A comparison of lidar, radar, and field measurements of canopy height in pine and hardwood forests of southeastern North America

Joseph O. Sexton a,*, Tyler Bax a, Paul Siqueira b, Jennifer J. Swenson a, Scott Hensley c

a Nicholas School of the Environment, Duke University, Durham, NC 27708-0328, USA
b Microwave Remote Sensing Laboratory, Department of Electrical and Computer Engineering, University of Massachusetts, Amherst, MA 01003-9284, USA
c National Aeronautics and Space Administration Jet Propulsion Laboratory, Pasadena, CA 91109-8099, USA

1. Introduction

Forest canopy structure is a key determinant of forest ecosystem processes, and its measurement is essential for ecosystem monitoring, modeling, and management. As the primary attribute of vertical structure, canopy height affects plant community dynamics and composition (e.g., Welden et al., 1991; Kruger et al., 1997), boundary layer meteorology and microclimate (e.g., Raupach, 2004), and wildlife habitat value (e.g., James, 1971). Horizontal variations in canopy height (e.g., “edges” and “gaps”) also affect plant and animal communities and their dynamics (e.g., Andren and Angelstam, 1988; Matlack, 1994; Didham and Lawton, 1999). Through allometric relationships, canopy height can be used to estimate stand age and foliage-height profile (e.g., Aber, 1979), stand volume (e.g., Cadee et al., 1997), site productivity (e.g., Monserud, 1984; Doolittle, 1979), and biomass (Dubayah and Drake, 2000).

Estimation of carbon content in forest ecosystems is of increasing importance for compliance with international environmental treaties aiming to offset greenhouse gas emissions (Balzter et al., 2007). However, considerable uncertainty surrounds regional stocks of carbon (Schimel et al., 2001), limiting the potential to effectively monitor and manage global carbon and hydrological cycles. These quantitative uncertainties could be greatly reduced by accurate measurement of forest canopy height across broad spatial scales (Balzter et al., 2007).

Canopy height is a fundamental variable in allometric equations that estimate forest biomass and productivity (Pataud et al., 2005; Andersen et al., 2006). Many studies have shown the power of using vegetation height from traditional and remotely sensed measurements in ecosystem models to estimate aboveground
biomass and carbon stocks, from plot to landscape scales (Lefsky et al., 2002a,b; Drake et al., 2002; Hurtt et al., 2004; Balzter et al., 2007). More recently, vegetation height measurements are being combined with ancillary data to develop regional estimates of aboveground biomass in order to generate a cohesive biomass and carbon database on a national scale (e.g., Kellndorfer et al., 2006).

Comparable measurements of canopy height require a consistent definition of the canopy itself, as well as what is measured as its height. Semantic distinctions become important when methods with different sensitivities are employed to measure a structure as complex as a forest canopy. A forest canopy is a complex volume, composed of the stems, branches, leaves, and other solid plant parts, as well as the gaps between them. From this, canopy height refers to the vertical distance from the ground surface to the highest tree part (i.e., non-gap) at that horizontal location. Because of the three-dimensional roughness of tree crowns and the gaps between trees (which often extend to the ground), measurements of height are affected by the horizontal resolution at which data are collected.

Accurate estimates of forest canopy height are therefore dependent on measurements that capture both the horizontal cover and vertical structure of aboveground vegetation (Kellndorfer et al., 2004). Traditional forest inventory methods of measuring tree height with poles or trigonometric transformations of distance and angle measurements are technically simple and inexpensive over small areas, but are difficult to apply in closed stands and impossible to implement over large areas (Andersen et al., 2006). Alternatively, airborne and satellite remotely sensed data provide an economical and efficient means of obtaining height measurements over much larger areas (Naesset, 1997; Dubayah and Drake, 2000; Andersen et al., 2006).

Passive optical remote sensing systems are suited to measure horizontal vegetation cover, and recent advances in active remote sensors have made it possible to acquire reliable measures of vertical forest structure as well (Waring et al., 1995; Treuhaft et al., 1996; Lefsky et al., 2002a; Treuhaft and Siqueira, 2000). Active remote sensors used for forest canopy height estimation include light detection and ranging (lidar) and interferometric synthetic-aperture radar (InSAR) (Balzter et al., 2007; Brown and Sarabandi, 2003). Currently, airborne lidar sensors can estimate height with sub-meter vertical accuracy and spatial resolution (e.g., Brandtberg et al., 2003), but are economically and computationally constrained at broad regional scales (Kellndorfer et al., 2004). In contrast, due to their capacity to systematically scan large areas and record data through cloud cover, airborne and spaceborne InSAR missions have produced an abundance of landscape- and global-scale data that have proven useful in determining canopy height in a number of studies (Balzter et al., 2007; Brown and Sarabandi, 2003; Kellndorfer et al., 2004; Townsend, 2002; Walker et al., 2007).

Lidar remote sensing uses high-frequency pulses, or “posts” of laser light to measure the distance from the sensor to target objects. Lidar sensors are most frequently carried aboard aircraft; and some satellite-based systems have recently been developed for cryosphere applications as well, e.g., ICESat (http://icesat.gsfc.nasa.gov). Calculation of height from lidar requires precise positioning of the sensor by a Differential Global Position System (DGPS) and an Inertial Measurement Unit (IMU) carried aboard the aircraft, and is based on the time elapsed from the round-trip travel of each laser pulse between the sensor and the target. Each of these lidar returns is recorded as a point in three-dimensional space; height above the ground is calculated as the vertical difference between laser returns from the target and those from the ground, or “bare earth” surface.

Lidar sensors vary according to the electromagnetic wavelength of the laser (e.g., 500 nm for bathymetric vs. 900–1064 nm for terrestrial applications), by power, by duration and frequency of the pulse, and by the laser beam’s cross-sectional diameter, determined by its divergence angle and flight altitude (Lefsky et al., 2002a,b). Lidar sensors can be further divided into two types: discrete- (including multiple-) return vs. continuous waveform. Discrete-return lidar provides one or several height measurement, returns per post, and is the more common of the two. These systems can be tailored to specific applications by varying aircraft speed, altitude, pulse rate, and other parameters. Alternatively, continuous-waveform lidar provides measurements of light interception continuously along the vertical axis of each post and can be used to describe the vertical distribution of biomass or foliage within vegetation canopies. Forest characterization using small-footprint, continuous-waveform lidar is an active research field, and discrete return, small footprint lidar has been successfully used for some time to measure tree height within stands (Maclean and Krabill, 1986). While evolving the potential to record information from each post similar to continuous-waveform lidar (Carter et al., 2007), discrete-return lidar methods continue to be refined for many forest types as well as for detailed characterizations such as mapping of individual crowns (e.g., Roberts et al., 2005).

Interferometric SAR uses two radar receivers carried aboard air or spacecraft and separated by a known distance, termed the “baseline”. The two receivers may be two physical units or a single device passed over the target twice. These receivers record the phase of a radar signal reflected from the ground or other targets in the “range”, or “cross-track” direction, perpendicular to the flight-line of the receivers (i.e., the “azimuth” direction). This phase cycles over the travel time of the radar pulse and can be used to measure the round-trip distance between the radar transmitter, target, and receivers. Through knowledge of the observing geometry, the difference between phases of the signal received at the two ends of the baseline can be translated into a topographic height via geometric transformation (Rosen et al., 2000). In the presence of vegetation, the topography measured from interferometric SAR may be anywhere from the ground surface to the top of the canopy, depending on the signal scattering and extinction characteristics of the vegetation. These characteristics vary with radar frequency (and inversely, wavelength)—longer microwaves (30–100 cm and larger, commonly referred to as P-band or UHF) tend to penetrate deeper into the canopy to the ground surface, whereas shorter-wavelength signals, including C-band (~6 cm), X-band (~3 cm), and smaller, tend to reflect nearer the top of the canopy.

Methods for estimating canopy height from SAR vary. Some compare interferometric height with an independent measurement of the ground surface (Kellndorfer et al., 2006; Simard et al., 2006). Others, including polarimetric interferometric synthetic aperture radar (PolInSAR), use both interferometric height and correlation, along with multiple baselines and/or polarizations for estimating vegetation height directly (Cloude and Papathanassiou, 1999; Treuhaft and Siqueira, 2000). A third approach, applied in this study, uses the difference in sensitivity between multiple wavelengths, measuring interferometric heights at two frequencies and calculating height as the difference in elevation between the two measurements (Wheeler and Hensley, 2000).

Each of these methods – lidar, radar interferometry in the various wavelengths, and field measurement – responds uniquely to the physical structure of the forest canopy. Empirical studies are necessary to understand the methods’ sensitivities and comparability for various forest types, and to inform selection among the methods for specific purposes. Between February 2000 and October 2001, four datasets were acquired over the Duke Forest (North Carolina, USA); (1) digital elevation data from the Shuttle Radar Topography Mission (SRTM) C-band radar, (2) X- and P-band
radar interferometric heights from an experimental flight of the airborne Geographic Synthetic Aperture Radar (GeoSAR), (3) airborne small-footprint lidar measurements acquired in leaf-off (winter) conditions, and (4) stand-inventory field measurements. Although initially for various purposes (only the GeoSAR and field measurements were originally intended for measuring canopy height), their nearly simultaneous collection presents a rare opportunity for comparing four of the dominant methods for measuring canopy height.

With this dataset, we conducted three analyses: (1) evaluation of the scaling behavior of lidar-derived canopy heights in pine forests, (2) comparison of canopy height measurements produced by the methods in pine and hardwood forests, and (3) calibration of the globally available SRTM elevation data to extract canopy height for southeastern forests. These analyses illustrate the relationships between canopy height measurements from the four methods in forests characteristic of southeastern North America, provide methods to extract canopy height measurements from available datasets, and facilitate adoption of remotely sensed methods for mapping canopy height at regional scales.

2. Methods

2.1. Study area

The Duke Forest (36°N, 79°W) covers 7050 ha of the Southern Appalachian Piedmont Section of the Southeastern Mixed Forest Province (Bailey, 1995) in central North Carolina, USA (Fig. 1). Forest communities, characteristic of the region, are seral mixes of planted and natural pine (predominantly loblolly pine, *Pinus taeda*) and many species of deciduous hardwoods, including oaks (*Quercus* spp.), hickories (*Carya* spp.), elms (*Ulmus* spp.), red maple (*Acer rubrum*), sweetgum (*Liquidambar styraciflua*), and tulip poplar (*Liriodendron tulipifera*). Mature forest canopies are typically closed and two-storied with pines and hardwoods reaching heights >30 m in the overstory, and shorter hardwoods in the under- and mid-story, but this varies in response to site productivity, timber management, and natural disturbances. Ground elevation ranges from 90 to 210 m, and local relief is gentle to moderate.

Management of the Forest is partitioned into stands of homogeneous forest composition and structure, ranging from 0.9 to 14.7 ha. Data were analyzed at two levels: (1) 112 stands covered by all four data sets (SRTM, lidar, GeoSAR, and field), including four treeless herbaceous stands digitized from orthoimagery to represent the shortest (near-zero) canopy heights in both the pine and hardwood types, and (2) 139 forest inventory plots, including five located in recently cleared stands (with remaining stumps, slash, and debris) to represent zero-height pine and hardwood types.

2.2. Data

2.2.1. Field measurement

A timber inventory of the Duke Forest was conducted in the summer and fall of 2000. Mensuration plot centers were spaced ~400 m apart on a regular sampling grid, and inclusion of trees within plots was determined based on a prism tally. At each plot, trees were selected with a 10 basal-area-factor (BAF) prism pivoted 1.3 m above the ground at the plot center. Border cases were decided by measuring and comparing the distance between the tree and plot center to the tree’s diameter at breast height (DBH); if the distance from the plot center to the tree (in feet) was less than the DBH (in inches) ÷ 2.75, then the tree’s height was measured and recorded. Height was measured with a Haga™ altimeter (http://www.haga-metallwaren.de/en_altimeter.htm) by an observer positioned 66 feet (~20.1 m) from the base of each tallied tree.

An index of pine dominance was used to quantify the contribution of evergreen (i.e., predominantly pine species and occasionally eastern redcedar, *Juniperus virginiana*) canopies to...
each plot’s height measurements. The index was calculated as:

\[
pineDom = \frac{\sum \text{height}_{\text{pine}}}{\sum \text{height}_{\text{all}}} \times 100.
\]  

(1)

where the numerator is the sum of heights for all pines in the tally and the denominator is the sum of heights for all trees in the tally. The index was used to categorize plots into discrete “pine” (pineDom ≥ 90) and “hardwood” (pineDom ≤ 5) types, which were then verified against aerial photos and the 2001 National Landcover Database Landcover Layer (Homer et al., 2004). Field measurements were focused on forest stands, and so did not include height measurements in short vegetation. To represent short vegetation in both pine and hardwood types, zero canopy height was assigned to five plots placed in areas of recently cleared forest, identified in 1998 aerial photos.

2.2.2. Lidar

In response to extensive damage from Hurricane Floyd (1999), the North Carolina Floodplain Mapping Program (NCFMP; http://www.ncfloodmaps.com) was established to update maps used to identify areas vulnerable to flooding. Two private engineering firms were contracted to collect elevation data for the purpose of updating statewide flood insurance maps. One of the firms, 3Di (http://www.3digitel.com/), acquired elevation measurements with small-footprint lidar over the Cape Fear River basin in February and March 2001. Because the primary intent of the program was to map the bare-earth surface, the measurements were acquired in winter, when deciduous hardwood canopies were in dormant, “leaf-off” condition. Measurements were acquired from Datis 1, Datis 2, Optech 1225 and Optech 1210 lidar sensors flying aboard aircraft at 7100–8300 m above sea level. Pulse rates were 45,000–50,000 pulses per second for Datis sensors and 50,000 pulses per second for Optech sensors. Average post-spacing was 4–6 m, beam diameter was 0.5–1.0 m, and swath widths were ~1000 m. Vertical accuracy of lidar measurements was <25 cm RMSE.

Sensors used by NCFMP acquired up to four returns per post. Heights of first and last returns were recorded, but were not differentiated by order (Hope Morgan, NCFMP, personal communication). To estimate canopy height from the undifferentiated returns, we subtracted the value of a bare-earth DEM from the elevation recorded at each return and selected the maximum height difference within pixels at a range of resolutions—from 3 to 30 m (i.e., between the resolutions of GeoSAR and SRTM measurements, or 0.5× to 5× the average lidar posting). The bare-earth DEM used for this and the SRTM canopy measurements was derived from the NCFMP lidar measurements by converting lidar returns representing local minima to a Triangulated Irregular Network (TIN) specifying known break lines, and then interpolating the TIN to raster at 6.096 m (i.e., 20 foot) resolution.

2.2.3. SRTM

In February 2000, the first space-borne, fixed-baseline interferometric synthetic-aperture radar (InSAR) was carried aboard the Space Shuttle Endeavor to compile a digital topographic database of Earth’s surface between 60°N and 57°S (Rabus et al., 2003). Acquisition occurred during winter in the northern hemisphere, when deciduous forest canopies were leafless. The Shuttle Radar Topography Mission (SRTM) used the SIR-C and X-SAR radar antennas to collect information in C-band (5.6 cm, 5.3 GHz) and X-band (3.1 cm, 9.6 GHz) wavelengths. The C-band data were processed at the NASA Jet Propulsion Laboratory (JPL) and released to the public at a resolution of 1-arc second (~30 m) within the United States and 3 arc seconds (~90 m) elsewhere (Hensley et al., 2000). Absolute vertical accuracy for the C-band DEM was specified as <16 m globally, with relative accuracies of <6 m in any 225 km × 225 km area (Rabus et al., 2003). Because of the short wavelength of the SRTM C-band sensor, the height response over vegetated terrain is influenced by integrated scattering from leaves, branches, and stems, measuring a phase center height that is higher than the underlying bare-earth surface in vegetated areas.

Vegetation canopy height can be derived from the SRTM DEM by subtracting from it another DEM that represents the ground elevation (Brown and Sarabandi, 2003; Kellndorfer et al., 2004; Dubayah et al., 2007). The method assumes negligible canopy penetration by the SRTM C-band radar—an assumption that is rarely met in vegetation canopies, and so calibration of the elevation differences to reference canopy heights is necessary (Kellndorfer et al. (2004). Whereas Kellndorfer et al. (2004) used the National Elevation Dataset (NED) as the bare-earth DEM, we subtracted the more precise bare-earth DEM derived from lidar from the SRTM DEM to improve comparability with our fully lidar-derived canopy height measurements.

2.2.4. GeoSAR

The GeoSAR instrument was initially constructed at NASA’s Jet Propulsion Laboratory (JPL), and is currently operated by Fugro, EarthData Inc. (http://www.fugroearthdata.com), a global geosciences mapping and remote sensing company. GeoSAR flies aboard a Gulfstream II aircraft at a nominal altitude of 10 km and consists of twelve autonomous interferometric systems—three on P-band (86 cm wavelength) and one on X-band (3 cm wavelength)—to cover a 10 km swath on either side. GeoSAR records several measurements (backscatter power, interferometric height, and interferometric correlation) for each of the two bands at 3 m resolution, with 1 m and 1–4 m vertical accuracy for X- and P-band DEMs, respectively.

In its early development stages, GeoSAR collected science data over a number of regions in the United States, including the Duke Forest and surrounding area on October 11, 2001. For this area, we calculated canopy height as the difference between X- and P-band interferometric heights. Due to low signal-to-noise ratios in the P-band in sparsely vegetated fields, missing height measurements in these areas were filled with lidar measurements for calibrating SRTM data (Section 2.3.3).

2.3. Analysis

2.3.1. Lidar scaling

We selected four pine stands to assess the effect of horizontal resolution on lidar-based canopy height measurement. The case studies represented stands with: (1) dense, homogeneous canopy cover; (2) small gaps; (3) narrow, linear gaps between planted rows; and (4) large, irregularly shaped gaps. Each stand boundary was internally buffered by 30 m to avoid errors due to spatial misregistration between datasets. Lidar heights were rasterized at resolutions equal to 0.5-, 1-, 2-, 3-, 4-, and 5-times the average posting (~6 m) and extracted from within the buffers. Pixel sizes therefore equaled the 3-m resolution of GeoSAR, the 6-m original lidar posting, and ranged up to the 30-m resolution of SRTM in increments of 6 m. Relative frequency histograms across the range of resolution were overlaid on one another to track changes in the distribution of measured height. Measures of central tendency (mean, median, and mode) were also calculated for each stand at each resolution and plotted over resolution.

2.3.2. Measurement comparison

Relative frequency distributions from all measurements were compared graphically for six representative hardwood stands and five representative pine stands. Results are shown using one stand of each type. Inventory plots were then divided into pine and
We calibrated the SRTM height measurements by least-squares regression, fitting functions of form:

\[ H = a + b H_{\text{SRTM}} \]  

and

\[ H = a + b H_{\text{SRTM}} + c (H_{\text{SRTM}})^{1/3}, \]  

where \( H \) is canopy height, \( a \) and \( b \) are the intercept and slope of a linear relationship between the SRTM C band (minus bare-earth DEM) measurements and \( H \), and \( c \) fits the cube-root-transformation of SRTM-based canopy height to model the effect of canopy penetration by the C band over increasing height. Negative or zero SRTM heights were recoded to an arbitrarily small, positive value (0.0001) to accommodate the cube-root transformation, and separate regressions were fit for pine and hardwood types, with lidar used as reference in pine and field data in hardwood plots.

3. Results

3.1. Lidar scaling

Lidar-based rasters of canopy height in pine stands were sensitive to the area over which returns were aggregated (Fig. 2). At a pixel-resolution equal to half the average lidar posting (3 m), the within-stand distributions of canopy height were heavily influenced by ground returns. This effect was greatest in sparsely canopied stands and weakest in densely canopied stands. The effect diminished quickly as resolution was coarsened beyond the average posting and was negligible in every stand at 3 \( \times \) the average posting (18 m). Whereas the ground effect was visible in the means, medians, and modes of the within-stand distributions, the mean was impacted most, with medians and modes stabilizing at relatively finer scales.

The scale at which the measures stabilized was impacted by horizontal variations in canopy cover, with a stand exhibiting large gaps (DU-39-18) showing a greater ground effect at fine scales than the more homogeneous stands (Fig. 3). A nearly closed-canopy stand with narrow (<5 m) linear gaps between planted rows (DU-39-18) was affected by ground returns more similarly to homogeneous, closed-canopy stands than to the stand with large, irregularly shaped gaps. Occasional tall trees – visible as secondary modes in the height distribution (Fig. 3) – also affected stand-level estimates, by increasing the estimates at coarse resolution. At both fine and coarse resolutions, the mode was least affected by heterogeneity, followed by the median, and then the mean.

In all stands, a range of resolutions existed in which the mean, median, and mode were similar. For homogeneous stands, this range was from 2 \( \times \) to 4 \( \times \) the lidar posting (i.e., 12–24 m resolution). In heterogeneous stands, the range was narrower, with canopy gaps increasing the minimum and emergent crowns decreasing the maximum of the range. Across all stands in this study, the mean, median, and mode were most similar at 3 \( \times \) the original posting (18 m).

3.2. Measurement comparisons

Each of the four methods produced internally consistent, interpretable estimates of canopy height when examined graphically, but to a lesser degree for SRTM in hardwoods (Figs. 3 and 4). However, direct (i.e., uncalibrated) comparison among methods was variable (Fig. 5). Both stand- and plot-level height distributions were more comparable overall in pine than in hardwood stands, with greater visual similarity between distributions (Fig. 4), approximately 1–2 m smaller RMSE and MAE (Table 1), and approximately 1.5–3 m smaller, more constant biases (Table 2) in

hardwood types (\( n = 30 \) and 50, respectively) based on the field-measured pine dominance index. Means of field-measured tree heights and lidar, GeoSAR, and SRTM rasters were extracted from within 30-m-radius buffers centered on each plot. Measurements were compared via their plot means following Willmott (1982), interpreting the term “error” simply as “difference” to avoid misrepresenting any of the measurements as truth a priori. The mean bias:

\[ B = N^{-1} \sum_{i=1}^{N} (x_i - y_i) \]  

quantifies the average vertical offset of one measurement (\( x \)) over another (\( y \)) across \( N \) plots. The standard deviation of bias:

\[ s_B = \left( \frac{(N-1)^{-1} \sum_{i=1}^{N} (x_i - y_i - B)^2}{N} \right)^{0.5} \]  

quantifies the variability around the bias. The root-mean-squared error:

\[ \text{RMSE} = \left( \frac{1}{N} \sum_{i=1}^{N} (x_i - y_i)^2 \right)^{0.5} \]  

and mean absolute error

\[ \text{MAE} = N^{-1} \sum_{i=1}^{N} |x_i - y_i| \]  

quantify the magnitude of difference between measurements; whereas MAE better represents the average magnitude of differences between measurements, RMSE is more sensitive to large differences and is included for comparison to previous studies. Willmott (1982) also provides metrics describing “systematic” and “unsystematic” differences between two measurements, \( x \) and \( y \). The systematic error:

\[ \text{MSE}_x = N^{-1} \sum_{i=1}^{N} (\hat{x}_i - x_i)^2 \]  

is the mean squared difference between the measurements \( x_i \) and their estimated values based on another measurement (\( y_i \)) that has been calibrated to \( x \) via a linear model \( \hat{x} = f(y|x) \). MSE quantifies the error remaining between two measurements after linear calibration. The unsystematic error:

\[ \text{MSE}_u = N^{-1} \sum_{i=1}^{N} (\hat{y}_i - y_i)^2 \]  

is the mean “residual” error between the original measurements \( y_i \) and those that have been calibrated to \( x \). MSE quantifies the variation in \( y \) removed by linear calibration to \( x \). The two components MSE and MSEu sum to MSE, with MSE \( \approx 0 \) and MSEu approaching MSE for a well-specified, well-fit model (Willmott 1982). We report their square roots (RMSE and RMSEu), which are in units of the response (m).

2.3.3. SRTM calibration

Stand canopy-height averages from each method were divided into pine and hardwood classes based on the 2001 National Land Cover Database (NLCD 2001), taking the landcover value of the nearest pixel to the plot center. To represent zero-height stands of each type, the four herbaceous stands were included in both hardwood and pine classes. After removing one outlier from the pine subset, the final sample included 37 pine and 84 hardwood plots.
pines than in hardwoods. Linear calibration between measurements was also more effective in reducing differences between measurements in pines than in hardwoods (Table 3).

In both pine and hardwood forests, field and lidar methods provided the tallest canopy height measurements, with lidar measurements slightly taller than those collected in the field. In order of decreasing height, these were followed by GeoSAR and then SRTM (Table 2). Because of this sorting, GeoSAR exhibited the greatest direct (i.e., uncalibrated) comparability to other measurements, followed by lidar, field, and SRTM (Table 1). Lidar

Fig. 2. Scaling characteristics of stand-level canopy height measurements (mean, median, and mode) across a range of pixel resolutions in pine stands of different structural characteristics: narrow, linear canopy gaps between planted rows (DU-33-10); large, irregularly shaped gaps (DU-39-18); small gaps (KO-04-04); and dense, homogeneous cover (KO-06-05).

Fig. 3. Effects of varying aggregation of lidar returns on within-stand distributions of canopy heights measured in pine stands representing: medium-height, patchy canopy cover (DU-39-18) and tall, homogeneous canopy cover (KO-04-04).
measurements were most directly comparable to field data, but this could only be evaluated in pine forests. Due to its negative bias and imprecise relationships to other methods, SRTM measurements were the most weakly comparable overall.

In terms of improvements measured by $R^2$, linear calibrations were most effective when based upon lidar (Table 3), with all other measurements serving similarly to one another in this regard. However, remaining “systematic” errors (RMSEs) were smallest when calibrations were based upon GeoSAR, followed by field data, lidar, and then SRTM. Likely due to shared sensitivity to within-canopy variation, linear calibration to GeoSAR provided the greatest overall reduction of RMSE in SRTM-based canopy height measurements.

3.2.1. SRTM calibration

SRTM exhibited a delayed response to canopy height in pine plots (Fig. 6), with reference (lidar) height rising more steeply than SRTM measurements in short stands where the ground surface

![Fig. 4. Comparison of within-stand distributions of canopy height measurements in a representative pine stand (DU-33-10) and a representative hardwood stand (EN-19-09).](image)

![Fig. 5. Comparison of field-measured and remotely sensed canopy heights among pine and hardwood plots.](image)
dominated scattering of the C band. In taller pine canopies (>10 m), where the signal’s penetration to the ground was reduced, SRTM increased in approximate 1:1 proportion to canopy height. The linear model accommodated for this pattern with a steep slope (1.48) and a high intercept (8.93). The cube-root model followed the data more consistently over its range, with an excessively shallow slope in the tallest stands.

The greatest difference between linear and cube-root calibrations occurred in the extremes of both pine and hardwood canopy height. Whereas the linear transformations over-predicted canopy height, the logarithmic corrections produced more even and accurate height estimates across the lower range of canopy heights, although the bias was not completely removed in hardwoods (Tables 4 and 5). At the high end of the range of canopy heights, the equations diverged as well, with linear functions producing higher estimates than cube-root calibrations. The effect was extreme in hardwood stands.

from the cube-root term were slight in hardwoods (Tables 4 and 5). Also in opposition to pine canopies, the cube-root term tended to over-correct the SRTM data in tall hardwood stands, forcing an excessively shallow slope in the tallest stands.

Fig. 6. Calibration of SRTM-derived canopy height based on lidar and GeoSAR measurements in pine (n = 48) and hardwood (n = 84) plots. Calibrated curves are overlaid on reference values measured with lidar in pine plots and by field methods (Haga alimeter) in hardwoods plots.

Table 3
Linear regression summaries for plot-level pine (n = 30) and hardwood (n = 50) canopy height measurements. MSE is mean “unsystematic”, or “residual” error remaining between calibrated and reference measurements (see text for full explanation).

<table>
<thead>
<tr>
<th>Model</th>
<th>Slope</th>
<th>Intercept</th>
<th>$R^2$</th>
<th>RMSE$_u$</th>
<th>RMSE$_s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lidar ~ Field</td>
<td>0.856</td>
<td>4.409</td>
<td>0.83</td>
<td>4.18</td>
<td>2.22</td>
</tr>
<tr>
<td>SRTM ~ Field</td>
<td>0.451</td>
<td>5.577</td>
<td>0.70</td>
<td>3.13</td>
<td>7.79</td>
</tr>
<tr>
<td>GeoSAR ~ Field</td>
<td>0.471</td>
<td>−1.845</td>
<td>0.54</td>
<td>4.67</td>
<td>13.42</td>
</tr>
<tr>
<td>SRTM ~ Lidar</td>
<td>0.534</td>
<td>3.106</td>
<td>0.87</td>
<td>2.05</td>
<td>8.20</td>
</tr>
<tr>
<td>SRTM ~ GeoSAR</td>
<td>0.570</td>
<td>−4.692</td>
<td>0.70</td>
<td>3.78</td>
<td>14.43</td>
</tr>
<tr>
<td>SRTM ~ GeoSAR</td>
<td>1.024</td>
<td>−7.384</td>
<td>0.73</td>
<td>3.53</td>
<td>7.04</td>
</tr>
</tbody>
</table>

Table 4
Regression summaries for plot-level calibration of SRTM- on lidar- and field-based canopy height in NLCD 2001 pine and hardwood types.

<table>
<thead>
<tr>
<th>Model</th>
<th>a</th>
<th>b</th>
<th>$R^2$</th>
<th>S.E. (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pine (n = 37)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H = a + bH_{Lidar}$</td>
<td>8.93</td>
<td>1.48</td>
<td>0.75</td>
<td>5.05</td>
</tr>
<tr>
<td>$H = a + bH_{Lidar} \cdot (H_{Lidar})^{1/3}$</td>
<td>0.435</td>
<td>9.92</td>
<td>0.83</td>
<td>3.99</td>
</tr>
<tr>
<td>Pine (n = 84)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H = a + bH_{Lidar}$</td>
<td>17.14</td>
<td>0.79</td>
<td>0.19</td>
<td>6.73</td>
</tr>
<tr>
<td>$H = a + bH_{Lidar} \cdot (H_{Lidar})^{1/3}$</td>
<td>10.92</td>
<td>6.67</td>
<td>0.23</td>
<td>6.48</td>
</tr>
</tbody>
</table>

Table 2
Mean bias (B) between paired averages of canopy height measurements within 30-m radius plots. Values for pine plots are in the upper-right half of the matrix and values for hardwood plots are in the lower left. Biases between pairs of measurements (e.g., field vs. GeoSAR) are reported as the difference of the first element of the pair along the diagonal over the second—e.g., the B (field, GeoSAR) is reported as height (field) – height (GeoSAR). Standard deviations of bias are in parentheses.

<table>
<thead>
<tr>
<th>Field</th>
<th>Lidar</th>
<th>GeoSAR</th>
<th>SRTM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pine</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.73</td>
<td>8.40</td>
<td>7.84</td>
<td>9.85</td>
</tr>
<tr>
<td>9.54</td>
<td>8.45</td>
<td>7.88</td>
<td>9.85</td>
</tr>
<tr>
<td>16.84</td>
<td>8.14</td>
<td>7.84</td>
<td>9.85</td>
</tr>
</tbody>
</table>

| Hardwood |       |        |      |
| GeoSAR   | 0.856 | 4.409  | 0.83 |
| SRTM     | 0.451 | 5.577  | 0.70 |
| GeoSAR   | 0.471 | −1.845 | 0.54 |
| SRTM     | 0.534 | 3.106  | 0.87 |
| GeoSAR   | 0.570 | −4.692 | 0.70 |
| SRTM     | 1.024 | −7.384 | 0.73 |

Table 1
Paired RMSE and MAE (in parentheses) between averages of canopy height measurements (m) in 30-m radius plots. Values for pine plots are in the upper-right half of the matrix and values for hardwood plots are in the lower left.

<table>
<thead>
<tr>
<th>Field</th>
<th>Lidar</th>
<th>GeoSAR</th>
<th>SRTM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pine</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.73</td>
<td>8.40</td>
<td>7.84</td>
<td>9.85</td>
</tr>
<tr>
<td>9.54</td>
<td>8.45</td>
<td>7.88</td>
<td>9.85</td>
</tr>
<tr>
<td>16.84</td>
<td>8.14</td>
<td>7.84</td>
<td>9.85</td>
</tr>
</tbody>
</table>

| Hardwood |       |        |      |
| GeoSAR   | 0.856 | 4.409  | 0.83 |
| SRTM     | 0.451 | 5.577  | 0.70 |
| GeoSAR   | 0.471 | −1.845 | 0.54 |
| SRTM     | 0.534 | 3.106  | 0.87 |
| GeoSAR   | 0.570 | −4.692 | 0.70 |
| SRTM     | 1.024 | −7.384 | 0.73 |
Ground signals were dominant in leafless stands and were not fully predominant in hardwood stands. In preliminary analyses, the sub-canopy and forest types (Dubayah and Drake, 2000).

Spatial heterogeneity of the canopy and lidar posting so that data measurements of fine-scale (i.e., sub-pixel) canopy heterogeneity. Ongoing studies are needed to show the interaction between the size and shape of canopy gaps, and so the scaling behavior of lidar measurements also has the potential to provide ancillary information about canopy height, including those of SRTM and GeoSAR.

Canopy height measured by lidar is sensitive to the scale at which the data are collected and aggregated into raster pixels, revealing an interaction between canopy heterogeneity and data resolution. By increasing the number of returns from which per-pixel maxima are calculated, coarsening raster resolution removes the effect of gaps on canopy height measurement. This effect varies with the size and shape of canopy gaps, and so the scaling behavior of lidar measurements also has the potential to provide ancillary measurements of fine-scale (i.e., sub-pixel) canopy heterogeneity. Ongoing studies are needed to show the interaction between spatial heterogeneity of the canopy and lidar posting so that data acquisition and processing parameters can be tuned to various forest types (Dubayah and Drake, 2000).

The winter acquisition of our lidar data restricted its utility in hardwood stands. In preliminary analyses, the sub-canopy and ground signals were dominant in leafless stands and were not fully removed even at 30-m resolution. High-density (~12 returns m⁻²), small-footprint lidar acquired in winter has been used to analyze vertical structure of deciduous trees at fine scales (Brandtberg et al., 2003), suggesting that regional characterizations are also possible at far coarser resolution than those studied here. However, as resolution was coarsened, hardwood pixels became increasingly contaminated by pine cover, which overshadowed the weak signal from the bare hardwood canopy. Because commercial lidar sensors typically operate at near-infrared wavelengths, the need for data filtering could possibly be met by using the reflection intensity of each return to discriminate returns from leaves vs. those from branches etc. (Lillesand et al., 2008).

Our lidar dataset was originally developed for the purpose of mapping bare-earth elevation, but our results also encourage further study and use of this statewide dataset for mapping structural characteristics of evergreen vegetation in North Carolina. Following national guidelines (FEMA, 2003), similar lidar mapping programs are being implemented in other states as well. The increasing coverage of these datasets shows great promise for mapping and studying evergreen canopy structure, biomass, growth, and habitat value at regional scales.

### 4.1. Measurement of canopy height by lidar

Our results corroborate studies confirming discrete-return lidar’s reliability for measuring canopy height in coniferous forests (Naesset, 1997; Dubayah and Drake, 2000; Means et al., 2000; Lefsky et al., 2002a,b; Roberts et al., 2005). Stand-level distributions of lidar height maxima provided intuitive representations of the vertical structure of pure loblolly pine stands of different sizes, including negligible mid-story cover and loss of apical dominance with increasing height (Fig. 2). Canopy heights measured by lidar also had greater accuracy and fewer systematic errors with respect to field measurement in pine forests than methods based on radar interferometry. Further, judged by their increased precision relative to field data (Table 3), lidar measurements are preferable over field data for calibrating interferometric measurements of canopy height, including those of SRTM and GeoSAR.

Canopy height measured by lidar is sensitive to the scale at which the data are collected and aggregated into raster pixels, revealing an interaction between canopy heterogeneity and data resolution. By increasing the number of returns from which per-pixel maxima are calculated, coarsening raster resolution removes the effect of gaps on canopy height measurement. This effect varies with the size and shape of canopy gaps, and so the scaling behavior of lidar measurements also has the potential to provide ancillary measurements of fine-scale (i.e., sub-pixel) canopy heterogeneity. Ongoing studies are needed to show the interaction between spatial heterogeneity of the canopy and lidar posting so that data acquisition and processing parameters can be tuned to various forest types (Dubayah and Drake, 2000).

The winter acquisition of our lidar data restricted its utility in hardwood stands. In preliminary analyses, the sub-canopy and ground signals were dominant in leafless stands and were not fully removed even at 30-m resolution. High-density (~12 returns m⁻²), small-footprint lidar acquired in winter has been used to analyze vertical structure of deciduous trees at fine scales (Brandtberg et al., 2003), suggesting that regional characterizations are also possible at far coarser resolution than those studied here. However, as resolution was coarsened, hardwood pixels became increasingly contaminated by pine cover, which overshadowed the weak signal from the bare hardwood canopy. Because commercial lidar sensors typically operate at near-infrared wavelengths, the need for data filtering could possibly be met by using the reflection intensity of each return to discriminate returns from leaves vs. those from branches etc. (Lillesand et al., 2008).

Our lidar dataset was originally developed for the purpose of mapping bare-earth elevation, but our results also encourage further study and use of this statewide dataset for mapping structural characteristics of evergreen vegetation in North Carolina. Following national guidelines (FEMA, 2003), similar lidar mapping programs are being implemented in other states as well. The increasing coverage of these datasets shows great promise for mapping and studying evergreen canopy structure, biomass, growth, and habitat value at regional scales.

### 4.2. Measurement of canopy height by Interferometric Synthetic Aperture Radar (InSAR)

Scattering of microwave energy is dependent on the size and arrangement of target objects, with strong signals returned to the sensor from objects with length approximately equal to the wavelength of radiation (Lillesand et al., 2008). Objects much larger than the wavelength do not transmit, but instead tend to reflect the signal away from the sensor. Within a forest canopy, short-wavelength (e.g., X-band) signals are dominated by leaves and twigs, intermediate wavelengths (e.g., C-band) by medium-sized branches, and long wavelengths (e.g., P-band) by large-diameter trunks and the ground surface. Due to the vertical distribution of canopy elements, a radar signal’s penetration into a forest canopy is likewise proportional to its wavelength, resulting in scattering heights of X, C, and P bands sorting vertically from top to bottom in the canopy (Jensen, 2000).

This leads to a sorting of elevation values measured by interferometry based on the different wavelengths, with heights measured by the X band representing nearly the top of the canopy, those from the P band representing the bare earth, and from the C-band intermediate between these two. By pairing a long-wavelength with a short-wavelength band, canopy height is estimated by subtracting bare-earth (BE) from top-of-canopy

---

**Table 5**

Validation of SRTM calibration for canopy height in pine and hardwood plots. Plots were stratified based on NLCD 2001 landcover types.

<table>
<thead>
<tr>
<th>Model</th>
<th>B (sd)</th>
<th>MAE</th>
<th>RMSE</th>
<th>RMSE_ε</th>
<th>RMSE_u</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pine (n = 37)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( H = a + bH_{\text{SRTM}} )</td>
<td>0.00 (0.00)</td>
<td>4.20</td>
<td>4.94</td>
<td>4.94</td>
<td></td>
</tr>
<tr>
<td>( H = a + bH_{\text{SRTM}} + c(H_{\text{SRTM}})^{1/3} )</td>
<td>0.15 (2.10)</td>
<td>3.11</td>
<td>3.90</td>
<td>0.47</td>
<td>2.10</td>
</tr>
<tr>
<td>Hardwood (n = 83)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( H = a + bH_{\text{SRTM}} )</td>
<td>0.00 (0.00)</td>
<td>5.13</td>
<td>6.65</td>
<td>0.00</td>
<td>6.65</td>
</tr>
<tr>
<td>( H = a + bH_{\text{SRTM}} + c(H_{\text{SRTM}})^{1/3} )</td>
<td>0.00 (0.00)</td>
<td>4.94</td>
<td>6.41</td>
<td>0.00</td>
<td>6.41</td>
</tr>
</tbody>
</table>

---

**Fig. 7.** Comparison of raw and calibrated SRTM measurements to reference height measurements in pine and hardwood plots. Reference heights in pine plots are from lidar measurements, and those in hardwood plots are from field measurements. Calibrated measurements are from a cubic fit in pine plots and a linear fit in hardwood plots.
than in hardwoods, suggesting that structural diversity plays a large role in measurement precision. SRTM’s precision was also likely affected, but this effect was overshadowed by the data’s dormant-season acquisition. Whether to avoid noise from structural or other variations, frequent calibration will likely improve precision when using InSAR to measure canopy height, especially over large, heterogeneous areas.

### 4.3. Calibration of SRTM for canopy height measurement

Uncalibrated differences between the SRTM DEM and bare-earth elevation measurements are known to underestimate canopy height of various forest types due to C-band radar’s penetration of the vegetation canopy (Kellndorfer et al., 2004). We also observed this bias in southeastern pine and hardwood forests, noting a substantially larger, more variable bias in dormant hardwood canopies. The increased bias in hardwood canopies shows that the effect of canopy penetration is exacerbated by the leafless condition of deciduous forests at the time of data acquisition.

Although uncalibrated SRTM-derived canopy heights underestimated true canopy height, a consistent physical relationship between SRTM measurements and canopy height allowed calibration of SRTM to more accurate measurements in pine stands. In pine forests, this relationship between canopy height ($H$) and the difference between the SRTM and a bare-earth DEM ($H_{SRTM}$) can be represented as:

$$ H = 0.435 \times H_{SRTM} + 9.92 \times (H_{SRTM})^{1/3}. $$

Unfortunately, although errors in both pine hardwood forests were reduced substantially by calibration (from 14.92 to 3.90 m in pine and from 16.84 to ~6.5 m in hardwood), such simple calibration does not appear nearly as effective in dormant deciduous as in fully canopied forests. Disagreement between calibrated SRTM- and lidar-derived canopy height in this study is greater than that between SRTM and field measurement in managed pine plantations on flat topography after outlier removal (Kellndorfer et al., 2004), but is more comparable to errors in mixed coniferous forest in the Sierra Nevada Mountains in Northern California (Dubayah et al., 2007). The picture emerging from these regional studies is that both forest structural complexity and underlying bare-earth topography effect the accuracy of SRTM-based canopy height measurements (Bourjine and Baghdadi, 2005; Bhang et al., 2007), but that calibration is possible in many situations.

In southeastern pine or structurally similar forests for which a reliable bare-earth DEM is available, linear or cube-root calibration of SRTM measurements substantially reduces vertical errors of SRTM-based canopy height estimates. Previous studies have recognized the potential for correcting SRTM underestimates of canopy height through simple linear regression against reference height measurements (Kellndorfer et al., 2004; Dubayah et al., 2007). Our addition of a cube-root term based on the effect of canopy penetration is important to prevent over-predicting the height of the shortest stands. However, in regions with a mixture of extremely tall and very short stands, or where bare-earth DEMs might be influenced by the forest canopy, calibrations will need to incorporate canopy penetration artifacts from both the SRTM (as top-of-canopy) and bare-earth DEMs. Canopy-height underestimation by SRTM data is a function of vegetation structure and will likely vary if vegetation varies considerably within an area of interest (Bhang et al., 2007; Dubayah et al., 2007). It would be beneficial for future efforts to focus on fitting context-dependent corrections to more accurately map vegetation height across large heterogeneous landscapes with SRTM or other C-band data.

---

**Fig. 8.** Schematic representation of the idealized response of X-, C-, and P-band radar interferometric heights to “true” canopy height; and the resulting canopy-sensitive ranges of GeoSAR and SRTM measurements. Axes are not drawn to scale.
4.4. Selection among methods

Consideration of sensitivity, scale, and comparability should inform selection of methods for estimating canopy height. Due to its expediency and low cost, in situ field measurement remains the best choice for measuring individual or small numbers of tree heights. However, field measurement requires extrapolation from small samples and is not capable of mapping canopy height over any but the smallest stands. Remote sensing provides an economically efficient means to map larger regions, with the various methods sorting along an axis of scale. Due to its fine resolution and customizable post-spacing, lidar is best suited at scales from several to several million hectares and is useful for calibrating broader-scale InSAR measurements. With larger swathes and systematic processing, airborne InSAR sensors such as GeoSAR achieve economy of scale for larger regions—from hundreds of thousands to millions of hectares. Especially in tall pine forests, GeoSAR’s precision should also support increased accuracy through calibration to lidar measurements. At the coarsest scales, an increasing number of spaceborne InSAR sensors—including the European ERS-2 and Envisat, the Japanese PALSAR, the German TerraSAR-X, the Italian COSMO SkyMed constellation, and Canadian RADARSAT series—promises the capacity to monitor forest resources at the global scale in the near future (Lillesand et al., 2008). The SRTM dataset is currently the only potential source of tree canopy height for the entire Earth. Although the SRTM DEM was collected only in 2001 and requires calibration with a separate bare-earth DEM to extract canopy height, it has the potential to provide a baseline measurement against which future measurements can be compared to measure forest growth. For this, calibrations stratified by ecoregion forest types should be conducted wherever circa-2001 canopy height datasets are available.

5. Conclusions

Forest canopies are complex volumes, and several methods—each with characteristic sensitivities, biases, and limitations of scale—are available to measure their height. Due to simpler structure and phenology, pine forests allow greater precision of height measurements than do hardwood forests in the southeastern United States. In southeastern pine forests, lidar measurements acquired in the dormant season have the highest precision of all methods studied, followed by those from field measurements and interferometric synthetic aperture radar (InSAR). Raster canopy-height surfaces produced from lidar returns are sensitive to the resolution at which they are aggregated. InSAR-derived canopy height measurements are subject to biases from factors including variable canopy penetration, but show potential for calibration based on lidar. X- and P-band interferometric radar measurements from airborne sensors (e.g., GeoSAR) provide more precise, accurate measurements than C-band data from spaceborne sensors (i.e., SIR-C/SRTM), but even the SRTM DEM can be calibrated to estimate canopy height across broad regions. Each of the remotely sensed methods studied produces reasonable and consistent depictions of canopy height that can be compared with data of similar provenance, but due to differences in underlying sensitivities between the methods, comparisons between measurements from various sources require cross-calibration and will be most useful at broad scales.

Acknowledgements

Lidar data were provided by the North Carolina Floodplain Mapping Program. Forest inventory measurements were provided by the Office of the Duke Forest, Duke University. John Kerkering performed and reported analyses of a pilot study, and John P. Fay assisted lidar data processing. Analyses were performed in the Duke University Landscape Ecology Laboratory, with technical assistance from Dean Urban, Ben Best and Ibrahim Almadidene. This research was supported by a NASA Earth System Science Fellowship, “Suburban forest dynamics: fusing remote sensing and ecological models” (J.O. Sexton, 2005–2008). Two anonymous reviewers provided suggestions that improved the manuscript greatly.

References


