Global Change Biology (2005) 11, 666–684, doi: 10.1111/j.1365-2486.2005.00936.x

Site-level evaluation of satellite-based global terrestrial gross primary production and net primary production monitoring

DAVID P. TURNER*, WILLIAM D. RITTS*, WARREN B. COHEN†, THOMAS K. MAEIRSPERGER*, STITH T. GOWER‡, AL A. KIRSCHBAUM‡, STEVE W. RUNNING§, MAOSHENG ZHAO§, STEVEN C. WOFSY¶, ALLISON L. DUNN¶, BEVERLY E. LAW*, JOHN L. CAMPBELL*, WALTER C. OECHEL‖, HYO JUNG KWON‖, TILDEN P. MEYERS**, ERIC E. SMALL††, SHIRLEY A. KURC†† and JOHN A. GAMON‡‡

*Department of Forest Science, Oregon State University, Corvallis, OR 97331-7501, USA, †USDA Forest Service, 3200 SW Jefferson Way, Corvallis, OR 97331, USA, ‡Department of Forest Ecology and Management, University of Wisconsin, Madison, WN 53706, USA, §School of Forestry, University of Montana, Missoula, MT 59812, USA, ¶Department of Earth and Planetary Science, Harvard University, Cambridge, MA 02138, USA, ‖Global Change Research Group, San Diego State University, San Diego, CA 92182, USA, **National Oceanic and Atmospheric Administration, Atmospheric Turbulence and Diffusion Division, Oak Ridge, TN 37831, USA, ††Department of Geological Sciences, University of Colorado, Boulder, CO 80309, USA, ‡‡Center for Environmental Analysis and Department of Biological Sciences, California State University, Los Angeles, CA 90032, USA

Abstract

Operational monitoring of global terrestrial gross primary production (GPP) and net primary production (NPP) is now underway using imagery from the satellite-borne Moderate Resolution Imaging Spectroradiometer (MODIS) sensor. Evaluation of MODIS GPP and NPP products will require site-level studies across a range of biomes, with close attention to numerous scaling issues that must be addressed to link ground measurements to the satellite-based carbon flux estimates. Here, we report results of a study aimed at evaluating MODIS NPP/GPP products at six sites varying widely in climate, land use, and vegetation physiognomy. Comparisons were made for twenty-five 1 km^2 cells at each site, with 8-day averages for GPP and an annual value for NPP. The validation data layers were made with a combination of ground measurements, relatively high resolution satellite data (Landsat Enhanced Thematic Mapper Plus at $\sim$ 30 m resolution), and process-based modeling. There was strong seasonality in the MODIS GPP at all sites, and mean NPP ranged from 80 g C m^{-2} yr^{-1} at an arctic tundra site to 550 g C m^{-2} yr^{-1} at a temperate deciduous forest site. There was not a consistent over- or underprediction of NPP across sites relative to the validation estimates. The closest agreements in NPP and GPP were at the temperate deciduous forest, arctic tundra, and boreal forest sites. There was moderate underestimation in the MODIS products at the agricultural field site, and strong overestimation at the desert grassland and at the dry coniferous forest sites. Analyses of specific inputs to the MODIS NPP/GPP algorithm – notably the fraction of photosynthetically active radiation absorbed by the vegetation canopy, the maximum light use efficiency (LUE), and the climate data – revealed the causes of the over- and underestimates. Suggestions for algorithm improvement include selectively altering values for maximum LUE (based on observations at eddy covariance flux towers) and parameters regulating autotrophic respiration.

Keywords: carbon, FPAR, global, gross primary production, light use efficiency, MODIS, net primary production, satellite remote sensing, scaling, validation

Received 6 July 2004; revised version received and accepted 25 August 2004

Correspondence: David P. Turner, tel. + 1 541 737 5043, fax + 1 541 737 1393, e-mail: david.turner@oregonstate.edu

Introduction

Regular monitoring of global terrestrial net primary production (NPP) and gross primary production (GPP) is needed for the purposes of evaluating trends in biospheric behavior (e.g. Nemani *et al.*, 2003), understanding the role of the biosphere in the global carbon cycle (Lucht *et al.*, 2002), and investigating large-scale patterns in food and fiber production (Running *et al.*, 2004). The Moderate Resolution Imaging Spectroradiometer (MODIS) sensor aboard the National Aeronautics and Space Administration (NASA) Terra polar orbiting satellite was designed in part for this purpose and continuous global NPP/GPP monitoring with satellite data has recently become operational (Running *et al.*, 2000, 2004). Satellite-based NPP/GPP products require validation to be useful for scientific purposes, but validation raises a host of scaling issues – notably the mismatch in scale between the 1 km grid of the MODIS products and the relatively small area over which NPP and GPP can be measured on the ground (Reich *et al.*, 1999; Turner *et al.*, 2004a). In this paper, we describe an approach to producing NPP/GPP validation data layers at the site level and report comparisons with MODIS products at six sites differing widely in vegetation physiognomy, phenology, and productivity.

Several design criteria can be envisioned for an effective scheme to validate MODIS NPP/GPP products at a given site. (1) The validation site should be large enough ($>9\,km^2$) to allow comparisons across multiple MODIS 1 km^2 cells so that sampling errors and georeferencing errors are minimized. (2) The validation data sets should be wall-to-wall surfaces, but anchored in ground measurements at specific georeferenced points. (3) The spatial and temporal resolution of the validation products should closely match those of the MODIS products. (4) The components of the MODIS NPP/GPP algorithm should be analyzed along with the products themselves so as to interpret possible errors or limitations. Of particular note is that GPP and autotrophic respiration (R_a) are estimated separately in the MODIS NPP algorithm and require separate assessment.

The specific components of the algorithm that produces the MODIS NPP/GPP products (Fig. 1), include climate- and satellite-based inputs as well as a look-up table for biome-specific parameters related to photosynthesis and autotrophic respiration (Running *et al.*, 2000). These parameters cover allometric variables such as specific leaf area, and ecophysiological variables such as maximum light use efficiency (LUE). The parameters are for the most part directly measurable and the degree to which site-specific values differ from biome-wide averages is of interest. The daily climate input data – incoming solar radiation, minimum temperature, and vapor pressure deficit (VPD) – are generated independently by a general circulation climate model that assimilates observations from satellites and a network of meteorological stations. Because the climate model is run at a coarse spatial resolution (1° latitude × 1.25° longitude), there is a significant loss of information at the 1 km scale of the MODIS products that must be evaluated. The primary inputs derived directly from the MODIS sensor itself are the fraction of photosynthetically active radiation that is absorbed by the canopy (FPAR) and leaf area index (LAI). These must be examined in terms of absolute magnitude, temporal patterns, and sensitivity to fine-scale spatial heterogeneity.

The BigFoot Project (http://www.fsl.orst.edu/larse/bigfoot/index.html) was organized specifically to address the MODIS product validation needs, and has formulated a scaling logic that largely meets the criteria noted above. The BigFoot NPP/GPP validation approach links ground measurements of NPP and GPP with high spatial resolution (30 m) remote sensing data and modeling. An earlier BigFoot paper (Turner *et al.*, 2003a) focused on validation of MODIS GPP at two

$$\mathbf{GPP} = \downarrow\mathbf{PAR} \times \mathbf{FPAR} \times (\varepsilon_{g\text{-}max} \times S_{Tmin} \times S_{VPD}),$$

$$\mathbf{NPP} = \mathbf{GPP} - R_a,$$

$$R_a = R_m + R_g$$

where

GPP	=	gross primary production (g C m^{-2} day^{-1})
↓PAR	=	incoming photosynthetically active radiation, from Data Assimilation Office (DAO) climate model
FPAR	=	fraction of ↓PAR absorbed by the plant canopy, from MODIS reflectances
$\varepsilon_{g\text{-}max}$	=	maximum light use efficiency (g C MJ^{-1}), from lookup table
S_{Tmin}	=	minimum temperature scalar (0–1), air temperature is from DAO
S_{VPD}	=	Vapor pressure deficit scalar (0–1), VPD is from DAO
NPP	=	net primary production (g C m^{-2} day^{-1})
R_a	=	autotrophic respiration (g C m^{-2} day^{-1})
R_m	=	maintenance respiration (g C m^{-2} day^{-1}), a function of biomass (derived from LAI) and temperature, summed across biomass compartments
R_g	=	growth respiration (g C m^{-2} day^{-1}), a function of biomass growth, summed across biomass compartments

Fig. 1 Components of the Moderate Resolution Imaging Spectroradiometer (MODIS) net primary production (NPP)/gross primary production (GPP) algorithm (Running *et al.*, 2000).

forested sites. In this paper, the MODIS/BigFoot comparisons are extended to a much broader range of biome types, and the partitioning of MODIS GPP into NPP and autotrophic respiration is examined.

Methods

Overview

The six sites described in this study are part of a network of sites representing the major biomes. As part of the BigFoot Project, a standard protocol of ground measurements and scaling is used at each site to evaluate the land cover, LAI, NPP, and GPP products from the MODIS Land Science Team (Justice *et al.*, 1998). The background and approach for the BigFoot Project are described in earlier publications (Campbell *et al.*, 1999; Reich *et al.*, 1999; Running *et al.*, 1999). Briefly, at each site a 5 km × 5 km study area is established which includes an eddy covariance flux tower. One hundred plots are laid out within the study area, all of which are measured seasonally for LAI and 50 of which are measured for above-ground NPP (ANPP). Land cover and LAI are then mapped at a 25 m resolution (resampled from 30 m) using these measurements and imagery from Landsat sensors (Cohen *et al.* 2003a, b). NPP and GPP are subsequently mapped using spatially distributed runs of the biogeochemical model Biome-BGC (Turner *et al.*, 2003a). Model inputs include daily meteorological data from the flux tower, the BigFoot land cover product, and the BigFoot LAI products. LAI is prescribed in the model runs to maximize use of information from the LAI measurements. Ground estimates of NPP are derived from the ANPP measurements plus a ratio of below- to above-ground production from a literature survey (Gower *et al.*, 1999). The NPP estimates are used in calibrating the Biome-BGC model. Model outputs of daily GPP are corroborated with flux-tower-based observations. BigFoot NPP/GPP products are then overlain with corresponding MODIS products for direct comparison (Cohen *et al.*, 2003b; Turner *et al.*, 2003a).

Site descriptions

Geographic coordinates and summary climate data for the sites are listed in Table 1. Vegetation descriptions, general locations, and acronyms are as follows. The agricultural fields site (AGRO) is located in the Midwest region of the United States (USA) and is composed of corn and soybean fields with small areas of urban development.

The temperate deciduous forest site (HARV) is predominantly closed hardwood forest with small areas of conifer forest, wetland, and urban development. HARV is within the Harvard Forest Long Term Ecological Research (LTER) site in Massachusetts, USA.

The temperate coniferous forest site (METL) is located on the eastern slope of the Cascade Mountains in Oregon, USA and is primarily open Ponderosa Pine (*Pinus ponderosa*) forests with areas of grassland and shrubland. The reference LAI and NPP data for this site were from an earlier study (Law *et al.*, 2001a, 2003). The boreal forest site (NOBS) is in northern Manitoba, Canada and is predominantly upland and open black spruce (*Picea mariana*) forest with areas of wetland and deciduous forests. The arctic tundra site (TUND) is low-stature coastal tundra vegetation with large areas of wetland and open waters. It is located near Barrow on the Arctic Coastal Plain in Northern Alaska, USA.

Table 1 Site location and long-term average climate variables

Code	Vegetation	Location	Precipitation (cm)	MAT (°C)
AGRO	Corn/soybean	Lat: 40.006658 Lon: −88.291535	99	11.23
HARV	Hardwood Forest	Lat: 42.528513 Lon: −72.172907	111	8.31
METL	Conifer Forest	Lat: 44.450722 Lon: −121.572812	44	7.75
NOBS	Boreal Forest	Lat: 55.885260 Lon: −98.477268	52	−3.20
TUND	Arctic Tundra	Lat: 71.271908 Lon: −156.613307	5	−10.91
SEVI	Desert Grassland	Lat: 34.350858 Lon: −106.689897	35	13.57

MAT, mean annual temperature; lat, latitude; lon, longitude.

Fig. 2 IKONOS image ('true color') of SEVI site showing the plot locations and flux tower location. The 1 km grid is indicative of the resolution of the Moderate Resolution Imaging Spectroradiometer gross primary production products.

The desert grassland site (SEVI) is predominantly perennial bunchgrasses, dominated with Black Grama (*Bouteloua eriopoda*) and Blue Grama (*Bouteloua gracilis*). Occasional cacti and shrubs (*Larrea tridentata*) are also present. Cattle grazing is not permitted. SEVI is in the Sevilleta LTER site in Central New Mexico, USA.

Sampling scheme for the field measurements

The 100 sample plots (25 m × 25 m) were separated into two classes. The first set was arrayed over a 1 km area surrounding the flux tower in a cyclic sampling design (Burrows *et al.*, 2002). The arrangement of plots in the cyclic sampling design is such that pairs of plots are separated by a range of distances. This design facilitated geostatistical analyses within the flux tower footprint. The second set of plots was distributed over the remaining area, either in a stratified random design (AGRO, HARV, NOBS) or based on the criterion of sampling the range of spectral variation found in the imagery from the Landsat Enhanced Thematic Mapper Plus (ETM +) sensor (e.g. Fig. 2).

LAI and ANPP measurements

The site-specific approaches to measuring LAI and ANPP are described in Campbell *et al.* (1999), Gower *et al.* (1997, 1999), Burrows *et al.* (2002), and Law *et al.* (2003). At TUND and SEVI, biomass was harvested and used to estimated leaf area after determining the specific leaf area. At AGRO, it was by reference to leaf area per plant and plant density. An allometric approach was employed at NOBS, and a light transmittance approach (LAI-2000) was used at METL and HARV. Measurements at METL were corrected for clumping and for wood interception (Law *et al.*, 2001b).

ANPP was based on the clipped green material (maximum green biomass) at AGRO, TUND and SEVI. At the three forested sites, stemwood production was from allometry based on either tree coring (NOBS, METL) or dendrometer bands (HARV). Foliage production at the forested sites was from litter baskets. Bryophyte production at NOBS was estimated with a crankwire approach (Bisbee *et al.*, 2001) and at TUND was estimated from the ratio of bryophyte to total ANPP from an earlier study (Miller *et al.*, 1980).

Mapping land cover and LAI

The approaches for mapping land cover and maximum LAI are described in Cohen *et al.*, (2003a, b). Both were mapped using imagery from the Landsat ETM + sensor. The original land cover classes were those from IGBP (Morisette *et al.*, 2002) and these were aggregated in some cases for the purposes of the NPP/GPP modeling. The general approach for LAI mapping was to establish empirical, cover-type-specific relationships of measured LAI to the Tasseled Cap (Crist & Cicone, 1984) spectral vegetation indices. The imagery was then used to map LAI.

The maximum LAI provided a mid-growing season reference point on the seasonal LAI trajectory (i.e. phenology). To construct the full seasonal trajectory of LAI for each BigFoot grid cell, a template was formulated for each cover type based on *in situ* observations. At HARV, the trajectory for the hardwood forest cover class was based on measurements of above- and below-canopy PAR (Wythers *et al.*, 2003). At all sites with conifer cover, the conifer LAI was assumed constant all year. Conifer LAI does vary seasonally to some degree (e.g. Gholz *et al.*, 1991) but these temporal patterns could not be consistently resolved in this analysis so were not treated here. At AGRO, weekly measurements of LAI were made at the flux tower. At TUND, a sprectroradiometer was used to monitor vegetation greenness from the beginning of spring snowmelt to the return of snow cover in the fall (Stow *et al.*, 2004). At SEVI, LAI measurements were made at four points during the growing season and the beginning of the growing season was based on the beginning of the summer monsoon rains, the mid-season maximum on the field LAI measurements, and

the end on the decreasing PAR and temperatures in late November. These trajectories were used as the templates, and at each 25 m cell, a unique seasonal LAI trajectory was created. This was accomplished by determining the ratio of the template LAI to the observed maximum LAI (from the remote sensing analysis) at mid-growing season, and applying that ratio each day to the relevant template LAI to get the full seasonal LAI trajectory for the grid cell.

Scaling GPP and NPP over the 25 km^2 study area

The BigFoot GPP/NPP scaling approach relies on running the Biome-BGC model (Running 1994; Kimball *et al.*, 2000) in a spatially distributed mode over a 5 km × 5 km grid of 25 m × 25 m cells (Turner *et al.*, 2003a). Biome-BGC is a daily time step ecosystem process model; thus, it simulates photosynthesis and autotrophic respiration as well as evapotranspiration. The algorithm for net photosynthesis is based on the Farquhar biochemical model, and GPP is calculated as the sum of net photosynthesis and daytime foliar respiration. The autotrophic respiration algorithm uses biomass nitrogen content and temperature, with a fixed base rate across all biomass compartments and biomes (as in Thornton *et al.*, 2002). Biomass compartments are leaves and roots for the grassland cover type and leaves, live stem, coarse roots, and fine roots for the other cover types. LAI is an important determinant of mass flux in Biome-BGC and as noted was prescribed spatially and temporally in these model runs.

Parameterization of the ecophysiological and allometric variables for each cover type was based on the literature review of White *et al.* (2000). Biomass carbon pools were determined allometrically by reference to the LAI. Leaf carbon was derived from LAI by way of the specific leaf area parameter. Soil depth was set at a uniform value, generally 0.5 m, except at METL where spatially varying soil depth was taken from an earlier study (Law *et al.*, 2004a). Daily time step meteorological data (↓PAR, precipitation, minimum temperature, maximum temperature, and VPD) were from observations at the flux towers. In the case of METL and HARV, where there was significant topographic variation, the meteorological data were interpolated to each 25 m cell using the MTCLM model Version 4.3 (Running *et al.*, 1987).

Two ecophysiological parameters in the Biome-BGC model were calibrated for each major cover type – the leaf carbon to nitrogen ratio (leaf C : N) and the fraction of leaf nitrogen as rubisco (FLNR). The purpose of the calibration was to remove bias in the model outputs relative to the NPP measurements. These variables were used because NPP is the net effect of photosynthesis and autotrophic respiration; FLNR strongly influences modeled photosynthesis whereas leaf C : N strongly influences modeled autotrophic respiration. For the calibration, the model was first run with default leaf C : N and FLNR values at the measurement plot locations (using BigFoot prescribed LAIs), and the root mean square error (RMSE) of simulated NPP vs. measured NPP was determined. The same procedure was then repeated with each combination of leaf C : N and FLNR over a prescribed range of values, with increments of 0.01 for FLNR and 1 for C : N. Ranges of potential leaf C : N and FLNR for the different cover classes were determined from the literature (White *et al.*, 2000). The combination with the lowest RMSE was selected for use in the spatial mode run. The calibration year was 2000 for AGRO, 2001 for HARV, NOBS, and METL, and 2002 for SEVI and TUND.

For the purposes of corroborating the model GPP estimates with observations at the flux tower, the model output for all 25 m cells within a 0.5 km radius of the flux tower were spatially averaged for each day and further temporally averaged over 8-day bins (the temporal resolution of the MODIS GPP product). This 0.7 km^2 area represented a crude approximation of the flux tower footprint. At the AGRO site, where the tower was situated in a soybean field, only soybean-occupied cells in the 0.5 km radius area were included in the averaging. The prevailing winds there are from the south to south-west where there is a 500 m fetch over soybean (Meyers & Hollinger, 2004). The comparisons of tower-based and modeled GPP were evaluated in terms of the phenological patterns, the mid-season maxima, and the intraseasonal variation.

Background information on the eddy covariance flux towers is listed in Table 2, with additional information available at the AmeriFlux (2004) and FLUXNET (2004) Internet sites. During the evaluation year, a continuous record of half-hourly net ecosystem exchange (NEE)

Table 2 Background information on the eddy covariance flux towers

Site	Investigator (institution)	Related publication
AGRO	T. Meyers (Oak Ridge National Laboratory)	Meyers & Hollinger (2004)
HARV	S. Wofsy (Harvard University)	Wofsy *et al.* (1993)
METL	B. Law (Oregon State University)	Anthoni *et al.* (2002)
NOBS	S. Wofsy (Harvard University)	Goulden *et al.* (1997)
TUND	W. Oechel (San Diego State University)	
SEVI	E. Small (University of Colorado)	Kurc & Small (2004)

was developed during the growing season from the tower measurements. Data gaps associated with equipment failures or unsuitable micrometeorological conditions were filled using standard techniques (Falge *et al.*, 2001). Half-hourly GPP was calculated during daylight hours as NEE minus ecosystem respiration (R_e), where the sign of R_e is negative (Goulden *et al.*, 1996a; Goulden *et al.*, 1996b; Turner *et al.*, 2003b). Half-hourly R_e was based on relationships of air temperature to night-time NEE during times of adequate turbulence, or NEE from chamber measurements under dark conditions at TUND. The continuous record of air temperature from the tower then provided a means to estimate half-hourly R_e.

The MODIS NPP/GPP products

The MODIS NPP/GPP data (Version 4.5) were produced at the University of Montana. The MODIS NPP/GPP algorithm (MOD17) is described in Running *et al.* (2000) and in the MOD17 User's Guide (Heinsch *et al.*, 2003). In production of the Version 4.5 NPP/GPP estimates, gaps in the MODIS Collection 4 LAI and FPAR products (inputs to MOD17) were filled by linear interpolation. In that procedure, the algorithm determines dates with a low data quality flag and looks forward and backward to find the closest dates with acceptable data-quality flags to do an interpolation. In addition, Data Assimilation Office (DAO) climate data at the original coarse resolution were interpolated to the 1 km resolution.

Before direct comparisons could be made, the BigFoot products in the Universal Transverse Mercator (UTM) coordinate system were reprojected to the native Sinusoidal Projection of the MODIS products (Cohen *et al.*, 2003b). To achieve spatial correspondence between the MODIS products and BigFoot products, the BigFoot NPP/GPP data were averaged over each 1 km^2 MODIS cell. These averages included zeros for nonvegetated grid cells. To achieve temporal correspondence, the BigFoot GPP data were averaged over the 8-day bin periods associated with the MODIS GPP products. The year of comparison was 2000 for AGRO and 2002 for the other sites.

As a follow-up to the direct comparison of MODIS and ground-based NPP/GPP, specific components of the MODIS GPP algorithm were examined. Interpolated meteorological data from DAO included $\downarrow$ PAR, daily minimum temperature, and VPD and these data were compared with meteorological observations from the flux tower. FPAR values used in generating the MODIS Version 4.5 NPP/GPP were compared with FPAR values derived from the BigFoot prescribed LAIs. The conversion of the prescribed LAIs to FPAR used a simple Beer's Law approach (Jarvis & Leverez, 1983):

$$\mathrm{FPAR} = 1 - (e^{(\mathrm{LAI}(-K))}), \quad (1)$$

where K is the canopy light extinction coefficient (also an ecophysiological parameter in Biome-BGC). K values were assumed to be 0.58 for broadleaf forests and 0.50 for all other vegetation types. More sophisticated transformation algorithms are available (Chen *et al.*, 1997; Gower *et al.*, 1999) but they could not be consistently parameterized across the wide range of vegetation types and conditions in this study. The ground-based FPAR values were averaged to get 8-day mean values over each 1 km^2 that could be compared directly to the MODIS values.

The daily LUE (ε_g) values from MODIS and BigFoot were also compared. ε_g is a key variable in the MODIS GPP algorithm and is calculated as is the quotient of GPP (in gC) and the PAR absorbed by the canopy (APAR in MJ). The MODIS daily ε_g was generated by running the MODIS NPP/GPP algorithm (Fig. 1) at the tower cell with standard inputs from the MODIS data stream. For the BigFoot values, daily ε_g was calculated as modeled GPP divided by modeled APAR. Results were spatially averaged for all 25 m cells in the 1 km^2 MODIS cell that included the flux tower.

Results

BigFoot products

There were between two and seven vegetation cover types at each site (Fig. 3). SEVI, TUND and METL were the simplest sites having greater than 90% of the vegetated cells in one class. AGRO had corn or soybeans in >90% of its area. Two of the forested sites were more heterogeneous, each having a mixture of one or more forest types along with other classes such as open shrubland or wetland. Classification accuracy was generally greater than 80% in the original BigFoot land cover products (Cohen *et al.*, 2003b) with somewhat higher accuracy for the more aggregated classes used in this study.

LAI at the six sites ranged from 0 to 8 m^2 m^{-2}, with highest values in the corn fields and deciduous forest (Fig. 4). Intermediate LAIs were found at METL where LAI is limited by site water balance (Law *et al.*, 2001b) and NOBS where LAI is limited by poor drainage and a short growing season. The TUND and SEVI sites had relatively low LAIs, with average mid-growing season values of less than 1.0 in each case. LAI mapping accuracy was checked with cross-validation using the LAI measurements plots, and results showed good agreement between observed and predicted LAIs (Cohen *et al.*, 2002, 2003a).

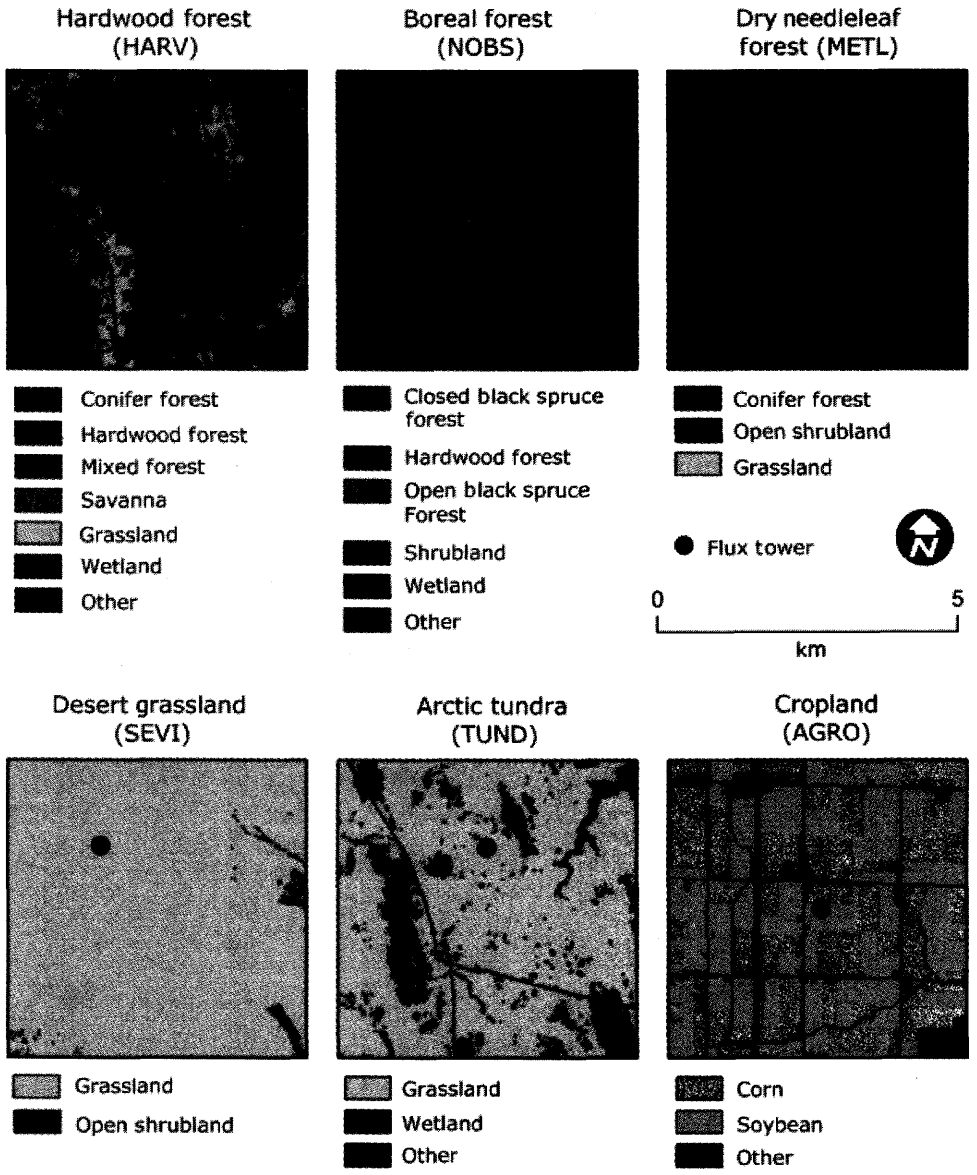

Fig. 3 Land cover at the sites based on Landsat Enhanced Thematic Mapper Plus.

For NPP, the calibration procedure brought the mean for simulated NPP at the field plots close to the mean for the measured NPPs at those plots (Table 3). However, the ratio of RMSE to NPP varied widely across sites. Lowest ratios were at SEVI and AGRO where measured LAI and NPP were well correlated and the error in mapping LAI with remote sensing was relatively small (Cohen *et al.*, 2003a). At TUND and METL, the ratios of RMSE to mean NPP were relatively high in part because those ratios were also relatively high for the mapped LAIs.

The BigFoot GPP (8-day mean) ranged from a maximum of about 12–13 $g C m^{-2} day^{-1}$ in the mid-growing season at HARV and AGRO, to intermediate values of 5–8 $g C m^{-2} day^{-1}$ at NOBS and METL, and low values of 2–3 $g C m^{-2} day^{-1}$ at SEVI and TUND (Fig. 5). As in Turner *et al.* (2003a), where similar comparisons were made for 2001, agreement with flux tower estimates of GPP was generally good at the HARV and NOBS site, showing close agreement at the beginning and end of the growing season. The mid-growing season maximum at NOBS was relatively high in the BigFoot product for 2002. At the AGRO site, BigFoot GPP tended to overestimate GPP late in the growing season, probably because of an observed decline in LUE that was not part of the model's NPP algorithm (Turner *et al.*, 2003b). The BigFoot GPP declined somewhat later than the tower observations at METL, a pattern most likely related to site water balance factors. Agreement of BigFoot and tower GPP

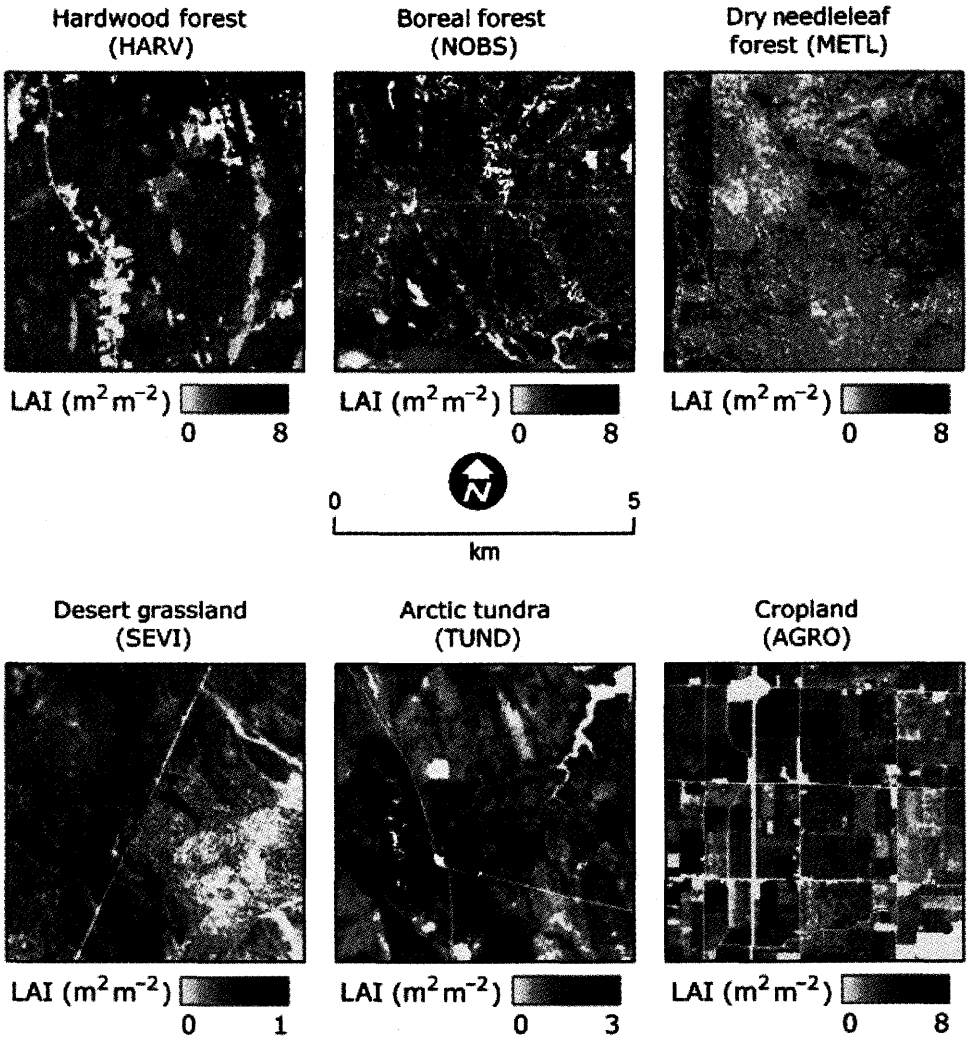

Fig. 4 Leaf area index at the sites based on Landsat Enhanced Thematic Mapper Plus.

was reasonable at TUND and SEVI, showing that the model captured the seasonal trajectories at these two climate extremes.

BigFoot/MODIS comparisons

With the exception of the SEVI site, the spatial heterogeneity of the MODIS GPP products was generally less than that found in the BigFoot GPP products (Fig. 6). Note that comparisons at TUND represent only the seven MODIS cells that were classified as vegetated in the MODIS land cover product. The mid-growing season maximum values were similar at METL and TUND, higher in the MODIS products at SEVI and NOBS, and lower in the MODIS products at AGRO and HARV. The agreement with regard to the beginning and end of the growing season was very good at NOBS and METL, fair at HARV and AGRO, and poor at SEVI and TUND. Total GPP matched well at HARV (though a function of counteracting errors) and NOBS, moderately well at METL and TUND, and most poorly at SEVI and AGRO (Fig. 7).

For NPP, there was closest agreement at TUND (Fig. 8). The MODIS product overestimated NPP at NOBS, METL, and SEVI and underestimated it at AGRO and HARV. There was generally little relationship in the pattern of spatial variation within an individual site, and the distributions of points in the one-to-one plots (i.e. for the twenty-five 1 km^2 cells) varied widely.

The cross-site comparison of the DAO meteorological data vs. the flux tower data revealed closest agreement for minimum temperature (data not shown), a somewhat low bias in VPD at high VPDs at the SEVI site, and site-specific variation in the bias for ↓ PAR (Fig. 9). ↓ PAR from DAO was notably low at TUND. The degree of agreement in 2002 was similar to that in 2001 at HARV and NOBS (Turner *et al.*, 2003a).

Table 3 Results of leaf carbon to nitrogen ratio (C : N) and fraction of leaf N in Rubisco (FLNR) calibration based on net primary production (NPP)

Site	C : N (ratio)	FLNR (%)	Mean NPP observations ($g C m^{-2} yr^{-1}$)	Mean NPP simulations ($g C m^{-2} yr^{-1}$)	RMSE ($g C m^{-2} yr^{-1}$)
AGRO					
Corn	10	0.18	839	833	105
Soybean	11	0.14	379	368	91
HARV					
Deciduous Forest	24	0.14	679	667	129
Conifer Forest	37	0.08	552	544	86
METL					
Conifer Forest	40	0.05	356	352	104
NOBS					
Upland Black Spruce	60	0.07	251	245	66
Open Black Spruce	50	0.05	181	183	37
TUND					
Grassland	25	0.15	64	64	34
SEVI					
Grassland	24	0.08	54	54	8

RMSE, root mean square error.

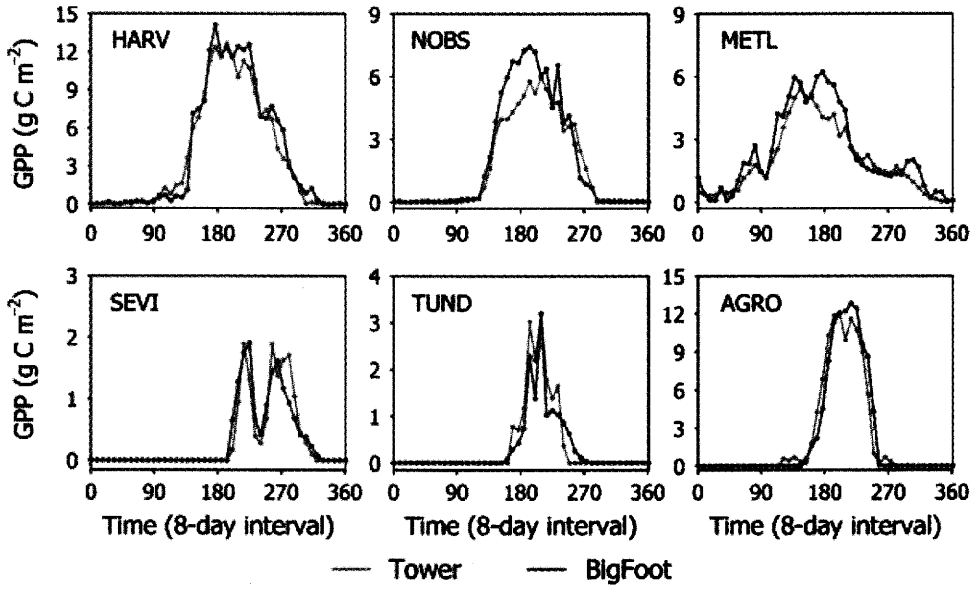

Fig. 5 Times series comparison (8-day means) of BigFoot and flux tower gross primary production (GPP). The BigFoot values are based on the mean of all Enhanced Thematic Mapper Plus resolution cells (25 m) in the 0.7 km^2 surrounding the flux tower.

Another factor strongly influencing the MODIS NPP/GPP products was the LAI/FPAR values. At HARV, NOBS, and AGRO, the MODIS FPAR was close to 1 at mid-season, in agreement with the BigFoot products (Fig. 10). At the other forest site, METL, MODIS FPAR tended to be overestimated in mid-season. MODIS FPAR was overestimated year round at SEVI and throughout the growing season at TUND.

The timing of the seasonal variation in FPAR showed close agreement at SEVI and AGRO. At all other sites the spring rise in the MODIS FPAR tended to precede the rise seen in the BigFoot products. The fall decline was well captured by the MODIS product at SEVI and AGRO but tended to be delayed at TUND. At METL, there was not an end of season decline in FPAR for either the MODIS or BigFoot products.

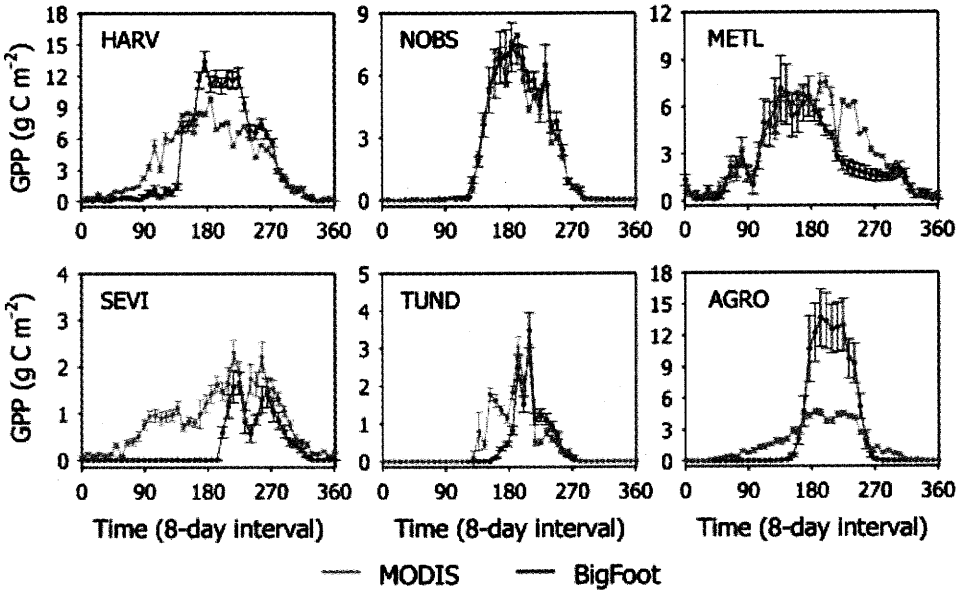

Fig. 6 Times series comparison (8-day means) of BigFoot and Moderate Resolution Imaging Spectroradiometer (MODIS) gross primary production. Values are means for twenty-five 1 km^2 cells. The bars are ± the standard deviation of the twenty-five 1 km^2 MODIS cells.

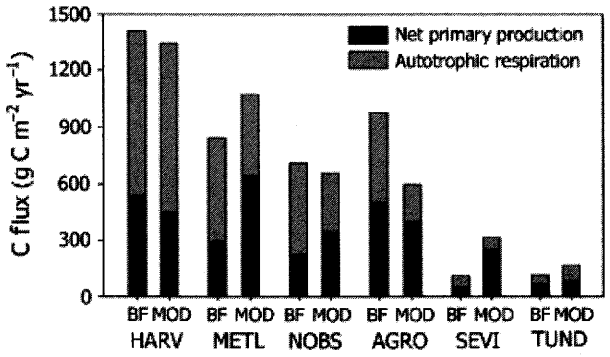

Fig. 7 Comparison of net primary production (NPP) and autotrophic respiration (R_a) from BigFoot and Moderate Resolution Imaging Spectroradiometer. Gross primary production is the sum of NPP and R_a.

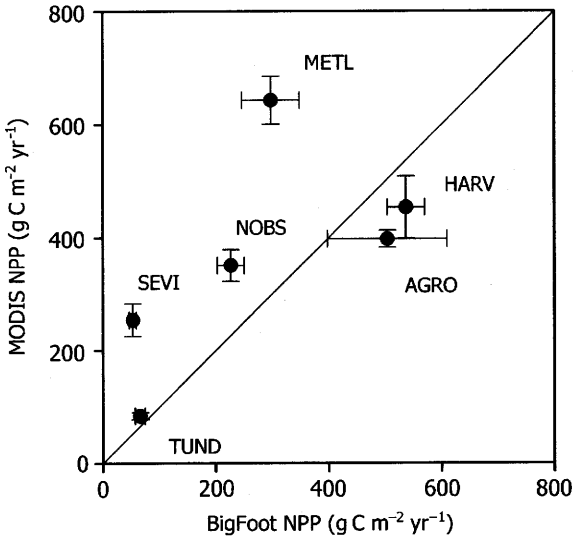

Fig. 8 Comparison of BigFoot and Moderate Resolution Imaging Spectroradiometer (MODIS) net primary production (NPP) estimates. The solid line is the one-to-one relationship. Values are means for twenty-five 1 km^2 cells. The bars are ± the standard deviation of the twenty-five 1 km^2 MODIS cells.

MODIS ε_g was very much lower than BigFoot ε_g at AGRO and TUND, and moderately lower at HARV, NOBS, and in the early growing season at SEVI. Agreement was good for the most part at METL (Fig. 11).

Discussion

NPP measurements

The measured NPPs in this analysis were used to calibrate the NPP process model. Consequently, the scaled NPPs were tightly constrained by the NPP measurements and retained the uncertainty present in those measurements. The measurement of NPP in any ecosystem is a formidable task (Gower *et al.*, 1999, 2001; Clark *et al.*, 2001) and it is likely that most NPP estimates are underestimates relative to the actual difference between GPP and autotrophic respiration (Clark *et al.*, 2001). Besides the uncertainty about turnover of foliage and fine roots, there are uncertainties with regard to herbivory and allocation to mycorrhizal fungi, rhizodeposition, and nonmethane hydrocarbons. Thus, current estimates of NPP must

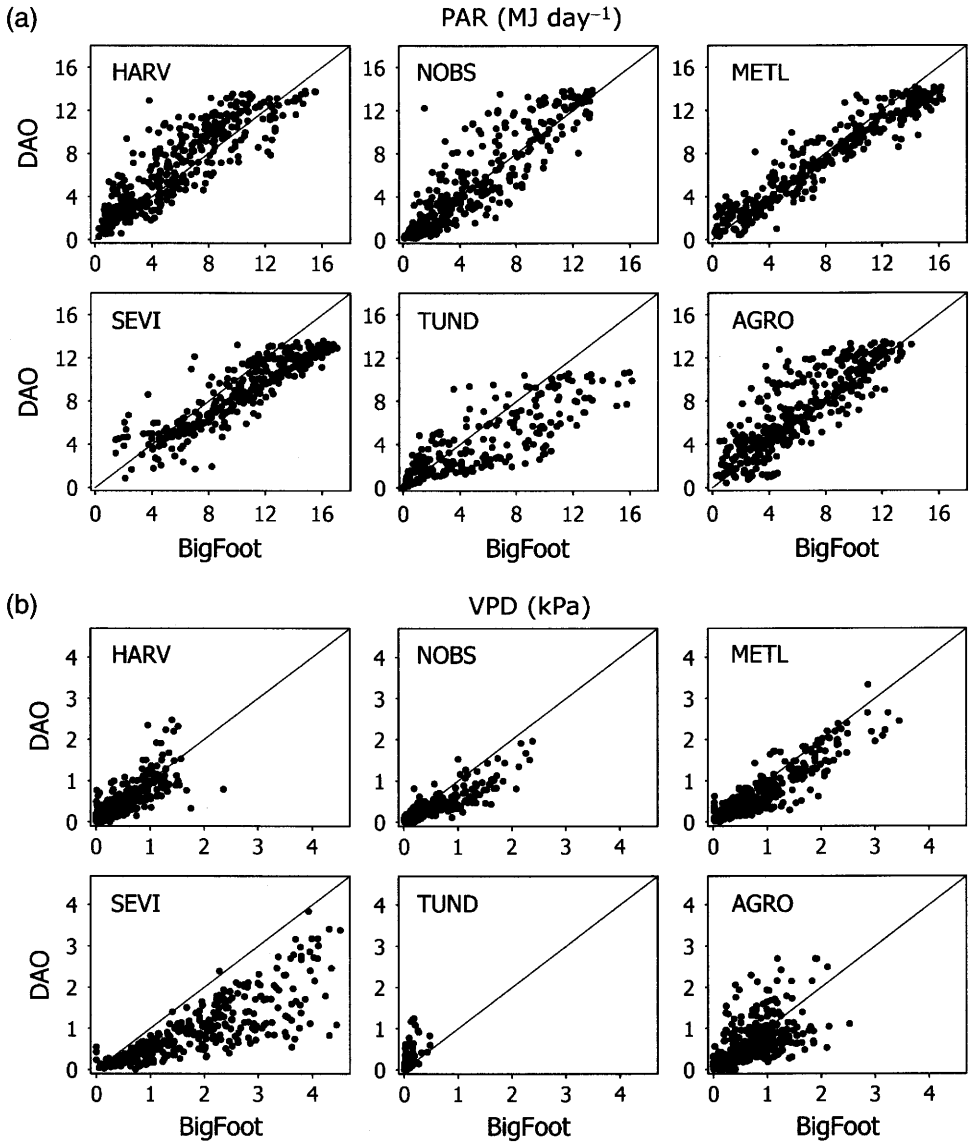

Fig. 9 Comparison of flux tower and interpolated Data Assimilation Office meteorological data (a = PAR, b = VPD).

be considered approximations that are gradually improving as new techniques and approaches are included in the measurement schemes.

That said, the sites in this study have been the subject of intensive investigations with regard to NPP and are relatively well characterized. The NOBS site was one of the core sites in the NASA-sponsored Boreal Ecosystem-Atmosphere Study (BOREAS) Project that was designed to improve understanding of boreal forest structure and function in the context of learning to scale up from plot-level measurements to regions (Sellers *et al.*, 1997). The allometric relationships used in this analysis were specifically developed at NOBS (Gower *et al.*, 1997), which reduces one potentially large source of uncertainty in estimation of forest NPP. Intensive studies of below-ground (Steele *et al.*, 1997) and bryophyte (Gower *et al.*, 1997) production have also been made at the NOBS site and the results of those studies informed our estimates of NPP. The HARV site is in the US LTER network and surrounds the longest continuously running eddy covariance flux tower in the world (Wofsy *et al.*, 1993; Goulden *et al.*, 1996b; Barford *et al.*, 2001). It has also been the subject of intensive studies of bolewood production, litter production, and below-ground production (Curtis *et al.*, 2002). The METL site has supported flux tower work since 1996 (Law *et al.*, 1999, 2000, 2003) and the surrounding vicinity has been the subject of intensive investigation of NPP components, including uncertainty analysis (Law *et al.*, 2001b, 2004a).

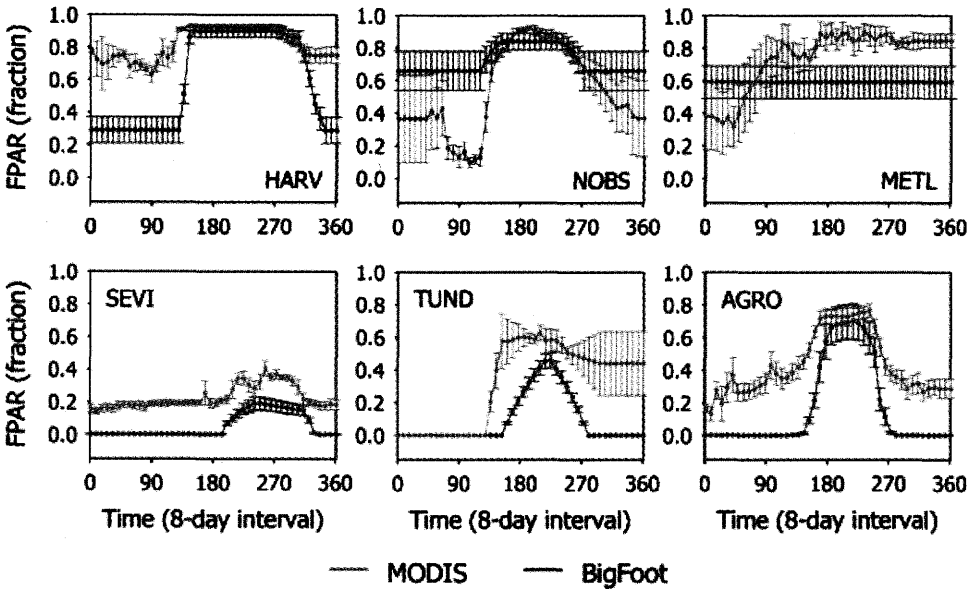

Fig. 10 Time series comparison of BigFoot and Moderate Resolution Imaging Spectroradiometer (MODIS) fraction of photosynthetically active radiation (FPAR). Values are means for twenty-five 1 km^2 cells. The bars are ± the standard deviation of the twenty-five 1 km^2 MODIS cells.

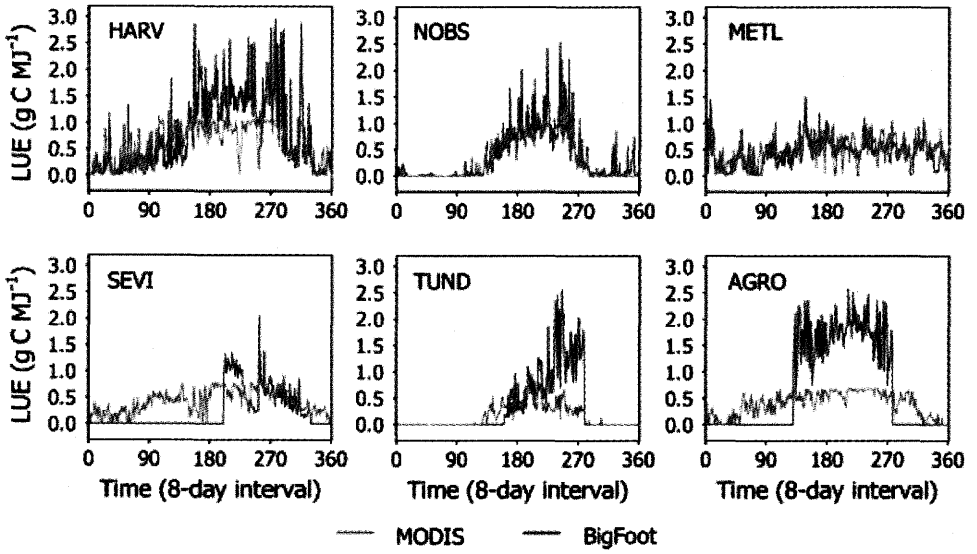

Fig. 11 Time series (daily) of light use efficiency for gross primary production (ε_g) for BigFoot and Moderate Resolution Imaging Spectroradiometer (MODIS) products at the MODIS cell occupied by the flux tower.

Measurement of ANPP in nonforest biomes is simpler in principle because the total biomass can be more readily evaluated at any given time. The measurements of ANPP at the corn and soybean fields (AGRO site) were probably the most accurate of all our sites because of the fine-scale homogeneity, the low likelihood of biomass loss from herbivores, and the sharp cutoff of productivity at the time of harvest. The mean of 500 g C m^{-2} yr^{-1} was towards the low end of the range of crop NPPs in the US Mid-west from Prince *et al.* (2001), which is partly a function of our average including nonvegetated areas like roads and urban development.

At both SEVI and TUND, the green vascular plant biomass at time of peak biomass undoubtedly underestimated foliage production because of herbivory, possible turnover of green matter before the sampling date, and possible turnover after the sampling data (Sala & Austin, 2000). Our NPP

estimates are in the low end of the ranges from earlier studies at or near those sites (Miller *et al.*, 1980; Whitford, 2002).

GPP measurements

Flux tower-based daily GPP estimates are now made at a large number of sites. However, there remain significant uncertainties, particularly in relation to estimation of daytime ecosystem respiration (Goulden *et al.*, 1996a; Turner *et al.*, 2003b). Chamber-based R_e estimates, as were made at TUND, are perhaps most reliable because periods of low turbulence are not an issue. At the other sites, R_e was based on daytime temperature and empirical relationships of NEE to air temperature at night during periods of adequate turbulence (Goulden *et al.*, 1996a). Comparison of our GPPs with estimates from other studies using micrometeorological approaches at similar sites showed good agreement in terms of maximum GPP for soybean (Rochette *et al.*, 1995), black spruce (Cienciala *et al.*, 1998), and dry coniferous forest (Goldstein *et al.*, 2000). The mixed deciduous forest at HARV had generally higher annual GPPs than at two beech forests in Europe (Valentini *et al.*, 1996; Granier *et al.*, 2000). At the TUND site, the annual GPP pattern compared well with purely chamber-based estimates (Oechel *et al.*, 1995). The GPP estimates here are the first published that we know of for a desert grassland site.

Tower GPP estimates are beginning to be used for model validation (Aber *et al.*, 1996; Cienciala *et al.*, 1998) but usually without dealing with issues of footprint movement or spatial heterogeneity in cover type and LAI within the footprint. Several flux tower sites have reported clear differences in the pattern of NEE depending on wind direction (Wofsy *et al.*, 1993; Aubinet *et al.*, 2002) and inspection of Figs 3 and 4 shows a significant degree of variation in land cover and LAI depending on compass heading for all the forested sites. This issue is less important for the sites with low stature vegetation (TUND, AGRO, SEVI), where the sensor height is low and footprint size relatively small. The assumption of a daytime average footprint of 0.7 km^2 centered on the tower, as used in this study, helped capture some of the spatial heterogeneity in the footprint, but could potentially be improved for model validation purposes by tracking of the footprint and evaluating spatial patterns in relation to variation in cover type (Schmid, 2002).

NPP/GPP scaling to 5 km × 5 km

The BigFoot protocol for scaling NPP to a 5 km × 5 km area at each site was designed to maximize use of available site-specific information and to produce flux estimates that could be aggregated to precisely match the spatial and temporal resolution of the MODIS NPP and GPP products (Reich *et al.*, 1999; Turner *et al.*, 2003a). Having an ecosystem process model at the center of the scaling protocol has the benefit of forcing consistency among the multitude of spatially and temporally varying observations and allowing for aggregation of results to relevant spatial and temporal scales (Turner *et al.*, 2004b).

The Landsat remote sensing analysis brought a great deal of information to bear on the analysis. The mapping of land cover produced a means to specify ecophysiological constants in the ecosystem process model. Ecophysiological differences between deciduous and coniferous species are well documented, so classification of forest type – as at the HARV site – was particularly useful. There, the spatial resolution of the BigFoot land cover data layer (25 m) was fine enough to capture the subkilometer scale patches of conifers. That 25 m resolution was also adequate for revealing the strong subkilometer scale impacts of management (e.g. clearcuts) at the METL site, and the significant areas of open water at TUND. The 25 m resolution was not fine enough to capture the heterogeneity associated the ice wedge polygons at TUND (Stow *et al.*, 2004). The scale of the dry hills and wet swales is on the order of several meters and the patches have very different vegetation and NPP. However, for this study that variation was lumped together in our 25 m cells. The five subsamples (1 m^2) that were clipped for each plot probably undersampled with respect to fine-scale heterogeneity with any 25 m × 25 m cell. On-going NPP scaling studies at TUND are being made with imagery from the IKONOS sensor (1 m resolution) which should more adequately address the spatial heterogeneity issues there.

Prescribing remote-sensing-based LAI spatially and temporally was also a significant source of local information. Across a broad range of ecosystems and within any given ecosystem, seasonal maximum LAI tends to be correlated with ANPP (Shaver *et al.*, 1996; Gower *et al.*, 2001; Asner *et al.*, 2003); thus, information on spatial variation in LAI is helpful in scaling NPP. Information on the spatial variation in LAI was particularly useful at SEVI where a large area on the west side of the study area that was still classified as grassland had been burned in a previous year and had not recovered full LAI and NPP (Figs 2 and 4). LAI within a cover class was often not normally distributed at our sites (data not shown), which also suggests the benefits of using LAI information on spatial variation in LAI within a cover type rather than simply prescribing a mean value by cover type.

Interannual variation in the seasonality of LAI is often correlated with interannual variation in GPP (e.g. Goulden *et al.*, 1996b), so the LAI temporal trajectory was an important input to the NPP scaling effort. At HARV, our temporal trajectory of LAI indicated by simple observations of above- and below-canopy PAR agreed well with independent observations of leaf expansion (Wythers *et al.*, 2003). Nevertheless, it was a gross simplification of the true canopy phenology as the canopy is very diverse and the species differ somewhat in leaf out phenology (Lechowicz, 1984). The prescribed LAI also missed the flush of vernal herbs at HARV because the below-canopy PAR sensor was well off the soil surface. Because change in growing season length is one of the expected consequences of projected global climate change and has putatively been detected by satellite remote sensing (Myneni *et al.*, 1997a), it will be important to carefully assess the ability of MODIS imagery to capture these changes (Zhang *et al.*, 2003). Given the subtle nature of interannual variation in leaf phenology, an adequate validation data set will require long-term ground-based monitoring with a suitable sampling scheme to capture spatial and temporal variation.

Because LAI is prescribed in the Biome-BGC model runs, the BigFoot NPP scaling protocol generally assumes there is a strong relationship between maximum LAI and NPP. This proved to be more so the case for the nonforest biomes than for the forest biomes. That result is partly driven by the fact that the LAI and NPP measurements are usually tightly coupled in the nonforest case (e.g. Shaver *et al.*, 1996). At TUND and SEVI, green biomass was used to estimate ANPP and was also converted to an LAI estimate with a specific leaf area parameter. In forests, the LAI estimates were from allometry or light interception and thus largely independent of the ANPP measurements. That independence introduced measurement error into the relationship, which was carried through into the calibration. Other factors potentially limiting the strength of the LAI/NPP relationship are soil drainage patterns, foliar nutrient concentrations (or photosynthetic capacity), and stand age. It is difficult to capture soil drainage effects in a distributed process model because of the complexities of mapping soil water holding capacity, in modeling subsurface water flow patterns, and modeling physiological response to saturated soils. There has been continued progress with mapping foliar nitrogen using high spectral resolution remote sensing in recent years (Smith *et al.*, 2002), and if the approach proves operational that information would improve model performance in some areas. NPP has been shown to decrease in late succession for many forest types, in some cases without significant reduction in LAI (Gower *et al.*, 1996). Landsat-scale remote sensing can be used to age forest stands based on change detection for young stands (Cohen *et al.*, 2002) or classification for older stands (Cohen *et al.*, 1995). Stand age data could thus be prescribed in spatial mode applications and used in model parameterization to help capture age effects on NPP (Law *et al.*, 2004b).

The meteorological observations at the flux tower provided specificity in the temporal dimension of the BigFoot NPP/GPP scaling approach. Uncertainties associated with the measurements of temperature, precipitation, $\downarrow$ PAR, and vapor pressure were relatively low and the data provided a strong signal when used as input to the process model. The daily meteorological data allow the model to simulate day-to-day changes in GPP and R_a and permitted estimation of daily LUE, a critical variable in the MODIS NPP/GPP algorithm. Interannual variation in meteorology at some of these sites can be large and the BigFoot scaling approach will ultimately permit assessment of the effectiveness of the MODIS NPP/GPP product in capturing interannual variation in NPP and GPP.

In comparing the BigFoot GPPs with flux tower GPPs, the only sites with obvious differences were METL and NOBS. In the METL simulation, soil water was largely depleted by early July and stomatal conductance, hence GPP, began to be strongly constrained. The GPP estimates from the tower, and local measurements of conductance at the leaf level (Irvine *et al.*, 2004), also indicate mid- to late growing season water stress, but coming on somewhat slower. Simulated maximum transpiration rate was higher than observations (based on a sap flow technique, Irvine *et al.*, 2004) so additional attention to parameterization of stomatal conductance may be needed at that site. Characterizing soil water availability is also problematic at METL as some areas are accessing water deeper than 0.8 m (Irvine *et al.*, 2002). At NOBS, simulated GPP was high relative to tower GPP during the month of June. This difference did not occur in 2001 (Turner *et al.*, 2003a) and may be related to relatively cool temperatures early in the growing season. Minimum temperatures in May averaged over 5 °C cooler in 2002 compared with 2001, which may have induced physiological responses that were not accounted for in the model.

NPP/GPP scaling to the global domain at 1 km resolution

As the spatial domain of interest expands from the landscape scale of the BigFoot products to the global scale of the MODIS products, there must inevitably be compromises in the scaling approach. The MODIS

NPP/GPP algorithm uses a simple LUE approach to estimating GPP rather than a full-process model as in the BigFoot approach. One consequence of this simplification is a reduced ability to detect drought stress as the LUE model is not simulating the water balance. The 1 km spatial resolution of the MODIS products is a compromise between the desire for fine spatial resolution to capture effects of climatic gradients, as well as, land use (Justice *et al.*, 1998), and the appeal of frequent coverage for detecting interannual variation in regional phenology. Other compromises relate to the quality of the meteorological inputs, the need for a relatively simple land cover classification scheme, and the need for a generalized radiation transfer algorithm for LAI and FPAR (Myneni *et al.*, 1997b) rather than a site-specific empirical approach to characterizing the vegetation (Cohen *et al.*, 2003a; Xiao *et al.*, 2004a). The key comparisons for the purposes of evaluating the consequences of these compromises are of the seasonal trajectory of the GPP, the total annual GPP, the annual NPP, and the spatial pattern in the NPP.

GPP seasonal trajectory. In the MODIS algorithm, the seasonal trajectory of GPP is highly dependent on the seasonal trajectory of ↓ PAR, FPAR, and the minimum temperature scalar (S_{Tmin}). The six sites in this study are all found at moderate-to-high latitude, so just the signal from daily ↓ PAR introduces a significant degree of seasonality into the MODIS GPP (see Figure 2 in Turner *et al.* (2003a) for ↓ PAR plot at HARV and NOBS). At all of these sites, GPP is reduced to near zero by S_{Tmin} during some part of the annual cycle (data not shown).

To initiate the growing season, the MODIS algorithm relies on some combination of S_{Tmin} increase and FPAR increase. At NOBS, both FPAR and S_{Tmin} are helping to initiate the beginning of the growing season, albeit that the FPAR increase may be responding to snow melt rather than LAI as much of the LAI is evergreen conifer. At TUND, SEVI, AGRO, and HARV there are problems in the MODIS GPP with anticipation of the beginning of the growing season. The MODIS FPAR appears to be too high in all cases early in the growing season (Fig. 10). At TUND, the MODIS FPAR begins an abrupt increase on day of year (DOY) 135 and reaches near its peak value for the growing season by DOY 160. This pattern closely matches an observed rise in Normalized Difference Vegetation Index (NDVI) (a spectral vegetation index) from a downward looking spectroradiometer on the ground (unpublished data). However, once the snow is gone the ground-based NDVI then continues to rise, presumably capturing the actual green up. MODIS FPAR at SEVI and AGRO was generally greater than 0.2 during the winter so that when temperatures began to rise, the simulated GPP became artificially high. At HARV, both BigFoot simulations and the flux tower GPP showed a later flushing of GPP than did the MODIS GPP trajectory. The simulations of Xiao *et al.* (2004b), where phenology was driven by a spectral vegetation index from the VEGETATION sensor, also showed a later GPP increase at HARV.

At the end of the growing season, the expected decline in MODIS FPAR is delayed at TUND, suggesting there may be artifacts associated with snow cover or cloud cover. Nevertheless, the GPP is shut down correctly by S_{Tmin}. At other nonconifer sites, the FPAR decline at the end of the growing season is helping decrease GPP in agreement with flux tower and BigFoot GPP. Neither the MODIS GPP algorithm nor the BigFoot scaling approach account for changes in photosynthetic capacity that have been observed towards the end of the growing season at some sites (Wilson *et al.*, 2001); thus, they would tend to overestimate GPP towards the end of the growing season. At the coniferous forest site (METL), FPAR is stable after the growing season but daily GPP decreases because of the decreasing ↓ PAR.

The mid-growing season dips in GPP in the BigFoot products are driven most frequently by low ↓ PAR (e.g. TUND). The MOD17 algorithm is effective when the drop in ↓ PAR is strong enough, but under partly cloudy conditions the algorithm tends to over-respond to a decrease in ↓ PAR because it does not account for the increase in LUE that is commonly observed under overcast conditions (Gu *et al.*, 1999; Turner *et al.*, 2003b). The Biome-BGC model used in the BigFoot scaling has an asymptotic relationship of photosynthesis to ↓ PAR so it more closely tracks tower GPP.

The MOD17 VPD scalar helped to capture a mid-season drop in GPP at SEVI. The scalar dropped to 0.2 on around DOY 240 which brought GPP down in agreement with the tower GPP. At HARV there were two short periods in July and August when the VPD scalar dropped below 0.2 thus bringing the MODIS GPP down sharply. However, tower GPP did not show this drop, which is consistent with leaf-level studies at HARV showing little response to VPD (Bassow & Bazzaz, 1998). These observations suggest that the MOD17 parameterization is oversensitive to VPD at that site. At METL, the VPD scalar brought MODIS GPP down sharply about DOY 180 which agrees well with the flux tower observations. The VPD scalar was also effective on occasional days at TUND, NOBS, and AGRO. An alternative to using the VPD scalar for tracking drought stress is the use of canopy water content indices based on shortwave infrared and near infrared reflectance (Xiao *et al.*, 2004a). These indices are

under investigation and if they prove to be effective then issues with characterizing VPD and parameterizing the VPD scalar would be minimized.

The MODIS FPAR was generally stable in the summer at all locations (Fig. 10). A notable exception was at SEVI where FPAR dropped from 0.35 to 0.28 during the mid-growing season in parallel with tower GPP. However, there was not an obvious decrease in green leaf biomass on the ground at the time (J. Carney, personal communication). Additional studies with hand held spectroradiometers are needed at low LAI sites to quantitatively show if the MODIS FPAR is responding specifically to vegetation light absorption under these circumstances. Such studies could also be used to explore the efficacy of spectral vegetation indices that have been proposed for direct tracking of canopy LUE (Gamon *et al.*, 1997).

Total GPP. The greatest underestimate of total GPP between the BigFoot and MODIS products was at AGRO where MODIS GPP was only two-thirds of BigFoot GPP (Fig. 7). The underprediction was driven primarily by an artificially low $\varepsilon_{g\text{-}max}$. Across the other sites, the MODIS $\varepsilon_{g\text{-}max}$ appears to be about right under clear sky conditions when LUE is relatively low (Fig. 11). LUE models such as MOD17 could potentially be modified to reflect an increasing ε_g under overcast conditions. The total annual GPP overestimates at SEVI and TUND are driven by an artificially high FPAR during particular parts of the year (Fig. 10). At HARV, the agreement in total GPP was good but was driven by counteracting errors in the MODIS products (i.e. an artificially long growing season but artificially low maximum GPP).

Annual NPP. The pattern of over- or underestimation of NPP generally followed that for GPP. The largest underestimation in the MODIS product was at AGRO. This relatively low value was typical for the whole region (data not shown). The underprediction at AGRO was primarily a problem with underestimating GPP. Maximum 8-day GPP at the flux tower was $\sim 13\,g\,C\,m^{-2}\,day^{-1}$ for soybeans (and probably higher for corn), whereas maximum MODIS GPP was $4\,g\,C\,m^{-2}\,day^{-1}$. This low GPP in mid-growing season when $\downarrow$ PAR, FPAR, and the VPD scalar were all high is indicative of a low value of $\varepsilon_{g\text{-}max}$ in the biome properties lookup table. The MODIS value was $0.68\,g\,C\,MJ^{-1}$ whereas estimated $\varepsilon_{g\text{-}max}$ at the AGRO site based on tower flux measurements is on the order of $3\,g\,C\,MJ^{-1}$ (Turner *et al.*, 2003a). It seems reasonable to conclude that the MODIS algorithm is significantly underestimating NPP in the America Mid-west. As croplands are usually fertilized and maintain relatively high rates of potential photosynthesis (WullschLeger, 1993), this problem could be addressed by raising $\varepsilon_{g\text{-}max}$. Note that as $\varepsilon_{g\text{-}max}$ is raised to reflect observations at the flux tower, it becomes increasingly important to introduce a modifier for clear sky vs. overcast conditions.

The largest overprediction of NPP was at SEVI (by a factor of 5). The NPP to GPP ratio in the MODIS product was 0.8 at SEVI (Fig. 7). However, the upper range of a physiologically realistic NPP to GPP ratio extends only to about 0.65 based on known rates of maintenance and growth respiration (Amthor, 2000), and that rate is expected in a cropland situation where stress is minimal. The high ratio at SEVI is mostly a problem with the overprediction of GPP, which was primarily associated with the artificially high FPAR in the off growing season period. Simulated leaf and fine root biomass at SEVI were similar for the BigFoot and MODIS products, as was the estimate for total R_a for the year ($\sim 50\,g\,C\,m^{-2}\,yr^{-1}$).

At NOBS, the overprediction in MODIS NPP is a problem of underestimating R_a rather than overestimating GPP (see also Turner *et al.*, 2003a). Intensive field studies associated with the BOREAS campaign gave an estimate of ~ 0.3 for NPP:GPP at NOBS site (Ryan *et al.*, 1997) and that is close to the ratio in the BigFoot product. The estimate from the MODIS products was about 0.5, which tends to cause the MODIS NPP to be overestimated. This may be an issue with the base rate as ecophysiological studies suggest relatively high respiration rates at a fixed temperature in plants grown in cool environment (Larigauderie & Korner, 1995) and this pattern is not reflected in the MODIS algorithm parameterization. The MODIS LAI is also a component of the R_a calculation but at NOBS it is too high (Cohen *et al.*, 2003b) which would suggest even more strongly that the base rate for respiration is too low. R_a may also be underestimated at METL. Chamber-based estimates suggest stem R_a is 33% of total foliage R_a (Law *et al.*, 1999) whereas the MOD17 value is 7%. Because of the difficulty of estimating livewood mass, it might be desirable in the case of forests to make stemwood R_a a fixed proportion of total R_a in the MOD17 algorithm.

Conclusions

Evaluation of the GPP and NPP estimates from coarse resolution sensors such as MODIS is greatly facilitated by application of a spatially distributed ecosystem process model at fine spatial resolution. This approach permits incorporation of site-level data on land cover, LAI, daily meteorology, and measurements of NPP and GPP. Spatial and temporal aggregation of model outputs permits rigorous comparisons with MODIS

products as well as analysis of the performance of the MODIS NPP/GPP algorithm. At six sites ranging widely in climate and vegetation characteristics, there was a broad array of agreement/disagreement between the ground-based and MODIS-based products, with notable limitations in the parameterization of LUE and in the seasonality of the MODIS FPAR at some sites. Continued site-level studies will support the rapid evolution of globally applied NPP/GPP algorithms and the products that underlie our emerging capability to monitor the biosphere.

Acknowledgements

This study was supported by the NASA Terrestrial Ecology Program. Flux tower measurements were funded by the Department of Energy, NOAA, NASA, and NSF. Data available through AmeriFlux, FLUXNET, and the ORNL DAAC Mercury Data System were essential to this study. Special thanks to the personnel responsible for flux tower operation at all the sites.

References

Aber JD, Reich PB, Goulden ML (1996) Extrapolating leaf CO_2 exchange to the canopy: a generalized model of forest photosynthesis validated by eddy correlation. *Oecologia*, **106**, 267–275.

AmeriFlux. 2004 http://public.ornl.gov/ameriflux/

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

Anthoni PM, Unsworth MH, Law BE *et al.* (2002) Seasonal differences in carbon and water vapor exchange in young and old-growth ponderosa pine ecosystems. *Agricultural and Forest Meteorology*, **111**, 203–222.

Asner GP, Scurlock JMO, Hicke JA (2003) Global synthesis of leaf area index observations: implications for ecological and remote sensing studies. *Global Ecology and Biogeography*, **12**, 191–205.

Aubinet M, Heinesch B, Longdoz B (2002) Estimation of the carbon sequestration by a heterogeneous forest, night flux corrections, heterogeneity of the site and inter-annual variability. *Global Change Biology*, **8**, 1053–1071.

Barford CD, Wofsy SC, Goulden ML *et al.* (2001) Factors controlling long- and short-term sequestration of atmospheric CO_2 in a mid-latitude forest. *Science*, **294**, 1688–1691.

Bassow SL, Bazzaz FA (1998) How environmental conditions affect canopy leaf-level photosynthesis in four deciduous tree species. *Ecology*, **79**, 2660–2675.

Bisbee K, Gower ST, Norman JM *et al.* (2001) Environmental controls on ground cover species composition and productivity in a boreal black spruce forest. *Oecologia*, **129**, 261–270.

Burrows SN, Gower ST, Clayton MK *et al.* (2002) Application of geostatistics to characterize leaf area index (LAI) from flux tower to landscape scales using a cyclic sampling design. *Ecosystems*, **5**, 667–679.

Campbell JL, Burrows S, Gower ST *et al.* (1999) BigFoot: Characterizing Land Cover, LAI, and NPP at the Landscape Scale for EOS/MODIS Validation. Field Manual 2.1. Environmental Sciences Division Publication 4937. Environmental Sciences Division, Oak Ridge National Laboratory: Oak Ridge, TN. http://www.daac.ornl.gov/eos_land_val/BigFoot/index_bigfoot.htm

Chen JM, Rich PM, Gower ST *et al.* (1997) Leaf area index of boreal forests: theory, techniques, and measurements. *Journal of Geophysical Research*, **102**, 29,429–29,444.

Cienciala E, Running SW, Lindroth A *et al.* (1998) Analysis of carbon and water fluxes from the NOPEX boreal forest: comparison of measurements with FOREST-BGC simulations. *Journal of Hydrology*, **212–213**, 62–78.

Clark DA, Brown S, Kicklighter DW *et al.* (2001) Measuring net primary production in forests: concepts and field methods. *Ecological Applications*, **11**, 356–370.

Cohen WB, Maiersperger TK, Gower ST *et al.* (2003a) An improved strategy for regression of biophysical variables and Landsat ETM + data. *Remote Sensing of Environment*, **84**, 561–571.

Cohen WB, Maiersperger TK, Yang Z *et al.* (2003b) Comparisons of land cover and LAI estimates derived from ETM + and MODIS for four sites in North America: a quality assessment of 2000/2001 provisional MODIS products. *Remote Sensing of Environment*, **88**, 233–255.

Cohen WB, Spies TA, Alig RJ *et al.* (2002) Characterizing 23 years (1972–95) of stand replacement disturbance in western Oregon forests with Landsat imagery. *Ecosystems*, **5**, 122–137.

Cohen WB, Spies TA, Fiorella M (1995) Estimating the age and structure of forests in a multi-ownership landscape of western Oregon, U.S.A. *International Journal of Remote Sensing*, **16**, 721–746.

Crist EP, Cicone RC (1984) A physically-based transformation of thematic mapper data – the TM tasseled cap. *IEEE Transactions in Geosciences and Remote Sensing*, **GE-2**, 256–263.

Curtis PS, Hanson PJ, Barford P *et al.* (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.

Falge E, Baldocchi D, Olson R *et al.* (2001) Gap filling strategies for long term energy flux data sets. *Agricultural and Forest Meteorology*, **107**, 43–69.

FLUXNET. 2004 http://daacl.esd.ornl.gov/FLUXNET/

Gamon JA, Serrano L, Surfus JS (1997) The photochemical reflectance index: an optical indicator of photosynthetic radiation use efficiency across species, functional types, and nutrient levels. *Oecologia*, **112**, 492–501.

Gholz HL, Vogel SA, Cropper WP Jr *et al.* (1991) Dynamics of canopy structure and light interception in *Pinus elliottii* stands, North Florida. *Ecological Monographs*, **61**, 33–51.

Goldstein AH, Hultman NE, Fracheboud JM *et al.* (2000) Effects of climate variability on the carbon dioxide, water, and sensible heat fluxes about a ponderosa pine plantation in the Sierra Nevada (CA). *Agricultural and Forest Meteorology*, **101**, 113–129.

Goulden ML, Daube BC, Fan S-M *et al.* (1997) Physiological responses of a black spruce forest to weather. *Journal of Geophysical Research*, **102**, 28987–28996.

Goulden ML, Munger JW, Fan S *et al.* (1996a) Measurements of carbon sequestration by long-term eddy covariance: methods and a critical evaluation of accuracy. *Global Change Biology*, **2**, 169–182.

Goulden ML, Munger JW, Fan S-M *et al.* (1996b) Exchange of carbon dioxide by a deciduous forest: response to interannual climate variability. *Science*, **271**, 1576–1578.

Gower ST, Krankina O, Olson RJ *et al.* (2001) Net primary production and carbon allocation patterns of boreal forest ecosystems. *Ecological Applications*, **11**, 1395–1411.

Gower ST, Kucharik CJ, Norman JM (1999) Direct and indirect estimation of leaf area index, fAPAR and net primary production of terrestrial ecosystems. *Remote Sensing of Environment*, **70**, 29–51.

Gower ST, McMurtrie RE, Murty D (1996) Above ground net primary production decline with stand age: potential causes. *Trends in Ecology and Evolution*, **11**, 378–382.

Gower ST, Vogel JC, Norman JM *et al.* (1997) Carbon distribution and above ground net primary production in aspen, jack pine, and black spruce stands in Saskatchewan and Manitoba, Canada. *Journal of Geophysical Research*, **102**, 29029–29041.

Granier A, Ceschia E, Damesin C *et al.* (2000) The carbon balance of a young beech forest. *Functional Ecology*, **14**, 312–325.

Gu L, Fuentes JD, Staebler RM *et al.* (1999) Responses of net ecosystem exchanges of carbon dioxide to changes in cloudiness: results from two North American deciduous forests. *Journal of Geophysical Research*, **104**, 31421–31434.

Heinsch FA, Reeves M, Bowker CF (2003) User's Guide, GPP and NPP (MOD 17A2/A3) Products, NASA MODIS Land Algorithm. http://www.forestry.umt.edu/ntsg/

Irvine J, Law BE, Anthoni PM *et al.* (2002) Water limitations to carbon exchange in old-growth and young ponderosa pine stands. *Tree Physiology*, **22**, 189–196.

Irvine J, Law BE, Kurpius MR *et al.* (2004) Age-related changes in ecosystem structure and function and effects on water and carbon exchange in ponderosa pine. *Tree Physiology*, **24**, 753–763.

Jarvis PG, Leverenz JW (1983) Productivity of temperate deciduous and evergreen forests. In: *Ecosystem Processes: Mineral Cycling, Productivity, and Man's Influence. Physiological Plant Ecology, New Series*, **Vol. 12D** (eds Lange OL, Nobel PS, Osmond CB, Ziegler H), pp. 233–280. Springer-Verlag, New York.

Justice CO, Hall DK, Defries R *et al.* (1998) The moderate resolution imaging spectroradiometer (MODIS): land remote sensing for global change research. *IEEE Transactions on Geosciences and Remote Sensing*, **36**, 1228–1249.

Kimball JS, Keyser AR, Running SW *et al.* (2000) Regional assessment of boreal forest productivity using an ecological process model and remote sensing parameter maps. *Tree Physiology*, **20**, 761–775.

Kurc SA, Small EE (2004) Dynamics of evapotranspiration in semiarid grassland and shrubland during the summer monsoon season, central New Mexico. *Water Resources Research*, **40**, 1–15 (W0305, doi: 1029/2004WR003068).

Larigauderie A, Korner C (1995) Acclimation of leaf dark respiration to temperature in alpine and lowland plant species. *Annals of Botany*, **76**, 245–252.

Law BE, Ryan MG, Anthoni PM (1999) Seasonal and annual respiration of a Ponderosa Pine ecosystem. *Global Change Biology*, **5**, 169–182.

Law BE, Sun OJ, Campbell J *et al.* (2003) Changes in carbon storage and fluxes in a chronosequence of ponderosa pine. *Global Change Biology*, **9**, 510–524.

Law BE, Thornton PE, Irvine J *et al.* (2001a) Carbon storage and fluxes in ponderosa pine forests at different developmental stages. *Global Change Biology*, **7**, 1–23.

Law BE, Turner DP, Lefsky M *et al.* (2004b) Carbon fluxes across regions: observational constraints at multiple scales. In: *Scaling and Uncertainty Analysis in Ecology: Methods and Applications* (eds Wu J, Jones B, Li H, Loukes O), Columbia University Press, New York (in press).

Law BE, Turner DP, Sun O *et al.* (2004a) Disturbance and climate effects on carbon stocks and fluxes across western Oregon USA. *Global Change Biology*, **10**, 1–16.

Law BE, Van Tuyl S, Cescatti A *et al.* (2001b) Estimation of leaf area index in open-canopy ponderosa pine forests at different successional stages and management regimes in Oregon. *Agricultural and Forest Meteorology*, **108**, 1–14.

Law BE, Waring RH, Anthoni PM *et al.* (2000) Measurements of gross and net ecosystem productivity and water vapor exchange of a *Pinus ponderosa* ecosystem, and an evaluation of two generalized models. *Global Change Biology*, **6**, 155–168.

Lechowicz MJ (1984) Why do temperate deciduous trees leaf out at different times? Adaptation and ecology of forest communities. *American Naturalist*, **124**, 821–842.

Lucht W, Prentice IC, Myneni RB *et al.* (2002) Climatic control of the high-latitude vegetation greening trend and Pinatubo effect. *Science*, **296**, 1687–1689.

Meyers TP, Hollinger SE (2004) An assessment of storage terms in the surface energy balance of maize and soybean. *Agricultural and Forest Meteorology*, **125**, 105–115.

Miller PC, Webber PJ, Oechel WC *et al.* (1980) Biophysical processes and primary production. In: *An Arctic Ecosystem, The Coastal Tundra at Barow, Alaska* (eds Brown J, Miller PC, Tiezen LL, Bunnell FL), pp. 66–101. Dowden, Hutchinson and Ross, Inc., Stroudsburg, PA.

Morisette J, Privette J, Justice C (2002) A framework for the validation of MODIS land products. *Remote Sensing of Environment*, **83**, 77–96.

Myneni RB, Keeling CD, Tucker CD *et al.* (1997a) Increased plant growth in the northern high latitudes from 1981–1991. *Nature*, **386**, 698–702.

Myneni RB, Nemani RR, Running SW (1997b) Estimation of global leaf area index and absorbed PAR using radiative transfer models. *IEEE Transactions on Geosciences and Remote Sensing*, **35**, 1380–1393.

Nemani RR, Keeling CD, Hashimoto H *et al.* (2003) Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science*, **300**, 1560–1563.

Oechel WC, Vourlitis GL, Hastings SJ *et al.* (1995) Change in arctic CO_2 flux over two decades: effects of climate change at Barrow, Alaska. *Ecological Applications*, **5**, 846–855.

Prince SD, Haskett J, Steininger M *et al.* (2001) Net primary production of U.S. Midwest croplands from agricultural harvest yield data. *Ecological Applications*, **11**, 1194–1205.

Reich PB, Turner DP, Bolstad P (1999) An approach to spatially-distributed modeling of net primary production (NPP) at the landscape scale and its application in validation of EOS NPP products. *Remote Sensing of Environment*, **70**, 69–81.

Rochette P, Desjardins RL, Pattey E *et al.* (1995) Crop net carbon dioxide exchange rate and radiation use efficiency in soybean. *Agronomy Journal*, **87**, 22–28.

Running SW (1994) Testing FOREST-BGC ecosystem process simulations across a climatic gradient in Oregon. *Ecological Applications*, **4**, 238–247.

Running SR, Baldocchi DD, Turner DP *et al.* (1999) A global terrestrial monitoring network integrating tower fluxes, flask sampling, ecosystem modeling and EOS satellite data. *Remote Sensing of Environment*, **70**, 108–128.

Running SW, Nemani R, Heinsch FA *et al.* (2004) A continuous satellite-derived measure of global terrestrial primary production. *BioScience*, **54**, 547–560.

Running SW, Nemani RR, Hungerford RD (1987) Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evapotranspiration and photosynthesis. *Canadian Journal of Forest Research*, **17**, 472–483.

Running SW, Thornton PE, Nemani R *et al.* (2000) Global Terrestrial Gross and Net Primary Productivity from the Earth Observing System. In: *Methods in Ecosystem Science* (eds Sala OE, Jackson RB, Mooney HA, Howarth RW), pp. 44–57. Springer-Verlag, New York.

Ryan MG, Lavigne MB, Gower ST (1997) Annual carbon cost of autotrophic respiration in boreal forest ecosystems in relation to species and climate. *Journal of Geophysical Research*, **102**, 28871–28884.

Sala OE, Austin AT (2000) Methods of estimating aboveground net primary productivity. In *Methods in Ecosystem Science* (eds Sala OE, Jackson RB, Mooney HA, Howarth RW), pp. 31–43. Springer-Verlag, New York.

Schmid HP (2002) Footprint modeling for vegetation atmosphere exchange studies: a review and perspective. *Agricultural and Forest Meteorology*, **113**, 159–183.

Sellers PJ, Hall FG, Kelly RD *et al.* (1997) BOREAS in 1997: experiment overview, scientific results, and future directions. *Journal of Geophysical Research*, **102**, 28731–28769.

Shaver GR, Laundre JA, Giblin AE *et al.* (1996) Changes in live plant biomass, primary production, and species composition along a riverside toposequence in arctic Alaska, USA. *Arctic and Alpine Research*, **28**, 363–379.

Smith ML, Ollinger SV, Aber JD *et al.* (2002) Direct estimation of aboveground forest productivity through hyperspectral remote sensing of canopy nitrogen. *Ecological Applications*, **12**, 1286–1302.

Steele SJ, Gower ST, Vogel JG *et al.* (1997) Root mass, net primary production and turnover in aspen, jack pine and black spruce forests in Saskatchewan and Manitoba, Canada. *Tree Physiology*, **17**, 577–587.

Stow DA, Hope A, McGuire D *et al.* (2004) Remote sensing of vegetation and land-cover change in arctic tundra ecosystems. *Remote Sensing of Environment*, **89**, 281–308.

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

Turner DP, Ollinger SV, Kimball JS (2004b) Integrating remote sensing and ecosystem process models for landscape to regional scale analysis of the carbon cycle. *BioScience*, **54**, 573–584.

Turner DP, Ollinger S, Smith ML *et al.* (2004a) Scaling net primary production to a MODIS footprint in support of Earth Observing System product validation. *International Journal of Remote Sensing*, **25**, 1961–1979.

Turner DP, Ritts WD, Cohen WB *et al.* (2003a) Scaling gross primary production (GPP) over boreal and deciduous forest landscapes in support of MODIS GPP product validation. *Remote Sensing of Environment*, **88**, 256–270.

Turner DP, Urbanski S, Wofsy SC *et al.* (2003b) A cross-biome comparison of light use efficiency for gross primary production. *Global Change Biology*, **9**, 383–395.

Valentini R, DeAngelis P, Matteucci G *et al.* (1996) Seasonal net carbon dioxide exchange of a beech forest with the atmosphere. *Global Change Biology*, **2**, 199–207.

White MA, Thornton PE, Running SW *et al.* (2000) Parameterization and sensitivity analysis of the BIOME-BGC terrestrial ecosystem model: net primary production controls. *Earth Interactions*, **4**, 1–85.

Whitford WG (2002) *Ecology of Desert Systems*. Academic Press, San Diego.

Wilson KB, Baldocchi DD, Hanson PJ (2001) 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.

Wofsy SC, Goulden JW, Munger S-MF *et al.* (1993) Net exchange of CO_2 in a mid-latitude forest. *Science*, **260**, 1314–1317.

WullschLeger SD (1993) Biochemical limitations to carbon assimilation in C3 plants. A retrospective analysis of A/Ci curves from 109 species. *Journal of Experimental Botany*, **44**, 907–920.

Wythers KR, Reich PB, Turner DP (2003) Predicting leaf area index from scaling principles: corroboration and consequences. *Tree Physiology*, **23**, 1171–1179.

Xiao X, Hollinger D, Aber J *et al.* (2004a) Satellite-based modeling of gross primary production in an evergreen needleleaf forest. *Remote Sensing of Environment*, **89**, 519–534.

Xiao X, Zhang Q, Braswell B *et al.* (2004b) Modeling gross primary production of temperate broadleaf forest using satellite images and climate data. *Remote Sensing of Environment*, **91**, 256–270.

Zhang X, Friedl MA, Schaaf CB *et al.* (2003) Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, **84**, 471–475.