

# Variation among solar radiation data sets for the Eastern US and its effects on predictions of forest production and water yield

John D. Aber\*, Rita Freuder

Complex Systems Research Center, University of New Hampshire, Durham, New Hampshire 03824, USA

**ABSTRACT:** Spatially explicit modeling of ecosystem response to environmental change is often constrained as much by the accuracy of the data planes used to drive the model as by the differences in structure and parameterization of the models employed. In humid zones models of forest production at the site level are often more sensitive to solar radiation inputs than to other climate drivers. Radiation data are the most sparsely measured of routine climatological data. Still, we identified 3 different Web-based sources, containing 5 long-term mean radiation data sets, which differed in method of development and extrapolation. Comparisons are presented among the 5 data sets for estimated monthly mean total radiation as a fraction of total potential. The 5 vary by more than 20% in mean monthly solar radiation and result in differences in predicted woody biomass production of as much as 21%, up to 12% in total net primary production. We selected 1 data set to develop a statistical model of mean monthly radiation in units relevant to photosynthesis models (photosynthetically active radiation [PAR],  $\mu\text{mol m}^{-2} \text{s}^{-1}$ ) for the Eastern US, as part of the USDA Forest Service Eastwide global change effort (a joint effort of the Northern and Southern Global Change Programs). A simple multiple linear model predicted spatial PAR values with a standard error of the estimate of less than 6% in all months, and with lowest values (<3%) in mid-year.

**KEY WORDS:** Wood production · NPP · Water yield · VEMAP · NOAA · National Solar Radiation Data Base · Percent sunshine

## INTRODUCTION

Solar radiation is one of the key environmental parameters driving photosynthesis, transpiration and net energy exchange at the Earth's surface. Most models of carbon, water and energy fluxes in terrestrial ecosystems require solar radiation as an input (Cramer et al. 1999). In the absence of water stress, availability of photosynthetically active radiation (PAR) to the canopy may become the most limiting climatic factor acting within the constraints established by nutrient availability. In humid regions, model-based estimates of net primary production (NPP) are often quite sensitive to changes in solar radiation inputs (e.g. Aber et al. 1995, Pan et al. 1996). Absorbed PAR is strongly re-

lated to predicted NPP in global model applications (Ruimy et al. 1999), although different models, which treat water balances differently, may display different sensitivities (e.g. Jenkins et al. 1999).

The accuracy of spatially explicit models of ecosystem response to environmental change is often constrained as much by the accuracy of the input data planes as by differences in structure and parameterization of the models employed (e.g. Jenkins et al. 1999). A goal of the USDA Forest Service Eastwide analysis of global change impacts on forest ecosystems (a joint project of the USDA Forest Service Northern and Southern Global Change Programs; [www.fs.fed.us/ne/global/](http://www.fs.fed.us/ne/global/) and [www.sgcp.ncsu.edu](http://www.sgcp.ncsu.edu), respectively) is to make such spatially explicit predictions in order to assess not only total changes in forest production for the entire region, but also the spatial pattern. One approach to meeting this goal is through applications

\*E-mail: john.aber@unh.edu

of the PnET model (Aber et al. 1995, 1996, 1997), a simple, extensively tested, lumped parameter model of forest ecosystems. Running this model for the Eastern US requires the development of spatial data sets for climate drivers across the region, including solar radiation.

Our search for a data set on mean monthly solar radiation across this region revealed 3 different sources and a total of 5 data sets, with significant differences among them. The purpose of this paper is to present the results of: (1) a comparison of these 5 solar radiation data sets at 12 locations across the Eastern US, (2) a test of the sensitivity of NPP and water yield predictions from the PnET model to these differences, and (3) a statistical analysis of the spatial variation in 1 of these data sets resulting in generalized equations for predicting monthly mean values across the Eastern US.

## METHODS AND DATA SETS

**Solar radiation data sets.** Of the 3 primary climatic drivers of ecosystem production (temperature, precipitation, radiation), radiation is measured with the least spatial and temporal intensity. A sparse radiation measurement network was operated by the US Department of Commerce between 1955 and 1972, with different stations collecting data for different lengths of time within that period. While paper copies of collected data are to be found in government research offices, we could not locate this data set on the Web. In 1993, the National Oceanographic and Atmospheric Administration (NOAA) initiated the Surface Radiation Budget Network (SURFRAD), which measures all the major components of the surface radiation budget. However, as of June 1995, only 4 stations had been established, with 3 additional locations identified as desirable for future stations (Deluisi & Augustine 1995). The NOAA National Data Center, through which standard data products can be ordered (<http://nndc.noaa.gov/onlinestore.html>), does not display any radiation data, only 'sunrise/sunset information'.

As a result, existing solar radiation data planes for the US are all derived through relationships with other measured variables. The most direct and extensive data sets relating to radiation are the measurements of 'percent of possible sunshine' or 'sunshine duration' in hours per day. These data have long been a standard part of basic weather data collection around the world, and remain part of the World Meteorological Organization's standard data set (WMO 1984, <http://www.ncdc.noaa.gov/ol/climate/research/normals/wmo-normals.html>). This method involves the use of simple physical systems to record the length of time during a day that

the intensity of sunlight is above a certain threshold. For the US, data records for selected sites go back more than 100 yr. Sunshine duration data can be converted to energy units by regression where simultaneous measurements are available. To date, regression development has been carried out generally at the site level, but some regional extrapolations have been conducted (e.g. Coulson 1975, McEntee 1980, Iqbal 1983, Goodale et al. 1998). Other methods predict various solar radiation parameters by regressing site-level radiation data, where measured in energy units, against routinely measured climate parameters such as minimum temperature, maximum temperature and precipitation (e.g. Bristow & Campbell 1984, Running et al. 1987, Glassy & Running 1994). Again, these algorithms can only be developed where radiation measurements are available.

Several such synthesized radiation data sets covering the US are available through the Web. The Web is used increasingly as a source for environmental information, and the wealth of data available through this medium has greatly increased the exchange of basic data required to map and model environmental change. However, users of Web-based data are often left to their own devices for intercomparison of data from different sources, and quality assurance/quality control of Web-based data can be variable between sources. As a result, models using data from different Web sources might come to different conclusions based on differences in input data alone, and not on model structure (cf. Jenkins et al. 1999).

Our Web-based search for solar radiation data for the USDA Forest Service Eastwide region (all locations east of 94° W longitude, running roughly from Duluth, MN, and Des Moines, IA, through St. Louis, MO, to Lake Charles, LA) located 3 different sources with 5 contrasting radiation data sets. These included the following:

*National Solar Radiation Data Base (NSRDB):* Collocated by the National Renewable Energy Laboratory (NREL) in Golden, CO, this source contains data summarized daily, monthly and yearly for 239 stations across the US for the 30 yr period from 1961 to 1990. Direct normal, diffuse horizontal and global horizontal data are included. We use global horizontal data for this analysis as representing total downwelling radiation. We chose a total of 70 NSRDB stations within the USDA Forest Service Eastwide analysis region. Data are expressed in  $\text{Wh m}^{-2} \text{d}^{-1}$ .

This database is derived from the direct measurements of radiation made between 1955 and 1972 through the Department of Commerce as described above. Because of the sparse nature of this record, all data sets are incomplete. Missing data have been filled in using statistical models specific to each site, gener-

ally based on relationships with observed or interpolated cloud cover and aerosol optical depth. The NREL includes estimates of uncertainties with all data sets, and a description of the methods used to extrapolate between measurements. Nearly all of the data from the 12 sites used here were in class 4 for uncertainty (range of 6 to 9%), with some in class 5 (9 to 13%). Thus the full 30 yr mean is largely synthetic, but represents a consistent set of extrapolations based on measured relationships between radiation and climate variables at each site or subregion.

Web address: [http://rredc.nrel.gov/solar/old\\_data/nsrdb/dsf](http://rredc.nrel.gov/solar/old_data/nsrdb/dsf)

*National Climate Data Center (NCDC)*: Compiled by NOAA, this source contains the mean monthly number of hours that the intensity of sunshine is above a threshold value, expressed as a fraction of the maximum potential value (the period from sunrise to sunset). The same climate stations were used as for the NSRDB data set, but the data were collected using autographic records and sunshine duration totalizers, which record the fraction of a day above the threshold intensity. Records using these simpler technologies extend much farther in time and record lengths vary from 5 (Pensacola, FL) to 112 (Blue Hills, MA) yr.

Web address: <http://www.ncdc.noaa.gov/ol/climate/online/ccd/avgsun.html>

*Vegetation/Ecosystem Modeling and Analysis Project (VEMAP)*: VEMAP, funded by several federal agencies, has been responsible for coordinating data set and model development for prediction of global change effects on ecosystems across the US (Kittel et al. 1996). Using a  $0.5 \times 0.5^\circ$  grid, data sets have been developed expressing both mean and transient climatic conditions over the past century. More recently, predictions from different general circulation models (GCMs) have been interpolated to the VEMAP grid to generate data fields for the next 100 yr.

Two sets of mean radiation data are available through VEMAP, from the first and second phase of the program (VEMAP 1 and VEMAP 2). For VEMAP 1, radiation data for all sites were estimated using a universal algorithm which computes atmospheric transmittance from the diurnal temperature range (maximum to minimum) and then multiplies this by total solar radiation at the top of the atmosphere (Glasse & Running 1994). This approach assumes a stable atmosphere with minimal advection. In VEMAP 2, the basic equation for predicting transmittance remains the same, but data are extrapolated in a different way (Thornton et al. 1997). In this case, data are provided for each month, beginning in 1895. We computed average monthly values for the 1961 to 1990 period from this data stream.

Web address: <http://www.cgd.ucar.edu/vemap/>

**Analyses.** Our analysis here is in 3 parts. To examine the uncertainty in the use of any one existing solar radiation data set, we compare estimates for 12 sites in the Eastern US (Table 1). The 12 sites were selected to represent different subregions within the larger Eastwide area. As differences in input data planes are only important in a regional modeling context if the model used is shown to be sensitive to those differences, we compare predictions from the PnET model obtained for these 12 locations using each of the 5 identified radiation databases. Finally, based on analyses presented below, we selected the NSRDB database for derivation of the Eastwide radiation data plane and used data from all 70 sites located in the Eastwide region to create a simple statistical model of variation in mean monthly solar radiation across the region.

*Data conversions for comparison of data sets:* We used fraction of total possible radiant energy received as the units for comparing the 5 data sets as this involved the fewest conversions to put all the data in the same units.

Both VEMAP data sets contain a data plane expressing solar radiation inputs as a fraction of potential (the FSR data plane). As VEMAP is a gridded data set, we used data from the grid cells which included the 12 cities. While the mean location and elevation of the grid cell may differ from the actual location of the city, the relatively slow and continuous change in radiation with latitude, longitude and elevation (see 'Results and discussion') suggests that this approximation does not introduce significant error.

The NCDC data are expressed as a fraction as well, but the basic measurement is sunshine duration (see discussion above) rather than total energy received. It has long been recognized that fractional sunshine duration is not equivalent to the fraction of possible solar radiation received expressed in energy units. However, the abundance of sunshine duration data

Table 1. The 12 sites used to determine model sensitivity to differences in solar radiation input data sets

Site	Latitude	Longitude	Elevation (m)
Asheville, NC	35.4	82.5	362
Atlanta, GA	33.7	84.4	315
Burlington, VT	44.5	73.2	104
Columbia, SC	33.9	80.0	69
Columbus, OH	40.0	82.9	249
Little Rock, AR	34.7	92.2	81
Miami, FL	25.8	80.3	2
Madison, WI	43.1	89.3	264
Minneapolis, MN	44.9	93.2	255
Portland, ME	43.0	70.3	19
Richmond, VA	37.5	77.3	54
Syracuse, NY	43.1	76.1	124

and the paucity of direct radiation measurements has lead to efforts to develop correlations between the two (see discussions in Coulson 1975, Iqbal 1983). Empirical relationships have been developed for many sites and are generally of the form

$$I/I_0 = a + bPS \quad (1)$$

where  $I/I_0$  = fraction of potential solar radiation (where both  $I$  and  $I_0$  are expressed in energy units), PS = percent sunshine, and  $a$  and  $b$  are empirical coefficients. Not surprisingly, the values for  $a$  and  $b$  differ from site to site depending on climatological regime, especially humidity and degree of cloudiness. For this exercise we used a general conversion equation ( $a = 0.18$ ,  $b = 0.62$ ) presented by Iqbal (1983). However, some ecosystem models use radiation inputs described as a fraction of the possible total and hence there is the potential for the misuse of raw NCDC percent sunshine data in this regard. To examine the consequences of this misuse, we used the original percent sunshine data (NCDC) and the adjusted data (using the above equation, NCDC adj.) in the modeling analysis.

For the 12-site comparison, we converted NSRDB data from  $\text{Wh m}^{-2} \text{d}^{-1}$  to a fraction by dividing each value by the maximum potential value listed for each site and month in the NSRDB database. Maximum potential represents radiation received at the top of the canopy and is calculated from latitude, day of year, and the solar 'constant'.

*Model applications:* Differences in input data planes are only important in a regional modeling context if the model used is shown to be sensitive to those differences. To test for this sensitivity, we used the 5 different radiation data sources as input to the PnET-II model, a simple, lumped parameter model of forest carbon, water and N interactions (Aber et al. 1995, Ollinger et al. 1998). The PnET-II model uses foliar nitrogen concentrations to estimate maximum potential photosynthesis. Standard relationships for the effects of light attenuation, temperature and water stress are used to reduce this potential to realized photosynthesis. Assuming a linear relationship between realized photosynthesis and leaf conductance, actual transpiration is a function of photosynthetic rate and vapor pressure deficit. Empirical relationships are used to drive phenology of canopy display and the allocation of fixed carbon to biomass production and respiration. The model has been tested against measured data for net primary production, woody biomass production and total water yield (drainage or stream flow) at several intensive study sites and gauged watersheds across the northeastern US (Aber et al. 1995, Ollinger et al. 1998).

In this application, we ran PnET-II for the 12 selected sites (Table 1) using standard vegetation parameters

for a broad-leaved deciduous forest, a pine forest and a spruce-fir forest (Aber et al. 1995). The purpose of using these 3 different vegetation types was to see if sensitivity to variation in radiation data was altered by the photosynthetic capacity of the stand. Under constant conditions, potential photosynthesis and growth rate decline in the sequence deciduous > pine > spruce-fir. PnET-II requires monthly mean maximum and minimum temperature and precipitation in addition to radiation data. These other parameters were obtained from the VEMAP 1 data set by locating each of the 12 sites in a VEMAP pixel and accessing the mean monthly climate data for that pixel. A previous analysis has shown that precipitation and temperature data planes derived by VEMAP and by direct, local statistical analysis for the northeastern US were indistinguishable (Jenkins et al. 1999). All sites were run with a soil water holding capacity of 12 cm. Thus the 5 different PnET-II runs for each of the location/vegetation combinations varied only with respect to the radiation data plane used. A total of 180 runs were performed (12 sites  $\times$  3 vegetation types  $\times$  5 radiation data sets).

The PnET model uses solar radiation inputs expressed as PAR  $\mu\text{mol m}^{-2} \text{s}^{-1}$ . VEMAP (SR data plane) data are expressed in  $\text{kJ m}^{-2} \text{d}^{-1}$  and the NSRDB data are in  $\text{Wh m}^{-2} \text{d}^{-1}$ , while NCDC data are expressed as a fraction of potential. Both of the NCDC data sets were multiplied by potential radiation from the VEMAP data set to give predicted actual values. All data were then converted to PAR using the following equation:

$$\frac{\mu\text{mol}}{\text{m}^2 \text{s}} = \frac{\text{Wh}}{\text{m}^2 \text{d}} \cdot \frac{3.6 \text{kJ}}{\text{Wh}} \cdot \frac{2.05 \text{mol}}{10^3 \text{kJ}} \cdot \frac{\text{h}}{3600 \text{s}} \cdot \frac{\text{d}}{\text{daylength (h)}} \cdot \frac{10^6 \mu\text{mol}}{\text{mol}} \quad (2)$$

where 2.05 mol of photons per kJ radiation is derived from a comparison based on data from the Harvard Forest, Petersham, MA (Aber et al. 1996), and daylength is the number of hours between sunrise and sunset. The conversion from Wh to kJ was only required for the NSRDB data set.

*Derivation of Eastwide US solar radiation data plane:* For reasons explained in the 'Results and discussion', we selected the NSRDB data set for use in developing a spatial data plane for solar radiation across the Eastern US. To predict monthly mean solar radiation we used data from all 70 NSRDB locations in the region, converted to PAR, and regressed this value against latitude, longitude and elevation individually and in a multiple linear regression to derive equations applicable to any location within the region. The multiple linear equation was used in conjunction with an existing (VEMAP) 2.5' resolution digital elevation model of the Eastern US to generate spatial maps of

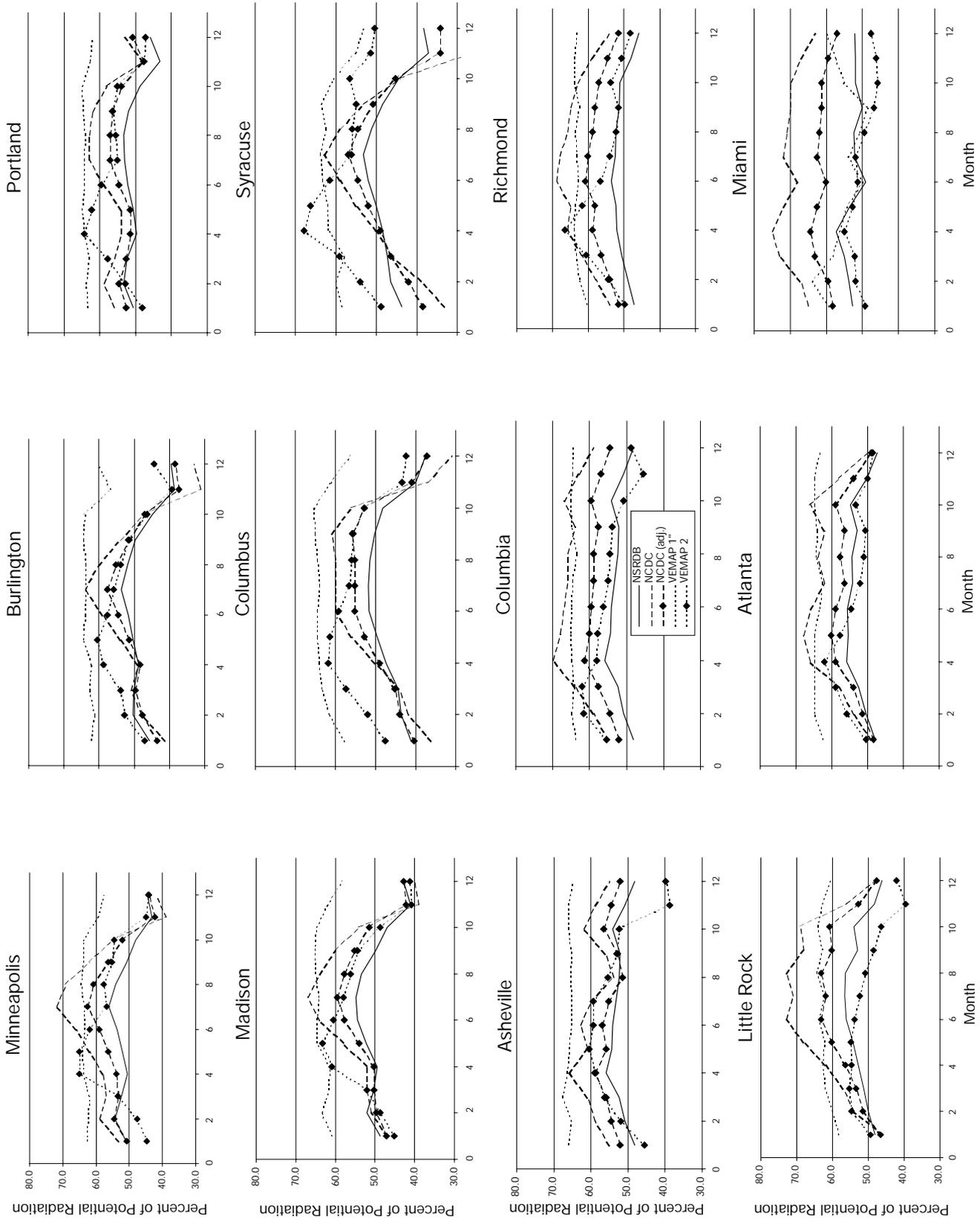


Fig. 1. Monthly mean solar radiation data, expressed as a percent of total potential, for 12 sites in the Eastern US, from 3 different databases. NCDC: National Climate Data Center, NSRDB: National Solar Radiation Data Base, VEMAP: Vegetation/Ecosystem Mapping and Analysis Program

PAR for different months. The same equations can be applied at any spatial resolution.

## RESULTS AND DISCUSSION

### Solar radiation data sets

Variation in the estimated fraction of total solar radiation received varied widely between data sources both spatially and temporally (Fig. 1). The NSRDB data set was generally the lowest of the 5, and showed the least seasonal variation. The NCDC data set was consistently higher than NSRDB and showed the most extreme seasonal variation, with highest values in summer and lowest in winter. Conversion of the NCDC percent sunshine data to energy units (NCDC adj.), reduced the seasonal variability considerably and brought the values closer to the NSRDB data. The VEMAP 1 data set showed similar temporal variation in comparison with NSRDB, but all values were considerably higher, except for the Miami data set. The VEMAP 2 data show increased seasonal variation in comparison with VEMAP 1, but lower overall averages. VEMAP 2 and NSRDB data are similar in many locations and months, especially in mid-summer and more southern locations.

Summarized for the entire year (Fig. 2), the NSRDB data set yields the lowest overall average fraction of solar radiation received. NCDC adj. and VEMAP 2 are next, followed by NCDC, with VEMAP 1 giving the highest average. NCDC data show a larger difference between annual mean and mid-year mean (April to September), such that NCDC and VEMAP 1 both have

the highest average value for this period, followed by VEMAP 2 and NCDC adj., with NSRDB having again the lowest values.

### Model applications

Results of the PnET-II runs using different solar radiation inputs showed no interaction with forest type. Fractional changes in productivity and water yield with different radiation inputs were consistent across forest type, so no further discussion on this potential source of variation in results is included. Those runs using data sets with higher solar radiation inputs showed increases in net photosynthesis and transpiration, yielding higher forest productivity and slightly lower water yield (Fig. 3). In all cases, regression analyses showed that the relationships between PnET-II predictions using NSRDB data and those using other data sets were linear with a slope not significantly different from 1 ( $p > 0.05$ ). Regression intercepts were significant in all cases. Combined, these 2 results demonstrate a constant percent bias in predictions due to the use of the different solar radiation data sets which can be examined as ratios (Fig. 4).

Differences between PnET-II predictions resulting from different radiation data sets were most pronounced for woody biomass production, ranging from 8 to 21% higher than values obtained with the NSRDB data set (Fig. 4). Differences in total NPP ranged from 5 to 12% and in water yield from -2 to -7%. In general, increases in production were linear with, and proportional to, increases in radiation expressed either as annual or mid-year averages (Fig. 5), although an asymptote may be suggested by the runs using the highest values for annual radiation data (VEMAP 1). The greater response in wood production compared with total NPP results from allocation routines in PnET which determine allocation to foliage based on an algorithm which optimizes total carbon gain, rather than using a constant proportion of total NPP (Aber et al. 1995).

While the range of variation in PnET-predicted NPP due to differences in solar radiation data sets lies within the range of error in field measurements, this is not simply random variation, but a consistent bias. The accumulation of such significant biases over entire regions could significantly alter predictions of carbon sequestration by forest ecosystems. Currently, the East-

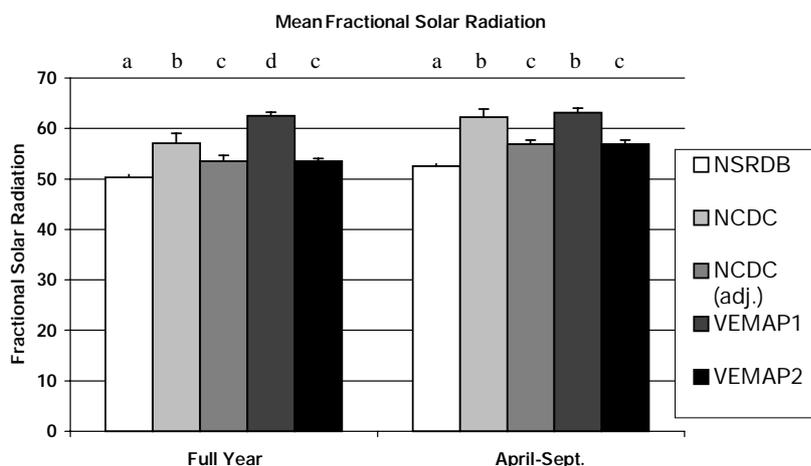


Fig. 2. Annual mean fractional solar radiation by data source. Data are averaged over the whole year and for a mid-year period (April to September). Small letters show significant differences ( $p < 0.05$ ) between data sets for the same time periods (e.g. between NSRDB annual and NCDC annual). Error bars are 1 SE of the mean

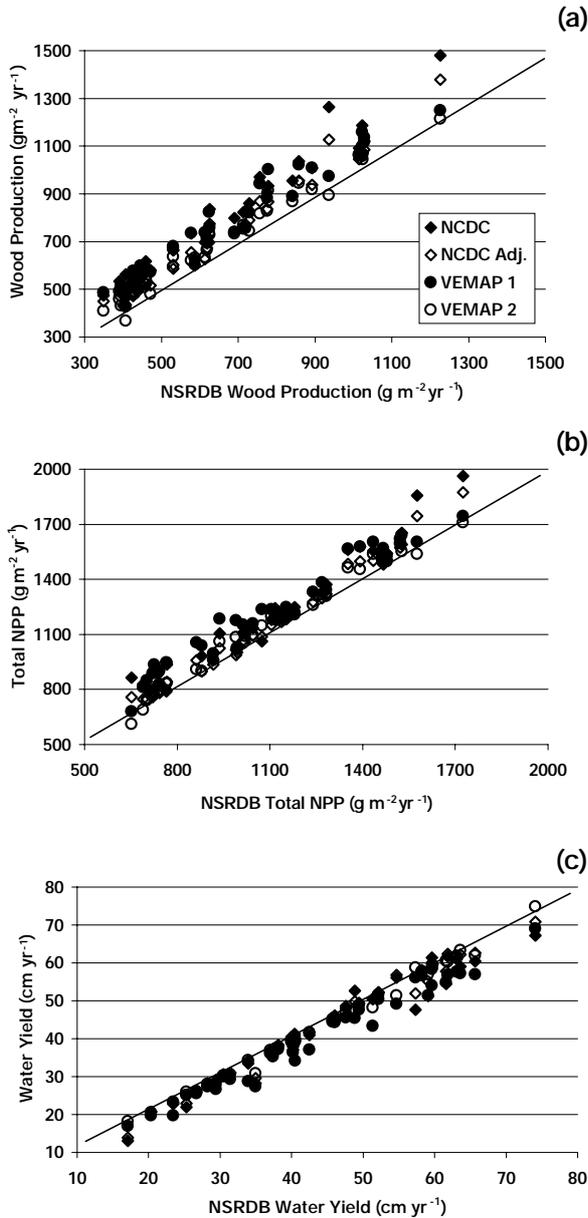


Fig. 3. Effects of different solar radiation data sets on forest ecosystem processes as predicted by the PnET-II model: (a) woody biomass production, (b) total net primary production, and (c) water yield. Values are expressed relative to results from runs using the NSRDB database. There are a total of 36 points per radiation data set (12 sites  $\times$  3 vegetation types)

ern US is identified as a potential major carbon sink (Fan et al. 1998), and eddy covariance measurements of net carbon sequestration in forests have been linked to the production of woody biomass (Goulden et al. 1996). If predicted woody biomass accumulation were to be used as a surrogate for measurements of C balance by eddy covariance where such measurements are not made, then an 8 to 21% high bias in this prediction could have significant consequences.

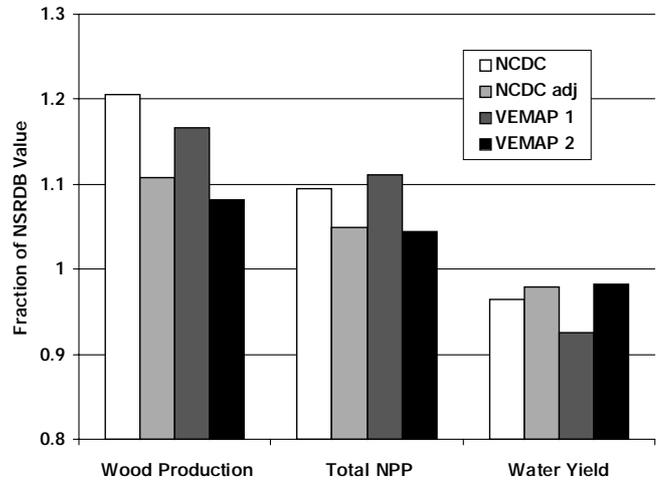


Fig. 4. Fractional change in ecosystem processes using different solar radiation data sets as compared with NSRDB. Based on the same data presented in Fig. 3

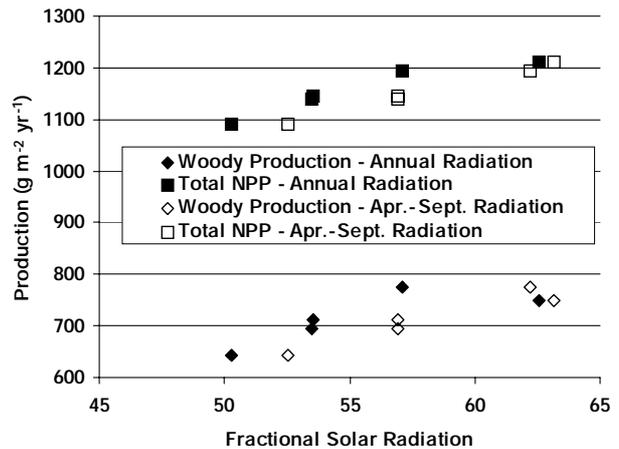


Fig. 5. Response of woody and total production to differences in radiation data sets expressed as annual or mid-year average fraction of total solar radiation

Use of different data sets can also affect testing or validation of models. For example, using the unadjusted NCDC data set as an estimate of fraction of solar energy received would lead to overestimates of 21 and 9% in wood and total NPP, respectively, and an underestimate of 4% in water yield. This would significantly alter the results of any validation exercise.

### A regional solar radiation data plane

We selected the NSRDB data set to use as the standard for comparisons of model runs, and for development of the Eastwide regional radiation data plane. This data set is derived most directly from measurements of radiation inputs in energy units within the

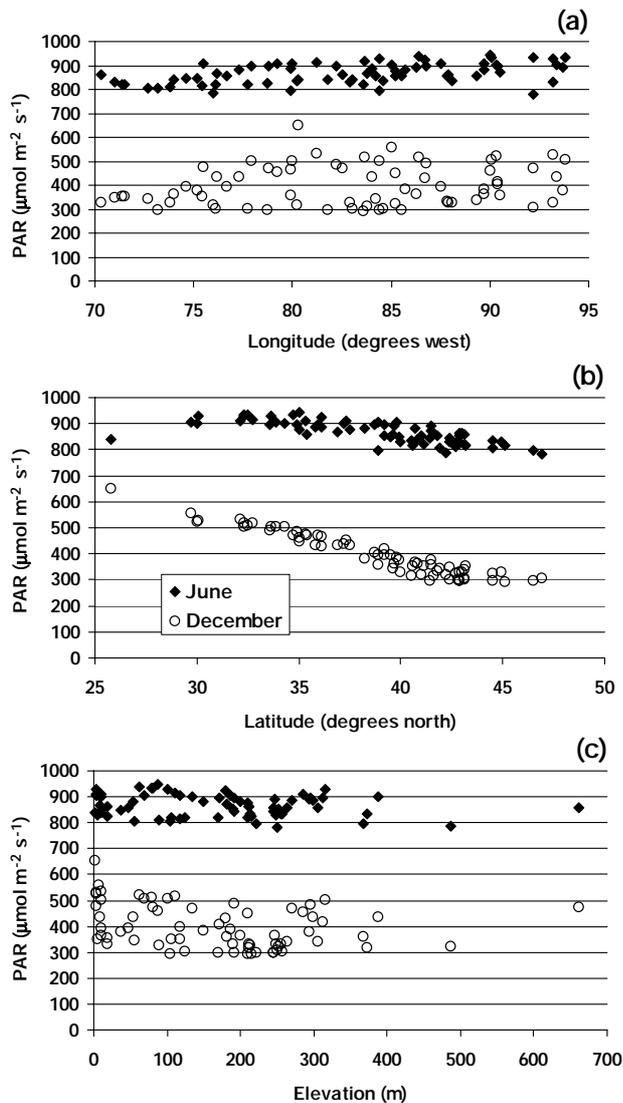


Fig. 6. Radiation data from the NSRDB database for 70 sites in the Eastern US converted to units of photosynthetically active radiation (PAR) in relation to: (a) longitude, (b) latitude, and (c) elevation. Closed diamonds are data for June, open circles are for December

region, even though those measurements are sparse in space and time. The adjusted NCDC values are close to NSRDB values on average, but the degree of similarity differs across the region (Fig. 1). If it were possible to derive site-specific or regionalized relationships (e.g. Eq. 1) between percent of possible sunshine from the NCDC data set and actual measured radiation loading, this combination might prove optimal for defining an absolute radiation data set. However, the lack of on-site radiation measurements rules out this approach. Given this, the NSRDB data set at least relies on statistical extrapolation from actual radiation measurements rather than on conversion from intensity duration of sunshine (NCDC) or on generalized

algorithms for converting other climatic variables into radiation data (VEMAP).

In previous regional applications, we have been able to develop simple multiple linear statistical models for the prediction of seasonal and spatial variation in climate data (Ollinger et al. 1993, 1995, Goodale et al. 1998). Across the Eastern US mean monthly radiation levels in the NSRDB data set vary consistently with latitude, longitude and elevation (Fig. 6). Not surprisingly, the strongest variation is with latitude in winter, but regression analyses show that both latitude and longitude are significant in all months of the year ( $p \ll 0.05$ ). For elevation, the relationship with radiation is stronger in winter than in summer, but is significant in all months except August (Table 2).

There is no indication of non-linearity in the relationship of radiation level with latitude, longitude or elevation (Fig. 6), so multiple linear regression equations were developed for predicting radiation levels across the Eastern US at monthly intervals. While  $R^2$  values for these equations are highest in winter and lowest in summer (Table 3), this results from the much narrower range of PAR values in summer (Fig. 6). Standard errors of the estimate (SEE) produced by the regressions range from 19.9 to 34.2  $\mu\text{mol m}^{-2} \text{s}^{-1}$  with no seasonal trend. As mean PAR values are much higher in the summer than in the winter, the ratio of SEE to the mean value across the region is lowest in mid-summer (Table 3). Applying the resulting equations for several months and comparing results with measured data demonstrates the goodness-of-fit, and lack of bias, in this regression approach (Fig. 7). This is not a validation of the spatial model as the data used to test the model are the same as those used to develop the equations, thus the goodness-of-fit description. The equations in Table 3 can be used in conjunction with a digital elevation model of any spatial resolution to generate maps of predicted mean solar radiation at monthly intervals for the Eastern US (e.g. 4 km resolution, Fig. 8).

## CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

We conclude that:

(1) Published solar radiation data sets for the Eastern US vary significantly. This is because they are developed from different sets of measurements (sparse radiation measurements or more intense ancillary measurements) and use different methods for spatial and temporal extrapolation. Users of any one of these data sets should be aware of this variation and its sources.

(2) Ecosystem model sensitivity to differences in solar radiation inputs can be high in humid regions where water limitations are moderate.

Table 2. Results from linear regression of NSRDB monthly solar radiation data against longitude, latitude and elevation (radiation in  $\mu\text{mol m}^{-2} \text{s}^{-1}$ )

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Constant	1014.4	1077.3	1206.3	1386.4	1247.4	1130.7	1029.32	1147.1	1209.5	1395.2	1280	1090.9
Latitude	-14.51	-13.16	-13.51	-15.5	-10.56	-6.803	-4.287	-8.291	-12.3	-19.97	-21.1	-17.88
R <sup>2</sup>	0.88	0.83	0.76	0.86	0.82	0.53	0.29	0.64	0.85	0.92	0.95	0.92
SEE	24.2	27.3	34.1	28.1	22.5	28.6	29.3	28	22.9	25.6	22.6	23.3
SEE/mean	0.053	0.048	0.050	0.036	0.027	0.033	0.034	0.034	0.031	0.041	0.049	0.058
Constant	267.5	381.7	546.7	598.8	604.3	610.3	612.2	544	506.5	297	217	191.8
Longitude	2.23	2.24	1.65	2.26	2.82	3.1	3.027	3.4	2.73	3.92	2.96	2.49
R <sup>2</sup>	0.03	0.04	0.009	0.02	0.11	0.22	0.31	0.21	0.07	0.06	0.02	0.02
SEE	68.4	63.9	69.2	74	49.5	36.8	29.1	41.2	57.5	90.35	96.1	82.6
Constant	476.8	593.5	711.4	814	858.2	877.7	870.1	838.5	755.6	658.3	500	431.2
Elevation	-0.143	-0.153	-0.166	-0.164	-0.115	-0.062	-0.04	-0.074	-0.131	-0.215	-0.218	-0.192
R <sup>2</sup>	0.057	0.079	0.082	0.066	0.067	0.02	0.007	0.027	0.07	0.08	0.07	0.07
SEE	67.4	62.4	66.6	72.5	50.6	41.2	34.8	45.8	57.7	89.6	93.9	80.3
p	0.026	0.011	0.009	0.018	0.017	0.03	0.05	0.089	0.017	0.012	0.015	0.0125

Table 3. Results of multiple linear regression of NSRDB solar radiation data against longitude, latitude and elevation (radiation in  $\mu\text{mol m}^{-2} \text{s}^{-1}$ )

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Constant	974.5	993.0	1187.3	1348.1	1068.5	880.9	746.1	886.7	1067.4	1224.2	1243.9	1055.0
Latitude	-14.29	-12.57	-13.12	-15.20	-9.64	-5.73	-3.03	-7.18	-11.54	-19.02	-20.79	-17.54
Longitude	0.408	0.835	0.140	0.374	1.835	2.604	2.940	2.712	1.450	1.742	0.346	0.343
Elevation	-0.0140	-0.0470	-0.0436	-0.0253	-0.0515	-0.0461	-0.0551	-0.0454	-0.0437	-0.0615	-0.0273	-0.0314
R <sup>2</sup>	0.88	0.83	0.76	0.86	0.86	0.66	0.53	0.75	0.87	0.94	0.95	0.92
SEE	24.36	27.03	34.17	28.40	19.89	24.35	23.93	23.26	21.51	23.51	22.75	23.39
Mean	452.50	567.50	683.20	786.20	838.60	867.20	863.30	826.00	733.20	621.70	462.90	398.50
SEE/mean	0.054	0.048	0.050	0.036	0.024	0.028	0.028	0.028	0.029	0.038	0.049	0.059

(3) Within the Eastern US, a simple multiple linear statistical model can be developed which can predict mean monthly solar radiation, as presented in the NSRDB database, to within  $\pm 5\%$  of measured values.

We also offer 2 suggestions for increasing the accuracy and usefulness of data sets of this critical environmental parameter:

(1) Efforts should be made to extend the NOAA radiation monitoring activity to increase the number of sites at which downwelling radiation in both energy units and as PAR are measured. Data should be posted to the Web as soon as quality assurance and control are completed.

(2) An investigation of the relationship between 'sunshine duration' and daily total radiational energy input should be undertaken using data from a large number of very different environments with the goal of developing generalized equations for converting between these 2 parameters. If such

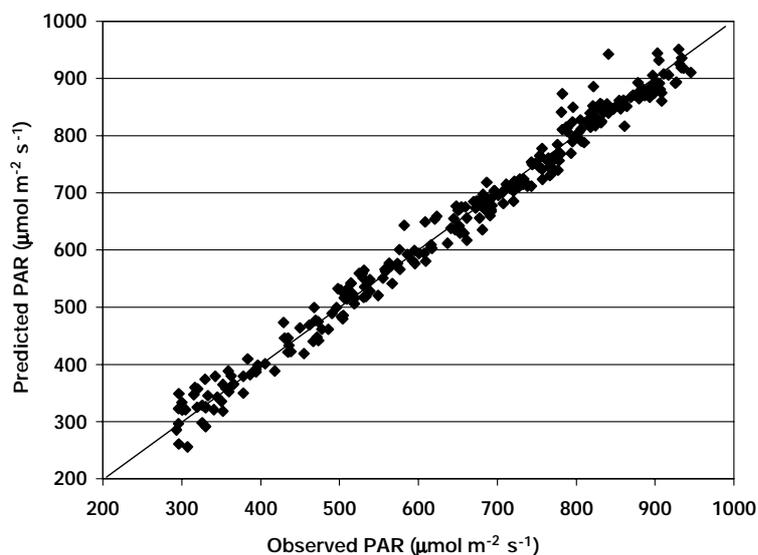


Fig. 7. Predicted versus observed mean monthly PAR using the multiple linear equations in Table 3. This is not a validation of the equations as the observed data were used to derive the equations. Rather, this demonstrates goodness-of-fit for the equations

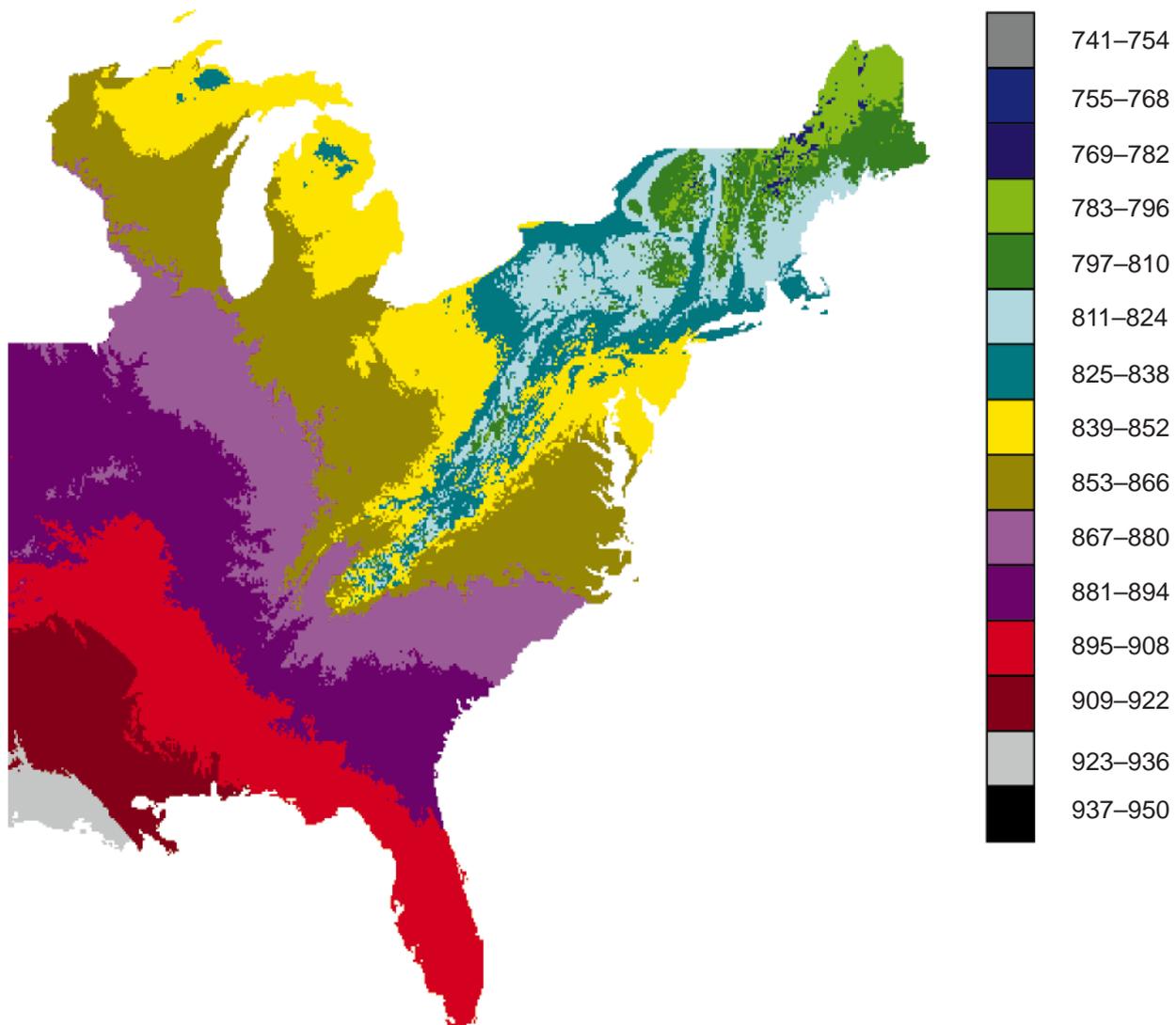


Fig. 8. Spatial distribution of mean monthly PAR values for the Eastern US for July as predicted by the multiple linear regression equation in Table 3. Values in  $\mu\text{mol m}^{-2} \text{s}^{-1}$

relationships prove both generalizable and accurate, then the existing and extensive time series data on 'sunshine duration' can be converted to units required for ecosystem models.

*Acknowledgements.* This research was supported by the USDA Forest Service Northern and Southern Global Change Programs through cooperative agreements with the University of New Hampshire. We thank Scott Ollinger for comments on the manuscript.

#### LITERATURE CITED

- Aber JD, Ollinger SV, Federer CA, Reich PB, Goulden ML, Kicklighter DW, Melillo JM, Lathrop RG (1995) Predicting the effects of climate change on water yield and forest production in the northeastern US. *Clim Res* 5:207–222
- Aber JD, Reich PB, Goulden ML (1996) Extrapolating leaf  $\text{CO}_2$  exchange to the canopy: a generalized model of forest photosynthesis validated by eddy correlation. *Oecologia* 106:257–265
- Aber JD, Ollinger SV, Driscoll CT (1997) Modeling nitrogen saturation in forest ecosystems in response to land use and nitrogen deposition. *Ecol Model* 101:61–78
- Bristow KL, Campbell GS (1984) On the relationship between incoming solar radiation and daily maximum and minimum temperature. *Agric For Meteorol* 31:159–166
- Coulson, KL (1975) *Solar and terrestrial radiation*. Academic Press, New York
- Cramer W, and 43 others (1999) Comparing global models of terrestrial net primary productivity (NPP): overview and key results. *Global Change Biol* 5(Suppl 1):1–15
- DeLuisi J, Augustine JA (1995) Monitoring surface radiation for climate, health, and the environment. Project title: the Surface Radiation Budget Network (SURFRAD). National Climate Data Center, Asheville, NC (accessed: April

- 1999). Available at <http://www.ncdc.noaa.gov/ol/climate/research/ogp/papers/deluisi.html>
- Fan S, Gloor M, Pacala S, Sarmiento J, Takahashi T, Tans P (1998) A large terrestrial carbon sink in North America implied by atmospheric and oceanic carbon dioxide data and models. *Science* 282:442–446
- Glassy JM, Running SW (1994) Validating diurnal climatology logic of the MT-CLIM model across a climatic gradient in Oregon. *Ecol Appl* 4:248–257
- Goodale CL, Aber JD, Ollinger SV (1998) Mapping monthly precipitation, temperature, and solar radiation for Ireland with polynomial regression and a digital elevation model *Clim Res* 10:35–49
- Goulden ML, Munger JM, Fan SM, Daube, Wofsy SC (1996) Exchange of carbon dioxide by a deciduous forest: response to interannual climate variability. *Science* 271:1576–1578
- Iqbal M (1983) An introduction to solar radiation. Academic Press, Toronto
- Jenkins JC, Kicklighter DW, Ollinger SV, Aber JD, Melillo JM (1999) Sources of variability in NPP predictions at a regional scale: a comparison using PnET-II and TEM 4.0 in northeastern US forests. *Ecosystems* 2:555–570
- Kittel TGF, Rosenbloom NA, Painter TH, Schimel DS, Fisher HH, Grimsdell A, VEMAP Participants, Daly C, Hunt ER Jr (1996) The VEMAP phase I database: an integrated input dataset for ecosystem and vegetation modeling for the coterminus United States. CD-ROM, Fort Collins, CO (accessed: April 1999); available from <http://www.cgd.ucar.edu/vemap/>
- McEntee MA (1980) A revision of the equation relating sunshine hours to radiation income for Ireland. *Irish J Agric Res* 19:119–125
- Ollinger, SV, Aber JD, Lovett GM, Millham SE, Lathrop RG (1993) A spatial model of atmospheric deposition for the northeastern US. *Ecol Appl* 3:459–472
- Ollinger SV, Aber JD, Federer CA, Lovett GM, Ellis J (1995) Modeling physical and chemical climatic variables across the northeastern US for a Geographic Information System. General Technical Report. NE-191, US Forest Service, Radnor, PA
- Ollinger SV, Aber JD, Federer CA (1998) Estimating regional forest productivity and water yield using an ecosystem model linked to a GIS. *Landscape Ecol* 13:323–334
- Pan Y, McGuire AD, Kicklighter DW, Melillo JM (1996) The importance of climate and soils for estimates of net primary production: a sensitivity analysis with terrestrial ecosystem model. *Global Change Biol* 2:5–23
- Ruimy A, and 43 others (1999) Comparing global models of terrestrial net primary productivity (NPP): analysis of differences in light absorption and light-use efficiency. *Global Change Biol* 5(Suppl 1):56–64
- Running SW, Nemani RR, Hungerford RD (1987) Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evapotranspiration and photosynthesis. *Can J For Res* 17:472–483
- Thornton PE, Running SW, White MA (1997) Generating surfaces of daily meteorological variables over large regions of complex terrain. *J Hydrol* 190:214–251
- World Meteorological Organization (1984) Technical regulations. WMO Publication No. 49, World Meteorological Organization, Geneva

*Editorial responsibility: Brent Yarnal,  
University Park, Pennsylvania, USA*

*Submitted: June 17, 1999; Accepted: November 11, 1999  
Proofs received from author(s): March 6, 2000*