosynthetic potential (i.e., maximum, light-saturated photosynthesis). Quantum efficiency is defined by intercellular CO_2 level, leaf temperature, and specificity of ribulose 1,5-bisphosphate carboxylase/oxygenase (rubisco) for CO_2 (Farquhar and von Caemmerer 1982). Photosynthetic potential is defined by leaf nitrogen content (which is distributed among leaves according to vertical position in the canopy), CO_2 level, and a species-dependent empirical parameter. Photorespiration is calculated from photosynthetic rate and rubisco specificity for CO_2. Leaf isoprene emissions are also calculated, based on absorbed PAR and leaf temperature. Most canopy physiological processes are estimated iteratively. (See Amthor et al. [1994] for the same approach applied to a big-leaf canopy.)

Bole and branch growth are based on time of the year (linked to timing of leaf growth), temperature, and substrate availability. Respiratory costs of leaf growth, leaf maintenance, and phloem loading are estimated according to standard principles (Amthor 2000a). Wood maintenance respiration is a function of temperature. Herbivores consume leaf tissue and convert it to CO_2 based on specified base rates and temperature.

Physically based equations are used to transfer heat and water (liquid and vapor) vertically between adjacent layers in the soil, and between the air above the forest floor and the top soil layer. The physical properties of the soil (clay, silt, sand, and rock fractions) in each layer are used to define thermal and hydraulic characteristics in that layer. Fine roots absorb or release (through "hydraulic lift") water in soil layers containing roots (layers containing roots are specified as input). Net water uptake by the root system is dictated by canopy transpiration during each hour (i.e., no water storage in the plant is accounted for). Water moving horizontally through the soil is lost to streams, and water moving down out of the bottom soil layer is considered deep drainage. Root growth in each soil layer containing roots is calculated from soil temperature, fine root density in the layer (amount of root per volume of soil), water content of the layer, and substrate availability. Respiration supporting root growth, maintenance (a function of soil temperature), and ion uptake from the soil solution are calculated each hour. Fine root death (turnover) rate in each layer is based on temperature and soil water content.

Death of bole, stump, and branches contribute to a coarse woody debris pool. That pool decomposes each hour based on temperature and moisture in the top soil layer. Decomposition of soil organic matter and the forest floor is based on an empirical rate specified for the site. That rate is in turn modulated by temperature and moisture of the top two soil layers.

LINKAGES v2.1 (Stan D. Wullschleger)

LINKAGES v2.1 (Wullschleger et al. 2003a) is derived from LINKAGES (Pastor and Post 1985) to study the effects of climate change (i.e., temperature and precipitation) and inter- and intra-annual variations in climate on long-term forest dynamics. It is a daily time-step model with a single big-leaf canopy and a multilayered soil. For this study, only pools and fluxes for the water cycle were provided. LINKAGES v2.1 was modified to incorporate a more physiology-based representation of plant and soil controls on potential and actual evapotranspiration (ET) over that found in the original LINKAGES. Modifications include replacing the Thornthwaite and Mather (1957) monthly calculation of potential ET with a daily scheme in which evaporation from the soil surface and canopy transpiration are treated separately (Shuttleworth and Wallace 1985). A maximum leaf conductance to water vapor is specified for the stand and modified according to daily radiation, temperature, vapor pressure deficit, and extractable soil water. Interception losses are determined for the canopy based on leaf area and stem area index (Federer 1995). Multiple soil layers (12) have been added to the model with water required for transpiration and evaporation extracted sequentially from each layer. Seedlings, saplings, and mature trees occupy specific soil layers and size classes, and thus experience differential stress during the season. Reductions in diameter increment due to drought are accomplished via the concept of ''stress days,'' whereby days on which soil water content falls below the permanent wilting point are accumulated. LINKAGES v2.1 retains all other components of the original LINKAGES model, which was based on the individual tree model FORET (Shugart and West 1977). Particularly, LINKAGES v2.1 retains litter production, decomposition, and associated nitrogen dynamics similar to those in the FORTNITE model (Aber and Melillo, 1982). Model predictions of species composition, basal area, and stems per hectare generated by LINKAGES v2.1 have recently been validated for the Walker Branch Watershed study site (Bugmann et al. 2001). Although LINKAGES provides estimates of forest wood increment by species, components of the carbon budget are not provided.

LoTEC (Anthony W. King)

LoTEC (Local Terrestrial Ecosystem Carbon) is an hourly time-step model with a single big-leaf canopy and a multilayered soil. It calculates pools and fluxes for both the carbon and water cycle. LoTEC is the ecosystem C cycle model implemented in each grid cell of the global model GTEC 2.0 (Global Terrestrial Ecosystem Carbon). It describes C and water dynamics of local, homogeneous vegetation stands at scales of several square meters to perhaps a hectare. It is a generic ecosystem simulator, with no features specific to temperate deciduous forests. LoTEC litter and soil C dynamics are a modification of the Rothamsted model (Jenkinson 1990) used in GTEC 1.0 (King et al. 1997, Post et al. 1997). The statistical NPP model in GTEC 1.0 (Lieth 1975) was replaced with a process-based model including big-leaf canopy physiology and plant growth/senescence. Hourly simulations of big-leaf CO_2 and water vapor fluxes are used, but plant growth and soil C dynamics are modeled with daily time steps. LoTEC predicts gross photosynthesis, plant growth, plant growth and maintenance respiration, litter production, decomposition, transpiration and precipitation interception losses, and soil water balance.

MAESTRA (Yiqi Luo and Dafeng Hui)

MAESTRA is a three-dimensional, hourly time-step model, originally developed by Wang and Jarvis (1990), to examine forest canopy radiation absorption, photosynthesis, and transpiration. The canopy is represented by an array of tree crowns whose positions and dimensions are specified. Each crown is divided into 6 (maximum of 10) horizontal layers with each layer divided into 12 grid points of equal volume. Each layer is specified by a number of physical and physiological properties, including radiation, temperature, LAI, and leaf nitrogen content. Radiation absorption is calculated for a ''target crown'' in the canopy. Positions and dimensions of trees surrounding the target crown are used to determine the amount of radiation incident on the target crown after passing through the neighboring crowns. Radiation penetration to each grid point is calculated for three wavebands (i.e., PAR, near-infrared, and thermal radiation), including consideration of direct, diffuse, and scattered radiation. Photosynthesis and transpiration at each grid point are calculated from the absorbed radiation. Leaf photosynthesis is estimated by the Farquhar photosynthesis model (Farquhar et al. 1980) coupled to the Ball-Berry stomatal conductance model (Ball et al. 1987). Transpiration is calculated by applying the Penman-Monteith formula (Jarvis and McNaughton 1986) to each grid point. Environmental variables driving model simulations are radiation, air temperature, air humidity, wind speed, and atmospheric CO_2 concentration above the canopy. The model assumes that air humidity, temperature, and CO_2 concentration are uniformly distributed within the canopy.

The model has previously been applied to study canopy carbon and water fluxes of *Picea sitchensis* (Wang and Jarvis 1990), *Pinus radiata* (McMurtrie and Wang 1993), *Betula pendula* (Wang et al. 1998), and *Pinus taeda* (Luo et al. 2001). More details of the model description can be found in Wang and Jarvis (1990) and Luo et al. (2001).

MAESTRA does not simulate belowground root or decomposition processes or soil water depletion, and therefore has a limited ability to capture the response of ecosystems to drought (see *Discussion*).

NuCM (Dale W. Johnson)

NuCM is a daily time-step model with a single big-leaf canopy and multiple soil layers. It only calculates water and nutrient cycle pools and fluxes. The NuCM model was designed by a team of investigators in the Integrated Forest Study (see Johnson and Lindberg 1992), and the code was written by Tetra-Tech, Incorporated, Pasadena, California, USA (Liu et al. 1992). NuCM depicts the cycling of N, P, K, Ca, and Mg at a stand level but also includes the fluxes of major cations (Al^{n+}, H^+, NH_4^+, Ca^{2+}, Mg^{2+}, K^+, Na^+) anions (NO_3^-, SO_4^{2-}, ortho-phosphate, Cl^-, HCO_3^-, organic anions), and Si in precipitation, throughfall, and soil solution. Because NuCM was designed primarily for simulating the effects of atmospheric deposition on nutrient cycling processes, its construction emphasizes soil and soil solution chemistry (Liu et al. 1992). The ecosystem is represented as a series of vegetation and soil components. The overstory consists of one generic conifer and one generic deciduous species of specified biomass and nutrient concentration (foliage, branch, bole, roots). For mixed-species stands, average values for biomass and nutrient concentration by component must be used. NuCM also allows an understory that can be divided into canopy, bole, and roots. Maximum potential growth in the model is defined by the user and is constrained in the model by the availability of nutrients and moisture. The forest floor is simulated from litterfall inputs and litter decay. The user defines bulk density, cation exchange capacity, exchangeable cations, adsorbed phosphate and sulfate, and four soil minerals and their composition. These inputs define the initial soil exchangeable/adsorbed pools and total pools. Initial total soil N pools are simulated from litterfall and decay and user-defined C/N ratios. Vegetation, litter, and soil pools change over a simulation in response to growth, litterfall, and decomposition, and nutrient fluxes via deposition, leaching, and weathering, as described below. Although NuCM executes calculations daily, only annual water cycle outputs are available for comparison with other models. Predictions of annual nutrient availability and flux available from the NuCM model (Johnson et al. 1998, 2002, 2003) are not presented here because they have no analogue from the other 12 models.

PnET-II (Ge Sun and Steven G. McNulty)

PnET-II is a lumped-parameter, monthly time-step model with a multilayered canopy and a single soil layer. It describes carbon and water dynamics in mature forests. It simulates both carbon and water cycles in a forest ecosystem using simplified algorithms that describe key biological and hydrologic processes. The PnET-II model was originally developed for studying forest ecosystem processes in northern forests (Aber and Federer 1992). PnET-II has been validated with field data from northern deciduous upland hardwood forests (Aber et al. 1995, 1996) and southern pine forests (McNulty et al. 1996, Sun et al. 2000), and it has been applied at a regional scale to study the potential effects of climatic change on U.S. forests (U.S. Global Change Research Program 2000).

Input parameters for vegetation, soil and site locations, and climate may be derived from the literature or measured from a local study site. Stand-level vegetation parameters include those regulating the physiological and physical processes such as photosynthesis, light attenuation, foliar nitrogen concentration, plant and soil respiration, and rainfall interception. Only one soil parameter, soil water-holding capacity (field capacity in percentage by rooting depth), is required. Climate input variables include minimum and maximum monthly air temperature, total monthly photosynthetically active radiation (PAR), and total monthly precipitation.

The model simulates the carbon cycle by tracking absorbed carbon during photosynthesis, allocation to foliage, wood, and root, and respiration from leaf, stem, and roots. PnET-II calculates the maximum amount of leaf area that can be supported on a site based on the soil, the climate, and parameters specified for the vegetative type. The model assumes that leaf area is equal to the maximum amount of foliage that could be supported due to soil water-holding capacity, species, and climate limitations. Predicted NPP equals total gross photosynthesis minus growth and maintenance respiration for leaf, wood, and root compartments. PnET-II calculates respiration as a function of the current month's minimum and maximum air temperature. Changes in water availability and plant water demand also place limitations on leaf area produced, so total leaf area decreases as vapor pressure deficit and air temperature increase above optimal levels. Reduced leaf area decreases total carbon fixation and altered ecosystem hydrology. The hydrologic cycle is simulated by the water balance equation. The input component of soil water storage is represented by net precipitation (i.e., precipitation − canopy interception), and outputs consist of canopy interception, plant transpiration, fast or macropore flow representing water not available for extraction by plant roots, and lateral and deep drainage. Soil evaporation is neglected in fully stocked forest ecosystems. Evapotranspiration is defined as the sum of plant transpiration and canopy interception. The model assumes that water that is not subjected to evapotranspiration eventually flows to streams as runoff. Transpiration is directly linked to forest photosynthesis and forest carbon gain processes by modeling transpiration as a function of water use efficiency and vapor pressure deficit. Therefore, PnET-II closely integrates forest hydrology with the biological processes.

SPA (Anne Hartley and Mat Williams)

The Soil–Plant–Atmosphere model (SPA, Williams et al. 1996) is a process-based model that simulates

ecosystem photosynthesis and water balance at fine temporal and spatial scales (30 minute time step, 10 canopy and soil layers). SPA includes a multilayered canopy divided into sun and shade components and a multilayered soil. The version of SPA used in this intercomparison provides complete water cycle data and data for annual gross primary production, but does not estimate ecosystem respiration components. The scale of parameterization (leaf level) and prediction (canopy level) were designed to allow the model to diagnose eddy covariance data, and to provide a tool for scaling up leaf-level processes to canopy and landscape scales (Williams et al. 2001b).

SPA employs a detailed radiative transfer scheme that determines the time-varying transmittance, reflectance, and absorption of longwave, near-infrared and direct and diffuse photosynthetically active radiation (PAR) by canopy layers and the soil surface. Absorption of PAR in each canopy layer is partitioned between sunlit and shaded foliage fractions. The SPA model employs some well-tested theoretical representations of eco-physiological processes, such as the Farquhar model of leaf-level photosynthesis (Farquhar and von Caemmerer 1982), and the Penman-Monteith equation to determine leaf-level transpiration. These two processes are linked by a novel model of stomatal conductance that optimizes daily carbon (C) gain per unit leaf nitrogen (N), within the limitations of canopy water storage and soil-to-canopy water transport. The maximum flux rate of water through vegetation is determined by the difference between soil water potential and the minimum sustainable leaf water potential, and by the hydraulic resistance of the soil–root–leaf pathway. Stomata adjust to equalize evaporative losses with the maximum hydraulic supply, minimizing the risk of cavitation.

SPA contains a detailed representation of soil hydrology and thermal dynamics. From the estimated transmission of radiation through the canopy, SPA determines the down-welling radiation at the soil surface, and then solves the surface energy balance by estimating the soil surface temperature, and partitioning net radiation into sensible, latent, and ground heat fluxes. The soil is divided into 10 layers of varying thickness, each with a specified organic matter and mineral content. The flux of heat through the soil profile is determined on the basis of the ground heat flux, the thermal gradient between soil layers, and the soil thermal conductivity and thermal heat content of each layer. The thermal parameters are dependent on soil organic matter and mineral fractions and soil water content, and phase transitions between liquid water and ice. The field capacity of each layer is determined according to soil texture and soil water retention curves. Heat is redistributed through the soil profile according to water movement, and from patterns of freezing and thawing. The ice content of each soil layer is determined daily. Root water uptake is explicitly linked with soil water

potential and soil hydraulic conductivity, as determined from soil water retention curves, through the plant hydraulic model outlined above. Roots are distributed through the upper soil layers and water is withdrawn from the layers with greatest moisture content. Precipitation inputs to soils are calculated after canopy interception, drainage, and evaporation from the canopy water store, and infiltration through the soil surface. A snow submodel tracks the dynamics of the snowpack and its effects on soil temperature.

The model is readily applied to different ecosystems, as there are relatively few parameters to be changed. The most critical are LAI and foliar N (accounting for phenological changes), plant hydraulic conductance, minimum leaf water potential, rooting depth, and soil texture. The SPA model has been applied in ecosystems ranging from 70° N to 2° S latitude. The SPA model has been extensively tested against independent eddy covariance data for the temperate oak–maple forest at the Harvard Forest (Wofsy et al. 1993, Williams et al. 1996) where the model was able to explain >90% of the variability in measured daily gross primary productivity (Williams et al. 1997).

METHODS

Ecosystem description

The deciduous forest ecosystem serving as the case study for this model–measurement comparison was the Walker Branch Watershed (35°57′ N; 84°17′ W; 250–330 m elevation) located on the U.S. Department of Energy's National Environmental Research Park near Oak Ridge, Tennessee (Johnson and Van Hook 1989). This forested watershed is largely upland oak in composition, yet it is representative of much of the eastern deciduous hardwood forest. Oak-dominated forests occupy ~617 000 km^2 or 42% of the total forested area in the eastern United States (Powell et al. 1992).

Long-term (50-year) mean annual precipitation was 1352 mm and mean annual temperature is 14.2°C. The acidic forest soils (pH 3.5 to 4.6) are primarily typic Paleudults. Plant extractable water (water held between 0 and −2.5 MPa) for the upper meter of soil is ~183 mm. A large fraction of this water (44%) is held in the upper 0.35 m of the soil profile. This surface layer is the location of 60% of all fine roots in the 0–0.90 m soil profile (Joslin and Wolfe 1998). Depth to bedrock is ~30 m and deep rooting may be a source of some water.

The TDE experimental site (Hanson et al. 2001a, b, 2003a) provided site-specific characteristics for forest structure and composition for the model simulations. Early aerial photographs show that the TDE site was forested in the late 1930s, but several large dominant trees show open growth characteristics, an indication of some harvesting before that time. *Quercus* spp. and *Acer* spp. are the major canopy dominants. *Liriodendron tulipifera* L. is a canopy dominant on the lower

TABLE 3. Initial conditions in 1993 for the dry matter pools, canopy leaf area index, midseason leaf mass per unit area (LMA), leaf nitrogen (N), and the percentage of sapwood in boles and branches.

Component	Initial value	Units
Initial mass		Units are dry mass/ground area
Foliage	496	g/m^2
Branches	2958	g/m^2
Boles	10 717	g/m^2
Stumps	1431	g/m^2
Saplings	858	g/m^2
Coarse roots	1844	g/m^2
Fine roots†	2879	g/m^2
O-layer litter	1856	g/m^2
Wood litter	1552	g/m^2
Soil organic matter from 0 to 0.9 m	6389	g/m^2
Leaf area index (LAI)	6.2	m^2/m^2
Species contribution to LAI		
Q. prinus	20.2	%
Q. alba	24.1	%
Quercus/Carya sp.	7.8	%
Acer sp.	25.9	%
Liriodendron	2.7	%
Nyssa/Oxydendrum	14.7	%
Cornus	0.7	%
Miscellaneous	3.7	%
Pinus	0.2	%
Upper canopy LMA	74	g/m^2
Middle canopy LMA	55	g/m^2
Lower canopy LMA	36	g/m^2
Upper canopy N	1.7	%
Middle canopy N	1.4	%
Lower canopy N	0.8	%
Percentage of wood as sapwood		
Branches	100	%
Bole	29	%

† Root data for the TDE site are for the 0–1 m depth increment (Joslin and Wolfe 1998).

slope positions, and *Nyssa sylvatica* Marsh. and *Oxydendrum arboreum* [L.] D.C. are the predominant species occupying midcanopy locations (Hanson et al. 2001*a*). In March of 1994, stand basal area averaged 21 m^2/ha. By December 1999, mean basal area had increased to 22.8 m^2/ha. The number of saplings (trees <0.1 m dbh) averaged 3073 trees/ha in 1994 and 2112 trees/ha in 1999. Saplings (predominantly *Acer rubrum* L. and *Cornus florida* L.) contributed an additional 3 and 2.6 m^2/ha to total stand basal area in 1994 and 1999, respectively (Hanson et al. 2001*a*).

Organization and model parameterization

All participants were instructed to use standardized environmental and model parameterization datasets for their model runs to avoid difficulties in interpretation associated with nonstandardized parameterization (Amthor et al. 2001) or scenarios (Weinstein et al. 1998). Site-specific hourly environmental data including incident photosynthetically active radiation (PAR, in micromoles per square meter per second), mean incident shortwave radiation (watts per square meter), within- or above-canopy air temperature (°C), soil temperature at 10 cm (°C), relative humidity (percent), rainfall (millimeters), and wind speed (meters per second), were available from an online database (Hanson

et al. 2001*b*). Species composition was described by Hanson et al. (2001*a*), and LAI and seasonal leaf phenology from Hanson et al. (2001*b*, 2003*c*). Soil physical and chemical conditions were as described by Peters et al. (1970).

Unless a given model was designed to run with ''generic'' settings, the participants were directed to parameterize their models from site- or regionally specific publications on leaf photosynthesis and conductance (Harley and Baldocchi 1995, Sullivan et al. 1996, Augé et al. 2000, Wilson et al. 2000*a, b*), foliar respiration (Bolstad et al. 1999, Mitchell et al. 1999, Amthor 2000*b*), stem respiration (Edwards and Hanson 1996), soil respiration (Hanson et al. 1993), litter decomposition (O'Neill and Norby 1996), ozone response (Hanson et al. 1994, Wullschleger et al. 1996), whole-plant water use (Wilson et al. 2001*b*, Wullschleger et al. 1998, 2000, 2001), and nonmethane hydrocarbon emissions (Hanson and Hoffman 1994, Harley et al. 1997). Two books were also noted as good sources for local parameterization information (Reichle 1980, Johnson and Van Hook 1989).

Participants were instructed to initiate their models for 1993 using the dry matter pools, LAI, and leaf mass per unit area values listed in Table 3, or to approximate
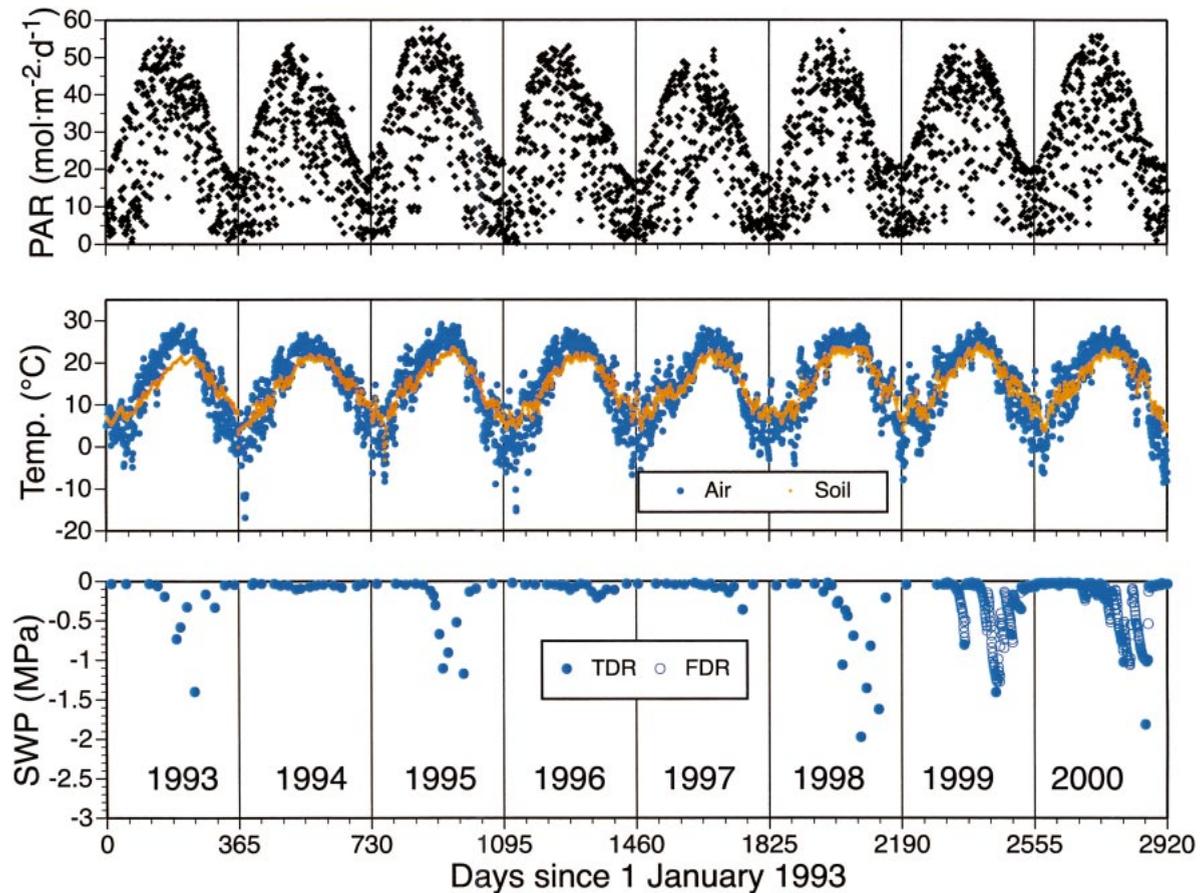
FIG. 2. Daily measures of incident photosynthetically active radiation (PAR), air and soil (at 10 cm) temperature (Temp.), and soil water potential (SWP) obtained from periodic time domain reflectometry (TDR) or frequency domain reflectometry (FDR) are plotted from 1993 to 2000.

those values if the model required initial simulations to be ''spun-up'' to 1993 initial conditions.

The simulation period

The simulations were conducted for the period from 1993 through 2000, during which substantial interannual differences in precipitation were observed (Hanson et al. 2003a). Growing-season precipitation was near normal in 1994, 1999, and 2000, but it was 26–38% less during the drought years of 1993, 1995, and 1998. Growing-season precipitation was 47 and 22% higher than normal in 1996 and 1997, respectively. Mean annual air temperature and annual incident solar radiation were not as variable as annual precipitation, but mean annual air temperatures in 1998 and 1999 were warmer than in the other years (1–2°C). Cumulative annual incident solar radiation at the site was similar across years, ranging from 2643 to 3155 MJ/m². A plot of daily total photosynthetically active radiation (PAR), mean canopy and soil temperature, and mean soil water potential from 1993 to 2000 illustrates inter- and intra-annual differences over the simulation period (Fig. 2).

Model outputs

Participants were requested to provide the following annual, daily, and hourly outputs for 1993 to 2000 if appropriate to their model.

1) Annual carbon cycle components including gross primary production (GPP), net primary production (NPP), net ecosystem exchange or production (NEE$_a$ or NEP), leaf respiration, wood (branch, bole, stump) respiration, and total soil respiration (or the root and heterotrophic components).

2) Annual water cycle components including evapotranspiration (ET$_a$), transpiration (T_a), evaporation from the canopy and/or soil surface (E_a), drainage and/or surface runoff (Drainage).

3) Annual growth of leaves, wood, and roots.

4) Daily time-step outputs for soil water content (by depth), evapotranspiration (ET$_d$), transpiration (T_d), and net ecosystem exchange (NEE$_d$).

5) Hourly estimates of net ecosystem exchange (NEE$_h$) and stand transpiration (T_h) for specific two-week drought and wet periods in 1998 and 1999, respectively. In this paper, uptake of carbon by the forest

(NPP and GPP) is defined as a positive flux, and positive and negative values of NEE represent carbon uptake or loss (respiration) from the forest, respectively.

Experimental observations for testing models

Data from studies on Walker Branch Watershed (Hanson et al. 1998, 2001*b*, 2003*a*, Wilson and Baldocchi 2001) were obtained to test the efficacy of model predictions at hourly, daily, and annual time steps.

Eddy covariance data for NEE_h from the Walker Branch AmeriFlux tower site (Wilson and Baldocchi 2001) were selected for two-week periods during a drought (days 238 to 251 of 1998) and during wet periods having optimum soil water availability (days 206 to 219 of 1999). These data were carefully selected for optimum eddy covariance conditions, because eddy-covariance-based NEE_h measurements for Walker Branch are biased because of difficult site and environmental conditions, i.e., sloping terrain and low night turbulence (Baldocchi et al. 2000, Curtis et al. 2002, Hanson et al. 2003*d*). Hourly transpiration (T_h) for the same two-week periods was extrapolated from sap flow probes according to the approach of Wullschleger et al. (2000, 2001, 2003*b*). Although the sap-flow estimates of transpiration are known to underestimate T_h (Wilson et al. 2001*b*), the daily dynamics and relative response to soil water deficits are appropriately represented. NEE_h and T_h are plotted against incident PAR under both wet and drought conditions for evaluation of the "light-response-curves." Response curves between T_h and vapor pressure deficit (VPD) are also provided.

Daily measurements of ET_d, NEE_d, mean surface soil water content, and soil respiration were available for comparisons with model outputs. Measured ET_d from eddy covariance (Wilson and Baldocchi 2001) was available for the period from 1995 to 2000. Gap-filled NEE_d eddy covariance data for Walker Branch based on look-up tables (Falge et al. 2001) were available from 1995 to 1998. Measured soil water content data were available from periodic observations of 100 vertically integrated (0–0.35 m) time-domain reflectometry rod pairs (TRASE, Soil Moisture Equipment Corporation, Santa Barbara, California, USA; Hanson et al. 2001*b*) from 1993 to 2000, and from horizontally installed frequency domain reflectometry probes placed between 0 and 0.35 m (CS615, Campbell Scientific, Logan, Utah, USA) for the period from 1998 to 2000 (Hanson et al. 2003*a*). Soil respiration measurements were available from periodic chamber observations on Walker Branch (Hanson et al. 2003*b*) from 1993 to 1999, and from nocturnal understory-eddy-covariance measurements from 1999 and 2000 (1.5 m tower; Wilson and Meyers 2001).

Independent measurements for comparison to annual model outputs were available for ET_a, watershed drainage, growth (roots, stems, leaves), NPP, and NEE_a. Annual ET_a for 1995 to 2000 was from integrated-annual and gap-filled eddy covariance data (Wilson and Baldocchi 2000, Wilson et al. 2001*b*), and from watershed-level calculations of the difference between precipitation inputs and stream flow outputs for 1993 to 2000 (Wilson et al. 2001*b*). Walker Branch stream flow outputs for 1993 to 2000 were used as a measure of drainage through the soil profile (P. J. Mulholland, *personal communication*). Growth measurements for 1993 to 2000 were from litter collection baskets for leaves (Hanson et al. 2003*c*), allometric relationships for stems and branches (Hanson et al. 2003*d*), and minirhizotron tubes and ingrowth cores for roots (Joslin and Wolfe 2003). Annual NPP data were estimated from independent biometric measurements (Curtis et al. 2002, Hanson et al. 2003*d*). NEE_a was estimated as NPP minus heterotrophic respiration, where heterotrophic respiration was assumed to be half of annual soil respiration (Hanson et al. 2000). Calculations based on soil respiration data from Hanson et al. (2003*b*) yielded a low estimate of NEE_a (Biometric Low), and calculations using the mean modeled soil respiration gave a higher NEE_a value (Biometric High).

Statistical evaluations

Several approaches were used to evaluate the "goodness-of-fit" of an individual model's hourly, daily, or annual predictions through comparisons of model output with independent field data. The data associated with wet (optimum moisture) or drought conditions were included to contrast a model's ability to capture measured data under extreme or adverse conditions (Loehle 1997). Analyses were limited to those models capable of providing input. For example, models based on monthly and daily time steps were not used in the comparison of hourly model outputs.

Hourly and daily data.—Scatter plots of modeled hourly stand NEE_h as a function of PAR and T_h as a function of VPD were generated for visual comparison with eddy covariance-based NEE_h and sapflow-based T_h observations, respectively. In addition, model predictions and observations for hourly and daily time steps were contrasted with a number of statistical approaches. Linear regression slopes, intercepts, and R^2 outputs were provided as a common initial comparison between observations and predictions, even though Mitchell (1997) appropriately critiques the use of regression for model testing. For linear regression, a slope of 1, an intercept of 0, and a high R^2 value were assumed to represent good agreement between model predictions and measurements. Plots of the difference between measured and model-predicted data (difference plots) were also generated to evaluate patterns of agreement over diurnal or annual time periods for the hourly or daily data, respectively.

We calculated mean bias (Bias) and mean absolute bias (ABS) from the following equations (Reynolds 1984, Walters 1994):

$$\text{Bias} = \frac{\sum (\hat{y}_i - y_i)}{n} \quad (1)$$

$$\text{ABS} = \frac{\sum |\hat{y}_i - y_i|}{n}. \quad (2)$$

Bias provides a direct measure of the tendency for over or under prediction (positive or negative values, respectively). The ABS value is a measure of the mean deviation from the observed values (i.e., the amount of scatter). To further evaluate model bias as a function of time of day (hourly predictions) or day of the year (daily predictions) scatter plots of the difference between predictions and observations were provided.

A modeling efficiency statistic (EF) proposed as an improvement over linear regression by Mayer and Butler (1993) is calculated from the following equation:

$$\text{EF} = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2} \quad (3)$$

where an EF value of 1 represents a perfect fit and values range from 1 to infinitely negative. EF values of 1, however, are not expected due to natural variability and experimental error associated with data collection. The EF value is similar to the calculated R^2 from the linear regression, but it uses the one-to-one line rather than the regression line as the reference point. As a point of reference, EF values much less than 0.5 were considered in poor agreement with measured data in this paper. Measurements of Bias, ABS, and difference plot deviations from diurnal or seasonal patterns need to be considered to evaluate the nature of differences suggested by low EF values.

Annual data.—For comparing annual summary statistics (e.g., NEE_a, and ET_a), the average annual model output from all simulated annual years predicted by each model were calculated for a range of variables. Box-and-whisker plots of the mean and interannual range of each model's annual predictions of carbon and water budget components were plotted with independent measurements of each variable (e.g., eddy covariance estimates of NEE_a). When independent variables were not available from direct measurements (e.g., annual stem or leaf respiration on a ground area basis), the mean output from all models provided a point of comparison for judging a models' performance. Anderson and Burnham (2001) described the application of a ''model-averaging'' approach for prediction purposes. While the goal of this paper is not the development of predictions from combined model outputs, mean model outputs were used as a reference point for evaluating model differences.

RESULTS

Hourly simulations

Observed NEE_h data from eddy covariance measurements along with predicted NEE_h as a function of incident PAR for 7 of the 13 models are plotted in Fig. 3. Within the figure, data are stratified into hours for a two-week wet period (optimum soil water) in 1999, or hours from a two-week severe drought in 1998. The measured data show light-saturated NEE_h (NEE_{max}) of 1.03 g $C \cdot m^{-2} \cdot h^{-1}$ and stand-level respiration of 0.15 g $C \cdot m^{-2} \cdot h^{-1}$ under optimum (or wet) soil water conditions (Table 4). NEE_h was reduced under drought conditions by 46% to 0.56 g $C \cdot m^{-2} \cdot h^{-1}$.

Under wet soil conditions the *ecosys* model overestimated NEE_{max}, the MAESTRA model came closest to the measured value, and the other models provided lower estimates (Fig. 3, Table 4). The CANOAK model had the lowest bias and highest EF value for the NEE_h data under conditions of optimum soil water availability. Under drought conditions, MAESTRA overestimated NEE_h, the CANOAK and *ecosys* models come closest to the observed values, LaRS and LoTEC produced slight underestimates of NEE_h, and INTRASTAND and EALCO substantially underestimated NEE_{max} during drought (Fig. 3 and Table 4). The difference plot for NEE_h (Fig. 4) shows that positive and negative bias were largely driven by daytime phenomena. This suggests that model over and under prediction associated with the hourly time step were a function of assumptions about the photosynthetic process.

Hourly observations and predictions of transpiration (T_h) as a function of VPD for 8 of the 13 models are plotted in Fig. 5. The data are stratified into hours from a two-week wet period in 1999 (i.e., optimum soil water), or from a two-week severe drought in 1998. The integrated-sapflow observations show maximum T_h under optimum soil water conditions to be 0.4 mm/h, falling to <0.2 mm/h under severe drought. The slope of the relationship between T_h and VPD ($R^2 = 0.66$ wet and 0.77 dry) was a strong indicator of the drought response changing from 0.18 to 0.06 $mm \cdot h^{-1} \cdot kPa^{-1}$ for the wet and dry periods, respectively (Table 5). ET_h-VPD relationships measured by eddy covariance ($R^2 = 0.45$ wet and 0.25 dry) showed an analogous response to drought, with the slope changing from 0.23 to 0.07 $mm \cdot h^{-1} \cdot kPa^{-1}$ under wet and dry conditions, respectively (data not shown). Under optimum soil water conditions, all models with the exception of LoTEC had low bias and high EF values, but INTRASTAND, LaRS, and MAESTRA produced especially good fits. Under drought conditions, differences among models were apparent. CANOAK had a substantial positive slope and bias, suggesting that it did not simulate the full impact of drought. Conversely, as with NEE_h, INTRASTAND overestimated the impact of drought with a low slope, negative bias, and poor EF value. The best performing model for T_h was LaRS, with an EF value of 0.85. The other five models (EALCO, *ecosys,* LoTEC, MAESTRA, SPA) exhibited intermediate performance, with EF values from 0.54 to 0.68. Because T_h falls near zero at night, it is not surprising that the
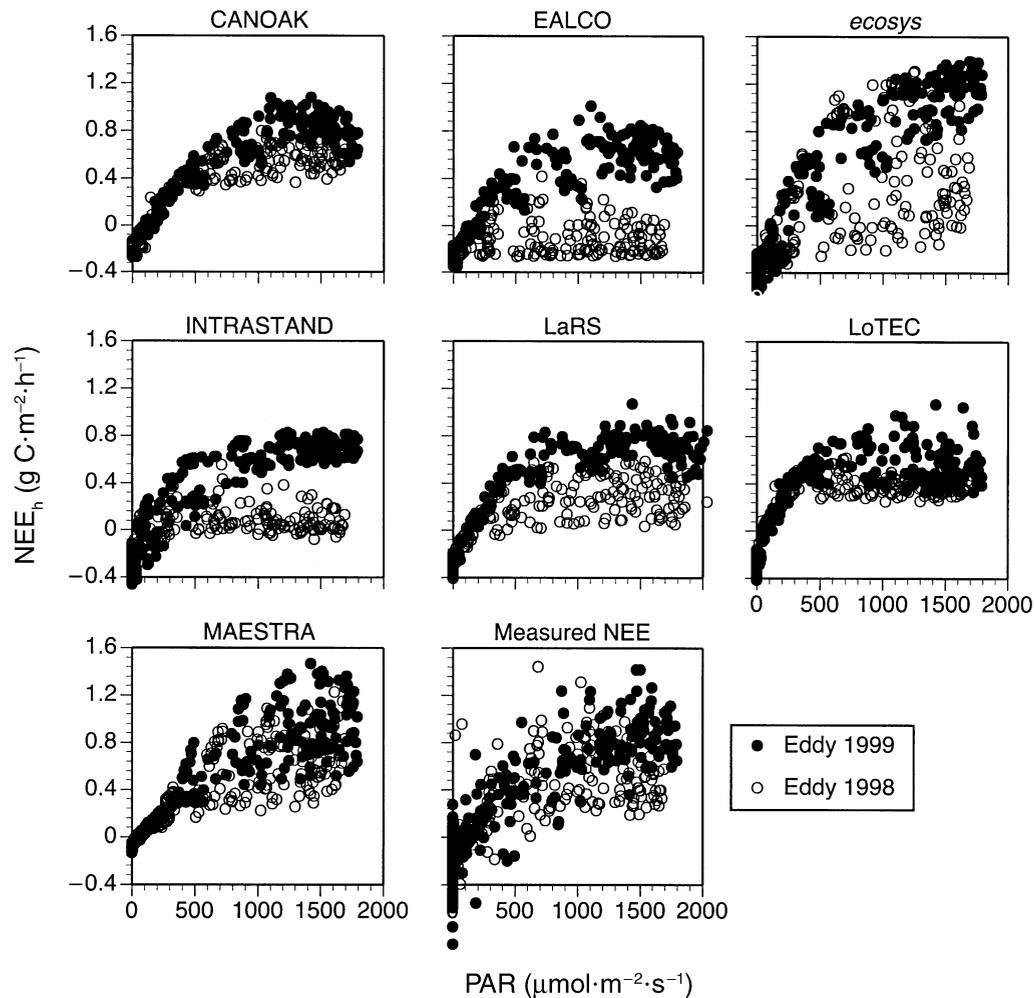
Fig. 3. Hourly net ecosystem exchange (NEE_h) as a function of incident light (PAR) for two-week severe-drought (open circles) or wet (solid circles) periods in 1998 and 1999, respectively. Data for seven models are provided along with eddy covariance NEE measurements for comparison.

difference plots (Fig. 6) showed virtually all bias during daylight hours.

The linear nature of the $T = f$ [VPD] relationship for LoTEC in Fig. 5 suggested that it used a very different mechanism from the other seven models. Further analysis indicated that the stomatal conductance simulated by LoTEC is explicitly limited by leaf water potentials more negative than a parameterized critical value, and leaf water potential is modeled in LoTEC using a simple whole-plant hydraulic resistance model. LoTECs relative insensitivity to VPD in 1999 (Fig. 5) is a direct consequence of the hydraulic resistance model's frequent calculation of leaf water potentials more negative than the critical value even in wet conditions. When LoTEC uses soil water potential rather than leaf water potential to constrain stomatal conductance (bypassing the hydraulic resistance model), the relationship between T_h and VPD for LoTEC in 1999 is very similar to those of CANOAK and SPA.

Daily simulations

Daily net ecosystem carbon exchange (NEE_d), evapotranspiration (ET_d), and soil water content (SW) from direct measurements, from 11 models, and for the mean daily model response were plotted in Fig. 7 for a 40-day wet period. Figs. 8 and 9 show analogous plots for a 40-day dry period, and for a 40-day period transitioning from wet to dry, respectively. Measured data for ET_d and SW were available for all dates and show good correspondence with the mean model prediction. Daily values of NEE_d were obtained from gap-filled eddy covariance data (Falge et al. 2001) known to have nighttime bias (Wilson and Baldocchi 2001) that underestimated respiratory losses from the forest and therefore overestimated NEE_d. Nevertheless, NEE_d measurements captured inter- and intra-annual dynamics associated with leaf production and loss and the occurrence of seasonal drought, and represent a useful data source for model evaluations.

TABLE 4. Characteristics of the stand-level, light-response curve of net ecosystem carbon exchange (NEE_h) for eddy covariance (EC) observations and model simulations, and expressions of model performance for estimates of NEE_h for a two-week dry period in 1998 and a two-week wet period in 1999.

	Stand-level light response variables			Observed vs. simulated regression			Goodness-of-fit variables		
Year and model	NEE_{max} (g C·m^{-2}·h^{-1})	R_{stand} (g C·m^{-2}·h^{-1})	LCP (μmol· m^{-2}·s^{-1})	Slope	Intercept (g C· m^{-2}·h^{-1})	R^2	Bias (g C· m^{-2}·h^{-1})	ABS (g C· m^{-2}·h^{-1})	EF
1999 wet period									
Observations, EC	1.03 ± 0.13	-0.15 ± 0.03	108 ± 22	NA	NA	NA	NA	NA	NA
CANOAK	0.88 ± 0.02	-0.26 ± 0.02	121 ± 8	0.93	-0.02	0.82	-0.04	0.16	0.81
EALCO	0.65 ± 0.02	-0.31 ± 0.02	158 ± 13	0.81	-0.10	0.81	-0.15	0.21	0.68
ecosys	1.40 ± 0.10	-0.41 ± 0.02	201 ± 16	1.29	-0.12	0.82	-0.04	0.28	0.77
INTRASTAND	0.73 ± 0.03	-0.32 ± 0.02	150 ± 12	0.86	-0.09	0.77	-0.12	0.21	0.70
LaRS	0.72 ± 0.02	-0.32 ± 0.01	100 ± 6	0.90	-0.06	0.80	-0.09	0.19	0.77
LoTEC	0.57 ± 0.02	-0.35 ± 0.02	82 ± 8	0.78	-0.09	0.71	-0.14	0.24	0.60
MAESTRA	1.09 ± 0.06	-0.09 ± 0.02	61 ± 14	0.87	0.10	0.76	0.07	0.17	0.73
1998 drought period									
Observations, EC	0.56 ± 0.05	-0.08 ± 0.03	47 ± 17	NA	NA	NA	NA	NA	NA
CANOAK	0.61 ± 0.03	-0.18 ± 0.01	97 ± 7	0.89	-0.03	0.77	-0.04	0.13	0.74
EALCO	-0.1	-0.21	60	0.20	-0.02	0.37	-0.34	0.35	-0.52
ecosys	0.46 ± 0.08	-0.36 ± 0.05	193 ± 56	1.21	-0.24	0.79	-0.20	0.27	0.64
INTRASTAND	0.08 ± 0.01	-0.20 ± 0.02	112 ± 34	0.35	-0.13	0.53	-0.24	0.26	-0.40
LaRS	0.32 ± 0.02	-0.28 ± 0.02	118 ± 18	0.75	-0.14	0.75	-0.18	0.21	0.47
LoTEC	0.38 ± 0.01	-0.23 ± 0.01	49 ± 4	0.71	-0.06	0.67	-0.11	0.18	0.48
MAESTRA	0.71 ± 0.06	-0.09 ± 0.02	58 ± 15	0.93	0.03	0.78	0.02	0.11	0.77

Notes: Features of the light response curve (95% CI) include light-saturated NEE_h (NEE_{max}), stand respiration (R_{stand}), and the light-compensation point (LCP) from Hanson et al. (1987). Measures of performance include the slope and intercept from linear regression between observed and simulated data, the mean bias (Bias), the mean absolute bias (ABS; where values close to zero represent good agreement), and the model efficiency factor (EF; where a perfect fit equals 1). NA = not applicable.

With respect to NEE_d simulations, most models captured the daily temporal dynamic of the gap-filled observations and the mean model response (Figs. 7, 8, and 9). NEE_d observations, however, exceeded most model predictions and the mean model response under both wet (Fig. 7) and dry (Fig. 8) conditions, but the difference was greatest under drought. MAESTRA provided the highest NEE_d predictions under both wet and dry conditions, but captured appropriate temporal dynamics (Figs. 7, 8, and 9). Conversely, BIOME-BGC consistently underpredicted the mean model response and observations, and showed little of the temporal
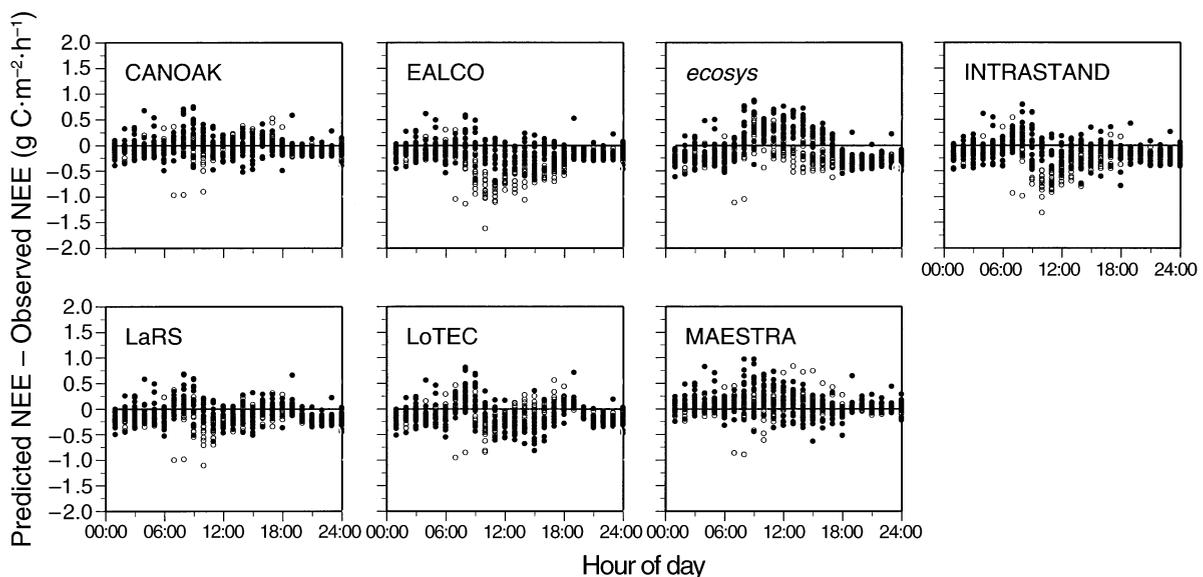
FIG. 4. Difference plot of simulated minus observed hourly net ecosystem exchange, NEE_h (measured by eddy covariance) for seven models for a two-week dry period (open symbols) and wet period (solid symbols) in 1998 and 1999, respectively.
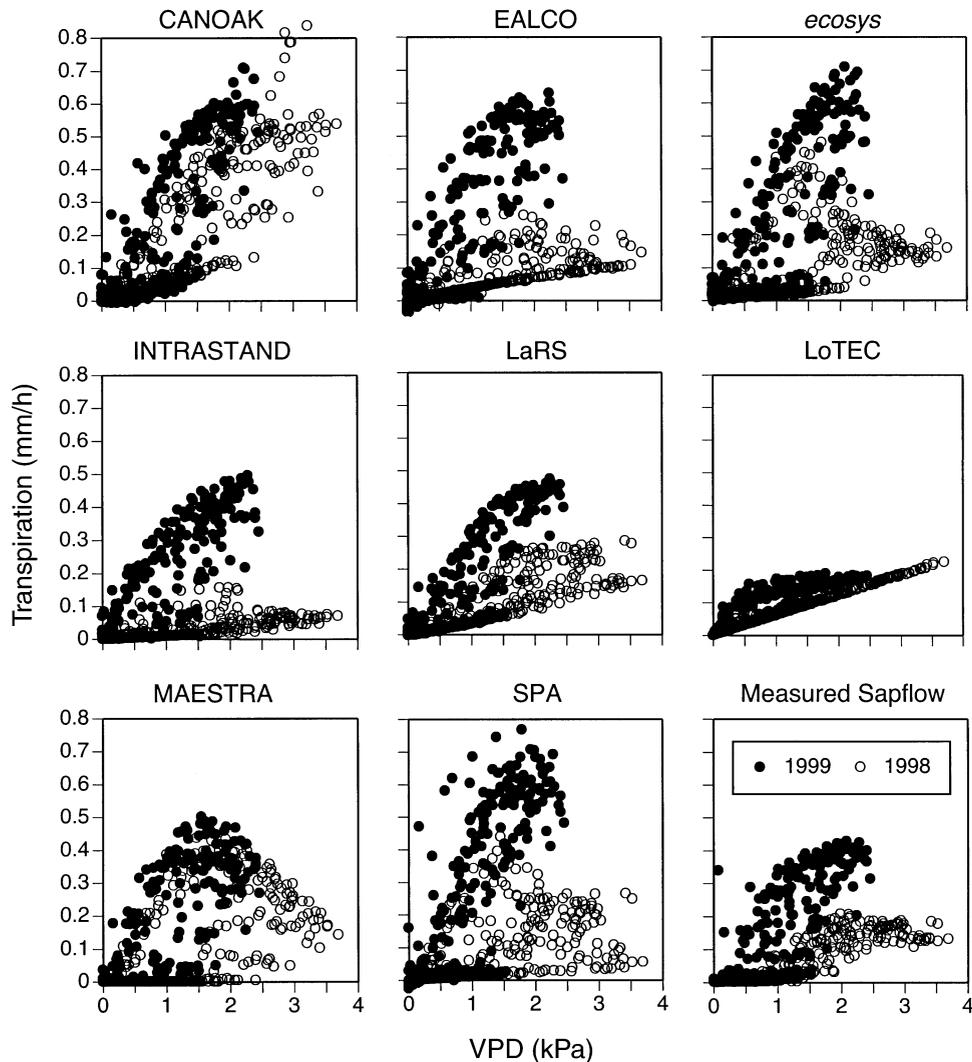
FIG. 5. Hourly transpiration (*T*) as a function of vapor pressure deficit (VPD) for a two-week dry period (1998) and a two-week wet period (1999). Data for eight models and measured sapflow data scaled to the stand level are presented.

dynamics captured in the other models (Figs. 7 and 8). Periodic underpredictions of the mean model NEE$_d$ values were also observed for INTRASTAND (Figs. 7 and 8) and BGC++ (Fig. 8). During the transitional dry-down example (Fig. 9), EALCO and BGC++ under-predicted NEE$_d$. A difference plot contrasting observed and predicted NEE$_d$ for all days from 1995 to 1998 (Fig. 10) and the model performance statistics in Table 6 showed that most individual models and the mean model predictions had a negative bias with respect to the gap-filled eddy covariance data. MAESTRA was the only model with a positive bias. CANOAK and LaRS simulations had the best agreement with the gap-filled EC data, yielding EF values of 0.73 and 0.64, respectively. The mean model prediction was not the best approximation of NEE$_d$. The difference plot (Fig. 10) showed the majority of model deviations from sea-

sonal NEE$_d$ patterns to occur during the growing season.

Most models did a good job tracking the magnitude and daily temporal dynamic of observed ET$_d$, and mean model predictions were a very good approximation of the observed ET$_d$ from eddy covariance measurements (Figs. 7, 8, and 9). The models showing the greatest deviation from ET$_d$ observations were not the same as for NEE$_d$. LoTEC underestimated ET$_d$ under wet conditions (Fig. 8) and during the wet-to-dry transition (Fig. 9), but was close to the mean model response and observations at the depth of the drought (Fig. 8). Under dry conditions CANOAK showed a large overprediction of ET$_d$ (Fig. 8). BGC++, LaRS, and LINKAGES sporadically overpredicted ET$_d$ during wet conditions. A difference plot contrasting observed and predicted ET$_d$ for all days from 1993 to 2000 (Fig. 11) and the

TABLE 5. Slope (and 95% CI) of a linear relationship between hourly transpiration (T_h) from sapflow and vapor pressure deficit (VPD) for observations and model simulations for a two-week severe drought period in 1998 and a two-week period of optimum soil water content in 1999.

Year and model	T_h vs. VPD; slope (mm·h^{-1}·kPa^{-1})	Observed vs. simulated regression			Goodness-of-fit variables		
		Slope	Intercept (mm/h)	R^2	Bias (mm/h)	ABS (mm/h)	EF
1999 optimum soil water							
Observations	0.179 ± 0.014	NA	NA	NA	NA	NA	NA
CANOAK	0.277 ± 0.020	1.47	0.00	0.98	0.06	0.07	0.81
EALCO	0.260 ± 0.021	1.43	0.01	0.97	0.06	0.07	0.81
ecosys	0.254 ± 0.023	1.45	0.00	0.96	0.06	0.06	0.81
INTRASTAND	0.192 ± 0.017	1.07	0.00	0.96	0.01	0.02	0.95
LaRS	0.211 ± 0.014	1.08	0.01	0.96	0.02	0.03	0.94
LoTEC	0.082 ± 0.004	0.36	0.04	0.78	−0.04	0.08	−1.12
MAESTRA	0.178 ± 0.019	1.08	−0.01	0.94	0.00	0.03	0.94
SPA	0.301 ± 0.025	1.62	0.01	0.95	0.09	0.09	0.72
1998 maximum drought period							
Observations	0.060 ± 0.004	NA	NA	NA	NA	NA	NA
CANOAK	0.185 ± 0.012	2.84	0.02	0.89	0.13	0.13	0.36
EALCO	0.046 ± 0.004	0.77	0.01	0.74	0.00	0.03	0.68
ecosys	0.061 ± 0.011	1.26	0.02	0.60	0.04	0.04	0.52
INTRASTAND	0.022 ± 0.002	0.38	0.00	0.68	−0.03	0.04	−0.55
LaRS	0.074 ± 0.004	1.18	0.01	0.93	0.02	0.03	0.85
LoTEC	0.060 ± 0.001	0.74	0.04	0.74	0.02	0.03	0.57
MAESTRA	0.088 ± 0.013	1.77	0.00	0.73	0.05	0.06	0.54
SPA	0.056 ± 0.011	1.16	0.01	0.62	0.02	0.04	0.59

Notes: Also provided are expressions of model performance for predictions of hourly transpiration (T_h) for the same periods. Measures of performance include the slope and intercept from linear regression between observed and simulated data, the mean bias (Bias), the mean absolute bias (ABS; where values close to zero represent good agreement), and the model efficiency factor (EF; where a perfect fit equals 1). NA = not applicable.

model performance statistics in Table 7 demonstrate that the mean model prediction is the best representation of observed ET_d with an EF value of 0.84 and minimal bias. Across 8 years, all of the individual models yielded EF values between 0.61 and 0.79 except for BGC++ and LoTEC. BGC++ and LoTEC exhibited substantial deviation from the observed data during the growing season, but the dominant trend for BGC++ was a positive bias, while LoTEC had a negative bias (Fig. 11).
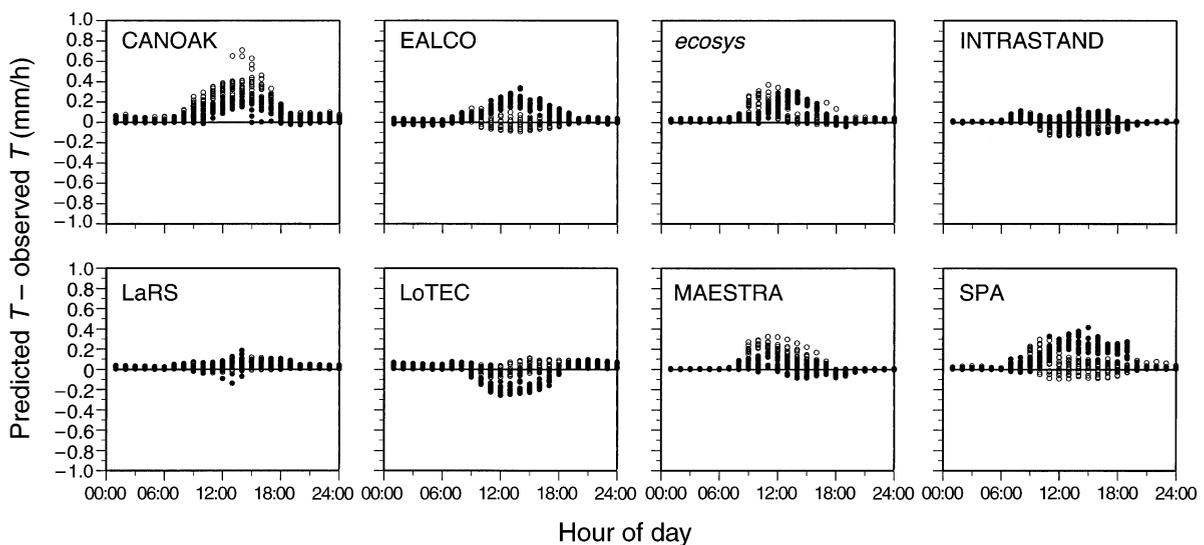
FIG. 6. Difference plot of simulated minus observed hourly transpiration (T) measured by heat dissipation sapflow sensors for seven models for a two-week dry (open symbols) and wet (solid symbols) period in 1998 and 1999, respectively.
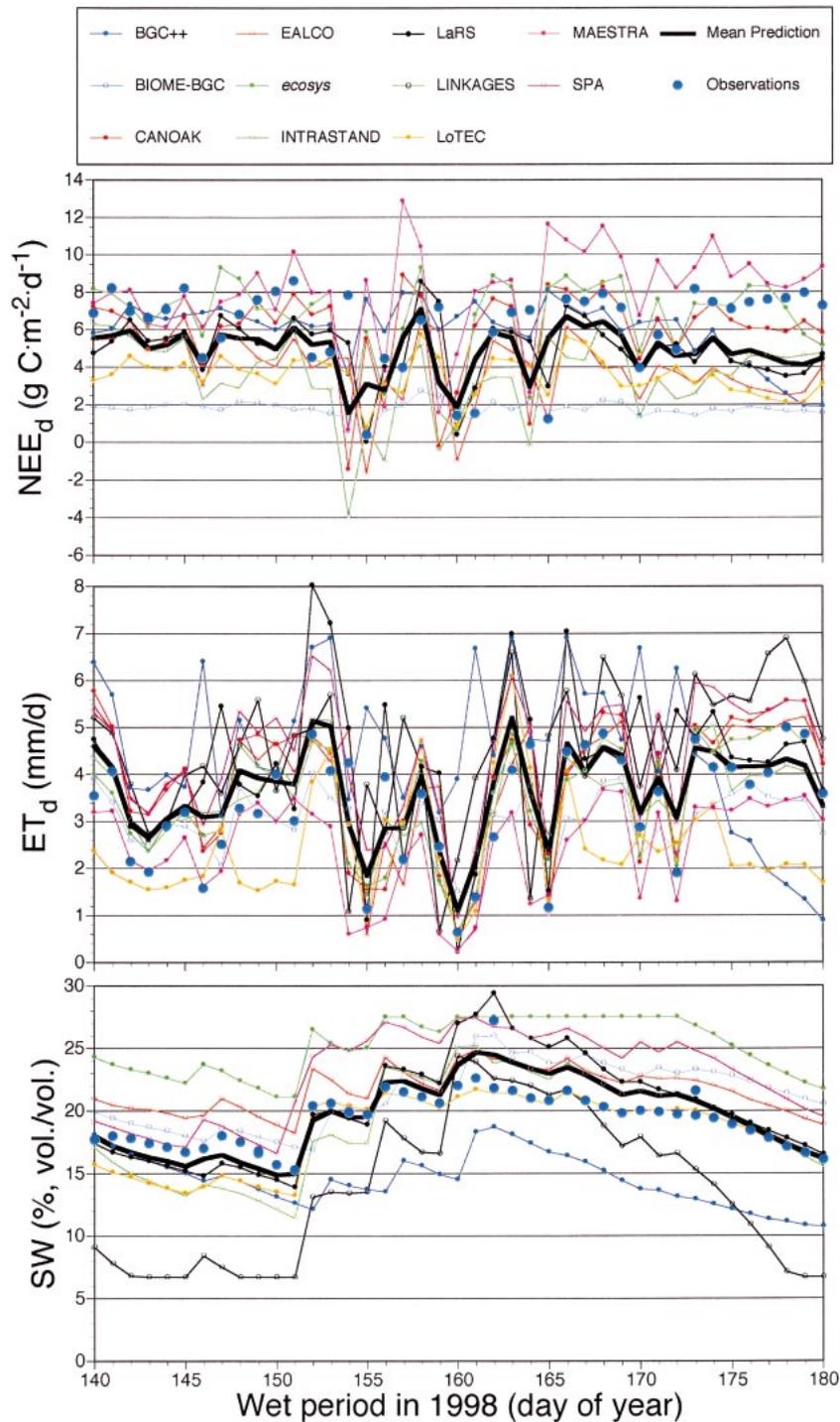
FIG. 7. Daily net ecosystem carbon exchange (NEE$_d$), evapotranspiration (ET$_d$), and soil water content (SW) from observations, selected models, or the mean daily model response for a 40-day ''wet'' period of nonlimiting soil water content in 1998.

With a few exceptions, most models predicted the extremes of wet and dry SW conditions shown in Figs. 7, 8, and 9. LINKAGES had the lowest minimum SW of all models and showed the greatest rate of drying

(Figs. 7 and 9). The highest SW values were typically generated by *ecosys* (Figs. 7 and 9), but under severe drought (Fig. 8) BGC++ maintained the highest SW values. Difference plots contrasting predicted vs. ob-
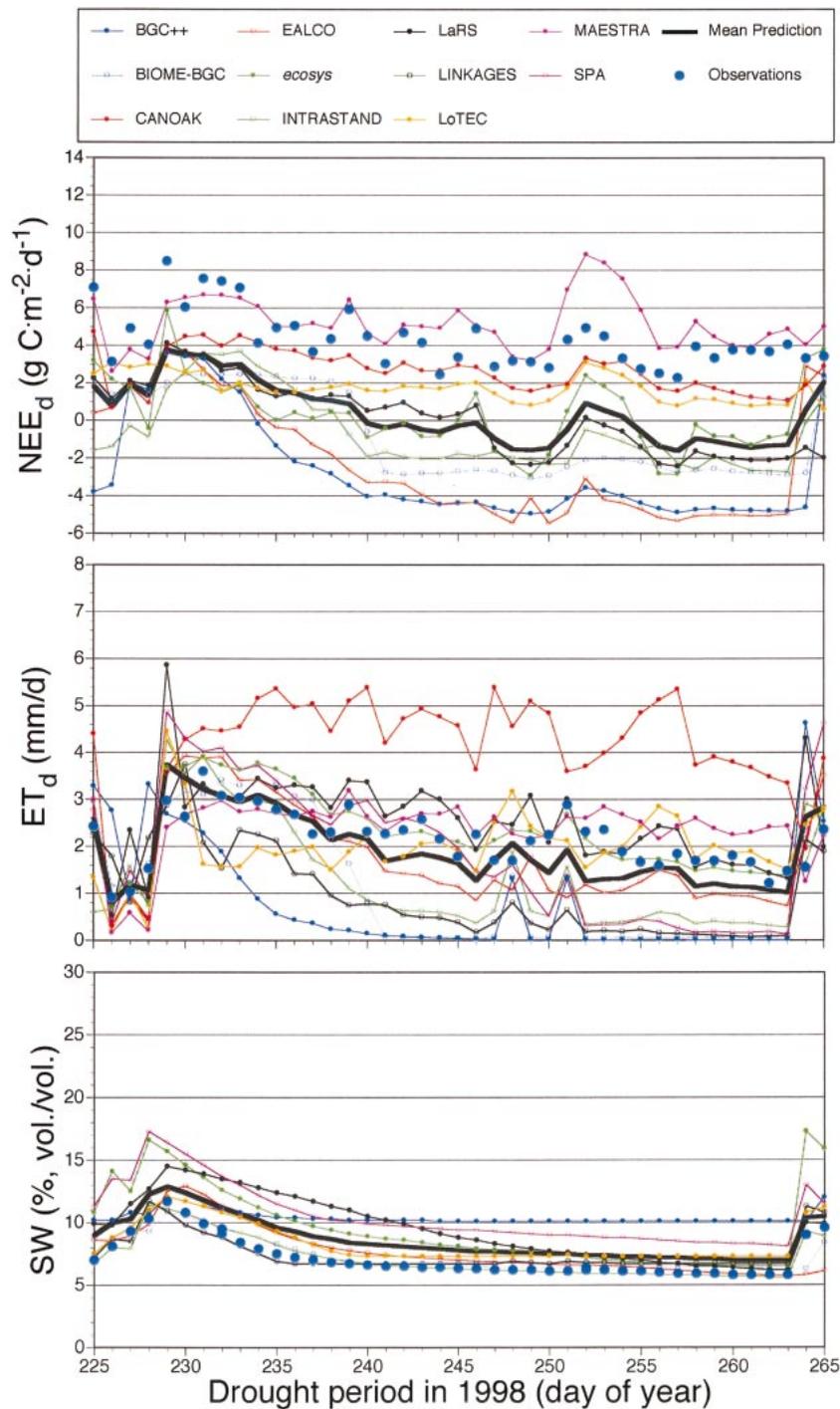
FIG. 8. Daily net ecosystem carbon exchange (NEE_d), evapotranspiration (ET_d), and soil water content (SW) from observations, selected models, or the mean daily model response for a 40-day ''dry'' period of limiting soil water contents in 1998.

served SW based on periodic time domain reflectometry (TDR) data (Fig. 12) and the performance statistics (Table 8) show good agreement with all models (EF values >0.5) and especially good agreement with BIOME-BGC (EF = 0.81), EALCO (EF = 0.83), IN-

TRASTAND (EF = 0.78) and LaRS (EF = 0.78). Difference plots for the continuous frequency domain reflectometry (FDR) data set (Fig. 13) and the performance statistics (Table 8) emphasized the inability of the models to capture day-to-day dynamics of soil wa-
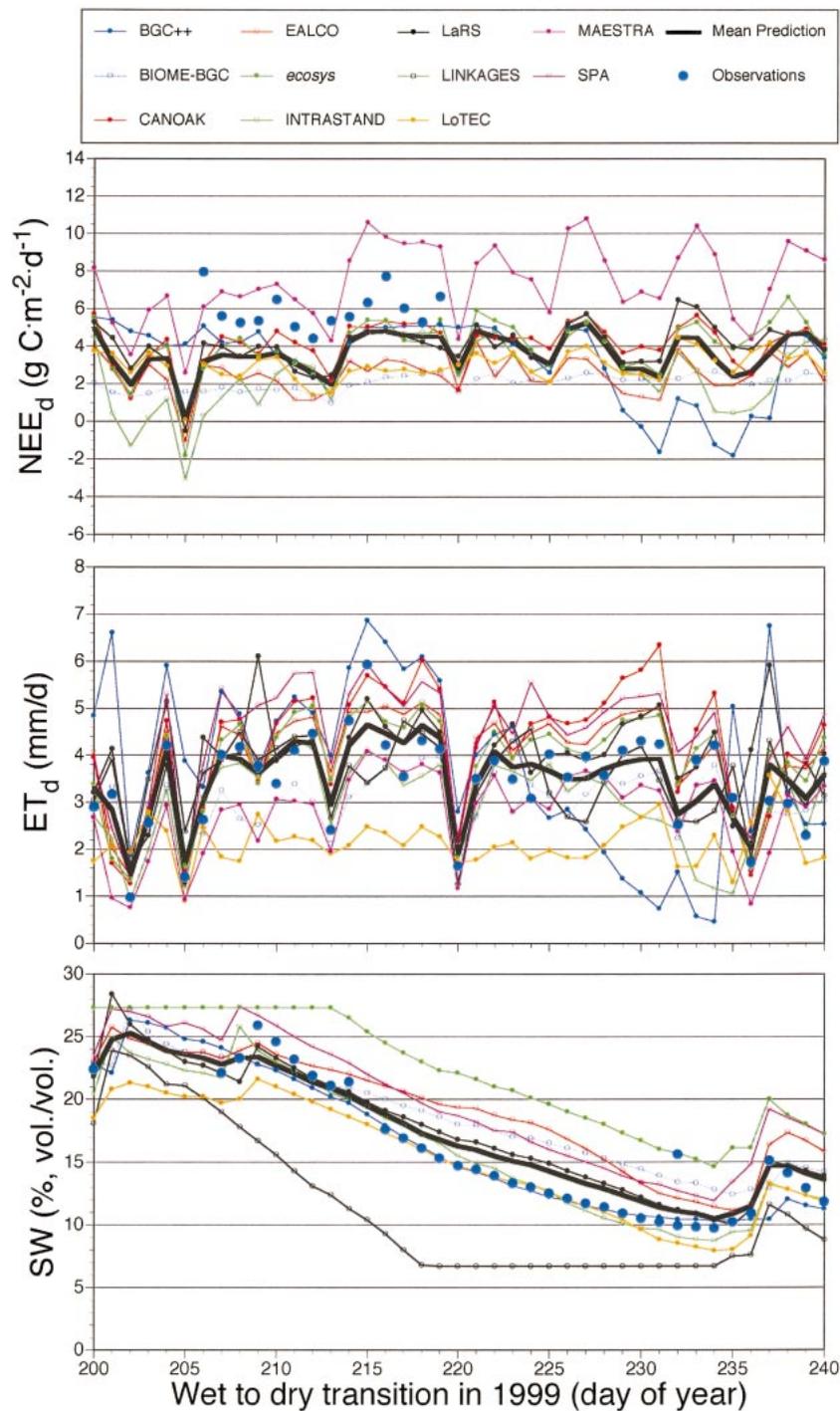
FIG. 9. Daily net ecosystem carbon exchange (NEE$_d$), evapotranspiration (ET$_d$), and soil water content (SW) from observations, selected models, or the mean daily model response for a 40-day transitional period going from nonlimiting to limiting soil water contents in 1999.

ter content changes. Those models demonstrating especially good performance against TDR also did well predicting data from FDR with EF values typically greater than 0.6.

Data from chamber and sub-canopy eddy covariance measurements of soil respiration (R_{soil}) were used to evaluate model predictions. Following rain events, the chamber-based measurements provided estimates of
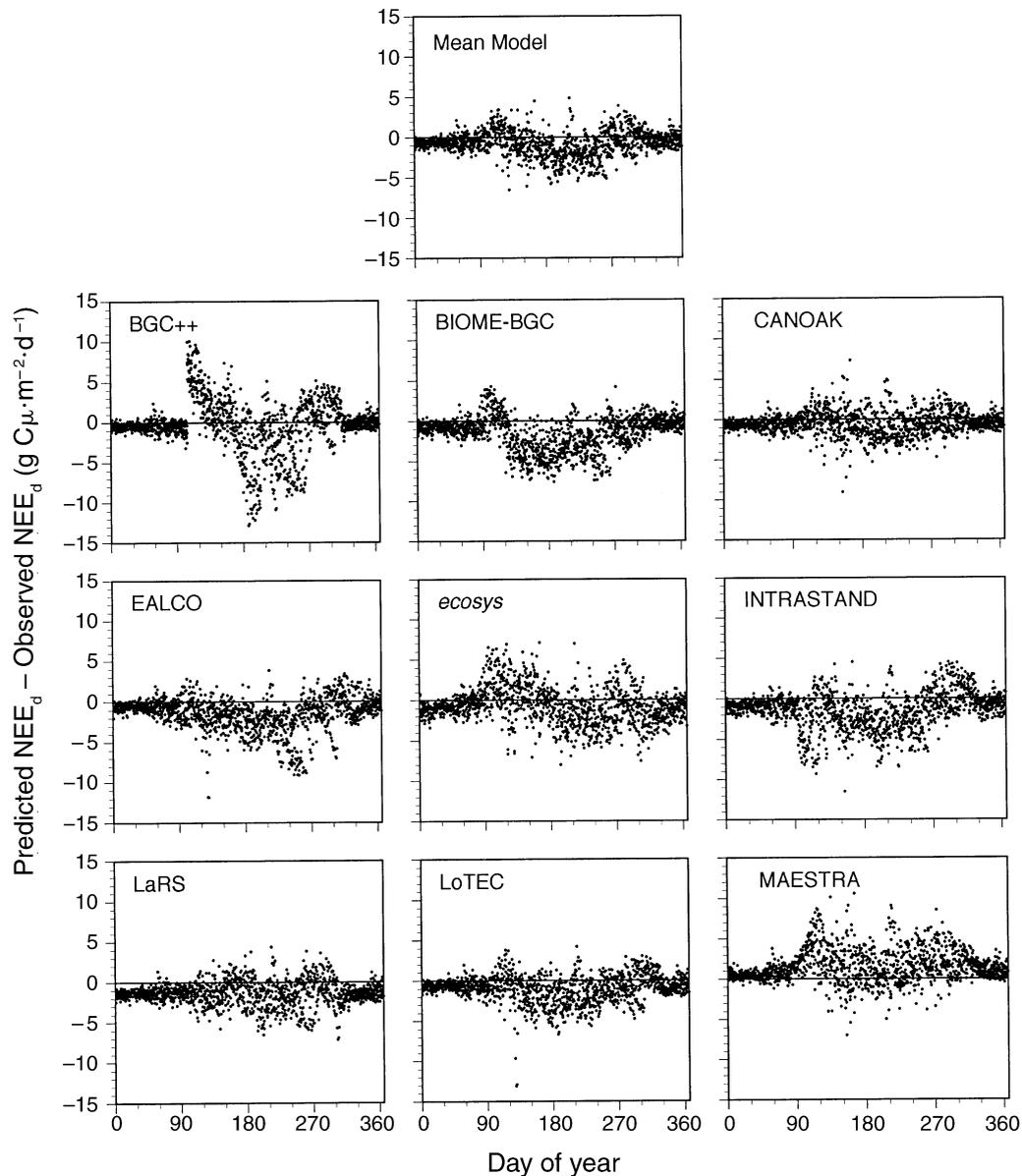
FIG. 10. Difference plot of simulated minus observed daily net ecosystem exchange, NEE_d (measured by eddy covariance, gap-filled) for nine models and the daily mean model output for the period from 1995 to 1998.

R_{soil} greater than those from sub-canopy eddy covariance measurements due to enhanced litter-layer decomposition (Hanson et al. 2003b). A plot of measured and model outputs for the growing season of 1999 (Fig. 14) revealed substantial differences between measured data sets and model simulations. Unlike the previous observations, the mean model prediction was not the best fit to the chamber-based measurements (EF = −0.30; Table 9), but it was nearly the best for the eddy covariance-based R_{soil} observations (EF = 0.47; Table 9 and Fig. 15). BIOME-BGC and CANOAK model predictions had slightly higher EF values for the eddy covariance-based R_{soil} measurements (0.49 and 0.48,

respectively). Most individual models and the mean model prediction did not capture the dynamics of the chamber R_{soil} data (Table 9 and Fig. 16). The INTRA-STAND model, specifically developed to capture litter-layer decomposition dynamics, does the best with an EF value of 0.57.

Daily mean model predictions for NEE_d, ET_d, R_{soil}, and SW are plotted in Fig. 17 along with available test data. For the models evaluated in this paper, the mean prediction appears to provide a robust approximation of available measurements for a number of variables, across multiple years, and over a wide range of environmental conditions. The agreement between mean

TABLE 6.   Model performance for predictions of daily net ecosystem exchange (NEE$_d$) from 1995 to 1998.

Model	Observed vs. simulated regression			Goodness-of-fit variables		
	Slope	Intercept (g C·m^{-2}·d^{-1})	R^2	Bias (g C·m^{-2}·d^{-1})	ABS (g C·m^{-2}·d^{-1})	EF
BGC++	0.65	0.19	0.30	−0.53	2.13	0.21
BIOME-BGC	0.44	−0.56	0.58	−1.71	2.04	−0.17
CANOAK	0.85	−0.43	0.80	−0.74	1.20	0.73
EALCO	0.64	−0.86	0.57	−1.60	1.89	0.29
ecosys	0.88	−0.66	0.62	−0.90	1.83	0.57
INTRASTAND	0.72	−1.09	0.55	−1.68	2.10	0.35
LaRS	0.92	−1.32	0.80	−1.50	1.72	0.64
LoTEC	0.67	−0.47	0.73	−1.16	1.49	0.45
MAESTRA	0.88	1.58	0.67	1.33	1.70	0.56
Mean model	0.74	−0.40	0.77	−0.94	1.35	0.59

Notes: Measures of performance include the slope and intercept from linear regression between observed and simulated data, the mean bias (Bias), the mean absolute bias (ABS; where values close to zero represent good agreement), and the model efficiency factor (EF; where a perfect fit equals 1).

model predictions and measured data suggests that annual integration of average predictions for variables lacking test data (e.g., GPP and autotrophic respiration) is a logical reference point for the evaluation of individual model performance.

Annual simulations

The efficacy of annual predictions from individual models was judged against mean model predictions and independent measurement data. A successful model prediction was considered to be one that captured the mean and interannual range of the multiyear response. Box-and-whisker plots were used as a simple mechanism for the comparisons of water and carbon cycle component predictions (Figs. 18–22). Tables 10–13 provide the annual and multiyear mean values from each model and measurement approach for the following variables: ET$_a$ (Table 10), Drainage (Table 11), NEE$_a$ (Table 12), and NPP (Table 13).

Water cycle components.—Annual model predictions for evapotranspiration (ET$_a$), transpiration (T$_a$), evaporation from canopy and soil surfaces (E$_a$), and drainage of excess water through the soil profile are plotted in Fig. 18. The annual, mean-model prediction for ET$_a$ ranged from 596–706 mm/yr with a multiyear mean of 650 mm/yr. Mean annual ET from watershed balance or eddy covariance observations yielded slightly lower values (613 and 601 mm/yr, respectively). The two measurement approaches had different multiyear ranges (492–778 and 515–658 mm/yr for watershed balance and eddy covariance, respectively).

SPA, BGC++, and CANOAK had higher, and LoTEC had lower than average multiyear ET$_a$. The interannual range in ET$_a$ documented in Fig. 18 shows good agreement between measurement and model values. However, individual years associated with minimum and maximum ET$_a$ were not the same for models and observations (Table 10). Watershed-based observations showed maximum and minimum ET$_a$ in 2000 and 1998, respectively, and eddy covariance based measurements showed maximum and minimum ET$_a$ in 1999 and 1995, respectively. The majority of models (7 of 13) simulated maximum ET$_a$ to be in 1997, but 1996 (BGC++, INTRASTAND), 1999 (CANOAK, EALCO, LaRS), and 2000 (SPA) were also projected to have maximum annual ET$_a$. Most models showed minimum ET$_a$ to be in the drought years of 1993, 1995, or 1998 in agreement with measured minimums. However, CANOAK had minimum ET$_a$ in 1996 and 1997, and LaRS had minimum ET$_a$ in 1994. The 1994, 1996, and 1997 years were all characterized as ''wet'' years (Hanson et al. 2003a) and were not expected to have minimum ET$_a$ values.

Nine models provided predictions of E$_a$, and the mean model prediction for E$_a$ was 200 mm/yr and had a limited range from 181 to 226 mm/yr over 8 years. Predicted E from BGC++ and LaRS exceeded this range. Predicted E$_a$ from CANOAK, LoTEC, and PnET-II was below the mean range.

The mean model prediction of T$_a$ from 10 models was 444 mm/yr and ranged from a low of 407 mm/yr in 1993 to a high of 486 mm/yr in 2000. CANOAK predictions of T$_a$ exceed this range, and LoTEC underpredicted the mean model T$_a$ values. Although the mean annual T$_a$ estimate from BGC++ (466 mm/yr) was very close to the multiple model mean, the interannual range of T$_a$ from BGC++ was over five times as great.

Ten models provided estimates of soil drainage. The interannual mean-model prediction covered a broad range from 530 to 1035 mm/yr with a multiyear mean of 710 mm/yr. Mean annual drainage from watershed balance measurements had a nearly identical range from 491 to 1020 mm/yr and a multiyear mean of 720 mm/yr. All models reported a similar interannual range for drainage. Mean annual drainage for LoTEC and
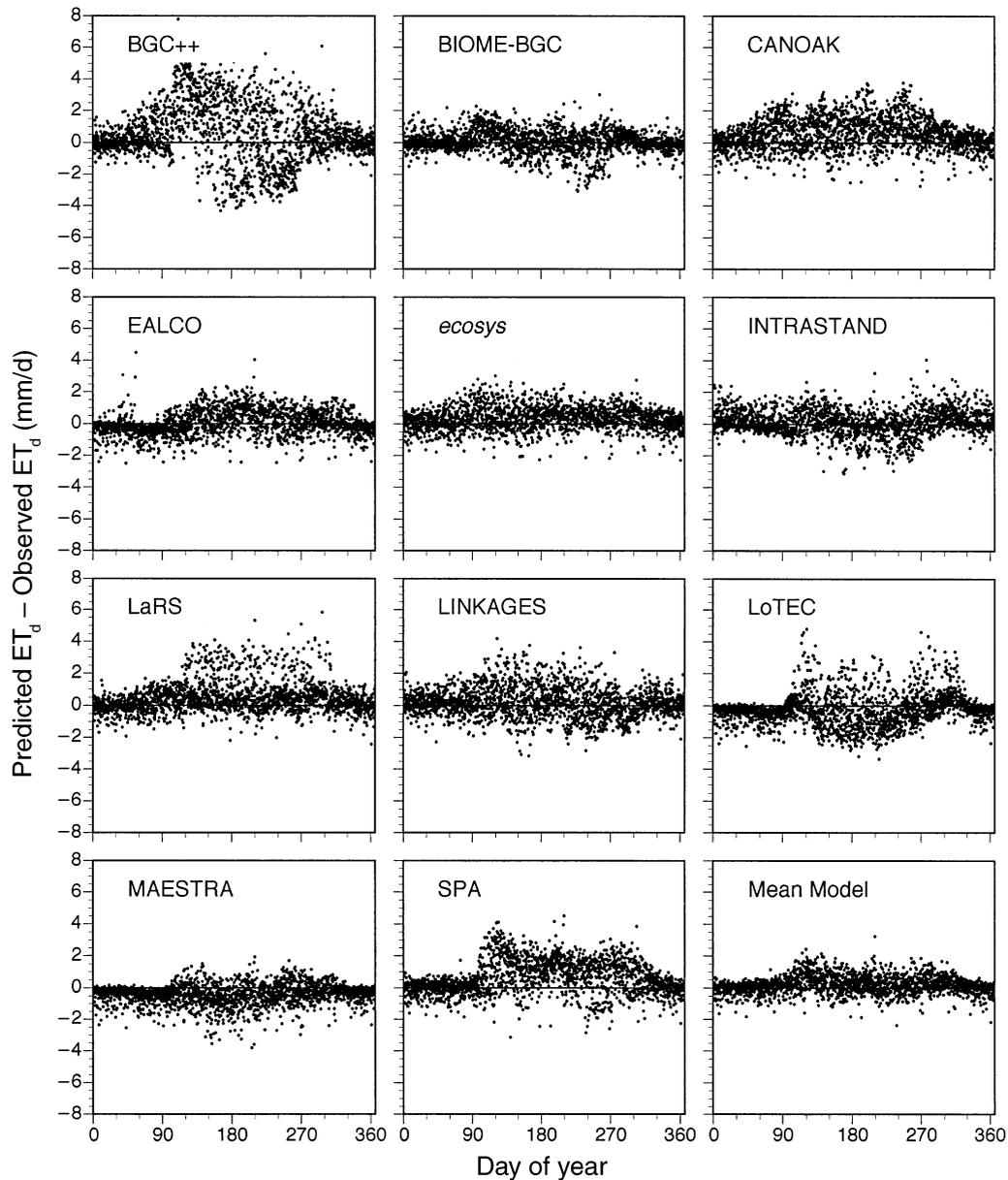
FIG. 11. Difference plot of simulated minus observed daily evapotranspiration (ET_d) measured by eddy covariance for 11 models and the daily mean model output (lower right graph) for the period from 1995 to 2000.

NuCM was more than 100 mm/yr greater than the "reference" value. BGC++ had annual drainage estimates that were >100 mm/yr lower than the references values. All models simulated maximum drainage in 1994 in agreement with the watershed balance measurement, and the watershed measurements showed minimum drainage in 2000 in agreement with 7 of 11 models (Table 11).

Carbon cycle components.—In contrast to the reasonable agreement among models for water cycle component predictions, estimates of NEE_a covered a wide range (Fig. 19). The mean model prediction for NEE_a was 281 g $C \cdot m^{-2} \cdot yr^{-1}$ with an interannual range from 150 to 368 g $C \cdot m^{-2} \cdot yr^{-1}$. The mean model predictions were higher than the high and low biometric estimates with an annual mean of 249 and 187 g $C \cdot m^{-2} \cdot yr^{-1}$, respectively. The biometric estimates of NEE_a had an overall range from 79 to 368 g $C \cdot m^{-2} \cdot yr^{-1}$. Gap-filled and nighttime-modeled estimates of NEE_a from eddy covariance data on Walker Branch were much higher, with a range from 528 to 710 g $C \cdot m^{-2} \cdot yr^{-1}$, than the mean model prediction or the biometric estimates of NEE_a. Of the nine models contributing annual estimates of NEE_a, only five models (BIOME-BGC, EALCO, IN-

TABLE 7. Model performance for predictions of daily evapotranspiration (ET) from 1995 to 2000.

Model	Observed vs. simulated regression			Goodness-of-fit variables		
	Slope	Intercept (mm/d)	R^2	Bias (mm/d)	ABS (mm/d)	EF
BGC++	1.00	0.50	0.34	0.50	1.38	0.32
BIOME-BGC	0.85	0.22	0.73	−0.03	0.51	0.71
CANOAK	1.17	0.21	0.73	0.49	0.83	0.67
EALCO	1.14	−0.28	0.80	−0.06	0.60	0.79
ecosys	1.03	0.21	0.79	0.26	0.58	0.77
INTRASTAND	0.83	0.32	0.67	0.03	0.59	0.64
LaRS	1.12	0.16	0.72	0.36	0.67	0.68
LINKAGES	0.92	0.23	0.64	0.10	0.70	0.63
LoTEC	0.75	0.07	0.52	−0.35	0.79	0.43
MAESTRA	0.82	−0.23	0.76	−0.53	0.65	0.61
NuCM	NA	NA	NA	NA	NA	NA
PnET-II	NA	NA	NA	NA	NA	NA
SPA	1.34	−0.04	0.79	0.52	0.87	0.69
Mean model	1.00	0.12	0.85	0.12	0.42	0.84

Notes: Measures of performance include the slope and intercept from linear regression between observed and simulated data, the mean bias (Bias), the mean absolute bias (ABS; where values close to zero represent good agreement), and the model efficiency factor (EF; where a perfect fit equals 1). NA = not available.
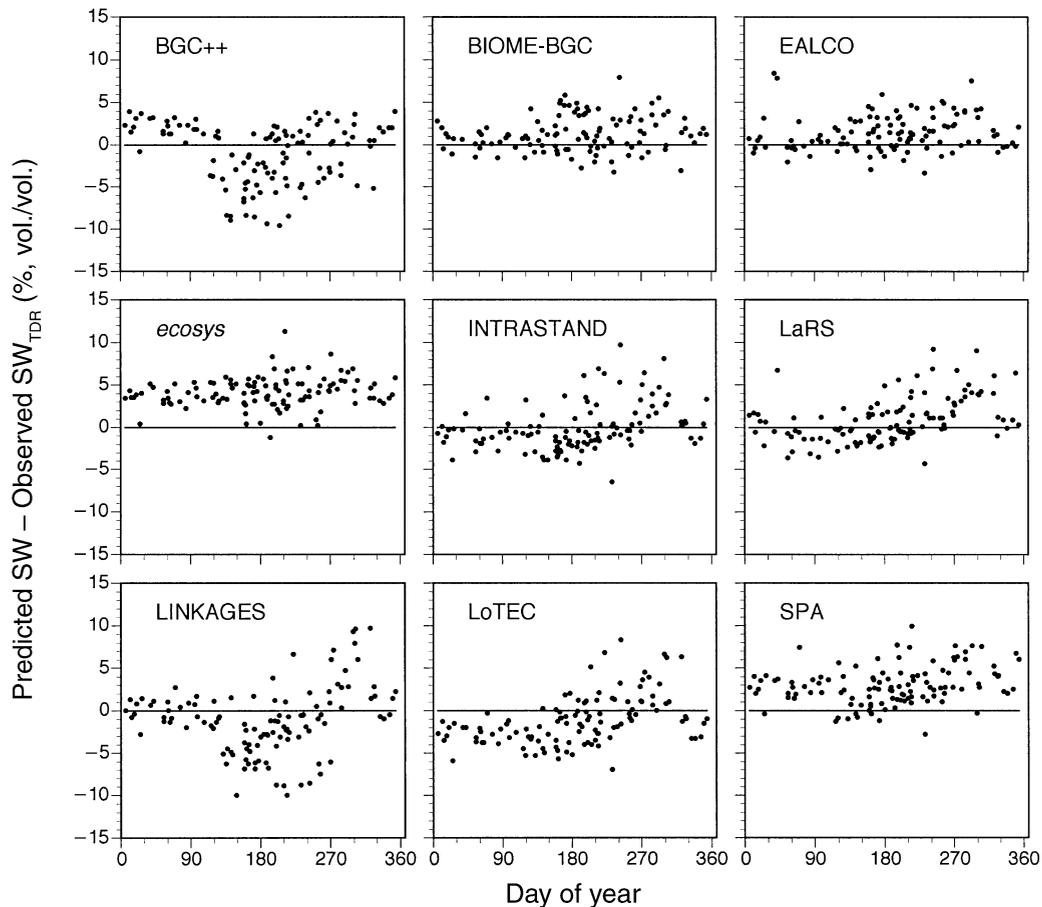
FIG. 12. Difference plot of simulated minus observed daily soil water content measured by time domain reflectometry (TDR) for nine models for the period from 1993 to 2000.

TABLE 8. Model performance for predictions of daily soil water content of the surface layer (0–35 cm) from 1993 to 2000 using two measurement approaches: periodic time domain reflectometry (TDR; $n = 100$) and hourly frequency domain reflectometry (FDR; $n = 6$).

| | Soil water content, measured using TDR | | | | | | Soil water content, measured using FDR | | | | | |
| | Observed vs. predicted regression | | | Goodness-of-fit variables | | | Observed vs. predicted regression | | | Goodness-of-fit variables | | |
Model	Slope	Intercept	R^2	Bias (%)	ABS (%)	EF	Slope	Intercept	R^2	Bias (%)	ABS (%)	EF
BGC++	0.90	0.8	0.70	−1.2	3.0	0.67	0.82	2.9	0.78	−0.2	2.5	0.74
BIOME-BGC	0.95	2.3	0.85	1.4	2.0	0.81	0.83	3.9	0.79	1.1	2.7	0.74
CANOAK	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
EALCO	0.9	3.2	0.88	1.3	1.8	0.83	0.79	4.7	0.86	1.2	2.5	0.77
ecosys	0.96	4.7	0.92	4.0	4.0	0.64	0.85	6.6	0.86	4.1	4.1	0.59
INTRASTAND	0.87	2.3	0.80	−0.2	2.1	0.78	0.81	3.9	0.80	0.7	2.5	0.75
LaRS	0.87	3.2	0.81	0.8	2.0	0.78	0.81	5.1	0.76	2.0	3.2	0.66
LINKAGES	0.91	0.4	0.66	−1.3	3.2	0.63	0.77	3.9	0.56	0.0	4.0	0.51
LoTEC	0.67	5.0	0.79	−1.3	2.6	0.55	0.65	5.4	0.76	−0.4	2.7	0.54
MAESTRA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NuCM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PnET-II	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SPA	0.94	4.0	0.87	2.8	3.0	0.71	0.88	5.6	0.83	3.6	3.8	0.63
Mean model	0.88	2.88	0.91	0.7	1.5	0.88	0.80	4.7	0.90	1.3	2.2	0.80

Notes: Measures of performance include the slope and intercept from linear regression between observed and simulated data, the mean bias (Bias), the mean absolute bias (ABS; where values close to zero represent good agreement), and the model efficiency factor (EF; where a perfect fit equals 1). Percentage bias and ABS are given on a volume/volume basis. NA = not available.

TRASTAND, LaRS, LoTEC) produced mean annual NEE$_a$ estimates that fell within the range of the mean model and biometric values. Three models (BGC++, CANOAK, and *ecosys*) produced higher estimates that were closer to the eddy covariance values. The PnET-II model's estimates of NEE$_a$ were unique with negative NEE$_a$ in 6 of 8 years.

The two measurements approaches disagreed on years having minimum or maximum NEE$_a$ (Table 12). Biometric NEE$_a$ values were highest in 1996 and lowest in 1998, while the eddy covariance method yielded the highest NEE$_a$ in 1999 and the lowest in 1995. The models showed no general agreement on the year having maximum NEE$_a$, but 7 of 10 models simulated minimum NEE$_a$ in 1998. Although LaRS predicted 1995 to be the year with lowest NEE$_a$, its value for 1998 was the second lowest. Only CANOAK and PnET-II showed 1998 to be a year of intermediate NEE$_a$.

Components of NEE$_a$ were plotted in an attempt to identify key processes responsible for the divergent estimates of NEE$_a$ (Figs. 20 and 21). The mean model prediction for GPP was 1693 g C·m^{-2}·yr^{-1} with a range of 1569–1817 g C·m^{-2}·yr^{-1}. The low NEE$_a$ estimates from the PnET-II model were explained by GPP estimates that were ~50% lower than GPP estimates from the other models. The highest GPP values were from the *ecosys* model with a mean of 2367 g C·m^{-2}·yr^{-1}. The mean model prediction for total autotrophic respiration (R_{auto}) was 919 g C·m^{-2}·yr^{-1} with a limited range from 880 to 959 g C·m^{-2}·yr^{-1}. Results from BGC++, BIOME-BGC, LaRS, and MAESTRA were closest to the mean model prediction for R$_{auto}$. EALCO, *ecosys,* INTRASTAND, and LoTEC had higher esti-

mates of R_{auto}, and CANOAK and PnET-II produced lower than average R_{auto} values.

The net difference between GPP and R_{auto} is NPP, and biometric estimates were available for comparison to NPP model simulations (Fig. 20). Mean annual NPP from biometric methods was 729 g C·m^{-2}·yr^{-1} with an interannual range from 604 to 840 g C·m^{-2}·yr^{-1}. The mean model prediction was similar with a mean NPP of 781 g C·m^{-2}·yr^{-1} and range from 624 to 878 g C·m^{-2}·yr^{-1}. As with NEE$_a$, the individual model estimates of NPP covered a wide range. The PnET-II model estimates of NPP were typically the lowest, and those for *ecosys* the highest. The simulated year with highest NPP was 2000 for 7 of 11 models, but the "wet" years 1996 (BGC++), 1997 (BIOME-BGC, LoTEC), and 1999 (CANOAK) were also years with maximum NPP (Table 13). The simulated year with lowest NPP was 1998 for 8 of 11 models. LaRS and PnET-II predicted the drought years of 1995 and 1993, respectively to be the years of lowest NPP. CANOAK was inconsistent with the other models simulating 1996 to be the year with lowest NPP.

Intermodel differences can be evaluated further by inspection of the leaf (R_{leaf}), stem (R_{stem}) and soil/root respiration (R_{soil}) plots in Fig. 21. Mean annual R_{leaf} at 321 g C·m^{-2}·yr^{-1} was consistently captured by most models. The *ecosys* model's estimate for R_{leaf} was more than twice the mean prediction at 662 g C·m^{-2}·yr^{-1}. BGC++ and CANOAK simulated low mean annual R_{leaf} values of 96 and 94 g C·m^{-2}·yr^{-1}, respectively. The mean model prediction for R_{stem} was 200 g C·m^{-2}·yr^{-1}, but MAESTRAs estimate for $R_{stem}$was over twice as high at 426 g C·m^{-2}·yr^{-1}. BIOME-BGC and PnET-II
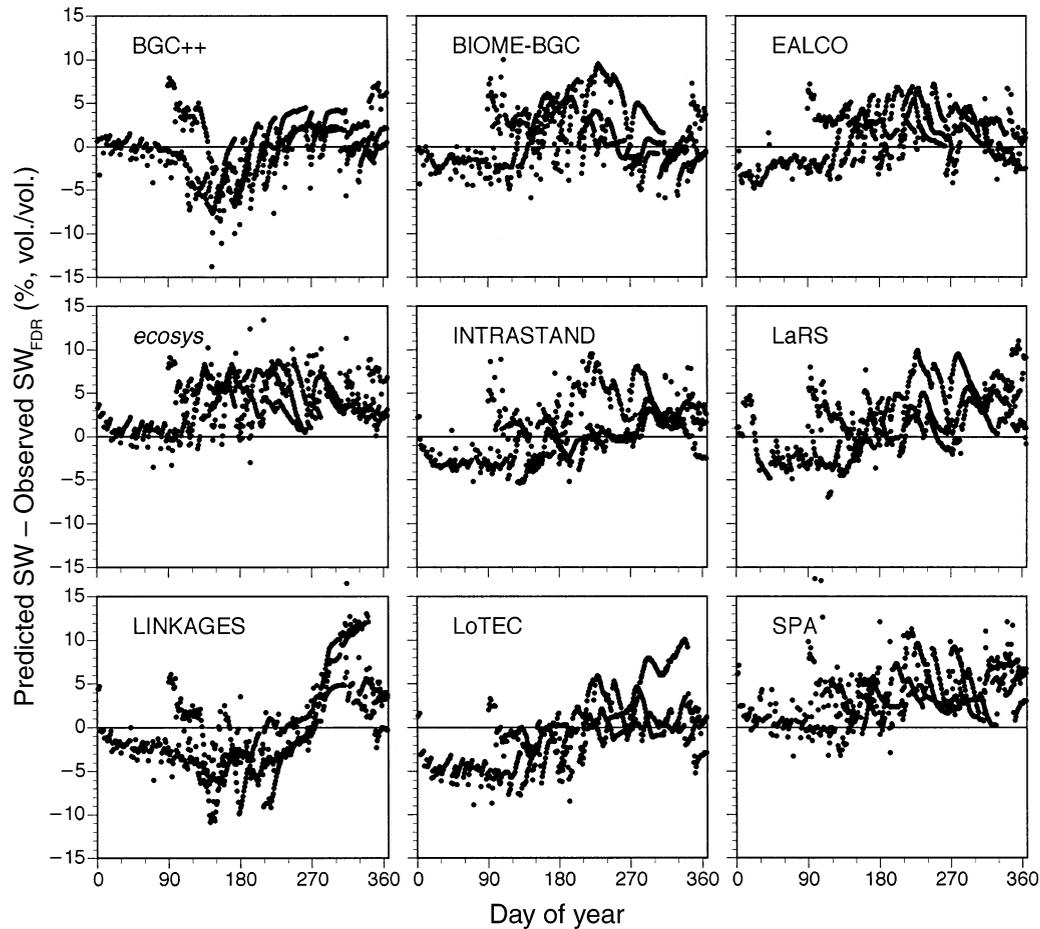
FIG. 13. Difference plot of simulated minus observed daily soil water content measured by frequency domain reflectometry (FDR) for nine models for the period from 1993 to 2000.

projected lower estimates of R_{stem} of 38 and 82 g C·m^{-2}·yr^{-1}, respectively. The mean annual model prediction for R_{soil} was 785 g C·m^{-2}·yr^{-1}. All model predictions, with the exception of BGC++ at 537 g C·m^{-2}·yr^{-1} and *ecosys* at 945 g C·m^{-2}·yr^{-1}, were within 115 g C·m^{-2}·yr^{-1} of that value.

Growth.—Independent measures of leaf, stem, and root growth were available for comparison with model predictions, but only half of the models provided estimates of growth for these components (Fig. 22). Observed annual leaf growth was 237 g C·m^{-2}·yr^{-1} and the mean model prediction was 201 g C·m^{-2}·yr^{-1}. BGC++, BIOME-BGC, *ecosys,* and LaRS simulated very similar leaf production, but PnET-IIs estimate of 111 g C·m^{-2}·yr^{-1} was only half the observed value. PnET-IIs dramatically reduced leaf production explains its unexpectedly low NEE$_a$ and GPP values (Figs. 19 and 20).

Measured annual stem growth ranged from 232 to 294 g C·m^{-2}·yr^{-1}, with the larger number including seed production along with aboveground wood. The interannual range of stem production varied ±63 g

C·m^{-2}·yr^{-1} around this average. Three models captured the magnitude and range of observed stem growth very well (LaRS, LoTEC, and PnET-II). BGC++ and *ecosys* predicted higher growth rates (377 and 341 g C·m^{-2}·yr^{-1}, respectively), but BGC++ had a dramatically larger range (±253 g C·m^{-2}·yr^{-1}). EALCO had mean annual stem growth similar to the observations (203 g C·m^{-2}·yr^{-1}), but a much wider range (±150 g C·m^{-2}·yr^{-1}). BIOME-BGC and LINKAGES gave lower stem growth rates (mean of 122 and 131 g C·m^{-2}·yr^{-1}, respectively).

Observed annual root growth was 153 g C·m^{-2}·yr^{-1}. Model predictions were either above (BGC++, BIOME-BGC, *ecosys,* LaRS) or below (LoTEC, PnET-II) the observed root growth data.

DISCUSSION

How good are the measured data for testing models?

A discussion of the reliability of field measurements is appropriate before the evaluation of model performance against field observations. Model parameteri-
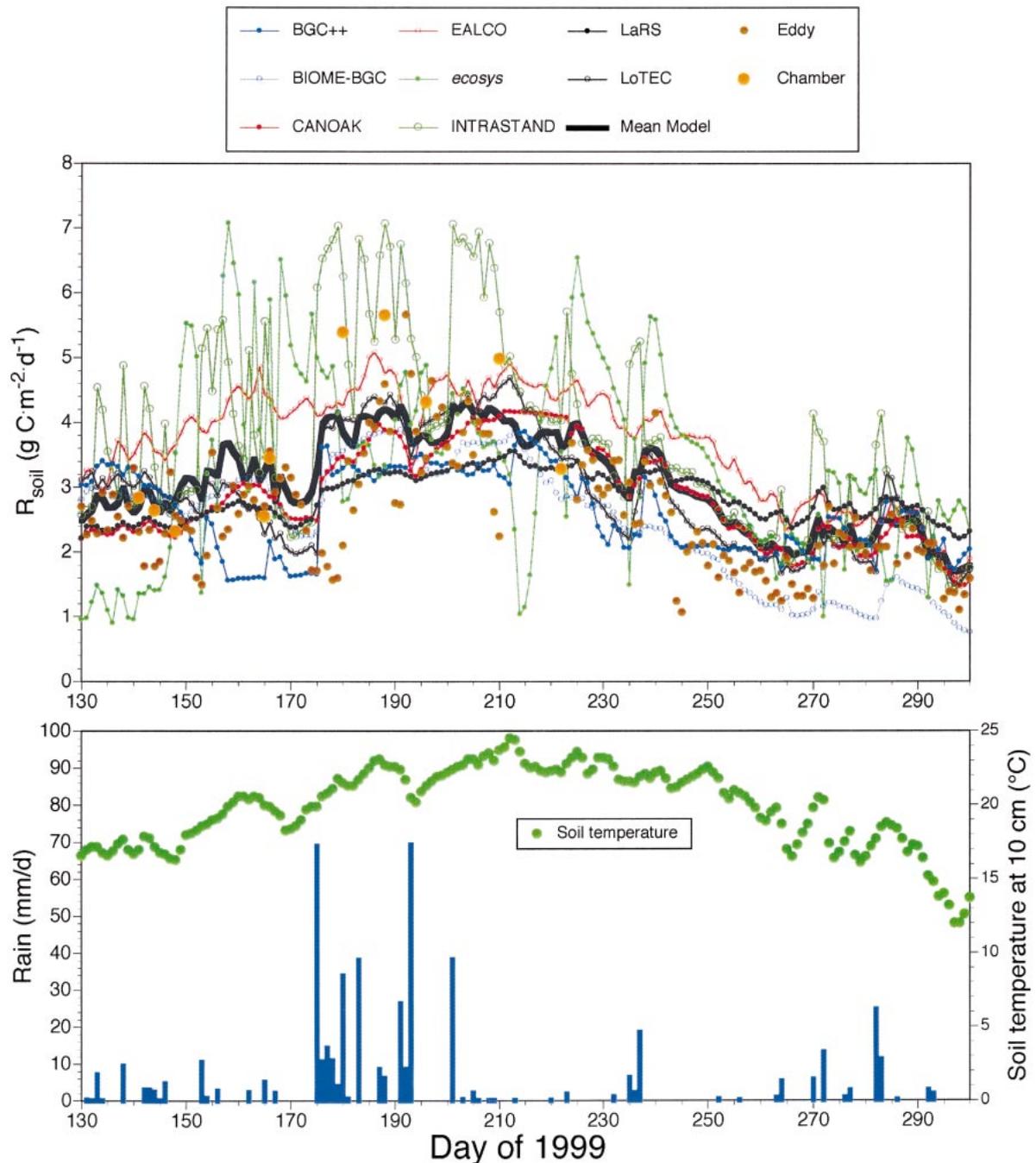
Fig. 14. Daily soil respiration (R_{soil}) from chamber observations (Chamber), 1.5 m eddy covariance observations (Eddy), selected models, or the mean daily model response for the growing season of 1999.

zation is typically based on data derived from the average of measurements derived from a single method, and the accuracy or uncertainty surrounding these methods is often ignored. Because alternative data are usually not available, investigators often assume that measured data are appropriate for model parameterization and/or testing and freely integrate such data over time. The model–data comparisons in this paper are not

immune to this assumption. However, where possible, more than one measurement approach was used to judge the utility of a model prediction. The following discussion provides information on known issues about the measured data before its application in the discussion of model performance.

The *Methods* description and *Results* section already suggested that eddy covariance data on NEE_h for Walk-

TABLE 9. Model performance for predictions of daily soil respiration (R_{soil}) based on 1.5 m eddy covariance ($n = 730$) or chamber-based observations ($n = 69$).

| | R_{soil} vs. eddy covariance | | | | | | R_{soil} vs. chamber | | | | | |
| | Observed vs. predicted regression | | | Goodness-of-fit variables | | | Observed vs. predicted regression | | | Goodness-of-fit variables | | |
Model	Slope	Inter-cept	R^2	Bias (g C·m^{-2}·d^{-1})	ABS (g C·m^{-2}·d^{-1})	EF	Slope	Inter-cept	R^2	Bias (g C·m^{-2}·d^{-1})	ABS (g C·m^{-2}·d^{-1})	EF
BGC++	0.67	0.53	0.38	−0.14	0.65	0.28	0.40	1.15	0.49	−0.84	1.01	−0.30
BIOME-BGC	0.79	0.40	0.53	−0.01	0.52	0.49	0.39	1.37	0.55	−0.65	0.92	−0.45
CANOAK	0.82	0.41	0.51	0.05	0.52	0.48	0.44	1.46	0.57	−0.82	1.19	−0.08
EALCO	1.06	0.49	0.43	0.62	0.99	0.35	0.52	1.94	0.48	0.35	0.84	0.06
ecosys	0.96	0.57	0.26	0.49	1.09	0.24	0.28	2.67	0.04	0.29	1.71	−0.25
INTRASTAND	1.16	0.34	0.40	0.67	0.96	0.33	0.87	0.92	0.65	0.49	0.78	0.57
LaRS	0.54	1.15	0.40	0.23	0.55	0.11	0.25	2.05	0.43	−0.42	0.96	−2.08
LINKAGES	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LoTEC	0.88	0.59	0.44	0.34	0.70	0.40	0.46	1.50	0.64	−0.27	0.74	−0.19
MAESTRA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NuCM	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PnET-II	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SPA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Mean model	0.86	0.56	0.53	0.28	0.58	0.47	0.45	1.63	0.67	−0.17	0.71	−0.30

Notes: Measures of performance include the slope and intercept from linear regression between observed and simulated data, the mean bias (Bias), the mean absolute bias (ABS; where values close to zero represent good agreement), and the model efficiency factor (EF; where a perfect fit equals 1). NA = not available.
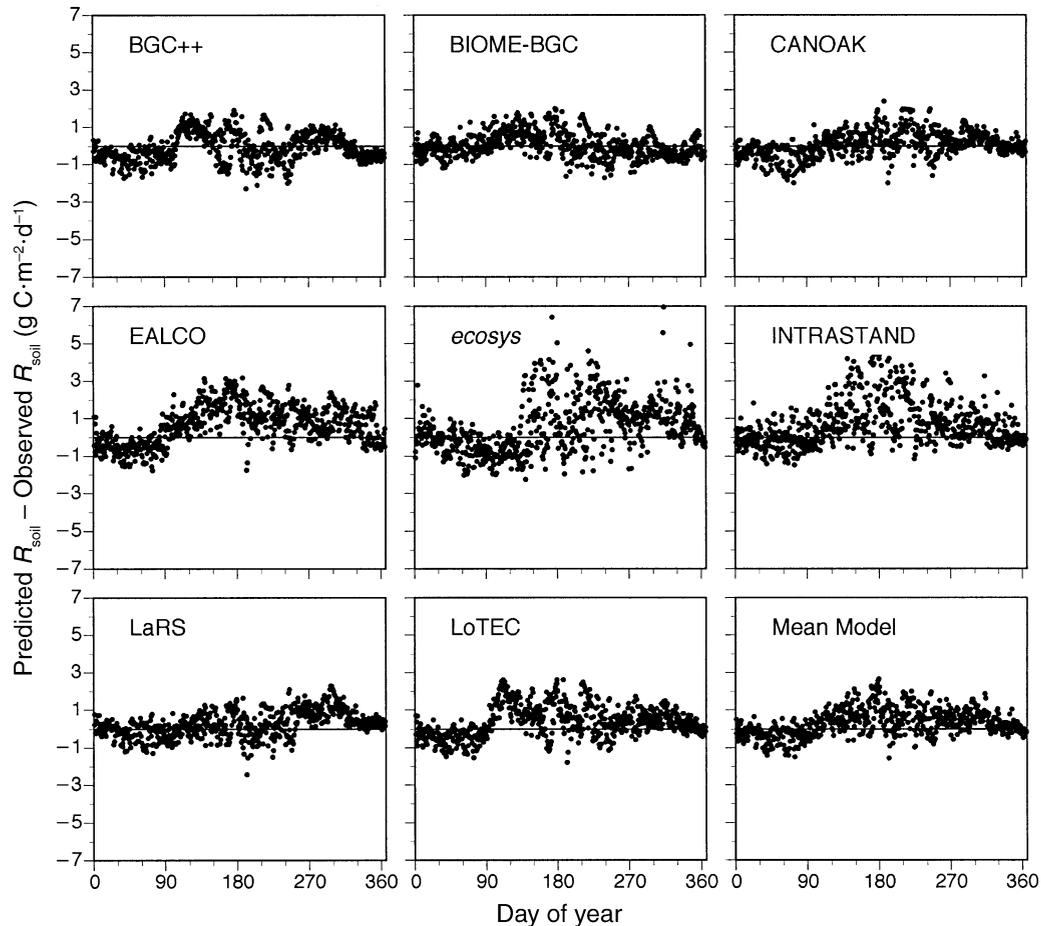
FIG. 15. Difference plot of simulated (predicted) minus observed daily soil respiration measured by 1.5 m eddy covariance for eight models and the mean model for the period from 1993 to 2000.
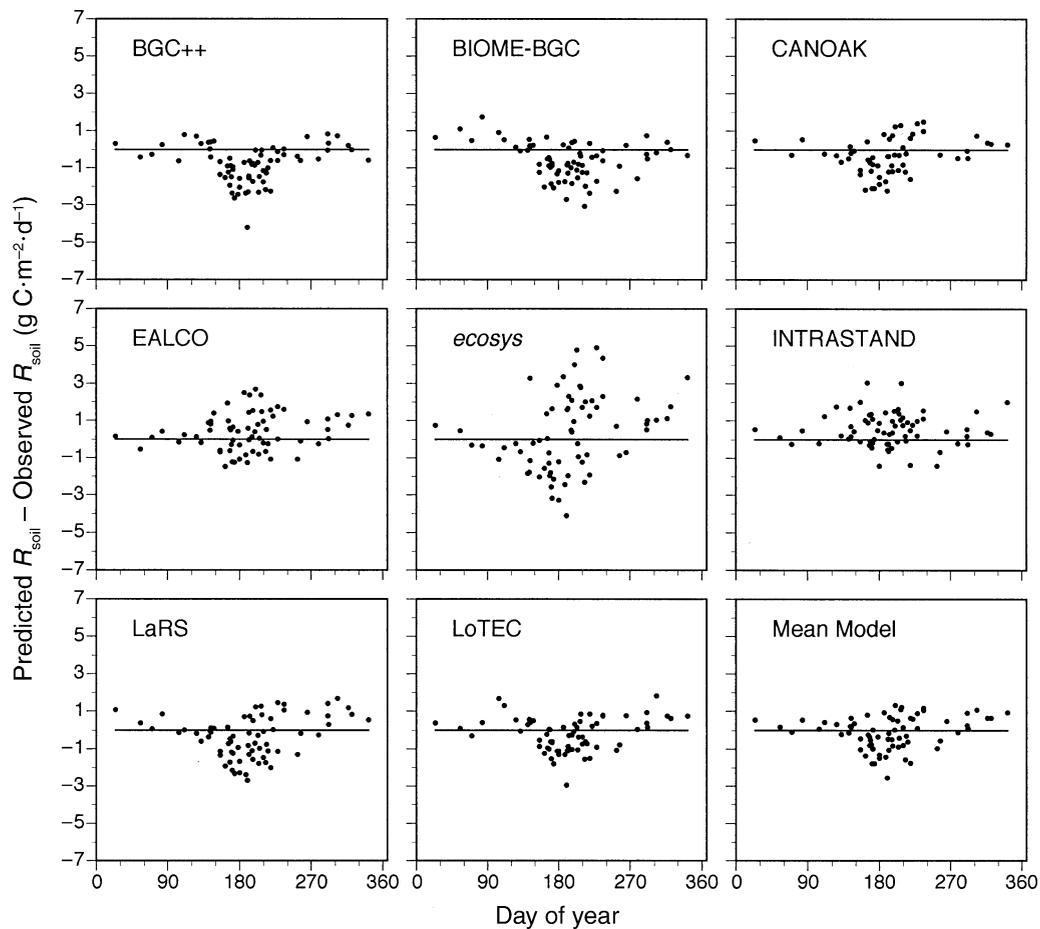
FIG. 16. Difference plot of simulated minus observed soil respiration measured by chamber methods for eight models and the mean model for the period from 1993 to 2000.

er Branch Watershed was subject to error due to low night turbulence and sloping terrain. The known problems for eddy covariance data on Walker Branch result in an underestimation of ecosystem respiration at night and an over estimate of integrated daily and annual NEE. The night turbulence and terrain issues associated with carbon flux are considered inconsequential to eddy-covariance-based estimates of ET because the error is introduced at night when ET values are near zero, thus minimizing any introduced error (Wilson and Baldocchi 2000). Therefore, while NEE_h measurements from Walker Branch eddy covariance are expected to yield overestimates of carbon uptake, the estimates for ET are interpreted as a reasonable and accurate approximation of water flux. Wilson et al. (2001b) showed that integrated annual ET from eddy covariance was in good agreement with watershed-scale hydrologic balance in support of this conclusion. In addition, Wilson et al. (2002) evaluated energy balance closure for 22 eddy covariance sites (including Walker Branch) and found that typically only 80% of the energy flux was captured. The missing 20% does suggest the po-

tential for the eddy covariance data to underestimate integrated daily or annual ET.

Sapflow-based estimates of T were derived from the summation of individual tree sapflow velocities multiplied by their species-specific estimates of stand sapwood area (Wullschleger et al. 2001). Wilson et al. (2001b) showed that this method underestimated actual transpiration for the Walker Branch forest. The relative response to drought, however, was captured by this method and is important for testing the ability of models to predict drought responses. Because of the underestimation of transpiration by the sapflow method, model-predicted T values show a positive bias when compared to the sapflow-based measurements of T (Table 5).

Chamber and understory eddy covariance-based measurements of R_{soil} are known to differ in magnitude and diurnal dynamics (Norman et al. 1997, Janssens et al. 2000, 2001, Law et al. 2001a). Some articles find better agreement than others, but all conclude that photosynthesis and respiration of aboveground plant material between the forest floor and the height of the
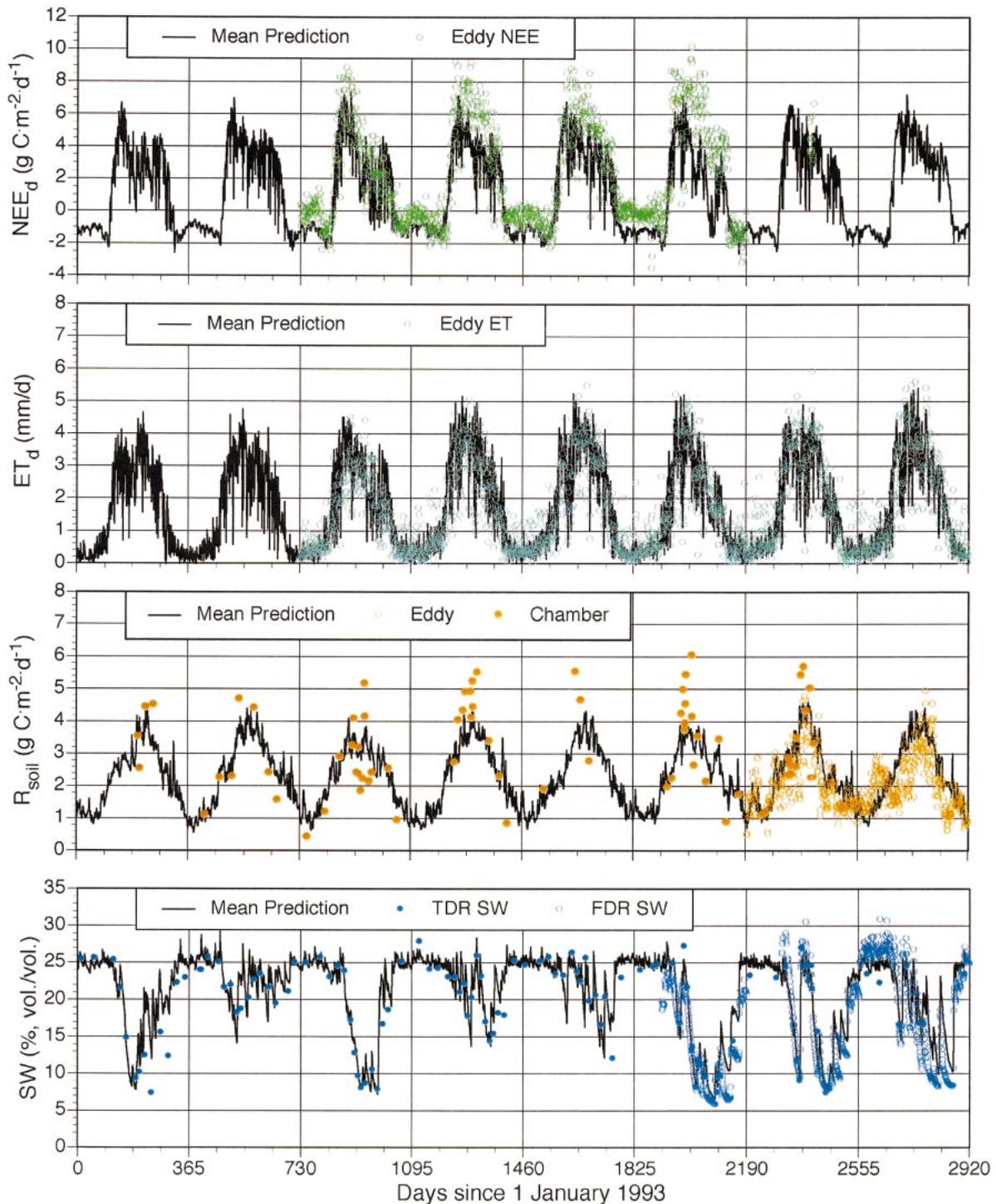
FIG. 17.   Mean model prediction for daily net ecosystem carbon exchange (NEE$_d$), evapotranspiration (ET$_d$), soil respiration (R$_{soil}$), and soil water content (SW), plotted together with available daily observations from 1993 to 2000.

short tower (typically 1–2 m) are responsible for observed differences between techniques. Janssens et al. (2001) found better agreement between modeled chamber data and nocturnal understory-eddy-covariance measurements. In this paper, both types of measure-

ments are used to judge model predictions of R$_{soil}$. Chamber measurements captured short-term dynamics of respiration of the litter layer driven by dynamic wetting and drying of the forest floor (Hanson et al. 2003b). This dynamic phenomenon was not routinely captured
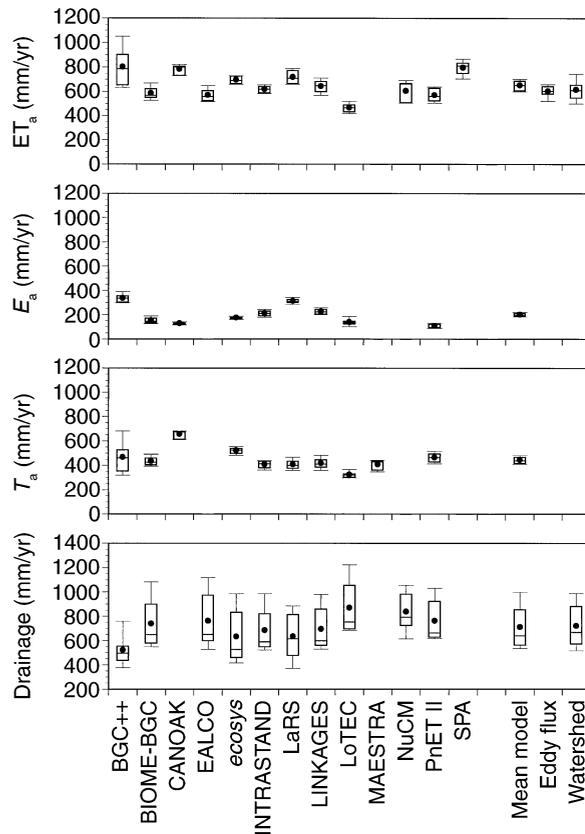
FIG. 18. Mean (solid circle), median (line), and range of annual evaporation (E_a), transpiration (T_a), evapotranspiration (ET_a), and drainage for Walker Branch from 1993 through 2000 for 9–12 models, the mean for all models, eddy covariance-based estimates of ET, and watershed balance estimates of ET.

by the nocturnal understory-eddy-covariance measurements because drying was largely a daytime phenomenon.

Measurements of mean soil water content by TDR and FDR were shown to accurately capture seasonal dynamics of the soil water content by volume (Hanson et al. 1998, 2003a). However, the depth increment or time step represented by the available measurements was not always in agreement with model assumptions. In addition, TDR and FDR are sensitive to rock volume (assumed to contain no water), and individual models may not be configured for the presence of rocks (i.e., rocks reduce the water storage capacity of a given soil layer). Soil water content measured by TDR in the field represents a point-in-time observation of water from 0 to a depth of 35 cm, but the model outputs could be skewed because they are daily averages and because a model may provide soil water values for different depth increments. The temporal wetting and/or drying dynamic captured by continuous FDR soil water measurements for a specific depth (i.e., horizontally installed probes) is responsible for lack-of-fit for nearly all model outputs (Fig. 12), because the models predict wetting and/or drying for a broader soil volume, not specific depths.

Biometric estimates of NEE_a (i.e., NEP) were based on direct measurements of NPP minus two alternative estimates of heterotrophic respiration. The direct measurement of NPP included measured canopy production (leaves, flowers, seeds), estimates of aboveground growth from site-specific allometry, and estimates of belowground production from a variety of techniques. The NPP estimate may be underestimated to some degree because it doesn't account for annual carbon production used in the growth and turnover of mycorrhizal fungi and root exudates. Correcting such an error would increase the biometric NEEa value, but the magnitude of such a correction is currently unknown. The low and

TABLE 10. Modeled and measured evapotranspiration (ET; mm/y) for 1993–2000. Numbers in brackets are the maximum and those underlined are the minimum annual ET values. NA = not available.

Model or measurement	1993	1994	1995	1996	1997	1998	1999	2000	8-year mean
BGC++	644	843	666	[1090]	964	628	773	800	801
BIOME-BGC	566	556	518	607	[682]	546	572	640	586
CANOAK	NA	NA	804	733	733	794	[823]	808	783
EALCO	516	529	523	572	577	543	[653]	639	569
ecosys	686	668	665	696	[730]	656	727	726	694
INTRASTAND	609	627	586	[657]	645	582	595	636	617
LaRS	666	657	668	710	760	706	[792]	786	718
LINKAGES	604	676	555	674	[720]	621	599	691	642
LoTEC	411	429	464	466	[522]	436	471	507	463
NuCM	NA	662	668	668	[695]	502	509	519	603
PnET-II	495	638	520	617	[636]	529	555	565	569
SPA	767	817	736	819	803	691	851	[873]	795
Mean model	596	646	614	692	[706]	603	660	683	650
Watershed estimate	581	629	592	671	518	492	642	[778]	613
Eddy covariance	NA	NA	515	584	624	583	[658]	644	601

TABLE 11. Modeled and measured drainage (mm/yr) for 1993–2000. Numbers in brackets are the maximum and those underlined are the minimum annual drainage values.

Model or measurement	1993	1994	1995	1996	1997	1998	1999	2000	8-year mean
BGC++	495	[831]	477	519	496	592	398	<u>366</u>	522
BIOME-BGC	576	[1118]	626	999	800	670	588	<u>537</u>	740
EALCO	624	[1145]	623	1052	897	675	<u>505</u>	575	762
ecosys	<u>400</u>	[1007]	475	931	737	567	488	448	632
INTRASTAND	<u>512</u>	[1030]	557	876	766	598	579	545	683
LaRS	607	[895]	558	769	862	622	355	<u>402</u>	634
LINKAGES	549	[981]	597	979	742	599	578	<u>519</u>	693
LoTEC	728	[1245]	<u>683</u>	1170	941	781	687	707	868
NuCM	NA	[1065]	999	793	926	775	710	<u>591</u>	837
PnET-II	634	[1036]	640	1020	827	686	628	<u>610</u>	760
Mean model	569	[1035]	624	911	799	657	552	<u>530</u>	710
Watershed estimate	574	[1020]	579	904	859	758	576	<u>491</u>	720

high estimates of heterotrophic respiration bracket measured annual CO_2 fluxes from decomposition and can be reconciled with alternate measures of soil carbon turnover reported for Walker Branch soils (Gaudinski and Trumbore 2003).

Notwithstanding these known issues associated with field measurements, observational data remain the basis for model development and parameterization, and independent measurements of key model processes or endpoints are the only viable means for judging model performance. Exact agreement between models and mean observations, however, is not a reasonable expectation. The variance and confidence interval around measured data typically represent a wide range of real-world possibilities, leaving ample room for acceptable model agreement. Only models that diverge widely from such data should be considered inappropriate and in need of improvement. Perhaps more important than quantitative agreement between data and model outputs is the ability of models to capture temporal dynamics

and interrelationships between modeled processes and environmental drivers (i.e., bias can be simply adjusted through parameterization). Therefore, good correlation between model predictions and observations is essential. The EF statistic used in this paper captures both magnitude and correlation issues associated with goodness of fit.

How well do models compare to one another and independent measurements?

Water cycle.—Hourly predictions of T varied between models. For the hourly transpiration data under wet conditions (Figs. 5 and 6 and Table 5), the magnitude of CANOAK, EALCO, *ecosys,* and SPA T_h simulations was higher than expected. LoTEC, a nonlayered big-leaf model, had T_h values lower than all of the other models (EF = −1.12; Table 5). Big-leaf models parameterized from individual leaf observations will produce biased model outputs if they don't account for natural gradients in conductance and VPD that oc-
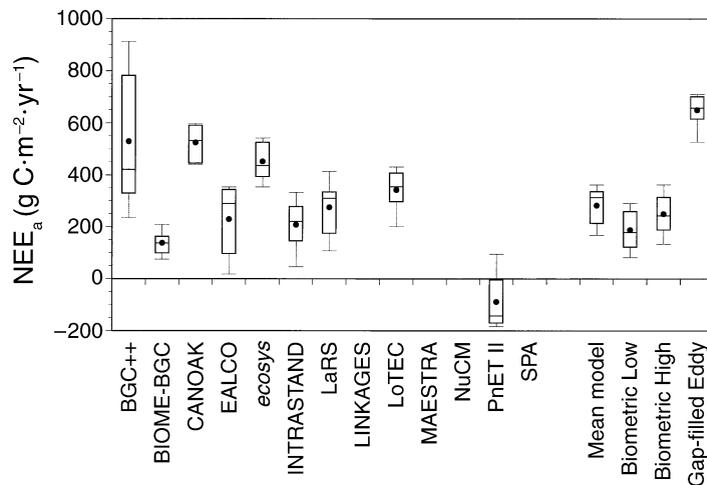
FIG. 19. Mean (solid circle), median (line), and range of annual net ecosystem production (NEE$_a$) for Walker Branch from 1993 through 2000 for nine models, the mean of all models, high and low biometric measurements, and eddy covariance observations. Biometric NEE$_a$ is an estimate based on measured net primary production minus high or low modeled estimates of soil respiration for the Walker Branch forest (see *Experimental observations for testing models*).

TABLE 12. Modeled and measured annual net ecosystem exchange (NEE$_a$; g C·m^{-2}·yr^{-1}) for 1993–2000. Numbers in brackets are the maximum and those underlined are the minimum annual NEE$_a$ values.

Model or measurement	1993	1994	1995	1996	1997	1998	1999	2000	8-year mean
BGC++	311	823	351	[951]	743	203	416	427	528
BIOME-BGC	155	114	86	142	[225]	70	133	171	137
CANOAK	NA	NA	525	440	447	538	[598]	591	523
EALCO	153	353	41	282	[354]	7	298	336	228
ecosys	371	453	420	418	518	346	[545]	534	451
INTRASTAND	[354]	271	201	132	239	8	162	285	207
LaRS	243	345	107	302	326	108	319	[443]	274
LoTEC	270	325	389	342	[434]	172	368	427	341
PnET-II	−173	−165	−142	−189	−145	−101	[96]	93	−91
Mean model	210	315	220	313	349	150	326	[368]	281
Biometric low	89	194	234	[287]	165	79	292	158	187
Biometric high	125	257	280	352	223	157	[368]	230	249
Eddy covariance	NA	NA	528	646	700	658	[710]	NA	648

Notes: Low and high biometric estimates are derived from two different methods of estimating soil respiration as described in the text. Gap-filled eddy covariance estimates are from Falge et al. (2001). NA = not available.

cur in forest canopies. Wullschleger et al. (2002) measured individual-leaf and canopy conductance in a closed-canopy *Liquidambar* stand and observed that canopy conductance expressed per unit leaf area was lower than one would expect from the extrapolation of individual leaf conductance.

All models with the exception of CANOAK showed substantial reductions in T$_h$ under severe drought conditions (Fig. 5 and Table 5). CANOAKs lack of a simulated soil water pool is the logical reason for its inability to capture drought feedbacks. MAESTRA, which also lacked a soil water feedback, did capture a drought response because of a more sensitive stomatal
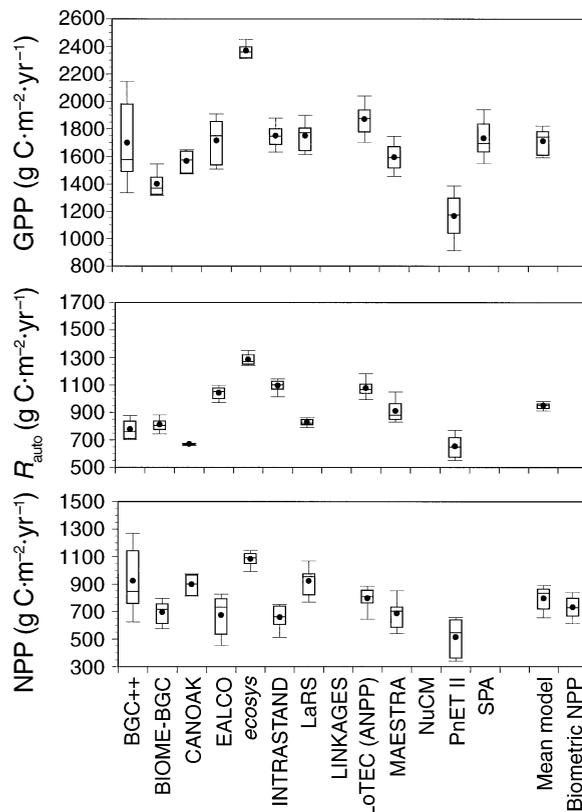
FIG. 20. Mean (circle), median (line), and range of the components of net primary production (NPP), including gross primary production (GPP) and autotrophic respiration (R_{auto}), for Walker Branch from 1993 through 2000 for 10 or 11 models, the mean of all models, and biometric measurements.
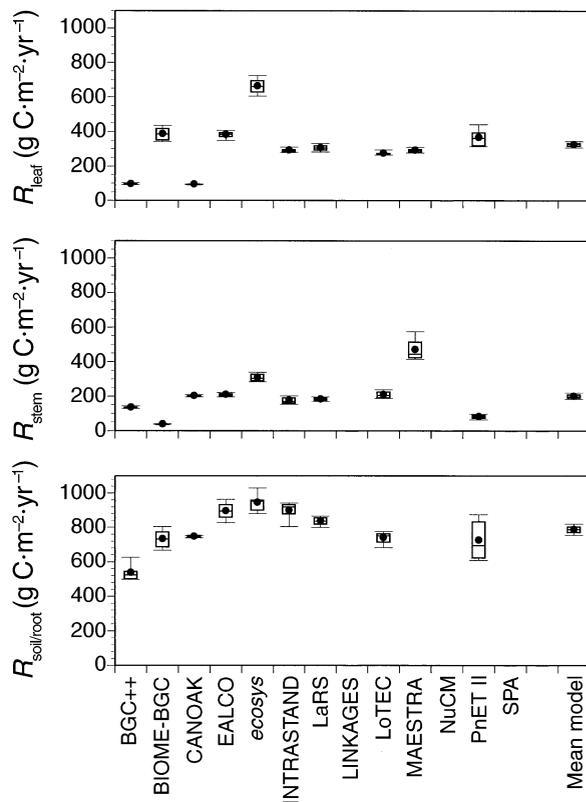
FIG. 21. Mean (solid circle), median (line), and range of the components of ecosystem respiration including leaf, stem, and the root/soil combination for Walker Branch from 1993 through 2000 for 10 models and the mean of all models.

TABLE 13. Modeled and measured annual net primary production (NPP; g C·m^{-2}·yr^{-1}) for 1993–2000. Numbers in brackets are the maximum and those underlined are the minimum annual NPP values. NA = not available.

Model or measurement	1993	1994	1995	1996	1997	1998	1999	2000	8-year mean
BGC++	893	1178	726	[1310]	1110	580	797	794	923
BIOME-BGC	640	759	586	750	[812]	570	681	753	694
CANOAK	NA	NA	894	812	817	909	[976]	968	896
EALCO	566	808	508	689	772	430	776	[834]	673
ecosys	1067	1116	1048	1056	1123	966	1125	[1155]	1082
INTRASTAND	747	732	626	587	698	479	627	[754]	656
LaRS	858	971	760	938	974	786	977	[1108]	922
LoTEC	734	846	795	826	[892]	604	791	864	794
MAESTRA	622	743	723	690	554	534	716	[895]	685
PnET-II	340	622	344	525	571	383	658	[657]	513
Mean model	718	864	701	818	832	624	812	[878]	781
Biometric estimate	604	762	742	833	713	638	[840]	703	729

responses to VPD. INTRASTAND, parameterized from leaf data, had the largest reductions in T_h under drought, a negative bias, and the lowest EF value (EF = −0.55; Table 5). A simple adjustment of the stomatal sensitivity to soil drying in line with the observed changes

FIG. 22. Mean (solid circle), median (line), and range of modeled vs. observed growth of leaf, stem, and root from 1993 through 2000 for 5–7 models, the mean of all models, and direct measurements. Observed leaf growth is from measured annual litter basket leaf mass adjusted for annual herbivory amounts.

in T_h was able to correct this bias (data not shown). The relative response of whole-plant measurements that integrate across sun and shade leaves and canopy position (e.g., sapflow), may be a better metric from which to parameterize drought sensitivity in stand-level models than instantaneous leaf chamber data.

Simulations for ET_d were in good agreement with measured data for most models especially under wet (i.e., optimum) moisture conditions (Figs. 7, 8, and 9) when daily values approached a maximum between 4 and 5 mm/d (Fig. 7). At the depth of drought most models simulated ET_d minimums between 0.5 and 2 mm/d (Fig. 8). CANOAK overestimated ET_d under extreme drought conditions, but achieved reasonable values otherwise (EF = 0.67; Table 7, Fig. 11). LoTEC with a negative bias and BGC++ with a positive bias had the least agreement with the observations (EF of 0.43 and 0.32, respectively, Table 7), and substantial deviations from the measured data during the growing season (Fig. 11). LaRS had one of the better EF values at 0.68 (Table 7), but the difference plot showed occasional spikes of positive bias (Fig. 11) that can also be seen in Fig. 7 (day 152), Fig. 8 (day 229), and Fig. 9 (day 209). Rapid water loss, associated with the evaporation of intercepted rainfall or dew accumulation, is captured by the LaRS model and can explain these deviations. The large deviations for BGC++ result from its parameterization being based on northern hardwood systems where large seasonal deviations have been observed. The operating concept behind BGC++ was to have a single parameter set for "generic" hardwoods that could be applied to broad regions. While this is a reasonable idea for global simulations, where simplifying assumptions are required, it undermines the performance of BGC++ when site-specific data are used to test model predictions.

With few exceptions, model simulations of annual water cycle components were in good agreement with measured values from eddy covariance or watershed-scale data (Fig. 18). The long-term mean evapotrans-

piration reported for Walker Branch watershed from 1969 to 1983 was 655 mm/yr (Luxmoore and Huff 1989) consistent with these data and simulations. Higher mean estimates of ET_a for BGC++ were associated with high E_a and lower than average drainage. ET_a for CANOAK and SPA was just above the mean model and observational data. For CANOAK, elevated T_a, not E_a, was responsible for the greater ET_a values. Higher ET_a for SPA is driven by its canopy conductance parameterization, but its breakdown among T_a and E_a could not be determined because those outputs were not provided by SPA. LoTEC, which had the lowest T_h stemming from its unusual T_h vs. VPD relationship (Fig. 5), also exhibited the lowest ET_a values and thus had higher than average drainage.
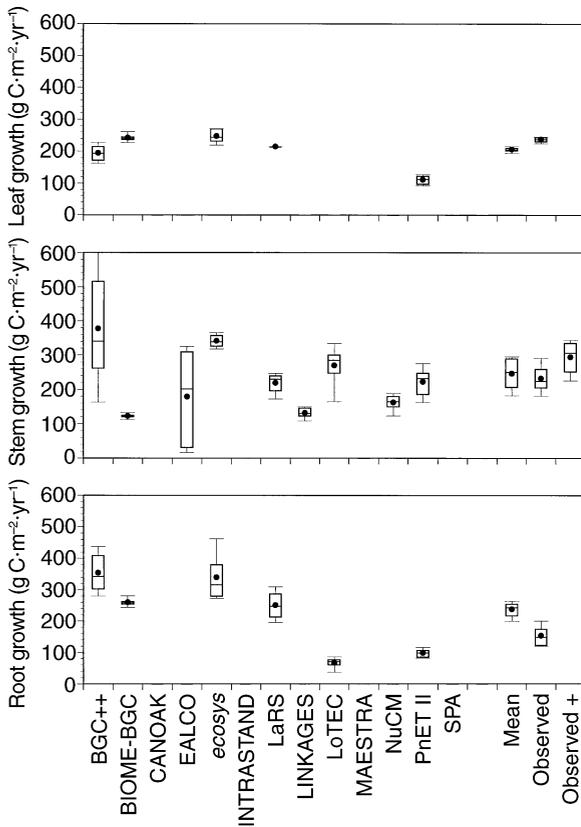
The year of maximum drainage (i.e., 1994) was correctly simulated by all models (Table 11). The dormant season of 1994 had precipitation inputs at least 200 mm greater than any other year (Hanson et al. 2003a), and it is encouraging that all models handled this "excess" precipitation in the appropriate manner. The year of minimum drainage was model dependent because it is the result of a balance between precipitation inputs, drainage, and water use (i.e., T). No models found the wet years of 1996 or 1997 to be associated with minimum drainage. Years with maxium or minimum ET_a were not consistent among models or data, but most models showed maximum ET_a in years with high growing-season precipitation (>500 mm in 1996, 1997, 1999 [Hanson et al. 2003a]). Simulated years having the lowest ET_a were distributed among the drought years of 1993, 1995, and 1998 for most models. LaRS had 1994 as year of minimum ET_a. CANOAKs years of minimum ET_a were the two wet years of 1996 and 1997 (Table 10). CANOAK predicted minimum ET_a in wet years in part because of limited radiation inputs and lower vapor pressure gradients associated with numerous rain events in those years. However, the primary reason CANOAK did not predict minimum ET_a for the drought years of 1995 or 1998 (a more common result) is that it lacked a soil water model necessary for predicting a drought feedback to simulated canopy transpiration.

Carbon cycle.—

1. *Hourly NEE.*—Under optimum water supply, NEE_h data were appropriately simulated by all models (EF > 0.5, Table 4). The seven models capable of providing data for NEE_h all provided reasonable estimates of NEE_{max} resulting from an asymptotic relationship between PAR and NEE_h in agreement with the observations. Not surprisingly, MAESTRA, a model that did not simulate soil respiration, produced less negative values for R_{stand} (Table 4). Overestimates of NEE_{max} by CANOAK and MAESTRA reflect their lack of soil water simulations and associated drought feedbacks.

Not all models captured the appropriate response to drought. Measured NEE_h data showed a 47% reduction in light-saturated NEE (NEE_{max}) under severe drought

(Fig. 3, Table 4), but two models overestimated the drought impact (EALCO EF = −0.51, INTRASTAND EF = −0.40; Table 4 and Fig. 4). As described above, INTRASTAND overestimated drought response via the inappropriate application of leaf characteristics to a stand-level response, and EALCOs hypersensitivity stemmed from inappropriate characterization of the soil water potential (see soil water discussion in *How good are the measured data for testing models?*). Because parameterization issues are simple to correct, we consider the simulations by most models to be in good agreement with the available data for this site. Conversely, inappropriate parameterization will remain a serious issue for future model applications to ecosystems without adequate reference data.

Daily simulations of NEE_d for most models showed a consistent negative bias with respect to NEE_d from eddy covariance and lower EF values than for ET_d (Figs. 7, 8, and 9, Table 6). Unlike the ET_d data where bias was limited at night, the NEE_d model predictions were all lower than measured NEE_d with the exception of the MAESTRA model, which lacked losses from soil/root respiration (Fig. 10). As discussed in *How good are the measured data for testing models?* the differences between simulations and eddy covariance measures of NEE_d were expected because of the nighttime bias associated with Walker Branch. Under both wet and dry conditions (Figs. 7, 8, and 9), MAESTRA had the highest simulated values for NEE_d because it did not attempt to model soil/root respiration. Clearly, adequate predictions of ecosystem NEE must include all significant carbon flux processes.

NEE_d simulations were in better agreement across models under wet (Figs. 7 and 9) vs. severe drought conditions (Fig. 8). Under severe drought CANOAK and LoTEC showed limited reductions in NEE_d, and BGC++ and EALCO had the greatest. Alternate models showed clear differences in the expression of drought impact on NEE_d, and most showed greater impacts than observed in the eddy covariance data. Difference plots of NEE_d (Fig. 10) show almost all lack of fit to be associated with the growing season. The initiation of leaf-out, the onset of leaf senescence, and the appropriate approximation of midseason drought responses were responsible for many model deviations from the observed data.

2. *Soil respiration.*—Based on our standard of good fit being an EF value >0.45, only BIOME-BGC (EF = 0.47) and CANOAK (EF = 0.48) models do a good job of capturing R_{soil} as measured by understory eddy covariance observations, and only one model, INTRASTAND (EF = 0.57), does a good job of simulating the chamber R_{soil} data (Fig. 14, Table 9). The causes for the disagreements between model simulations and the understory eddy covariance data are related to footprint differences between the conceptual eddy covariance footprint and the model simulations, which were largely parameterized from Walker Branch chamber

data (1990–1991 data from Hanson et al. 1993). The R_{soil} chamber data used here included measurements when the litter layer was wet and total R_{soil} elevated (Hanson et al. 2003b). The R_{soil} algorithms for most models did not allow for the impact of differential wetting and drying of the litter layer, and therefore could not capture the features of the chamber R_{soil} data. Improvements in the way ecosystem models handle R_{soil} are needed to account for differential contributions from the litter layer, roots, and changing pools of available carbon.

3. *Annual NEE, NPP, and components.*—In contrast to the intermodel and data-to-model agreement for components of annual water flux, there was less agreement between models and data for measures of annual stand carbon flux. Hanson et al. (2003d) reported that NEP for Walker Branch ranged from 79 to 292 g $C \cdot m^{-2} \cdot yr^{-1}$ from 1993 to 2000 (i.e., the biometric low data in Fig. 19 and Table 12), but recognized that the individual annual estimates might have a minimum error of ±37%. They also showed that the eddy covariance NEE_a data for Walker Branch could not be reconciled with biometric measurement metrics. Because the NEP estimates of Hanson et al. (2003d) are based on the soil respiration model imbedded within the INTRASTAND model, we included a different NEP estimate based on mean model soil respiration (biometric high data in Fig. 19 and Table 12). The high biometric NEP data typically did not account for extra heterotrophic carbon losses associated with the wetting and drying of the litter layer (Hanson et al. 2003d), and were on average 60 g $C \cdot m^{-2} \cdot yr^{-1}$ greater. For the evaluation of NEE_a predictions in this paper, we used both the biometric low and high estimates of NEP as a reference and consider an acceptable interannual model range for NEEa to be from 0 to 500 g $C \cdot m^{-2} \cdot yr^{-1}$ (i.e., extremes of the biometric estimates ±~37%). For the following models, in order from highest to lowest NEE_a, simulations fell within the expected range: LoTEC, LaRS, INTRASTAND, EALCO, and BIOME-BGC (Fig. 19). Three models produced simulations of NEE_a above the expected range (BGC++, CANOAK, and *ecosys*), and one model's NEE_a value was below this range (PnET-II). As suggested in *Annual simulation: Growth*, the low NEE_a simulations for PnET-II are the result of a lower-than-expected prediction of canopy leaf growth. Adjustments to the carbon allocation algorithms within PnET-II would help to solve this problem.

No single component of NEE_a was responsible for over- or underestimates of NEE_a. In the case of *ecosys*, higher than expected NEE was driven by exceptionally large estimates of GPP (Fig. 20), even though *ecosys* simulated the highest autotrophic and leaf respiration (Figs. 20 and 21, respectively). Similarly, low GPP was responsible for the low NEEa projections from PnET-II. The explanation for high NEE_a for BGC++ and

CANOAK is the result of lower than expected leaf respiration (Fig. 21).

Simulations of interannual variability of NEE_a or NPP (Tables 12 and 13) showed some consistency across models. Most models predicted 1998 to be the year of lowest NEE_a in agreement with the biometric estimate from (Hanson et al. 2003d). The alternate biometric NEP estimate (i.e., Biometric High Table 12) showed 1993 to be the lowest NEP year, but 1998 was nearly as low. Simulated years with greatest NEE_a were spread among years of ample growing season precipitation with the exception of INTRASTAND that yielded a high NEE_a value for 1993. The years of greatest NEE_a based on biometric data were 1996 or 1999 (Table 12). INTRASTAND predicted high NEE_a in 1993 because of lower than average litter and wood decomposition. If INTRASTAND carbon pools had been ''spun-up'' to equilibrium pool sizes prior to the 1993 model run, the spuriously high 1993 NEE_a prediction would have been avoided. Because biometric methods miss interannual changes in NEE_a driven by changes in internal or ''hidden'' nonstructural carbohydrate pools (Hanson et al. 2003d), exact interannual agreement between modeled NEE_a and NEP from biometric methods should not be expected.

With the exception of *ecosys,* all models fell within 10% of the 1993 to 2000 range of NPP values (540–925 g $C \cdot m^{-2} \cdot yr^{-1}$) reported by Hanson et al. (2003d). The year having the greatest NPP was simulated to be 2000 by 7 of 11 models, and 1998 was simulated to be the year of lowest NPP by 8 of 11 models. Biometric estimates of NPP showed 1996 and 1999 to be high years and 1993 to be the lowest year. Lack of agreement between simulated and observed NPP values may not be surprising given the difficulty of predicting carbon allocation to growth.

4. *Growth.*—Simulation of leaf, stem, and root growth should be a long-term goal of all ecosystem models. However, only 5, 8, and 6 of 13 models attempted simulations of leaf, stem, and root growth, respectively (Fig. 22). Other models used observed growth if necessary. Leaf growth was captured well by BGC++, BIOME-BGC, *ecosys,* and LaRS, but PnET-II underestimated leaf development. Stem growth simulations were much more diverse, but the mean of all model simulations was very close to the observed values. BGC++ simulated abnormally high stem growth over a much broader range than was measured from 1993 to 2000. EALCO also simulated a broad interannual range of stem growth, but the magnitude of the mean was closer to the observed values. No models captured the measured range of annual root growth. BGC++, BIOME-BGC, *ecosys,* and LaRS provided similar but higher root growth estimates, and LoTEC and PnET-II had similar but lower root growth estimates. Because direct measurement of root production is difficult and subject to large error (Joslin et al. 2000, Curtis et al. 2002, Hanson et al. 2003d) all of these

simulated root growth rates may be within an acceptable range.

Measured leaf growth showed little interannual variation and was reflected by most models, with LaRS representing the extreme of no interannual variation (data not shown). BGC++ simulated a large reduction in leaf production between 1997 and 1998 (208 to 173 g C·m^{-2}·yr^{-1}) that is not justified by field observations (Hanson et al. 2003c). Large swings in leaf production in BGC++ resulted from predicted differences in prior-year growth and nitrogen mineralization and their resulting carry-over effect on current year leaf growth. This error resulted from doing carbon and nitrogen allocation on an annual time step and has since been corrected (E. R. Hunt, *personal communication*). Stem growth had more interannual variation than leaf production, showing lower than mean growth in the 1993 and 1995 drought years. The most extreme drought year (1998) did not result in stem growth reductions because the drought occurred too late in the growing season and was disconnected from the annual stem growth cycle (Hanson et al. 2001a). LaRS did the best job of capturing the magnitude and interannual dynamic of measured stem growth. Interannual root growth dynamics showed peak growth in 1994 and 1996, and low growth in 1998. LaRS and BGC++ root growth simulations captured this pattern, but predicted nearly two and three times greater root production, respectively, than was measured. Although many ecosystem carbon and water cycling models avoid growth predictions because the factors controlling carbon allocation to growing tissues are poorly understood, it is encouraging to see that several of the models in this intercomparison are able to capture key features of stem, leaf, and root growth of an upland oak system.

Soil water.—Even though simulated soil water data from all models produced reasonable EF values when compared to both TDR and FDR observations (Table 8), complete agreement with the magnitude of daily SW content was difficult for most models. The general patterns of SW increases and decreases were captured by all models (Figs. 7, 8, and 9), but over- and underestimates of percent SW commonly differed by 2–5 units (percent water content; Figs. 12 and 13). The importance of a 2–5 unit difference between modeled and measured SW values is amplified when the data are translated to soil water potential via nonlinear equations in some models (e.g., INTRASTAND). Other models (e.g., LINKAGES) propagated the response of drought from the ''plant extractable water'' and were thus less dependent on attainment of the correct SW values.

EALCO provides an example of why the correct simulation of SW and the physical features important to SW must be captured by a model. EALCO failed to produce accurate simulations during severe drought conditions, such as those shown in Fig. 3 (NEE$_h$), Fig. 5 (T_h), and Fig. 8 (NEE$_d$). EALCO used canopy water potential for ET and carbon (e.g., photosynthesis) calculations. Canopy water potential was determined by the model as a result of water flow and transport processes for which soil water potential was a required input. To convert SW (volume per volume, as a percentage) to soil water potential, the measured moisture release curve (Hanson et al. 1998) was used, but the soil coarse fraction (Cf) was not used to adjust the size of the soil water pool. In the case of EALCO, this oversight had a minor impact when the soil was wet, but it made a large difference when the soil was very dry (i.e., a given SW value yielded much lower soil water potentials). During the two dry weeks used for the model intercomparison, the root zone soil water content simulated by EALCO (without including Cf) was ~1 MPa lower than it should have been. This error caused extremely low canopy water potential and prematurely shut down photosynthesis and transpiration. For those models whose drought responses are based on soil water potential, the soil characteristics including texture, water-holding capacity, and the soil water retention relationship must be appropriately defined. If not, the nonlinear nature of the conversion from soil water content to soil water potential may cause significant acceleration or deceleration of the onset of drought depending on the direction of the error.

Many models did not simulate the daily details of measured soil water content and associated water and carbon flux. Why then were they able to produce acceptable estimates of NEEa or ET$_a$ (Fig. 18)? The answer is tied to drought duration. In the deciduous forest region of the eastern United States interannual precipitation patterns result in a unpredictable occurrence of drought (Hanson and Weltzin 2000). When droughts do occur they tend to be of limited duration late in the growing season. Therefore, because drought is not a sustained phenomenon in the Walker Branch region, the net effects of drought on annual ET or NEE were of limited consequence to annual carbon or water flux. Nevertheless, it is important to simulate drought accurately, even if it is not a sustained phenomenon under current climate.

Mean model results.—The average output of multiple models consistently provided the best or nearly the best fit to measured data, as shown in Fig. 17. The utility of the mean response stems from the large number of models involved, and the averaging of over- and underestimation of parameters within individual models. The utility of averaged model outputs for weather forecasting has been recognized for some time (Tracton and Kalnay 1993, Hamill et al. 2000, Cheung 2001). The meteorological community has embraced the concept of ''ensemble mean'' model outputs for generating the best possible model predictions, and concluded that the improvements are the result of a reduction in errors associated with individual deterministic model runs (Hamill et al. 2000). Ensemble model runs can be generated from the outputs of multiple models, similar to

our mean output, or from more sophisticated Monte Carlo manipulations of a single model (Cheung 2001).

Even though the simple "ensemble mean" model response presented in this paper is a robust predictor for the current case study, it may be premature to conclude that simple averaging of all model responses will yield better future predictions. Models that generate consistent over- or under-predictions might logically be excluded from such calculations. Ensemble mean predictions can be generated from the averaging of multiple model outputs (our example), but the stochastic perturbation of a single mechanistically correct model avoids the technical difficulties of maintaining and operating multiple model codes (Richardson 2001). We recommend that the ecosystem modeling community recognize the advantages of stochastic ensemble mean predictions, and consider their use with improved models having robust and mechanistically correct representations of ecosystem processes.

Model structure and temporal resolution.—In this study, models having structural detail in the canopy and soil (i.e., those with many layers such as *ecosys* and LaRS) consistently provided hourly, daily, and annual outputs in agreement with measured data. A similar conclusion was made for a boreal forest model intercomparison (Amthor et al. 2001). Simulations from models with intermediate structural detail (EAL-CO, INTRASTAND, BIOME-BGC) were nearly as good. Conversely, the simple big-leaf models (BGC++ and LoTEC) and models without key ecosystem components (CANOAK minus soils, MAESTRA minus soils, LINKAGES, and NuCM without carbon flux) were unable to produce outputs in consistent agreement with the full range of data tested. Higher process resolution models are capable of capturing the details of ecosystem response to climate variables, but at the cost of intensive parameter requirements. Whenever site-specific parameters are not available, or models are not properly parameterized, higher process resolution models could be at risk of biased outputs. Aggregation studies have shown that simple big-leaf models of GPP can accurately reproduce the predictions of more complex models (Williams et al. 1997). So, although this study showed models with higher process resolution to be the better performers, it is likely that simpler models can be recalibrated, or their functions adjusted, to generate improved predictions. A potential result of the calibration of simplified models, however, is a reduction in their utility for application to new and/or different questions.

Good performance of a single model at one time step did not guarantee success at other time steps. For example, inability to capture salient features of the magnitude of the drought response for hourly transpiration (CANOAK and LoTEC had too little response, IN-TRASTAND too much response) had little impact on annual predictions. Good model simulations were produced by models operating at both the hourly and daily

time steps, but the monthly model, PnET-II, failed to capture annual carbon flux data consistent with observations. PnET-IIs lack of fit is the result of limited allocation of carbon to leaf growth, not its monthly time step. In eastern deciduous forests subject to random summer drought (Hanson and Weltzin 2000), correct simulation of water and carbon cycle responses to soil water deficit will benefit from the temporal resolution of a daily or hourly time step.

Are the results from general models different from those of site-calibrated models?

In order to understand the performance of a given model in the context of this or other intercomparisons, one must consider the original intent and scale of a model. (That is, is the model being used outside of its original "range?") The time step, structure, and the intended use (interpolation vs. prediction) that worked for the original system(s) may not function appropriately in the next. With the exception of CANOAK and INTRASTAND, all of the models were originally developed for other systems, yet most were quite capable of capturing many features of the carbon and water fluxes for the upland oak forest of Walker Branch. That is, their structure and assumptions were robust enough to be useful in other locations with appropriate parameterization.

Five models (BIOME-BGC, *ecosys,* LaRS, LINK-AGES, and PnET-II) all required a spin-up period and did not start with exactly the same initial conditions as specified in Table 3. Nevertheless, all except PnET-II provided good to excellent simulations of component water and carbon fluxes. These generic models for which all (or many) carbon, water, and nutrient pools are explicitly simulated attempt a more difficult task than those models developed primarily for flux interpolation (e.g., INTRASTAND). While greater use of prescribed site data within the generic models would have improved model agreement with measured data, such "model tuning" is done at the cost of wider applicability to a range of sites and is considered undesirable.

Several of the models (or their precursors) used in this comparison have been used to address issues of environmental change over space and/or time: BIOME-BGC (Hunt et al. 1996, White et al. 1997), LINKAGES (Bugmann et al. 2001, Wullschleger et al. 2003a), LoTEC (King et al. 1997), and PnET-II (Aber et al. 1995, McNulty et al. 1996, Ollinger et al. 2002). Most of the general models used in this paper were able to produce good simulations for one or more water or carbon flux variables, but BGC++ and PnET-II had difficulty simulating some variables consistent with the measured data. BGC++ predicted greater seasonal swings in ET_d, NEE_d, and a broad range of interannual NEE_a estimates (Figs. 10, 11, and 19) because its generic parameterization of a deciduous hardwood stand is based on NPP vs. precipitation relationships devel-

oped from spatial data (Lauenroth and Sala 1992). Although the substitution of space for time may be a logical approach for global model applications, in the current comparison, variability associated with space rather than time appears to inflate the intra- and inter-annual swings in BGC++ model predictions. The negative NEE$_a$ bias of the PnET-II model (Fig. 19) is the result of a dramatic under-prediction of GPP of the upland oak forest.

Can current model projections set logical bounds on ecosystem responses to environmental change appropriate to policy questions?

Models are our best and perhaps only realistic tools for evaluating the fate of ecosystem goods and services over a wide range of environmental change scenarios. Models can be used to test sensitivity of processes and/or simulate ecosystems responses to future conditions (e.g., Aber et al. 2001, Sampson et al. 2001, Lasch et al. 2002). In such a role, models are an excellent source of hypotheses for additional study. Unfortunately, as observed in this model intercomparison, it is too easy for models to be applied to circumstances for which they were not intended. The divergent simulation of ecosystem NEE$_a$ by PnET-II and BGC++ in this intercomparison are two examples that underscore the need to be careful in the application of ''established'' models to new situations. In addition, while CANOAK has been appropriately applied to the analysis of complex questions of water, carbon, and energy flux in a deciduous canopy (Baldocchi et al. 2002), the version of CANOAK used in this comparison failed to appropriately capture drought responses for the simple reason that soils and their water status were not part of the model. Conversely, the uniform ability of most models to capture key components of the water cycle suggests that the mechanisms related to water flux are understood, and water cycle outputs from these models may be appropriate for landscape extrapolations. This is especially true for those models containing full energy balance as an internal check on the limit of evapotranspiration (e.g., EALCO, *ecosys,* LaRS, SPA).

Model performance was typically better for optimum as opposed to drought conditions, suggesting that further work may be needed to find an acceptable approach for expressing mechanisms by which ecosystems perceive and respond to soil water deficits. The inability of many models to handle drought is not unexpected given that they were typically developed for well-watered conditions. Models capable of providing realistic projections for nonambient future conditions, however, are exactly what are needed to address policy questions relevant to environmental change issues. Carey et al. (2001) also concluded that ''models that describe stand-level processes accurately'' are needed for the future prediction of forest carbon sinks.

Homann et al. (2000) conducted an intermodel comparison of belowground processes with 14 biogeo-chemical cycling models and concluded that substantial improvement is required for current models to provide simulations applicable to public-policy decisions. The current study shows that model outputs are generally in good agreement with data over a range of temporal scales, but also points out examples of substantial deviation. Given the requisite use of models for environmental change assessment activities, and a common inability to ''get-it-right'' in this and other model tests, further validation exercises are encouraged for the development of robust ecosystem models.

Finally, policy makers often ask questions about options for the ''management'' of natural and man-made ecosystems for the goods and services that they provide. Unfortunately, such questions go well beyond the capabilities of more than half of the models evaluated here. Several models did simulate leaf, stem, and root growth, but the agreement between model estimates and observations is often inconsistent (Fig. 22). Comprehensive ecosystem models must go beyond predicted changes in carbon, water, energy, and nutrient cycles. To be of the greatest benefit in the policy arena, ecosystem models must translate changes in biogeo-chemical cycling processes into measurable changes in ecosystem goods and services such as wood production, soil carbon sequestration, clean water yield, and biodiversity.

CONCLUSIONS

A single model did not consistently perform the best at all time steps or for all variables considered. Inter-model comparisons showed good agreement for water cycle fluxes, but considerable disagreement among models for predicted carbon fluxes. The mean of all model outputs, however, was nearly always the best fit to the observations. In the absence of adequate independent test data, the mean of multiple model outputs is an acceptable reference point for the evaluation of individual model performance.

Detailed mechanistic models operating at hourly to daily time steps consistently performed the best in the current intercomparison. Some models intended for general application did not perform well, underscoring the need for caution when applying models to new ecosystems beyond those for which they were developed. Most models focused on the carbon and water cycle, but the best performance was obtained from models that included a complete energy budget.

The largest disagreements between model outputs, and between models and measured data, were related to a model's ability to simulate water deficits and propagate responses to water deficits. Further review of the manner in which models simulate and respond to drought is needed, including an evaluation of the importance of gradients of root uptake and soil water availability with depth.

A number of the models included measures of tree growth, but results were inconsistent with respect to

measured data, especially for leaves and roots. Future models must improve their ability to capture growth processes. Mechanisms for allocating carbon to growth of plant tissues are sorely needed. The detailed mechanistic models made no attempt to evaluate plant survival or estimate temporal changes in species composition, but such processes are important to society. Future model development might consider merging successional concepts from models like LINKAGES with the robust carbon, water, and energy cycles captured by the detailed mechanistic models. A model with some of these features has been applied to the European climate space (Smith et al. 2001).

Additional model testing and intercomparison activities are recommended to improve ecosystem models for policy applications. To enable future comparisons, a conscious effort on the part of experimental and/or measurement groups to archive detailed environmental and site-characteristic data is needed. Substantial effort is being expended to archive data from the flux networks (Baldocchi et al. 1996, 2001) for future analyses. Similar attention to the archiving of results and operational measurements from expensive, long-term studies such as the FACE experiments (Karnosky et al. 2001, Norby et al. 2001) should also be given high priority. Detailed model intercomparisons represent a nontrivial exercise that clearly warrants independent support. Improvements of ecosystem models, from periodic and rigorous testing, are needed if the research community is to provide the best possible tools for evaluating scenarios of environmental change relevant to the policy arena.

LITERATURE CITED

Aber, J. D. 1997. Why don't we believe the models? ESA Bulletin **78**:232–233.

Aber, J. D., and C. A. Federer. 1992. A generalized, lumped-parameter model of photosynthesis, evapotranspiration, and net primary production in temperate and boreal forest ecosystems. Oecologia **92**:463–474.

Aber, J. D., and J. M. Melillo. 1982. FORTNITE: a computer model of organic matter and nitrogen dynamics in forest ecosystems. University of Wisconsin Research Bulletin **R3130**.

Aber, J. D., R. P. Neilson, S. McNulty, J. M. Lenihan, D. Bachelet, and R. J. Drapek. 2001. Forest processes and global environmental change: predicting the effects of individual and multiple sensors. BioScience **51**:735–751.

Aber, J. D., S. V. Ollinger, C. A. Federer, P. B. Reich, M. L. Goulden, D. W. Kicklighter, J. M. Mello, and R. G. Lathrop, Jr. 1995. Predicting the effects of climate change on water yield and forest production in Northeastern U.S. Climate Research **5**:207–222.

Aber, J. D., P. B. Reich, and M. L. Goulden. 1996. Extrapolating leaf CO_2 exchange to the canopy: a generalized model of forest photosynthesis validated by eddy correlation. Oecologia **106**:257–265.

Amthor, J. S. 1996. Models of plant growth costs and their application to trees in Walker Branch Watershed. UCRL-AR-125547. Lawrence Livermore National Laboratory, Livermore, California, USA.

Amthor, J. S. 2000*a*. The McCree–de Wit–Penning de Vries–Thornley respiration paradigms: 30 years later. Annals of Botany **86**:1–20.

Amthor, J. S. 2000*b*. Direct effect of elevated CO_2 on nocturnal in situ leaf respiration in nine temperate deciduous tree species is small. Tree Physiology **20**:139–144.

Amthor, J. S., et al. 2001. Boreal forest CO_2 exchange and evapotranspiration predicted by nine ecosystem process models: intermodel comparisons and relationships to field measurements. Journal of Geophysical Research **106**: 33,623–33,648.

Amthor, J. S., M. L. Goulden, J. W. Munger, and S. C. Wofsy. 1994. Testing a mechanistic model of forest-canopy mass and energy exchange using eddy correlation: carbon dioxide and ozone uptake by a mixed oak–maple stand. Australian Journal of Plant Physiology **21**:623–651.

Anderson, D. R., and K. P. Burnham. 2001. Commentary on models in ecology. ESA Bulletin **82**:160–161.

Augé, R. M., C. D. Green, A. J. W. Stodola, A. M. Saxton, J. B. Olinick, and R. M. Evans. 2000. Correlations of stomatal conductance with hydraulic and chemical factors in several deciduous tree species in a natural habitat. New Phytologist **145**:483–500.

Bachelet, D., R. P. Neilson, J. M. Lenihan, and R. J. Drapek. 2001. Climate change effects on vegetation distribution

and carbon budget in the United States. Ecosystems **4**:164–185.

Baldocchi, D. D. 1997. Measuring and modeling carbon dioxide and water vapour exchange over a temperate broadleafed forest during the 1995 summer drought. Plant Cell and Environment **20**:1108–1122.

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

Baldocchi, D. D., J. Finnigan, K. B. Wilson, U. K. Paw, and E. Falge. 2000. On measuring net ecosystem carbon exchange over tall vegetation on complex terrain. Boundary-Layer Meteorology **96**:257–291.

Baldocchi, D. D., and P. C. Harley. 1995. Scaling carbon dioxide and water vapour exchange from leaf to canopy in a deciduous forest II. Model testing and applications. Plant Cell and Environment **18**:1331–1340.

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

Baldocchi, D. D., and C. A. Vogel. 1996. Energy and CO_2 flux densities above and below a temperate broad-leaved forest and a boreal pine forest. Tree Physiology **16**:5–16.

Baldocchi, D. D., and K. B. Wilson. 2001. Modelling CO_2 and water vapour exchange of a temperate broadleaved forest on daily to decadal time scales. Ecological Modelling **142**:155–184.

Baldocchi, D. D., K. B. Wilson, and L. Gu. 2002. How the environment, canopy structure and canopy physiological functioning influence carbon, water, and energy fluxes of a temperate broad-leaved deciduous forest—an assessment with the biophysical model CANOAK. Tree Physiology **22**:1065–1077.

Ball, J. T., I. E. Woodrow, and J. A. Berry. 1987. A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. Pages 221–224 in I. Biggins, editor. Progress in photosynthesis research. Volume 1V.5. Proceedings of the VII International Photosynthesis Congress. Martinus Nijhoff, Dordrecht, The Netherlands.

Bolstad, P. V., K. Mitchell, and J. M. Vose. 1999. Foliar temperature–respiration response functions for broadleaved tree species in the southern Appalachians. Tree Physiology **19**:871–878.

Bugmann, H. K. M., S. D. Wullschleger, D. T. Price, K. Ogle, D. F. Clark, and A. M. Solomon. 2001. Comparing the performance of forest gap models in North America. Climatic Change **51**:349–388.

Canadell, J. G., et al. 2000. Carbon metabolism of the terrestrial biosphere: a multitechnique approach for improved understanding. Ecosystems **3**:115–130.

Carey, E. V., A. Sala, R. Keane, and R. M. Callaway. 2001. Are old forests underestimated as global carbon sinks? Global Change Biology **7**:339–344.

Cheung, K. K. W. 2001. A review of ensemble forecasting techniques with a focus on tropical cyclone forecasting. Meteorological Applications **8**:315–332.

Curtis, P. S., P. J. Hanson, P. Bolstad, C. Barford, J. C. Randolph, H. P. Schmid, and K. B. Wilson. 2002. Biometric and eddy-covariance based estimates of annual carbon storage in five eastern North American deciduous forests. Agricultural and Forest Meteorology **113**:3–19.

Edwards, N. T., and P. J. Hanson. 1996. Stem respiration in a closed-canopy upland oak stand. Tree Physiology **16**:433–439.

Falge, E., et al. 2001. Gap filling strategies for defensible annual sums of net ecosystem exchange. Agricultural and Forest Meteorology **107**:43–69.

Fan, S., M. Gloor, J. Mahlman, S. Pacala, J. Sarmiento, T. Takahashi, and P. Tans. 1998. A large terrestrial carbon sink in North America implied by atmospheric and oceanic carbon dioxide data and models. Science **282**:442–446.

Farquhar, G. D., and S. von Caemmerer. 1982. Modelling of photosynthetic response to environmental conditions. Pages 549–587 in O. L. Lange, P. S. Nobel, C. B. Osmond, and H. Ziegler, editors. Encyclopedia of plant physiology, NS Volume 12B. Physiological plant ecology II. Springer-Verlag, New York, New York, USA.

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

Federer, C. A. 1995. BROOK90: a simulation model for evaporation, soil water, and streamflow, Version 3.1. Computer freeware and documentation. USDA Forest Service, Durham, New Hampshire, USA.

Gaudinski, J. B., and S. E. Trumbore. 2003. Soil carbon turnover. Pages 190–209 in P. J. Hanson and S. D. Wullschleger, editors. North American temperate deciduous forest responses to changing precipitation regimes. Springer, New York, New York, USA.

Gleick, P. H., and D. B. Adams. 2000. Water: the potential consequences of climate variability and change for the water resources of the United States. Report of the Water Sector Assessment Team of the National Assessment of the Potential Consequences of Climate Variability and Change. U.S. Global Change Research Program, Washington, D.C., USA.

Grant, R. F. 2001. A review of the Canadian ecosystem model *ecosys*. Pages 173–264 in M. Shaffer, editor. Modeling carbon and nitrogen dynamics for soil management. CRC Press, Boca Raton, Florida, USA.

Hamill, T. M., S. L. Mullen, C. Snyder, Z. Toth, and D. P. Baumhefner. 2000. Ensemble forecasting in the short to medium range: report from a workshop. Bulletin of the American Meteorological Society **81**:2653–2664.

Hanson, P. J., N. T. Edwards, C. T. Garten, Jr., and J. A. Andrews. 2000. Separating root and soil microbial contributions to soil respiration: a review of methods and observations. Biogeochemistry **48**:115–146.

Hanson, P. J., N. T. Edwards, T. J. Tschaplinski, S. D. Wullschleger, and J. D. Joslin. 2003d. Estimating the net primary and net ecosystem production of a southeastern upland *Quercus* forest from an 8-year biometric record. Pages 378–395 in P. J. Hanson and S. D. Wullschleger, editors. North American temperate deciduous forest responses to changing precipitation regimes. Springer, New York, New York, USA.

Hanson, P. J., and W. A. Hoffman. 1994. Emissions of non-CH_4 organic compounds and CO_2 from forest floor cores. Soil Science Society of America Journal **58**:552–555.

Hanson, P. J., R. E. McRoberts, J. G. Isebrands, and R. K. Dixon. 1987. An optimal sampling strategy for determining CO_2 exchange rate as a function of photosynthetic photon flux density. Photosynthetica **21**:98–101.

Hanson, P. J., E. G. O'Neill, M. L. S. Chambers, J. S. Riggs, J. D. Joslin, and M. E. Wolfe. 2003b. Soil respiration and litter decomposition. Pages 163–189 in P. J. Hanson and S. D. Wullschleger, editors. North American temperate deciduous forest responses to changing precipitation regimes. Springer, New York, New York, USA.

Hanson, P. J., L. J. Samuelson, S. D. Wullschleger, T. A. Tabberer, and G. S. Edwards. 1994. Seasonal patterns of light-saturated photosynthesis and leaf conductance for mature and seedling *Quercus rubra* L. Foliage: differential sensitivity to ozone. Tree Physiology **14**:1351–1366.

Hanson, P. J., D. E. Todd, and J. S. Amthor. 2001a. A six year study of sapling and large-tree growth and mortality

responses to natural and induced variability in precipitation and throughfall. Tree Physiology **21**:345–358.

Hanson, P. J., D. E. Todd, and M. A. Huston. 2003*a*. Walker Branch Throughfall Displacement Experiment (TDE). Pages 8–31 *in* P. J. Hanson and S. D. Wullschleger, editors. North American temperate deciduous forest responses to changing precipitation regimes. Springer, New York, New York, USA.

Hanson, P. J., D. E. Todd, M. A. Huston, J. D. Joslin, J. Croker, and R. M. Augé. 1998. Description and field performance of the Walker Branch Throughfall Displacement Experiment: 1993–1996. ORNL/TM-13586. Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA.

Hanson, P. J., D. E. Todd, and J. D. Joslin. 2003*c*. Canopy production. Pages 303–315 *in* P. J. Hanson and S. D. Wullschleger, editors. North American temperate deciduous forest responses to changing precipitation regimes. Springer, New York, New York, USA.

Hanson, P. J., D. E. Todd, J. S. Riggs, M. E. Wolfe, and E. G. O'Neill. 2001*b*. Walker Branch Throughfall Displacement Experiment Data Report: site characterization, system performance, weather, species composition and growth. ORNL/CDIAC-134, NDP-078A. Carbon Dioxide Information Analysis Center, U.S. Department of Energy, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA.

Hanson, P. J., and J. F. Weltzin. 2000. Drought disturbance from climate change: response of United States forests. Science of the Total Environment **262**:205–220.

Hanson, P. J., S. D. Wullschleger, S. A. Bohlman, and D. E. Todd. 1993. Seasonal and topographic patterns of forest floor CO_2 efflux from an upland oak forest. Tree Physiology **13**:1–15.

Harley, P. C., and D. D. Baldocchi. 1995. Scaling carbon dioxide and water vapor exchange from leaf to canopy in a deciduous forest: I. leaf model parameterization. Plant Cell and Environment **18**:1146–1156.

Harley, P., A. Guenther, and P. Zimmerman. 1997. Environmental controls over isoprene emission in deciduous oak canopies. Tree Physiology **17**:705–714.

Harley, P. C., R. B. Thomas, J. F. Reynolds, and B. R. Strain. 1992. Modelling photosynthesis of cotton grown in elevated CO2. Plant Cell and Environment **15**:271–282.

Homann, P. S., R. B. McKane, and P. Sollins. 2000. Belowground processes in forest-ecosystem biogeochemical simulation models. Forest Ecology and Management **138**:3–18.

Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, and D. Xiaosu, editors. 2001. Climate Change 2001: the scientific basis: Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, Cambridge, UK.

Huff, D. D., R. J. Luxmoore, J. B. Mankin, and C. L. Begovich. 1977. TEHM: a terrestrial ecosystem hydrology model. International Biological Program Technical Publication, EDFB/IBP-76/8, ORNL/NSF/EATC-27. Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA.

Hunt, E. R., Jr., M. B. Lavigne, and S. E. Franklin. 1999. Factors controlling the decline of net primary production with stand age for balsam fir in Newfoundland assessed using an ecosystem simulation model. Ecological Modelling **122**:151–164.

Hunt, E. R., Jr., S. C. Piper, R. Nemani, C. D. Keeling, R. D. Otto, and S. W. Running. 1996. Global net carbon exchange and intra-annual atmospheric CO_2 concentrations predicted by an ecosystem process model and three-dimensional atmospheric transport model. Global Biogeochemical Cycles **10**:431–456.

Hutchinson, B. A., and D. D. Baldocchi. 1989. Forest meteorology. Pages 21–95 *in* D. W. Johnson and R. I.

VanHook, editors. Analysis of biogeochemical cycling processes in Walker Branch Watershed. Springer-Verlag, New York, New York, USA.

Iverson, L. R., and A. M. Prasad. 2001. Potential changes in tree species richness and forest community types following climate change. Ecosystems **4**:186–199.

Janssens, I. A., A. S. Kowalski, and R. Ceulemans. 2001. Forest floor CO_2 fluxes estimated by eddy covariance and chamber-based model. Agricultural and Forest Meteorology **106**:61–69.

Janssens, I. A., A. S. Kowalski, B. Longdoz, and R. Ceulemans. 2000. Assessing forest soil CO_2 efflux: an *in situ* comparison of four techniques. Tree Physiology **20**:23–32.

Jarvis, P. G., and K. G. McNaughton. 1986. Stomatal control of transpiration: scaling up from leaf to region. Advances in Ecological Research **15**:1–49.

Jenkins, J. C., D. W. Kicklighter, and J. D. Aber. 2000. Regional impacts of increased CO2 and climate change on forest productivity. Pages 383–423 *in* R. H. Mickler, R. A. Birdsey, and J. Hom, editors. Responses of northern U.S. forests to environmental change. Springer-Verlag, New York, New York, USA.

Jenkinson, D. S. 1990. The turnover of organic carbon and nitrogen in soil. Philosophical Transactions of the Royal Society of London, Series B **329**:361–368.

Johnson, D. W., P. J. Hanson, and D. E. Todd. 2002. The effects of throughfall manipulation on soil leaching in a deciduous forest. Journal of Environmenal Quality **31**:204–216.

Johnson, D. W., P. J. Hanson, and D. E. Todd. 2003. Nutrient availability and cycling. Pages 396–414 *in* P. J. Hanson and S. D. Wullschleger, editors. North American temperate deciduous forest responses to changing precipitation regimes. Springer, New York, New York, USA.

Johnson, D. W., P. J. Hanson, D. E. Todd, Jr., R. B. Susfalk, and C. F. Trettin. 1998. Precipitation change and soil leaching: field results and simulations from Walker Branch Watershed, Tennessee. Water Air and Soil Pollution **105**:251–262.

Johnson, D. W., and S. E. Lindberg, editors. 1992. Atmospheric deposition and forest nutrient cycling. Springer-Verlag, New York, New York, USA.

Johnson, D. W., and R. I. VanHook. 1989. Analysis of biogeochemical cycling processes in Walker Branch Watershed. Springer-Verlag, New York, New York, USA.

Joslin, J. D., and M. H. Wolfe. 1998. Impacts of long-term water input manipulations on fine root production and mortality in mature hardwood forests. Plant and Soil **204**:165–174.

Joslin, J. D., and M. H. Wolfe. 2003. Fine root growth response to precipitation change. Pages 274–302 *in* P. J. Hanson and S. D. Wullschleger, editors. North American temperate deciduous forest responses to changing precipitation regimes. Springer, New York, New York, USA.

Joslin, J. D., M. H. Wolfe, and P. J. Hanson. 2000. Effects of shifting water regimes on forest root systems. New Phytologist **147**:117–129.

Karnosky, D. F., B. Gielen, R. Ceulemans, W. H. Schlesinger, R. J. Norby, E. Oksanen, R. Matyssek, and G. R. Hendrey. 2001. FACE systems for studying the impacts of greenhouse gases on forest ecosystems. Pages 297–324 *in* D. F. Karnosky, R. Ceulemans, and G. Scarascia-Mugnozza, editors. The impact of carbon dioxide and other greenhouse gases on forest ecosystems. CABI, Wallingford, UK.

King, A. W., W. M. Post, and S. D. Wullschleger. 1997. The potential response of terrestrial carbon storage to changes in climate and atmospheric CO_2. Climatic Change **35**:199–227.

Kirschbaum, M. U. F., and A. Fischlin. 1996. Climate change impacts on forests. Pages 95–129 *in* R. T. Watson, M. C.

Zinyowera, and R. H. Moss, editors. Climate change 1995 impacts, adaptations and mitigation of climate change: scientific–technical analysis. Cambridge University Press, New York, New York, USA.

Landau, S., R. A. C. Mitchell, V. Barnett, J. J. Colls, J. Craigon, K. L. Moore, and R. W. Payne. 1998. Testing winter wheat simulation models' predictions against observed UK grain yields. Agricultural and Forest Meteorology **89**:85–99.

Lasch, P., F. W. Badeck, M. Linder, and F. Suckow. 2002. Sensitivity of simulated forest growth to changes in climate and atmospheric CO_2. Forstwissenschaftliches Centralblatt **121**:155–171.

Lauenroth, W. K., and O. E. Sala. 1992. Long-term forage production in North American shortgrass steppe. Ecological Applications **2**:397–403.

Law, B. E., F. M. Kelliher, D. D. Baldocchi, P. M. Anthoni, J. Irvine, D. Moore, and S. VanTuyl. 2001a. Spatial and temporal variation in respiration in a young ponderosa pine forest during a summer drought. Agricultural and Forest Meteorology **110**:27–43.

Law, B. E., P. E. Thornton, J. Irvine, S. Van Tuyl, and P. M. Anthoni. 2001b. Carbon storage and fluxes in ponderosa pine forests at different developmental stages. Global Change Biology **7**:755–777.

Lieth, H. 1975. Modeling the primary productivity of the world. Pages 237–263 in H. Lieth and R. H. Whittaker, editors. Primary productivity of the Biosphere. Springer-Verlag, New York, New York, USA.

Liu, S., R. Munson, D. W. Johnson, S. Gherini, K. Summers, R. Hudson, K. Wilkinson, and L. F. Pitelka. 1992. The nutrient cycling model (NuCM): overview and application. Pages 583–609 in D. W. Johnson and S. E. Lindberg, editors. Atmospheric deposition and forest nutrient cycling. Springer-Verlag, New York, New York, USA.

Loehle, C. 1997. A hypothesis testing framework for evaluating ecosystem model performance. Ecological Modelling **97**:153–165.

Luo, Y., B. Medlyn, D. Hui, D. Ellsworth, J. F. Reynold, and G. Katul. 2001. Gross primary productivity in the Duke Forest: modeling synthesis of the free-air CO_2 enrichment experiment and eddy-covariance measurements. Ecological Applications **11**:239–252.

Luxmoore, R. J., and D. D. Huff. 1989. Water. Pages 164–196 in D. W. Johnson and R. I. VanHook, editors. Analysis of biogeochemical cycling processes in walker branch watershed. Springer-Verlag, New York, New York, USA.

Malcolm, J. R., A. Markham, R. P. Neilson, and M. Garaci. 2002. Estimated migration rates under scenarios of global climate change. Journal of Biogeochemistry **29**:835–849.

Mayer, D. G., and D. G. Butler. 1993. Statistical validation. Ecological Modelling **68**:21–32.

McMurtrie, R. E., and Y. P. Wang. 1993. Mathematical models of the photosynthetic response of tree stands to rising CO_2 concentrations and temperature. Plant Cell and Environment **16**:1–13.

McNulty, S. G., J. M. Vose, and W. T. Swank. 1996. Loblolly pine hydrology and productivity across the southern United States. Forest Ecology and Management **86**:241–251.

McNulty, S. G., J. M. Vose, and W. T. Swank. 1998. Predictions and projections of pine productivity and hydrology in response to climate change across the southern United States. Pages 391–405 in R. A. Mickler and S. Fox, editors. The productivity and sustainability of southern forest ecosystems in a changing environment. Springer, New York, New York, USA.

Melillo, J. M., T. V. Callaghan, F. I. Woodward, E. Salati, and S. K. Sinha. 1990. Effects on ecosystems. Pages 282–310 in J. T. Houghton, G. J. Jenkins, and J. J. Ephraums,

editors. Climate change: the IPCC scientific assessment. Cambridge University Press, New York, New York, USA.

Mitchell, K. A., P. V. Boldstad, and J. M. Vose. 1999. Interspecific and environmentally induced variation in foliar dark respiration among eighteen southeastern deciduous tree species. Tree Physiology **19**:861–870.

Mitchell, P. L. 1997. Misuse of regression for empirical validation of models. Agricultural Systems **54**:313–326.

Norby, R. J., K. Kobayashi, and B. A. Kimball. 2001. Rising CO_2—future ecosystems. New Phytologist **150**:215–221.

Norman, J. M. 1982. Simulation of microclimates. Pages 65–99 in J. L. Hatfield and I. J. Thomason, editors. Biometeorology and integrated pest management. Academic Press, New York, New York, USA.

Norman, J. M., C. J. Kucharik, S. T. Gower, D. D. Baldocchi, P. M. Crill, M. Rayment, K. Savage, and R. G. Striegl. 1997. A comparison of six methods for measuring soil-surface carbon dioxide fluxes. Journal of Geophysical Research **102**(D24):28771–28777.

Ollinger, S. V., J. D. Aber, P. B. Reich, and R. J. Freuder. 2002. Interactive effects of nitrogen deposition, tropospheric ozone, elevated CO_2 and land use history on the carbon dynamics of northern hardwood forests. Global Change Biology **8**:545–562.

O'Neill, E. G., and R. J. Norby. 1996. Litter quality and decomposition rates of foliar litter produced under CO_2 enrichment. Pages 87–103 in G. W. Koch and H. A. Mooney, editors. Terrestrial ecosystem responses to elevated CO_2. Academic Press, New York, New York, USA.

Pastor, J., and W. M. Post. 1985. Development of a linked forest productivity–soil process model. ORNL/TM-9519. Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA.

Peters, L. N., D. F. Grigal, J. W. Curlin, and W. J. Selvidge. 1970. Walker Branch Watershed Project: chemical, physical and morphological properties of the soils of Walker Branch Watershed. Pages 32–33 in Technical manual ORNL-TM-2968. Oak Ridge National Laboratory, Oak Ridge, Tennesee, USA.

Post, W. M., A. W. King, and S. D. Wullschleger. 1997. Historical variations in terrestrial biospheric carbon storage. Global Biogeochemical Cycles **11**:99–109.

Powell, D. S., J. L. Faulkner, D. Darr, Z. Zhu, and D. W. MacCleery. 1992. Forest resources of the United States, 1992. USDA Forest Service Rocky Mountain Forest and Range Experiment Station General Technical Report **RM-234**.

Reichle, D. E., editor. 1980. Dynamic properties of forest ecosystems. International Biological Programme 23. Cambridge University Press, London, UK.

Reynolds, M. R. 1984. Estimating the error in model predictions. Forest Science **30**:454–469.

Richardson, D. S. 2001. Ensembles using multiple models and analyses. Quarterly Journal of the Royal Meteorological Society **127**:1847–1864.

Running, S. W., and J. C. Coughlan. 1988. A general model of forest ecosystem processes for regional application. Ecological Modelling **42**:125–154.

Running, S. W., and S. T. Gower. 1991. FOREST-BGC: a general model of forest ecosystem processes for regional applications. II. Dynamic carbon allocation and nitrogen budgets. Tree Physiology **9**:147–160.

Running, S. W., and E. R. Hunt, Jr. 1993. Generalization of a forest ecosystem process model for other biomes, BIOME-BGC, and an application for global-scale models. Pages 141–158 in J. Ehleringer and C. Field, editors. Scaling physiological processes: leaf to globe. Academic Press, Orlando, Florida, USA.

Ryan, M. G., E. R. Hunt, Jr., R. E. McMurtrie, G. I. Ågren, J. D. Aber, A. D. Friend, E. B. Rastetter, W. M. Pulliam,

R. J. Raison, and S. Linder. 1996*a*. Comparing models of ecosystem function for temperate conifer forests. I. Model description and validation. Pages 313–362 *in* A. I. Breymeyer, D. O. Hall, J. M. Melillo, and G. I. Agren, editors. Global change: effects on coniferous forests and grasslands. John Wiley and Sons, New York, New York, USA.

Ryan, M. G., R. E. McMurtrie, G. I. Ågren, E. R. Hunt, Jr., J. D. Aber, A. D. Friend, E. B. Rastetter, and W. M. Pulliam. 1996*b*. Comparing models of ecosystem function for temperate conifer forests. II. Simulation of the effect of climate change. Pages 363–387 *in* A. I. Breymeyer, D. O. Hall, J. M. Melillo, and G. I. Agren, editors. Global change: effects on coniferous forests and grasslands. John Wiley and Sons, New York, New York, USA.

Sampson, D. A., I. A. Janssens, and R. Ceulemans. 2001. Simulated soil CO_2 efflux and net ecosystem exchange in a 70-year-old Belgian Scots pine stand using the process model SECRETS. Annals of Forest Science **58**:31–46.

Satterlund, D. R. 1972. Wildland watershed management. Ronald Press, New York. New York, USA.

Shugart, H. H., and D. C. West. 1977. Development of an Appalachian deciduous forest simulation model and its application to assessment of the impact of the Chestnut Blight. Journal of Environmental Management **5**:161–179.

Shuttleworth, W. J., and J. S. Wallace. 1985. Evaporation from sparse crops—an energy combination theory. Quarterly Journal of the Royal Meteorological Society **111**:839–855.

Smith, B., I. C. Prentice, and M. T. Sykes. 2001. Representation of vegetation dynamics in the modelling of terrestrial ecosystems: comparing two contrasting approaches within European climate space. Global Ecology and Biogeography **10**:621–637.

Sullivan, N. H., P. V. Bolstad, and J. M. Vose. 1996. Estimates of net photosynthetic parameters for twelve tree species in mature forests of the southern Appalachians. Tree Physiology **16**:397–406.

Sun, G., D. M. Amatya, S. G. McNulty, R. W. Skaggs, and J. H. Hughes. 2000. Climate change impacts on the hydrology and productivity of a pine plantation. Journal of the American Water Resources Association **36**(2):367–374.

Thornthwaite, C. W., and J. R. Mather. 1957. Instructions and tables for computing potential evapotranspiration and the water balance. Publications in Climatology **10**:183–311.

Thornton, P. E., B. E. Law, H. L. Gholz, K. L. Clark, E. Falge, D. S. Ellsworth, D. S. Goldstein, R. K. Monson, D. Hollinger, M. Falk, J. Chen, and J. P. Sparks. 2002. Modeling and measuring the effects of disturbance history and climate on carbon and water budgets in evergreen needleleaf forests. Agricultural and Forest Meteorology **113**:185–222.

Tracton, M. S., and E. Kalnay. 1993. Operational ensemble prediction at the national meteorological center—practical aspects. Weather Forecasting **8**:379–398.

U.S. Global Change Research Program. 2000. Climate change impacts on the United States: the potential consequences of climate variability and change. Overview. A Report to the National Assessment Synthesis Team. Cambridge University Press, Cambridge, UK.

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

Verma, S. B., D. D. Baldocchi, D. E. Anderson, D. R. Matt, and R. J. Clement. 1986. Eddy fluxes of CO_2, water vapor, and sensible heat over a deciduous forest. Boundary-Layer Meteorology **36**:71–91.

Walters, D. K. 1994. Evaluation methodology for forest ecosystem change models Volume I. Dissertation. University of Minnesota, St. Paul, Minnesota, USA.

Wang, S., R. F. Grant, D. L. Verseghy, and T. A. Black. 2001. Modelling plant carbon and nitrogen dynamics of a boreal aspen forest in CLASS—the Canadian Land Surface Scheme. Ecological Modelling **142**:135–154.

Wang, S., R. F. Grant, D. L. Verseghy, and T. A. Black. 2002*a*. Modelling carbon-coupled energy and water dynamics of a boreal aspen forest in a General Circulation Model land surface scheme. International Journal of Climatology **22**: 1249–1265.

Wang, S., R. F. Grant, D. L. Verseghy, and T. A. Black. 2002*b*. Modelling carbon dynamics of boreal forest ecosystems using the Canadian Land Surface Scheme. Climatic Change **55**:451–477.

Wang, Y. P., and P. G. Jarvis. 1990. Description and validation of an array model—MAESTRO. Agricultural and Forest Meteorology **51**:257–280.

Wang, Y. P., A. Rey, and P. G. Jarvis. 1998. Carbon balance of young birch trees grown in ambient and elevated atmospheric CO_2 concentration. Global Change Biology **4**: 797–807.

Weinstein, D. A., W. P. Cropper, Jr., and S. G. McNulty. 1998. Summary of simulated forest responses to climate change in the southeastern United States. Pages 479–500 *in* R. A. Mickler and S. Fox, editors. The productivity and sustainability of southern forest ecosystems in a changing environment. Springer, New York, New York, USA.

White, M. A., P. E. Thornton, and S. W. Running. 1997. A continental phenology model for monitoring vegetation responses to interannual climatic variability. Global Biogeochemical Cycles **11**:217–234.

White, M. A., P. E. Thornton, S. W. Running, and R. R. Nemani. 2000. Parameterization and sensitivity analysis of the BIOME-BGC terrestrial ecosystem model: net primary production controls. Earth Interactions **4**:1–85.

Whittaker, R. H. 1956. The vegetation of the Great Smoky Mountains. Ecological Monographs **26**:1–80.

Whittaker, R. H., and G. E. Likens. 1975. The biosphere and man. Pages 305–328 *in* H. Lieth and R. H. Whittaker, editors. Primary productivity of the biosphere. Springer-Verlag, Berlin, Germany.

Williams, M., B. J. Bond, and M. G. Ryan. 2001*a*. Evaluating different soil and plant hydraulic constraints on tree function using a model and sap flow data from ponderosa pine. Plant Cell and Environment **24**:679–690.

Williams, M., B. E. Law, P. M. Anthoni, and M. Unsworth. 2001*b*. Use of a simulation model and ecosystem flux data to examine carbon–water interactions in ponderosa pine. Tree Physiology **21**:287–298.

Williams, M., E. B. Rastetter, D. N. Fernandes, M. L. Goulden, G. R. Shaver, and L. C. Johnson. 1997. Predicting gross primary productivity in terrestrial ecosystems. Ecological Applications **7**:882–894.

Williams, M., E. B. Rastetter, D. N. Fernandes, M. L. Goulden, S. C. Wofsy, G. R. Shaver, J. M. Melillo, J. W. Munger, S.-M. Fan, and K. J. Nadelhoffer. 1996. Modelling the soil–plant–atmosphere continuum in a *Quercus-Acer* stand at Harvard Forest: the regulation of stomatal conductance by light, nitrogen and soil/plant hydraulic properties. Plant Cell and Environment **19**:911–927.

Wilson, K. B., and D. D. Baldocchi. 2000. Seasonal and interannual variability of energy fluxes over a broadleaved temperate deciduous forest in North America. Agricultural and Forest Meteorology **100**:1–18.

Wilson, K. B., and D. D. Baldocchi. 2001. Comparing independent estimates of carbon dioxide exchange over 5 years at a deciduous forest in the southeastern United

States. Journal of Geophysical Research **106**(D4):34167–34178.

Wilson, K. B., D. D. Baldocchi, and P. J. Hanson. 2000*a*. Quantifying stomatal and non-stomatal limitations to carbon assimilation resulting from leaf aging and drought in mature deciduous tree species. Tree Physiology **20**:787–797.

Wilson, K. B., D. D. Baldocchi, and P. J. Hanson. 2000*b*. Spatial and seasonal variability of photosynthetic parameters and their relationship to leaf nitrogen in a deciduous forest. Tree Physiology **20**:565–578.

Wilson, K. B., D. D. Baldocchi, and P. J. Hanson. 2001*a*. Leaf age affects the seasonal pattern of photosynthetic capacity and net ecosystem exchange of carbon in a deciduous forest. Plant Cell and Environment **24**:571–583.

Wilson, K. B., et al. 2002. Energy balance closure at FLUX-NET sites. Agricultural and Forest Meteorology **113**:223–243.

Wilson, K. B., and P. J. Hanson. 2003. Deciduous hardwood photosynthesis: species differences, temporal patterns and response to soil water deficits. Pages 35–47 *in* P. J. Hanson and S. D. Wullschleger, editors. North American temperate deciduous forest responses to changing precipitation regimes. Springer, New York, New York, USA.

Wilson, K. B., P. J. Hanson, P. J. Mulholland, D. D. Baldocchi, and S. D. Wullschleger. 2001*b*. A comparison of methods for determining forest evapotranspiration and its components across scales: sap-flow, soil water budget, eddy covariance, and catchment water balance. Agricultural and Forest Meteorology **106**:153–168.

Wilson, K. B., and T. P. Meyers. 2001. The spatial variability of energy and carbon dioxide fluxes at the floor of a deciduous forest. Boundary-Layer Meteorology **98**:443–473.

Winslow, J. C., E. R. Hunt, Jr., and S. C. Piper. 2001. A globally applicable model of daily solar irradiance estimated from air temperature and precipitation data. Ecological Modelling **143**:227–243.

Wofsy, S. C., M. L. Goulden, J. W. Munger, S.-M. Fan, P. S. Bakwin, B. C. Daube, S. L. Bassow, and F. A. Bazzaz. 1993. Net exchange of CO_2 in a mid-latitude forest. Science **260**:1314–1317.

Wullschleger, S. D., C. A. Gunderson, P. J. Hanson, K. B. Wilson, and R. J. Norby. 2002. Sensitivity of stomatal and canopy conductance to elevated CO_2 concentration—interacting variables and perspectives of scale. New Phytologist **153**:485–496.

Wullschleger, S. D., C. A. Gunderson, L. M. Tharp, D. C. West, and W. M. Post. 2003*a*. Simulated patterns of forest succession and productivity as a consequence of altered precipitation. Pages 433–446 *in* P. J. Hanson and S. D. Wullschleger, editors. North American temperate deciduous forest responses to changing precipitation regimes. Springer, New York, New York, USA.

Wullschleger, S. D., P. J. Hanson, and G. S. Edwards. 1996. Growth and maintenance respiration in leaves of northern red oak seedlings and mature trees after three years of ozone exposure. Plant Cell and Environment **19**:577–584.

Wullschleger, S. D., P. J. Hanson, and D. E. Todd. 2001. Transpiration from a multi-species deciduous forest as estimated by xylem sap flow techniques. Forest Ecology and Management **143**:205–213.

Wullschleger, S. D., P. J. Hanson, and D. E. Todd. 2003*b*. Forest water use as influenced by precipitation change. Pages 363–377 *in* P. J. Hanson and S. D. Wullschleger, editors. North American temperate deciduous forest responses to changing precipitation regimes. Springer, New York, New York, USA.

Wullschleger, S. D., P. J. Hanson, and T. J. Tschaplinski. 1998. Whole-plant water flux in understory red maple exposed to altered precipitation regimes. Tree Physiology **18**:71–79.

Wullschleger, S. D., K. B. Wilson, and P. J. Hanson. 2000. Environmental control of whole-plant transpiration, canopy conductance and estimates of the decoupling coefficient for large red maple trees. Agriculture and Forest Meteorology **104**:157–168.

Zheng, D., E. R. Hunt, Jr., and S. W. Running. 1993. A daily soil temperature model derived from air temperature and precipitation for continental applications. Climate Research **2**:183–191.