Estimating forest biomass in the USA using generalized allometric models and MODIS land products

Xiaoyang Zhang[1,2] and Shobha Kondragunta[1]

Received 31 January 2006; revised 4 April 2006; accepted 6 April 2006; published 11 May 2006.

[1] Spatially-distributed forest biomass components are essential to understand carbon cycle and the impact of biomass burning emissions on air quality. We estimated the density of forest biomass components (foliage biomass, branch biomass, and aboveground biomass) at a spatial resolution of 1 km across the Contiguous United States using foliage-based generalized allometric models and Moderate-Resolution Imaging Spectroradiometer (MODIS) land data. The foliage biomass for each forest type was calculated from MODIS leaf-area indices, land cover types, and vegetation continuous fields. The foliage-based models were developed using available diameter-based allometric equations and used to estimate branch biomass and aboveground biomass. The resultant aboveground biomass density matches well with the data from Forest Inventory and Analysis program at both state and pixel levels. **Citation:** Zhang, X., and S. Kondragunta (2006), Estimating forest biomass in the USA using generalized allometric models and MODIS land products, *Geophys. Res. Lett.*, *33*, L09402, doi:10.1029/2006GL025879.

1. Introduction

[2] Biomass components in forests have been used as important parameters in investigating forest-atmosphere carbon exchanges and biomass burning emissions [e.g., *Dixon et al.*, 1994; *Ito and Penner*, 2004]. To estimate tree biomass at field plot scales (usually less than one acre), a large number of studies have focused on the development of species and site specific allometric models depending on bole diameter at breast height [e.g., *Paster et al.*, 1984; *Ter-Mikaelian and Korzukin*, 1997]. The plot estimates of national forest inventories are commonly aggregated to represent forest biomass at national or regional scales [*Brown et al.*, 1999; *Jenkins et al.*, 2001]. However, the forest inventories only characterize the commercially valuable wood rather than all forest biomass and need many years to complete [*Brown et al.*, 1999].

[3] Remote sensing provides a method to develop spatially-distributed forest biomass from local to regional areas. Recently, vegetation biomass parameters have been directly associated with remotely sensed vegetation indices using empirical regression techniques. For example, field measurements are statistically related to Landsat TM data [e.g., *Lu*, 2005] and to AVHRR Normalized Difference Vegetation Index (NDVI) [*Dong et al.*, 2003]. To avoid the difficulties inherent in linear statistical models, nonparametric approaches have also been developed. Specifically, forest biomass is estimated from MODIS (Moderate Resolution Imaging Spectroradiometer) reflectances and ancillary variables (precipitation, temperature, and elevation) using a decision tree-based model [*Baccini et al.*, 2004] and from SPOT VEGETATION using an artificial neural network [*Fraser and Li*, 2002].

[4] These statistical methods suffer from several fundamental limitations. (1) Generally, a model developed and trained using data at a specified time or region is unlikely to work well at a different time and region. (2) A large amount of field measurements are required to build up and to test the models for each calculation. (3) Temporal and spatial resolutions in the training data surveyed in fields require matching satellite observations for developing algorithms.

[5] In this study, we estimated biomass in different forest components across the Contiguous United States (CONUS) from MODIS land data using allometric models. Specifically, we developed the foliage-based generalized allometric models from published species-specific equations to calculate forest biomass components. The foliage biomass was then functionally associated with vegetation leaf area index (LAI) and specific leaf area (SLA). Finally, the foliage biomass, branch biomass, and aboveground biomass in forest areas were calculated from MODIS LAI, land cover type, and tree continuous field at a spatial resolution of 1 km.

2. Methodology

2.1. Foliage-Based Generalized Allometric Models

[6] The tree allometric models are developed for determining tree biomass by regressing the biomass of entire trees or their components against some easily measured variables in the fields. These models are generally species-specific and site-specific, but they are also generalized to estimate biomass in mixed species across large regions [e.g., *Jenkins et al.*, 2003; *Wirth et al.*, 2004]. In the allometric models, the most commonly used form is to link biomass components to the diameter at breast height (DBH) [e.g., *Paster et al.*, 1984; *Ter-Mikaelian and Korzukin*, 1997]:

$$M_c = \alpha_1 D^{\beta_1} \tag{1}$$

$$M_f = \alpha_2 D^{\beta_2} \tag{2}$$

where M_c is the oven-dry weight (kg) of the biomass components of a tree, including branches (M_b) and total aboveground biomass (M_a); M_f is foliage biomass (kg); D represents DBH (cm); and α_1, α_2, β_1 and β_2 are coefficients.

[1]NOAA/NESDIS/Center for Satellite Applications and Research, Camp Springs, Maryland, USA.
[2]Also at Earth Resources Technology, Inc., Jessup, Maryland, USA.

0094-8276/06/2006GL025879$05.00

[7] While the DBH currently cannot be obtained from satellite data, canopy leaf properties are widely measured from spectral reflectance. Therefore, we derived a foliage-based allometric model by eliminating D in equation (1) using equation (2).

$$M_c = \delta M_f^{\gamma} \tag{3}$$

where γ and δ are coefficients.

[8] To determine parameters γ and δ, a set of samples are needed. Because of the lack of original field measurements, we simulated samples from the existing diameter-based models in literature. From the published models at various environments [e.g., *Gholz et al.*, 1979; *Ter-Mikaelian and Korzukin*, 1997] the biomass samples were derived for broadleaf forests and needleleaf forests, separately. In particular, the samples were simulated from each group of species-specific diameter-based models:

$$\begin{cases} M_f = \alpha_f D^{\beta_f} + \varepsilon_f \\ M_b = \alpha_b D^{\beta_b} + \varepsilon_b \\ M_a = \alpha_a D^{\beta_a} + \varepsilon_a \end{cases} \tag{4}$$

where f represents foliage; a is aboveground tree; b is branch; ε is uncertainty; D is the DBH; α and β are coefficients. When randomly selecting a D value varying within the range of original data sets and ε values within the standard errors (±SE) in the original models, we calculated a pair of M_f, M_b, and M_a. Five pairs of biomass samples, which were assumed to represent the equation characteristics, were randomly simulated from each group of models.

[9] Using the simulated data sets and least squares fitting, we estimated coefficients γ and δ in foliage-based allometric models for broadleaf forests and needleleaf forests respectively. The models were also designed separately for the eastern and western US because climate is humid in the eastern US while it is Mediterranean-like in the far western US.

2.2. Calculation of Foliage Biomass

[10] Foliage biomass density is a function of LAI and SLA. It can be calculated using the following formula [*Heinsch et al.*, 2003]:

$$M_f = LAI/SLA \tag{5}$$

LAI is a function of vegetation growing seasons and the maximum value can be retrieved from MODIS data (see details in the following sections). The SLA is defined in mass units of carbon and is converted to dry weight (m^2/kg). We determined the SLA values for various land cover types according to field measurements [*White et al.*, 2002] and global MODIS GPP (Gross Primary Productions) and NPP (Net Primary Productions) models [*Heinsch et al.*, 2003].

2.3. MODIS Land Products

[11] The MODIS LAI product (MOD15A2) provides green leaf area index at a spatial resolution of 1 km globally [*Myneni et al.*, 2002]. We collected LAI data at an interval of 8 days from 2002–2004 in the CONUS. To reduce the noise in the LAI time series, these data were composed to monthly LAI by averaging the LAI values that passed quality checks indicated by quality assessment flags in the LAI product. The monthly LAI within a year was compared to retrieve the maximum monthly LAI in order to minimize the impacts of LAI seasonality. Finally, to reduce the uncertainty induced by interannual climate change and other factors, the maximum monthly LAI in 2002, 2003, and 2004 were averaged to represent the maximum LAI.

[12] MODIS vegetation continuous field (MOD44B) algorithm produces percent tree cover, percent nontree vegetation (shrubs, crop, and herbaceous), and percent bare ground at a resolution of 500 m [*Hansen et al.*, 2003]. We collected the data set produced from MODIS data between November 2000 and December 2001, which is the only version currently available.

[13] We obtained MODIS land cover data (MOD12Q1) at a 1 km resolution for 2002, 2003, and 2004 [*Friedl et al.*, 2002]. To minimize the uncertainty in these three data sets, we compiled a land cover data set based on the highest confidence assessment value in each pixel. The land cover types were further stratified to needleleaf forests, broadleaf forests, mixed forests, savanna, shrublands, and grasslands.

2.4. LAI Data in Forests

[14] The maximum monthly LAI was assigned to different forest types in each pixel according to land cover type data. The percent tree cover in the land cover types of needleleaf forests and broadleaf forests was considered as needleleaf trees and broadleaf trees, respectively. The percent tree cover in other land cover types was considered to be equally mixed by broadleaf and needleleaf trees because it is hard to separate them.

[15] Because the maximum monthly LAI value in a pixel is generally for a mixture of tree and nontree vegetation, we retrieved the tree LAI in subpixels for needleleaf trees, broadleaf trees, and mixed trees, respectively. Specifically, we first extracted LAI values for each land cover type from the relatively pure pixels where the related percent cover was larger than 90%. These LAI values between 2002 and 2004 were then averaged to represent the LAI in a pure land cover type, which was applied to calculate the LAI ratios between each forest type and shrubs or grasses. Thus, the ratios combining with the percentages of tree and nontree vegetation were used to estimate the tree LAI values in subpixels by assuming that the LAI in a pixel was a linear mixture between tree and nontree vegetation.

[16] Using the LAI and SLA data, the foliage biomass density was determined from equation (5). To input the foliage biomass density in a pixel to allometric models for the calculation of biomass components, a scaling effect was involved since the models were derived from the measurements of individual trees. To reduce the potential scaling effect, the foliage biomass for a tree crown area was simply estimated from biomass density with the assumption that the mean radius of crown size is about 3 m [e.g., *Brown et al.*, 2000; *Popescu et al.*, 2003].

2.5. Data for Biomass Assessment

[17] Two data sets were obtained to evaluate our biomass estimates. Forest inventory database produced by National Forest Inventory Analysis (FIA) program in the US Department of Agriculture (USDA) Forest Service is available at field plot scales in the US (FIA, www.fia.fs.fed.us,

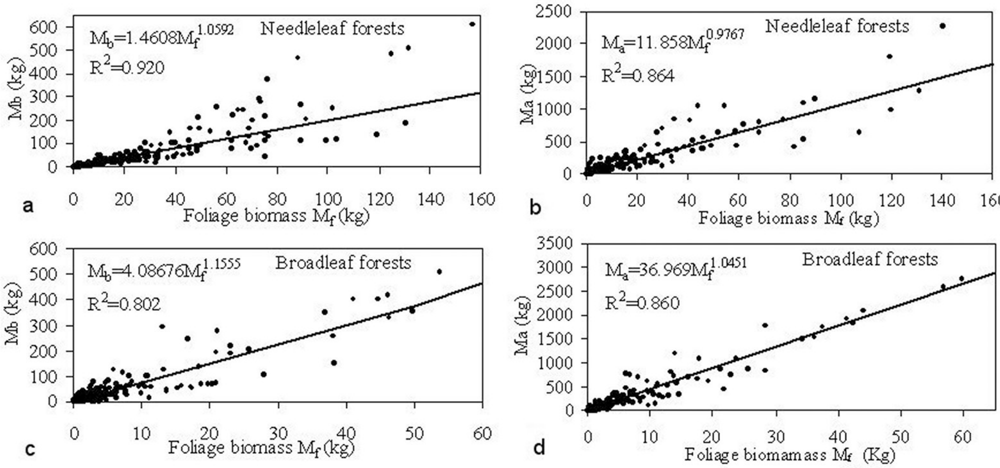

Figure 1. Foliage-based allometric models in the eastern US. The aboveground biomass (M_a) and branch biomass (M_b) vary with foliage biomass (M_f) in the (a and b) needleleaf forests and (c and d) broadleaf forests, respectively.

2003). Since an FIA plot represents one acre (0.004 km^2) sample area, the plot-based data are hard to directly compare with the estimates in MODIS pixels. Therefore, the average of plot biomass in each state of the eastern US [*Chojnacky et al.*, 2004] was adopted.

[18] The second data set is a forest biomass map with a spatial resolution of 30 m in the National Forests in California. This data set was generated by intersecting FIA-derived timber volume estimates with a forest cover map [*Franklin et al.*, 2000]. The timber volume data were then converted to the biomass values using an expansion factor coefficient [*Franklin et al.*, 2000; *Baccini et al.*, 2004]. This biomass data set was aggregated to compare with our MODIS estimates.

3. Results and Discussion

3.1. Relationships for Branch Biomass and Aboveground Biomass on Foliage Biomass

[19] Figure 1 shows variations in branch biomass and aboveground biomass with foliage biomass derived from simulated data in the eastern US. The branch biomass and aboveground biomass are significantly related to foliage biomass (R^2 > 0.8) although there are some outliers for samples with large foliage biomass. This suggests that the variance increases with foliage biomass (canopy size), which is also common in the species-specific models [*Grote*, 2002]. The results in the western US provide similar relationships (R^2 > 0.95).

[20] The foliage-based allometric models show that the coefficient δ varies greatly but γ is close to 1 in both aboveground biomass and branch biomass. This result implies that tree biomass components tend to be linearly related to foliage biomass. In other words, the foliage-based models are not very sensitive to the measurement scale of foliage biomass, especially in low density regions.

3.2. Spatial Patterns in Biomass Components

[21] Figure 2 displays spatial patterns in foliage biomass, branch biomass, and total aboveground biomass. The foliage biomass is higher in needleleaf forest areas such as the Pacific Northwest and boreal forest areas. Relatively, the values are small in broadleaf forest regions. This difference is mainly attributed to the SLA values which are about 4 times larger in needleaf forests than in broadleaf forests [*White et al.*, 2002]. The biomass density in both branch biomass and aboveground biomass are high in the Pacific Northwest regions and the eastern US (especially Appalachian Mountains). In contrast, the density is very low in the central and southwestern US where croplands and shrublands are mainly dominated.

[22] Figure 3 presents the forest biomass density for the states with forest cover larger than 10% of state territory. On average, the aboveground biomass density at state scales is generally larger than 100 tons/ha. It is larger than 170 tons/ha in Washington, Pennsylvania, and West Virginia, while the values are smaller in Colorado and Florida. The average biomass density across the CONUS is about 4.9 ± 1.2, 21.6 ± 4.3, and 141.1 ± 28.8 tons/ha for foliage, branch, and aboveground biomass, respectively. The total aboveground biomass varies considerably in different states, which is larger than 2.7 $\times$ 10^9 tons in Washington and Oregon.

3.3. Assessment of Biomass Estimates

[23] Comparison of aboveground biomass density indicates that estimates from MODIS data and allometric models match the average of FIA values very well at a state level in the eastern US (Figure 4a). The root mean square error (RMSE) is 21 tons/ha and the coefficient of determination (R^2) is 0.58. However, the aboveground biomass in Iowa is greatly underestimated from the MODIS data. This is likely due to the fact that trees are sparse in the 1 km pixels because the forest cover is only 0.4% in Iowa while the FIA data usually represent plots well covered by forests.

[24] The comparison in California provides assessment at a pixel level (Figure 4b). The result reveals that the MODIS estimates are associated with over 61% of the variation in FIA data with a RMSE of 46 tons/ha. Exclusion of 3% of outliers increases R^2 by 0.68 and decreases RMSE by 40 tons/ha. Similar to the estimates from the regression tree-based model

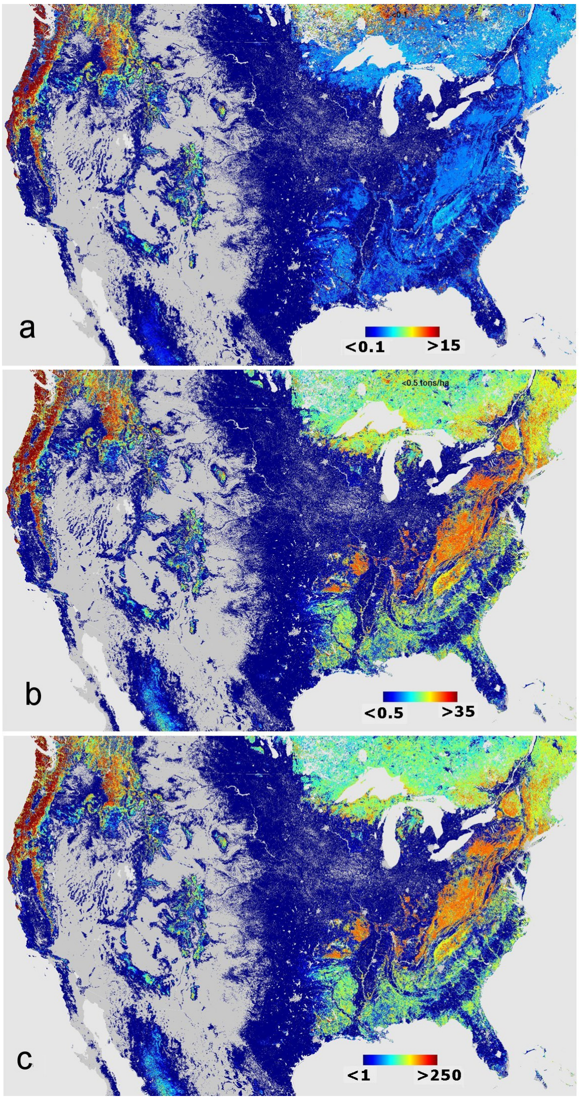

Figure 2. Biomass (tons/ha) components derived from MODIS land data and foliage-based generalized allometric models. (a) Foliage biomass, (b) branch biomass, and (c) aboveground biomass.

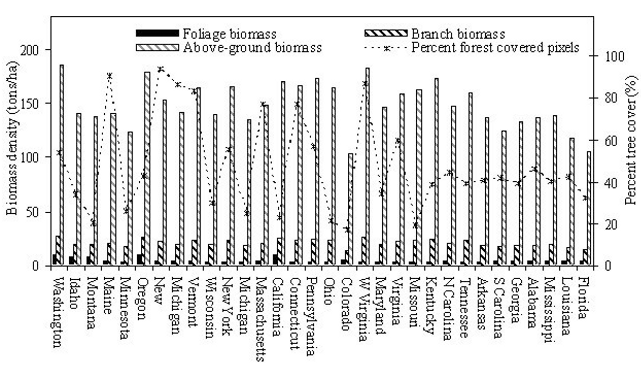

Figure 3. Average biomass density (foliage, branch, and aboveground) and percentage of forest-covered areas in different states.

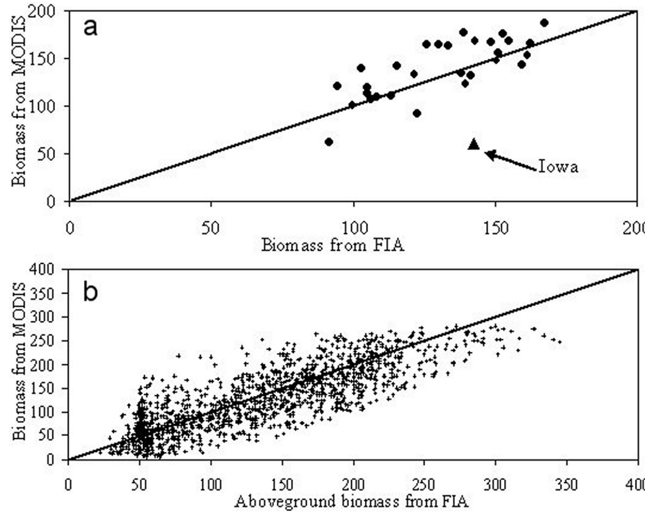

Figure 4. Evaluation of aboveground biomass density (tons/ha) measured from MODIS land data using FIA data. (a) Sate level in the eastern US, (b) pixel level in the National Forests in California.

[*Baccini et al.*, 2004], aboveground biomass is slightly underestimated for the areas where the biomass is larger than 250 tons/ha. This discrepancy is likely attributed to the factors including the sparse forests containing in MODIS pixels, the saturation of MODIS estimates, and the representations of different years in MODIS and FIA data.

4. Conclusions

[25] The results in this study suggest that MODIS data combined with foliage-based allometric models provide a robust tool to estimate biomass components across continental scales. Once the allometric models are well established, they can be directly applied to calculate and update biomass data easily in regional and global areas. Moreover, this method produces not only commonly used aboveground biomass, but also foliage biomass and branch biomass. These components are particularly important for biomass burning estimates.

[26] The generalized foliage-based allometric models indicate that foliage biomass accounts for more than 80% of variance in both branch biomass and aboveground biomass. On average, the forest biomass density across the CONUS is 5, 22, and 141 tons/ha for foliage, branch, and aboveground trees, respectively. The aboveground biomass produced in this study matches field data sets very well with an RMSE of 21 tons/ha on state average and 40 tons/ha at pixel scales, although the time periods and the measurement sizes in field data do not match those from satellite data.

[27] **Acknowledgments.** This research was supported by NOAA Joint Center for Satellite Data Assimilation. The authors wish to express their thanks to Alessandro Baccini for providing biomass data in California, to Wenze Yang for help in preparing LAI data, to Maoshen Zhao for discussions and suggestions, and to Dan Tarpley and Felix Kogan for helpful comments. The views, opinions, and findings contained in those works are these of the author(s) and should not be interpreted as an official NOAA or US Government position, policy, or decision.

References

Baccini, A., M. A. Friedl, C. E. Woodcock, and R. Warbington (2004), Forest biomass estimation over regional scales using multisource data, *Geophys. Res. Lett.*, *31*, L10501, doi:10.1029/2004GL019782.

Brown, S. L., P. Schroeder, and J. S. Kern (1999), Spatial distribution of biomass in forests of the eastern USA, *For. Ecol. Manage.*, *123*, 81–90.

Brown, P. L., D. Doley, and R. J. Keenan (2000), Estimating tree crown dimension using digital analysis of vertical photographs, *Agric. For. Meteorol.*, *100*, 199–212.

Chojnacky, D. C., R. A. Mickler, L. S. Meath, and C. W. Woodall (2004), Estimates of down woody materials in eastern US forests, *Environ. Manage.*, *33*, S44–S55.

Dixon, R. K., R. A. Houghton, A. M. Solomon, M. C. Trexler, and J. Wisniewski (1994), Carbon pools and flux of global forest ecosystems, *Science*, *263*, 185–190.

Dong, J. R., et al. (2003), Remote sensing estimates of boreal and temperate forest woody biomass: Carbon pools, sources, and sinks, *Remote Sens. Environ.*, *84*, 393–410.

Franklin, J., C. E. Woodcock, and R. Warbington (2000), Multi-attribute vegetation maps of forest service lands in California supporting resource management decisions, *Photogramm. Eng. Remote Sens.*, *66*, 1209–1217.

Fraser, R. H., and Z. Li (2002), Estimating fire-related parameters in boreal forest using SPOT VEGETATION, *Remote Sens. Environ.*, *82*, 95–110.

Friedl, M. A., et al. (2002), Global land cover mapping from MODIS: Algorithms and early results, *Remote Sens. Environ.*, *83*, 287–302.

Gholz, H. L., C. C. Grier, A. G. Campbell, and A. T. Brown (1979), Equations for estimating biomass and leaf area of plants in the Pacific Northwest, *Res. Pap. 41*, 39 pp., Oreg. State Univ., For. Res. Lab., Corvallis, Oreg.

Grote, R. (2002), Foliage and branch biomass estimation of coniferous and deciduous tree species, *Silva Fennica*, *36*, 779–788.

Hansen, M. C., R. S. DeFries, J. R. G. Townshend, M. Carroll, and C. Dimiceli (2003), Global percent tree cover at a spatial resolution of 500 meters: First results of the MODIS vegetation continuous fields algorithm, *Earth Interact.*, *7*, 1–15.

Heinsch, F. A., et al. (2003), User's guide GPP and NPP (MOD17A2/A3) products NASA MODIS land algorithm, Sch. of For., Univ. of Mont., Missoula, Mont. (Available at http://www.ntsg.umt.edu/modis/MOD17UsersGuide.pdf)

Ito, A., and J. E. Penner (2004), Global estimates of biomass burning emissions based on satellite imagery for the year 2000, *J. Geophys. Res.*, *109*, D14S05, doi:10.1029/2003JD004423.

Jenkins, J. C., R. A. Birdsey, and Y. Pan (2001), Biomass and NPP estimation for the mid-Atlantic region (USA) using plot-level forest inventory data, *Ecol. Appl.*, *11*, 1174–1193.

Jenkins, J. C., D. C. Chojnacky, L. S. Heath, and R. Birdsey (2003), National-scale biomass estimators for United States tree species, *For. Sci.*, *49*, 12–35.

Lu, D. (2005), Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon, *Int. J. Remote Sens.*, *26*, 2509–2525.

Myneni, R. B., et al. (2002), Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data, *Remote Sens. Environ.*, *83*, 214–231.

Paster, J., J. Aber, and J. M. Melillo (1984), Biomass prediction using generalized allometric regressions for some Northeast tree species, *For. Ecol. Manage.*, *7*, 265–274.

Popescu, S. C., R. H. Wynne, and R. F. Nelson (2003), Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass, *Can. J. Remote Sens.*, *29*, 564–577.

Ter-Mikaelian, M. T., and M. D. Korzukin (1997), Biomass equations for sixty-five North America tree species, *For. Ecol. Manage.*, *97*, 1–24.

White, M. A., P. E. Thornton, S. W. Runing, and R. R. Nemani (2002), Parameterization and sensitivity analysis of the BIOME_BGC Terrestrial Ecosystem Model: Net Primary production controls, *Earth Interact.*, *4*, 1–85.

Wirth, C., J. Schumacher, and E. Schulze (2004), Generic biomass functions for Norway spruce in Central Europe—A meta-analysis approach toward prediction and uncertainty estimation, *Tree Physiol.*, *24*, 121–139.

S. Kondragunta and X. Zhang, NOAA/NESDIS/ORA, 5200 Auth Road, Camp Springs, MD 20746, USA. (xiaoyang.zhang@noaa.gov; shobha.kondragunta@noaa.gov)