Available online at www.sciencedirect.com
SCIENCE DIRECT
Remote Sensing of Environment 102 (2006) 63 – 73
Remote Sensing of Environment
www.elsevier.com/locate/rse

Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy

Peter Hyde [a,*], Ralph Dubayah [a], Wayne Walker [b], J. Bryan Blair [c], Michelle Hofton [a], Carolyn Hunsaker [d]

[a] Department of Geography, University of Maryland, College Park, Maryland 20742, United States
[b] Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, Michigan 48109, United States
[c] Laboratory for Terrestrial Physics, NASA's Goddard Space Flight Center, Greenbelt, Maryland 20771, United States
[d] U.S. Forest Service, Department of Agriculture, Fresno, California 93729, United States

Received 9 September 2005; received in revised form 24 January 2006; accepted 30 January 2006

Abstract

Measurements of forest structure are important for wildlife habitat management. An optimal strategy for mapping forest structure would include detailed measurements of the vertical dimension, which are traditionally provided by field sampling, together with the broad spatial coverage afforded by remote sensing. While no single sensor is capable of delivering this at the present time, it should be possible to combine information from multiple sensors to achieve a reasonable approximation. In this study, we compare estimates of forest structural metrics derived from remote sensing to measurements obtained in the field (large tree maximum canopy height, mean canopy height, standard deviation canopy height, and biomass). We then statistically combine structural information from LiDAR, RaDAR, and passive optical sensors in an attempt to improve accuracy of our estimates. The results of this study indicate that LiDAR is the best single sensor for estimating canopy height and biomass. The addition of ETM+ metrics significantly improved LiDAR estimates of large tree structure, while Quickbird and InSAR/SAR improved estimates either marginally or not at all. The combination of all sensors was more accurate than LiDAR alone, but only marginally better than the combination of LiDAR and ETM+. Structure metrics from LiDAR and RaDAR are essentially redundant, as are ETM+ and Quickbird.

Keywords: LiDAR; RaDAR; SAR; InSAR; Fusion; Habitat; Forest structure; Canopy height; Biomass

1. Introduction

Measurements of forest structure are critical for many applications, including wildlife management and biodiversity studies, fire modeling, and carbon stock estimation. Canopy height and associated metrics of vertical heterogeneity (North et al., 1999), when considered together with site characteristics, are indicators of old-growth forest conditions and thus are of interest to researchers studying old-growth endemics. Canopy height is an important input for ecosystem and fire models and is highly correlated with biomass. Biomass is a key component of the carbon cycle, as forests represent large carbon sources and sinks (Skole & Tucker, 1993), and is also a surrogate for fuel loading estimation (Finney, 1998). Large trees, in particular, may provide essential habitat to California spotted owls (North et al., 1999) and are an important component of aboveground biomass.

Traditionally, these attributes have been measured in the field using hand-held equipment. Field-based methods can be highly accurate but are time-consuming and thus are typically limited in scope to either mapping at fine scales or sampling at the landscape scale. Multispectral (Hyyppä et al., 1998) and hyperspectral remote sensing (Pu & Gong, 2004) have been used to map structural metrics at moderate resolution and broad scales. However, passive optical sensors have difficulty penetrating beyond upper canopy layers (Weishampel et al., 2000) and are better suited for mapping horizontal structure, e.g., land cover type. Interferometric synthetic aperture radar (InSAR) can provide measures of vertical structure at

* Corresponding author.
E-mail address: phyde@geog.umd.edu (P. Hyde).

0034-4257/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2006.01.021

landscape scales at varying degrees of accuracy; however, at the present time these are best suited for structurally homogeneous forest types (Treuhaft & Cloude, 1999; Treuhaft & Siqueira, 2000). Full waveform-digitizing, large footprint LiDAR provides highly accurate measurements of forest structure at the footprint level of observation (Nelson et al., 1984, 1988; Nilsson, 1996; Lefsky et al., 1999a; Drake et al., 2002; Hyde et al., 2005); however, they are not capable of imaging entire landscapes. Due to the high cost of flight time, the need to limit scanning to near nadir in order to prevent ranging errors, and the presence of coverage gaps due to aircraft pitch and roll, a typical large footprint LiDAR mission acquires samples (albeit at a high frequency) instead of the wall-to-wall coverage provided by other sensors, such as RaDAR or passive optical sensors.

The optimal strategy for mapping forest structure would include the finely detailed measurements of the vertical dimension that field sampling provides as well as the broad spatial coverage of remote sensing. Although no single technology is currently capable of providing this level of forest structural information, advancements in InSAR and LiDAR will likely lead to broad-scale mapping of vertical structure in the near future. In the meantime, it is possible to map forest structure at intermediate scales by statistically combining or fusing information from multiple sensors to take advantage of the highly detailed vertical measurements provided by full waveform-digitizing LiDAR, the broad-scale mapping capabilities of passive optical sensors, and the coarse sensitivity to horizontal and vertical structure afforded by InSAR. Combining information from multiple sensors, or data fusion, has yielded promising results for the estimation of forest structural characteristics (Wulder et al., 2004). Hudak et al. (2002) combined regression and co-kriging models from LiDAR and multispectral data; the results were more accurate than either data set alone. Wulder and Seeman (2003) used texture metrics from Landsat TM images to improve LiDAR estimates of canopy height (from 61% to 67% variability explained). Moghaddam et al. (2002) found that combining Landsat TM and several RaDARs was more accurate in predicting ground-based measurements of forest structure than any single sensor alone. Slatton et al. (2001) combined LiDAR data with interferometric RaDAR to improve the estimates of vegetation heights.

1.1. Objectives

Previous work (Hyde et al., 2005) established that large footprint, waveform LiDAR could be used to map forest structure within our study area at the footprint level with a high degree of accuracy. LiDAR was also highly accurate at measuring maximum canopy height and biomass at the "stand" (defined as 1 ha) level of observation when at least 40% of the area of observation was sampled. Where LiDAR data are sparse, it is an open research question whether or not improvements can be made at the stand and landscape scales via the combination of data from multiple sensors. For this study, only large diameter (>76 cm dbh) trees will be considered because at the stand level only stems in this size class were measured during our field data collection. Furthermore, a comparison at the footprint level of observation is somewhat problematic due to the inconsistent geolocation accuracy and resolution of the various sensors used in this study.

The primary objective of this effort is to quantify and compare the predictive power of individual remote sensing data sets to estimate large tree canopy height and biomass at the landscape scale. The secondary objective is to combine large footprint, waveform LiDAR data with other remote sensing data sets to determine if there is either synergy or redundancy in predictive power when combining other remote sensing data sets and large footprint, waveform LiDAR data. The tertiary objective is to ascertain the optimal sampling regime for large footprint, waveform LiDAR, i.e., to determine how sparsely large footprint, waveform LiDAR can be sampled (and fused with other remote sensing data sets) and still achieve a reasonable degree of predictive power. The results will be used to create landscape scale maps of forest structure suitable for wildlife habitat analysis.

The paper is organized as follows. First we describe collection of field plot data and provide details of the remote sensing (LVIS, SAR/InSAR, ETM+, Quickbird, DEM) data acquisition, which took place over the Sierra Nevada. This is followed by a presentation of the methods used in the processing and analysis of both remote sensing and field data, including the estimation of canopy height and biomass. We then present the results of statistical comparisons between field-derived and remote sensing-derived forest structural attributes and the results of multi-sensor fusion. Finally, we discuss the significance of results relative to the retrieval of forest structure at the landscape level.

2. Data collection

The data used in this study include in situ observations of forest structure, LiDAR data sets (Hyde et al., 2005), and other remote sensing data sets (Fig. 1).

2.1. Site description

The study area is located in the Sierra Nevada mountains of California. This site is approximately 60,000 ha, with elevation ranging from 853 to 2743 m. Forest types include white fir (*Abies concolor*), red fir (*Abies magnifica*), Sierran mixed-conifer, ponderosa pine (*Pinus ponderosa*), giant sequoia (*Sequoiadendron giganteum*), and montane hardwood-conifer (for a complete description, see Hunsaker et al., 2002).

2.2. Field data

One hundred twenty plots were distributed through the northern part the study area (Sierra National Forest) using a modified stratified random sampling scheme. Although the plots were centered on laser footprints, the actual waveforms were not examined before the stratification (to prevent bias). Hence, no attempt was made to retain only waveforms that

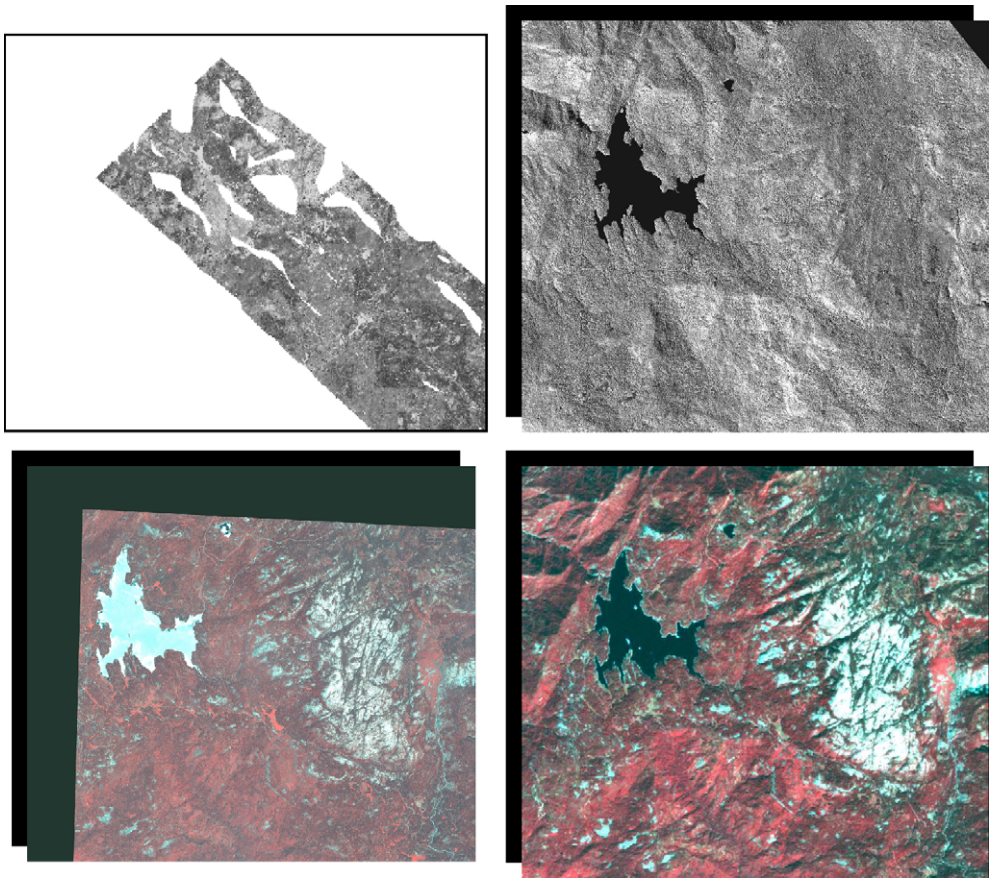

Fig. 1. Images depicting the remote sensing data sets in the Sierra Nevada used for this study. The top left image is from LiDAR, the top right image is from InSAR, the bottom left is Quickbird and the bottom right is ETM+. The images are cropped and enlarged to highlight the different structural features that each sensor is capable of resolving.

showed a strong ground return. The number of plots placed within each land cover type was proportional to their actual distribution within Sierra National Forest, with the exception of the red fir class; this vegetation type was oversampled because of its importance as remnant old-growth. Circular plots with an area of 1 ha (56.4 m radius) were established; these "stand" level plots were designed to be commensurate with existing USDA Forest Service Forest Inventory and Analysis plots.

Field plot data were collected during the summers of 2000 and 2001 and error-checked in 2002. All live stems with a diameter at breast height (dbh) $\geq$76 cm were inventoried and species type was recorded. The dbh of each stem was measured with a fiberglass tape. The height of the stem was measured with Impulse LR laser range finder (Laser Technology Inc., Englewood, CO). Summary statistics of the field plot data are provided in Table 1.

Table 1
Summary statistics of field plot measurements

	Minimum	Maximum	Mean	Standard deviation
Diameter at breast height (cm)	76.0	548.0 [a]	103.1	24.9
Height (m)	2.1 [b]	80.6 [a]	36.9	13.7
Biomass (kg)	1122.9 [b]	156698.0 [a]	8042.00	9224.3

[a] Giant sequoia.
[b] Snag.

2.3. Digital Elevation Model (DEM) data

A 10 m resolution Digital Elevation Model was acquired from the U.S. Geological Survey. This data set consists of "bald earth", ground elevation, or the elevation of the earth's surface devoid of vegetation or anthropogenic features. The vertical resolution is ±7 m.

2.4. LiDAR data

LiDAR data were collected by the Laser Vegetation Imaging Sensor (Blair et al., 1999) during October 1999 while deciduous trees were in leaf-on condition. LVIS is an airborne laser altimeter that records the time and amplitude of a laser pulse reflected off target surfaces. LVIS is a full waveform-digitizing system and records the vertical distribution of nadir-intercepted surfaces at 30 cm vertical resolution. The LVIS instrument flew aboard the NASA C-130H aircraft at about 7 km above ground level. LVIS is a large footprint LiDAR, recording spots or "footprints" illuminated within a 7° potential field of view. For the Sierra Nevada flights these footprints had a nominal radius of 12.5 m, nominally separated by 12.5 m across track and

continuous along track. Because of variations in altitude of the plane above the topography of the Sierra, actual footprint radii vary between 9 and 11 m. An area of about 175 km^2 was mapped. The fundamental observation of LVIS is a waveform that gives the vertical distribution of nadir-intercepted surfaces. The amplitude of the waveform at any height is proportional to the amount of reflective material intercepted at a particular height, the orientation of that material, and its reflectance (ignoring such effects as multiple scattering within the footprint). Initial processing of the data is required to remove various biases to permit accurate geolocation. The data are then further processed to find ground and canopy returns using automated methods, i.e., computer-coded algorithms (Hyde et al., 2005).

2.5. SAR/InSAR data

An X-band (3 cm wavelength, HV polarization) single-baseline RaDAR data set was selected for this study because the interferometric scattering phase center of X-band occurs relatively near to the top of the forest canopy. The RaDAR data set was generated by the STAR-3*i* system from Intermap Technologies (Englewood, CO). The STAR-3i instrument was flown aboard a Learjet in August 1999 and acquired a 3500 km^2 swath as well as a second overflight to minimize shadows and layover areas. Both a SAR backscatter intensity image and an InSAR "terrain model", that represents the elevation of the top (or near the top) of the canopy, were provided. The nominal spatial resolution of the data sets is 2.5 m and 10 m for the backscatter intensity image and the DEM, respectively. The reported accuracy for these products is 2.5 m RMSE horizontal and 3.0 m RMSE vertical. The intensity images are filtered to remove speckle.

2.6. ETM+ data

A Landsat ETM+ scene from October of 1999 covering Sierra National Forest and Sequoia National Park was obtained from USGS Eros Data Center. The scene consists of six visible and short-wave infrared bands with a nominal spatial resolution of 30 m (thermal and panchromatic bands were not used). The image was acquired during leaf-on conditions.

2.7. Quickbird data

Quickbird imagery was acquired from Satellite Imaging Corp (Houston, TX). Quickbird imagery consists of four spectral bands (3 visible, 1 near infrared) with a nominal spatial resolution of 2 m. Images were acquired in June 2002 and May 2003 in leaf-on conditions.

3. Data analysis

3.1. LiDAR data

The focus of this study is the estimation of several structural metrics: canopy height (mean, maximum, and standard deviation) and aboveground biomass. Canopy height is directly retrieved from waveform data using algorithms described below. It requires identification of a ground return in the waveform, and associated with this, the identification of the canopy portion of the waveform. Biomass is not directly measured by LVIS; rather, metrics derived from LiDAR waveforms, such as canopy height, and height of median energy, are correlated with canopy structure to compute biomass estimates.

Past studies, e.g., Lefsky (1997), have relied mainly on manual methods for identifying ground returns, especially where the returns are weak relative to the background noise level. While appropriate for validation studies with small numbers of waveforms, manual methods are impossible for the large number of LVIS waveforms being used here (ca. one million). Canopy height is determined relative to the ground, so accurately retrieving ground elevation is critical. An automated algorithm (described in Hyde et al., 2005) for finding both ground and canopy height was employed.

While LVIS does not measure biomass directly, metrics derived from LiDAR have proven effective in estimating forest biomass (Drake, 2001; Lefsky, 1997; Lefsky et al., 1999a, 2001, 1999b; Nelson et al., 1984, 1988; Nilsson, 1996). Canopy height by itself is sufficient for accurate biomass prediction in some more structurally simple biomes, as canopy height and biomass tend to be highly correlated (Lefsky et al., 1999b). In more structurally complex biomes, such as tropical and old-growth Western coniferous forests, some indication of the depth of the canopy is also useful for predicting biomass (Lefsky et al., 1999a; Drake, 2001). The metrics used in this study to estimate biomass include maximum canopy height, canopy cover, and height of the median energy return (HOME) (Table 2).

3.2. SAR/InSAR data

Elevation from a USGS DEM was used as the "ground" and subtracted from the InSAR DEM ("canopy top height") to produce a difference image reflecting the height of the InSAR scattering phase center, which will always be less than the

Table 2
Metrics derived from LiDAR waveforms

Metric	Description
MINMAXHT	Minimum height of the top of the canopy (m)
MAXMAXHT	Maximum height of the top of the canopy (m)
MEANMAXHT	Mean height of the top of the canopy (m)
STDEVMAXHT	Standard deviation of the height of the top of the canopy (m)
MINCOV	Minimum canopy cover (%)
MAXCOV	Maximum canopy cover (%)
MEANCOV	Mean canopy cover (%)
STDEVCOV	Standard deviation of canopy cover (%)
MINHOME	Minimum height of the median energy of the waveforms (m)
MAXHOME	Maximum height of the median energy of the waveforms (m)
MEANHOME	Mean height of the median energy of the waveforms (m)
STDEVHOME	Standard deviation of the height of the median energy of the waveforms (m)

observed canopy height, to produce a canopy height model. Summary statistics (minimum, maximum, mean, median, standard deviation) of this metric, as well as SAR backscatter intensity, were calculated from all pixels whose majority was within the 1 ha plots (Table 3).

3.3. ETM+ data

A USGS DEM was used to orthorectify the ETM+ images. Digital numbers (DN) were converted to at-sensor radiances (L) using published ETM+ calibration constants (ETM+ Data Users Guide) and the following equation:

$$L = a_0 + a_1 * \mathrm{DN}$$

where a_0 and a_1 refer to the sensor gain and bias, respectively.

3.4. Atmospheric and topographic correction

Radiance was converted to reflectance using top-of-atmosphere irradiance derived using the MODTRAN atmospheric correction model (Berk et al., 1989). A cosine correction was used to correct for the influence of topography on the amount of incident solar radiation. None of the field plots were obstructed by terrain, i.e., all field plots received direct and diffuse illumination at the time of image acquisition.

3.5. Metrics

Principle components and NDVI (NIR − red/NIR + red) were calculated at the stand level for the image. Summary statistics (minimum, maximum, mean, median, standard deviation) of these metrics were calculated from all pixels whose majority was within the 1 ha plots (Table 4).

3.6. Quickbird data

A USGS DEM was used to orthorectify the images. Digital numbers (DN) were converted to at-sensor radiances (L) using published calibration constants (Satellite Imaging Corp., Houston, TX).

3.7. Atmospheric and topographic correction

Radiance was converted to reflectance using top-of-atmosphere irradiance derived using the MODTRAN atmospheric correction model (Berk et al., 1989). A cosine correction was used to correct for the influence of topography on amount of incident solar radiation. None of the field plots were obstructed by terrain, i.e., all field plots received direct and diffuse illumination at the time of image acquisition.

3.8. Metrics

Principle components and NDVI (NIR − red/NIR + red) were calculated at the footprint level. Summary statistics (minimum, maximum, mean, median, standard deviation) of these metrics were also calculated from all pixels whose majority was within the 1 ha plots (Table 5).

Table 3
Metrics derived from SAR/InSAR

Metric	Description
MIN.SARBS	Minimum SAR backscatter intensity (dn)
MAX.SARBS	Maximum SAR backscatter intensity (dn)
MEAN.SARBS	Mean SAR backscatter intensity (dn)
STDEV.SARBS	Standard deviation of SAR backscatter intensity (dn)
MIN.SARHT	Minimum InSAR height (m)
MAX.SARHT	Maximum InSAR height (m)
MEAN.SARHT	Mean InSAR height (m)
STDEV.SARHT	Standard deviation of InSAR height (m)

Table 4
Metrics derived from ETM+ scene (October 1999)

Metric	Description
MIN.NDVI	Minimum NDVI
MAX.NDVI	Maximum NDVI
MEAN.NDVI	Mean NDVI
STDEV.NDVI	Standard deviation of NDVI
MIN.PCA1	Minimum first principle component
MAX.PCA1	Maximum first principle component
MEAN.PCA1	Mean first principle component
STDEV.PCA1	Standard deviation first principle component
MIN.PCA2	Minimum second principle component
MAX.PCA2	Maximum second principle component
MEAN.PCA2	Mean second principle component
STDEV.PCA2	Standard deviation second principle component
MIN.PCA3	Minimum third principle component
MAX.PCA3	Maximum third principle component
MEAN.PCA3	Mean third principle component
STDEV.PCA3	Standard deviation third principle component
MIN.PCA4	Minimum fourth principle component
MAX.PCA4	Maximum fourth principle component
MEAN.PCA4	Mean fourth principle component
STDEV.PCA4	Standard deviation fourth principle component

All values are unitless.

Table 5
Metrics derived from Quickbird

Metric	Description
MIN.QBNDVI	Minimum NDVI
MAX.QBNDVI	Maximum NDVI
MEAN.QBNDVI	Mean NDVI
STDEV.QBNDVI	Standard deviation of NDVI
MIN.QBPCA1	Minimum first principle component
MAX.QBPCA1	Maximum first principle component
MEAN.QBPCA1	Mean first principle component
STDEV.QBPCA1	Standard deviation first principle component
MIN.QBPCA2	Minimum second principle component
MAX.QBPCA2	Maximum second principle component
MEAN.QBPCA2	Mean second principle component
STDEV.QBPCA2	Standard deviation second principle component
MIN.QBPCA3	Minimum third principle component
MAX.QBPCA3	Maximum third principle component
MEAN.QBPCA3	Mean third principle component
STDEV.QBPCA3	Standard deviation third principle component

All values are unitless.

3.9. Field plot data

Field measurements at the footprint and stand levels include maximum canopy height, i.e., the height of the tallest stem within each plot, mean height of all stems within the plot, and standard deviation of the maximum height of all stems within the plot. Species-specific allometric equations relating stem biomass to height and dbh were obtained from the USDA Forest Service (Waddell & Hiserote, 2003). These equations were applied to the field data to calculate total standing (above-ground) biomass for each stem (≥76 cm dbh); the biomass of all stems within each plot was added to provide stand level totals.

4. Methods

4.1. Stand (1 ha) level

All possible subsets regression was used to compare the remote sensing data sets to field data. The best model, where all variables are significant at the $p<0.05$ level and Mallow's Cp most closely approximated the number of variables in the model, was selected. First, ETM+, Quickbird, LiDAR, and SAR/InSAR data sets were compared separately. Next, ETM+, Quickbird, and SAR/InSAR data sets were compared, in various combinations, to field data to produce models of canopy height and biomass. The combinations include: (1) LiDAR and Quickbird, (2) LiDAR and SAR/InSAR, (3) LiDAR and ETM+, (4) Quickbird, SAR/INSAR, and ETM+ and, (5) LiDAR, Quickbird, SAR/INSAR, and ETM+.

4.2. Landscape level

LiDAR data were randomly sampled at a variety of frequencies, corresponding to the following percentages of the total LiDAR data set: 1, 2, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100. Where LiDAR data were present, canopy height and biomass maps at each sample interval were produced by applying the field-validated models to resampled 1 ha pixels. Where LiDAR data were not present, only the other remote sensing data sets (and their corresponding field-validated model) were used. The result was a series of images at each sample interval that used LiDAR where present; the gaps where LiDAR was not present were "filled in" with the other remote sensing data sets. The mean canopy height and biomass of each image were then calculated. In addition, SAR/InSAR, Quickbird, and ETM+ metrics were sampled at the same frequency as the LiDAR data set (above) and put into regression models to predict the LiDAR observations of canopy height and biomass at the footprint level of observation (from Hyde et al., 2005).

5. Results

5.1. Stand (1 ha) level

For the single sensor models of maximum canopy height (Table 6), LiDAR ($r^2=0.76$, RMSE = 8.8 m) was more accurate than Quickbird ($r^2=0.66$, RMSE = 10.5 m), SAR/InSAR ($r^2=0.56$, RMSE = 11.9 m), ETM+ ($r^2=0.71$, RMSE = 9.6 m). The addition of either Quickbird or ETM+ to LiDAR resulted in slight ($r^2=0.79$, RMSE = 8.2) to moderate ($r^2=0.83$, RMSE = 7.5) improvement, respectively; the addition of SAR/InSAR resulted in no improvement. The combination of all sensors (LiDAR, ETM+, SAR/InSAR, and Quickbird) was the most accurate ($r^2=0.84$, RMSE = 7.3 m), while all other sensors excluding LiDAR (SAR/InSAR, ETM, and Quickbird) performed as well as LiDAR alone ($r^2=0.76$, RMSE = 8.8 m).

Table 6
Results of all possible subsets regression models relating remote sensing metrics and field-measured maximum canopy height

Sensor(s)	Coefficient of determination (r^2)	RMSE	Model
LiDAR	0.757	8.8	−2.4+(0.6*MEANMAXHT)+(0.9*SDMAXHT)+(0.5*MAXMAXHT)
QB	0.659	10.5	−419.1+(−0.07*MIN_QBPC1)+(0.06*STDEV_QBPC1)+(0.7*MEAN_QBPC3)+(0.9*STDEV_QBPC3)
InSAR	0.557	11.9	21.2+(0.8*MEAN_SARHT) *(3.1*STDEV_SARHT)
ETM	0.712	9.6	188.7+(−0.2*MIN_PCA1)+(0.2*MEAN_PCA2)+(−0.7*MEAN_PCA3)
LiDAR+QB	0.795	8.2	−11.7+(0.4*MEANMAXHT)+(0.5*MAXMAXHT)+(30.2*MAX_QBNDVI)+−(0.02*MIN_QBPC1)
LiDAR+InSAR	0.757	8.8	−2.4+(0.6*MEANMAXHT)+(0.9*SDMAXHT)+(0.5*MAXMAXHT)
LiDAR+ETM	0.827	7.5	162.8+(0.7*MAXMAXHT)+(−0.2*MEANCOVER)+(−0.1*MIN_PCA1)+(0.1*MAX_PCA2)+(−0.3*MIN_PCA3)
QB+InSAR+ETM	0.764	8.8	208.9+(−0.2*MIN_PCA1)+(−0.3*MEAN_PCA3)+(−0.4*MAX_PCA4)+(0.5*MAX_SARHT)+(−0.05*MIN_QBPC1)
LiDAR+QB+InSAR+ETM	0.835	7.3	245.6 (0.7*MAXMAXHT)+(−0.1*MEANCOVER)+(192.4*MIN_NDVI)+(−0.2*MIN_PCA1)+(−0.5*MIN_PCA3)

In all cases, $n=111$ and $p<0.001$.

For the single sensor models of mean canopy height (Table 7), LiDAR ($r^2=0.61$, RMSE = 7.3 m) was more accurate than Quickbird ($r^2=0.57$, RMSE = 7.8 m), SAR/InSAR ($r^2=0.45$, RMSE = 8.7 m), ETM+ ($r^2=0.60$, RMSE = 7.5 m). The addition of either Quickbird or ETM+ to LiDAR resulted in slight ($r^2=0.66$, RMSE = 6.9) to moderate ($r^2=0.72$, RMSE = 6.4) improvement, respectively; the addition of SAR/InSAR resulted in no improvement. The combination of all sensors (LiDAR, ETM+, SAR/InSAR, and Quickbird) was the most

accurate (r^2=0.72, RMSE=6.4 m), while all other sensors excluding LiDAR (SAR/InSAR, ETM, and Quickbird) performed slightly better than LiDAR alone (r^2=0.64, RMSE=7.1 m).

For the single sensor models of standard deviation canopy height (Table 8), LiDAR (r^2=0.59, RMSE=3.5) was more accurate than Quickbird (r^2=0.48, RMSE=3.9), SAR/InSAR (r^2=0.46, RMSE=4.0), ETM+ (r^2=0.52, RMSE=3.7). The addition of either Quickbird or ETM+ to LiDAR resulted in slight (r^2=0.60, RMSE=3.6) to moderate (r^2=0.63, RMSE=3.2) improvement, respectively; the addition of SAR/InSAR resulted in no improvement. The combination of all sensors (LiDAR, ETM+, SAR/InSAR, and Quickbird) was the most accurate (r^2=0.64, RMSE=3.2), while all other sensors excluding LiDAR (SAR/InSAR, ETM, and Quickbird) were less accurate than LiDAR alone (r^2=0.56, RMSE=3.5).

For the single sensor models of biomass (Table 9), LiDAR (r^2=0.77, RMSE=75.3 mg/ha) was more accurate than Quickbird (r^2=0.55, RMSE=105.0 mg/ha), SAR/InSAR (r^2=0.51, RMSE=109.1 mg/ha), ETM+ (r^2=0.49, RMSE=111.2 mg/ha). The addition of ETM+ to LiDAR resulted in slight (r^2=0.80, RMSE=71.3 mg/ha) improvement; the addition of SAR/InSAR or Quickbird resulted in no improvement. The combination of all sensors (LiDAR, ETM+, SAR/InSAR, and Quickbird) was the most accurate (r^2=0.83, RMSE=66.6 mg/ha), while all other sensors excluding LiDAR (SAR/InSAR, ETM, and Quickbird) performed worse than LiDAR alone (r^2=0.72, RMSE=84.2 mg/ha).

Table 7
Results of all possible subsets regression models relating remote sensing metrics and field-measured mean canopy height

Sensor(s)	Coefficient of determination (r^2)	RMSE	Model
LiDAR	0.611	7.3	3.7+(0.3*MEANMAXHT)+(0.4*MAXMAXHT)
QB	0.566	7.8	−183.2+(−0.01*MIN_QBPC1)+(0.07*STDEV_QBPC1)+(0.4*MEAN_QBPC3)
InSAR	0.454	8.7	17.3+(0.5*MEAN_SARHT)+(91.9*STDEV_SARHT)
ETM	0.603	7.5	222.6+(−148.8*MEAN_NDVI_1)+(−0.2*MIN_PCA1_1)+(−0.6*MEAN_PCA3)
LiDAR+QB	0.666	6.9	21.1+(0.2*MEANMAXHT)+(0.3*MAXMAXHT)+(−0.05*MIN_QBPC1)
LiDAR+InSAR	0.611	7.3	3.7+(0.3*MEANMAXHT)+(0.4*MAXMAXHT)
LiDAR+ETM	0.715	6.4	146.2+(0.4*MAXMAXHT)+(−0.1*MEANCOVER)+(0.1*MAX_PCA1_1)+(0.5*STDEV_PCA1)+(−0.1*MEAN MIN_PCA3)
QB+InSAR+ETM	0.643	7.1	45.5 (−0.1*MIN_PCA1_1)+(0.3*MEAN_SARHT)+(−0.04*MIN_QBPC1)+(0.05*STDEV_QBPC1)
LiDAR+QB+InSAR+ETM	0.715	6.4	112.3+(0.4*MAXMAXHT)+(−0.1*MEANCOVER)+(−0.1*MAX_PCA1)+(−0.6*STDEV_PCA1)+(0.5*MAX_PCA4)

In all cases, n=111 and p<0.001.

Table 8
Results of all possible subsets regression models relating remote sensing metrics and field-measured standard deviation canopy height

Sensor(s)	Coefficient of determination (r^2)	RMSE	Model
LiDAR	0.587	3.5	−3.7+(0.2*MEANMAXHT)+(0.2*MAXMAXHT)+(−0.5*MEANHOME)+(0.3*SDHOME)+(0.08*MINCOVER)
QB	0.482	3.9	−115.3+(−0.02*MEANQBPC1)+(0.04*STDEV_QBPC2)+(0.01*MAX_QBPC2)+(0.2*MAX_QBPC3)
InSAR	0.457	4.0	2.8+(−0.15*MIN_SARHT)+(0.4*MAX_SARHT)
ETM	0.517	3.7	37.5+(−187.0*STDEV_NDVI)+(−0.6*MIN_PCA1)
LiDAR+QB	0.595	3.6	−25.5+(0.2*MEANMAXHT)+(0.5*SDMAXHT)+(0.01*MEAN_QBPC2)+(0.2*STDEV_QBPC3)
LiDAR+InSAR	0.611	3.4	−2.7+(0.2*MEANMAXHT)+(0.2*MAXMAXHT)+(−.5*MEANHOME)+(0.08*MINCOVER)+(0.2*MAX_SARHT)
LiDAR+ETM	0.631	3.2	−34.0+(0.3*SDMAXHT)+(0.1*MAXMAXHT)+(−0.1*MEANHOME)+(−58.4*MAX_NDVI)+(0.1*MAX_PCA2)
QB+InSAR+ETM	0.563	3.5	7.8+(−0.04*MIN_PCA1)+(0.2*MAX_SARHT)+(0.2*STDEV_QBPCA)
LiDAR+QB+InSAR+ETM	0.640	3.2	24.3+(0.4*SDMAXHT)+(0.07*MINCOVER)+(−0.04*MINPCA1)+(0.1*MAX_SARHT)+(−0.03*MIN_QBPC3)

In all cases, n=111 and p<0.001.

5.2. Landscape scale

Images of mean large tree canopy height and biomass over the entire study area were calculated over the spatial extent of the 1 to 100% LiDAR samples using LiDAR where it existed and the other remote sensing data sets (Quickbird, ETM+, and SAR/InSAR) where LiDAR was not available. Mean large tree canopy height and biomass were calculated from each image; the results are plotted in Fig. 2. Mean canopy height of large trees remained about the same as LiDAR sample size increased (Fig. 2A), declining only about 1.5 m as the sample size approached 100%. Biomass of large trees (Fig. 2B) increased fairly substantially, from about 155 to 175 mg/ha.

Table 9
Results of all possible subsets regression models relating remote sensing metrics and field-measured biomass

Sensor(s)	Coefficient of determination (r^2)	RMSE	Model
LiDAR	0.773	75.3	−49.8+(6.6*MEANMAXHT)+(−7.5*MINMAXHT)+(18.9*MEANHOME)
QB	0.550	105.0	−3004.1+(−480.1*MIN_QBNDVI)+(−1.9*MIN_QBPC3)+(6.7*MEAN_QBPC3)
InSAR	0.509	109.1	38.1+(9.5*MEAN_SARHT)+(9.8*STDEV_SARHT)
ETM	0.490	111.2	1837.4+(−1.8*MEAN_PCA1)+(−6.6*MEAN_PCA3)
LiDAR+QB	0.773	75.3	−49.8+(6.6*MEANMAXHT)+(−7.5*MINMAXHT)+(18.9*MEANHOME)
LiDAR+InSAR	0.773	75.3	−49.8+(6.6*MEANMAXHT)+(−7.5*MINMAXHT)+(18.9*MEANHOME)
LiDAR+ETM	0.796	71.3	−367.2+(5.5*MEANMAXHT)+(−6.1*MINMAXHT)+(18.8*MEANHOME)+(−2.7*MINCOVER)+(−1671.1*MEAN_NDVI)
QB+InSAR+ETM	0.716	84.2	2448.1(−1866.6*MEAN_NDVI)+(−1.0*MEAN_PCA1+(11.6*MEAN_SARHT)+(−0.4*MIN_QBPC2)+(−3.1*MAX_QBPC3)
LiDAR+QB+InSAR+ETM	0.827	66.6	−950.1+(4.2*MEANMAXHT)+(−5.4*MINMAXHT)+(17.6*MEANHOME)+(−2.3*MINCOVER)+(−2464.0*MEAN_NDVI+(2.3*MEAN_PCA3_1)+(3.9*MEAN_SARHT)+(−2.8*STDEV_SARBS)

In all cases, $n=111$ and $p<0.001$.

5.3. Discussion

Once again, LiDAR has been shown to be effective in estimating forest structure. One unique aspect of this study is that LiDAR was shown to be effective in measuring height and biomass of large trees, which are important components of wildlife habitat and represent a key carbon reservoir. Quickbird and SAR/InSAR did not perform nearly as well; this is not surprising given that only LiDAR has the ability to provide an accurate return from both the top and bottom of the canopy and provide a volumetric response (via the height of the median energy return or HOME). While the backscatter from other, longer wavelengths RaDARs (e.g., C, L and P) are sensitive to canopy volume, the X-band wavelength used in this study is much less so. The InSAR height metric was less accurate at predicting height and biomass than LiDAR height, possibly because the X-band penetrated further into the top of the canopy, causing a foreshortening of the height measurement. It is also probable that at least some of the elevation values from the DEM were corrupted by the canopy; LiDAR ground elevation values are likely more reliable since we are confident that the last return is actually coming from the ground. However, other sensors did have significant utility in aiding LiDAR in the prediction of canopy height. In contrast, the other sensors did not contribute as much to the accuracy of the biomass predictions.

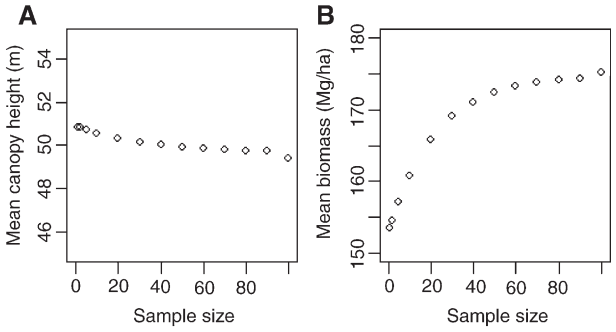

Fig. 2. Landscape scale canopy height (A) and biomass (B) of large trees as a function of LiDAR sample size, 1–100%.

ETM+, however, was surprisingly accurate at predicting canopy height. Additionally, ETM+ and LiDAR appear to be quite complementary, producing more accurate measurements of all variables together compared to any other single sensor or LiDAR/Quickbird and LiDAR/RaDAR combinations. ETM+ alone was fairly poor at predicting biomass, probably because of signal saturation at high biomass levels. It is necessary to remember that the structure metrics include only the largest (>76 cm dbh) trees; when compared to footprint level LiDAR (from Hyde et al., 2005) with known errors accounted for (and therefore in very close agreement with field data) that include all trees, the combination of ETM+, Quickbird, and RaDAR provides poor estimates of canopy height and even worse estimates of biomass (Fig. 3).

At the landscape scale, as LiDAR sample size increases, the estimate of the height of large trees stays about the same. If the image containing the 100% LiDAR sample is used as a baseline of height "truth", then the canopy height images from RaDAR, ETM+, Quickbird are reasonably accurate regardless of the size of the LiDAR sample. In other words, the gaps in LiDAR coverage are measured fairly accurately with the other sensors, providing a reasonable approximation of canopy height.

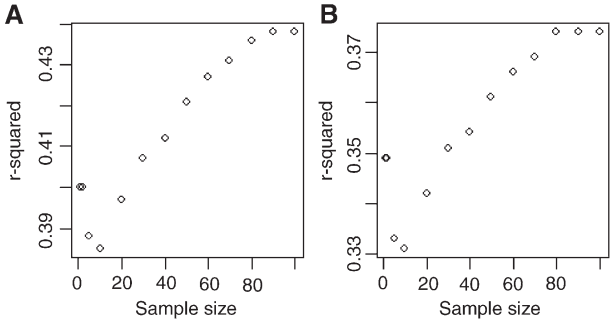

Fig. 3. Coefficients of determination (r^2) values from regression of footprint LiDAR observations of canopy height (A) and biomass (B) of all trees as function of LiDAR sample size, 1–100%.

However, this is not true of biomass; gaps in LiDAR coverage are not measured as accurately with the other remote sensing data sets and the net result is a greater than 10% reduction in total estimated biomass for the entire landscape.

Recall that the ultimate goal of this effort is to provide spatially continuous maps of forest structure at the landscape scale as a prerequisite for habitat suitability studies. Using the statistical relationship developed between field and LiDAR plus other remote sensing data sets for canopy height and biomass, we created maps of each of these over the domain of the data sets (Figs. 4 and 5). It is this mapping of forest structural characteristics at the landscape scale which we believe will be of great benefit to future habitat studies.

5.4. Conclusion

LiDAR provides accurate estimates of stand level canopy height and biomass at the stand (1 ha) level even though it is sampling, not mapping at this scale. Some improvements were achieved by adding some or all additional remote sensing data sets. However, these improvements come at a cost, including (1) cost of data acquisition, (2) cost of data processing, and (3)

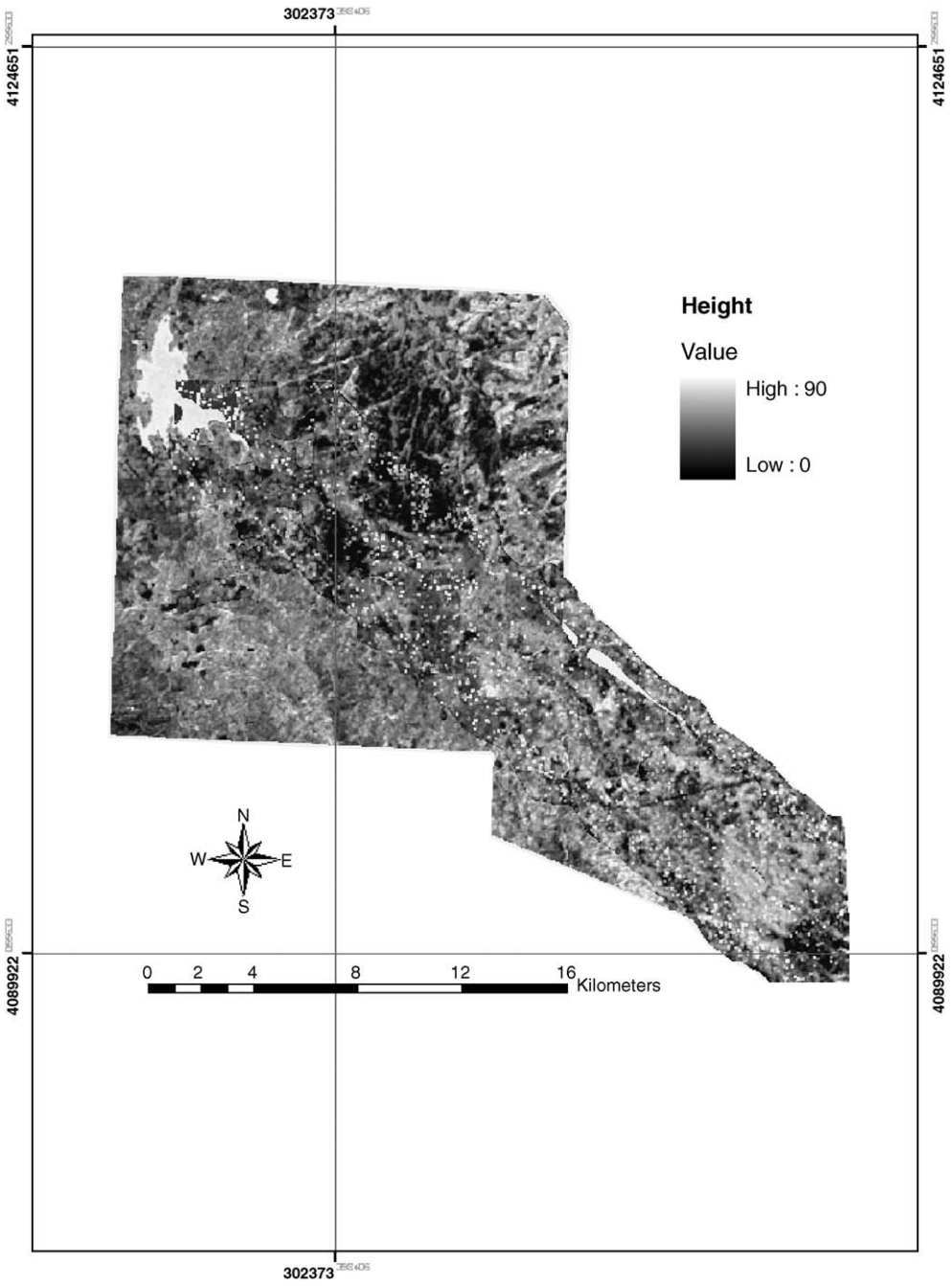

Fig. 4. Canopy height map from LiDAR and other remote sensing data sets. The units are in meters.

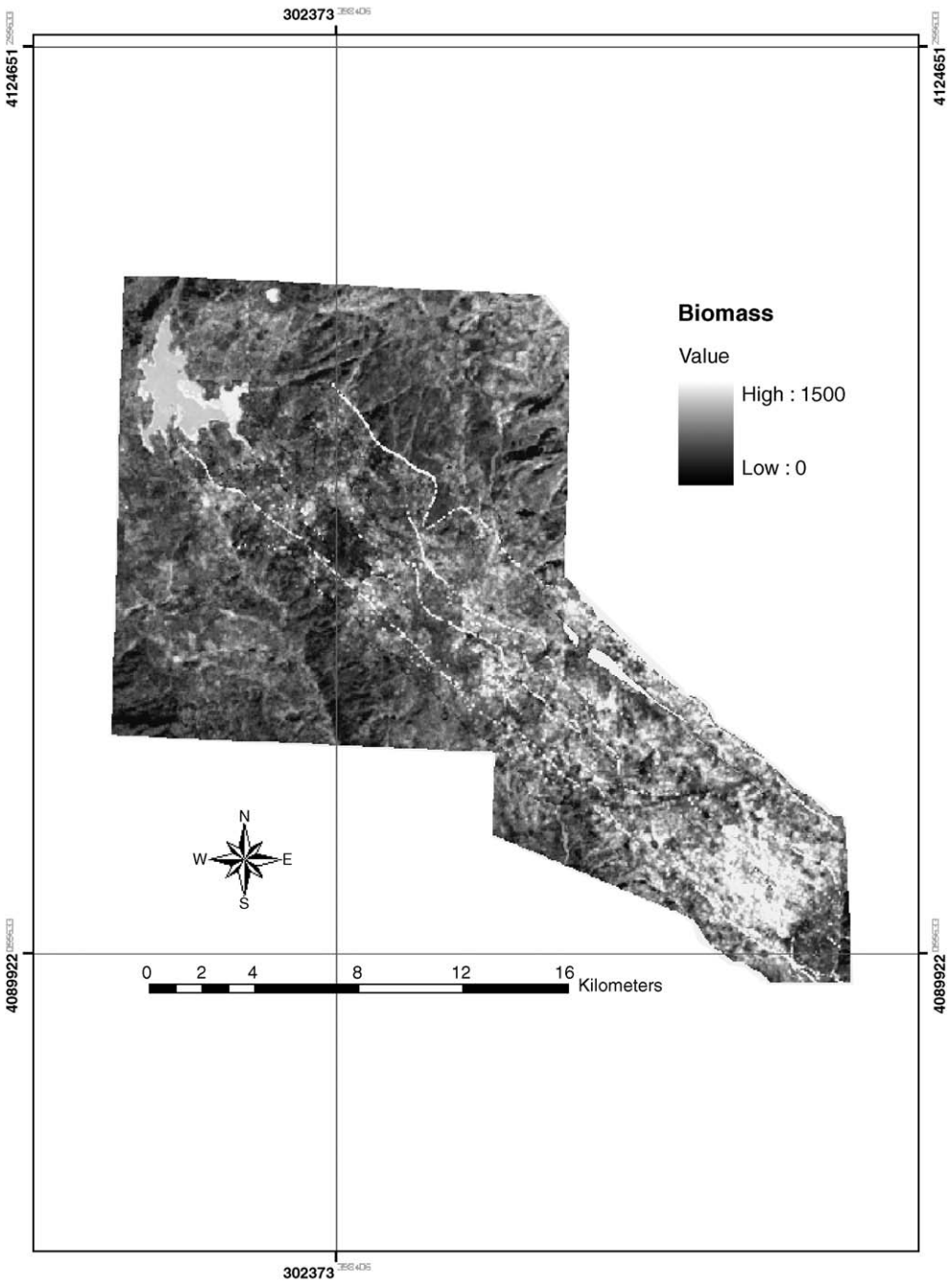

Fig. 5. Biomass map from LiDAR and other remote sensing data sets. The units are in mg/ha.

statistical certainty. The cost of acquiring and processing ETM+ is relatively minimal; the cost of acquiring (Quickbird) or processing (InSAR) is not. Atmospheric and topographic correction of passive optical data is also not trivial, requiring computing power, time, and expertise.

It is not clear at this point whether or not estimates of big tree height and biomass at the stand (1 ha) level have more utility for habitat modeling than do samples taken at the footprint scale. It is possible that these maps do not provide any additional predictive power for a given application, but there is no way to know this a priori. Future work on this question by the authors is forthcoming.

There are many other sensors available that were not used in this study but nonetheless have great promise for the retrieval of canopy structure. Only X-band radar was used in this study because of the lack of availability of other bands at a fine enough spatial resolution. Other longer wavelength bands would likely penetrate foliage and provide strong returns from the trunks of big trees. Multi-angle sensors such as POLDER or MISR would also provide information about canopy height, while hyperspectral data sets could improve biomass estimates. Also, only simple vegetation indices were computed from multispectral data; other, more sophisticated algorithms could be used, such as automated crown detection. Only simple

LiDAR metrics were used; there is a wealth of vertical structure information contained in waveforms that are largely underused by this, or any other, study. In addition, canopy volume, basal area, tree density, species diversity, and other such measurements of structure would be useful for wildlife habitat studies. Finally, a truly comprehensive and complete study of multisensor fusion would include more than statistical models; physical models should be developed and used to explain the causal mechanisms behind sensor returns.

References

Berk, A., Bernstein, L., & Robertson, D. (1989). *MODTRAN: A moderate resolution model for LOWTRAN7. GL-TR-89-0122.* Bedford, MA: U.S. Air Force Geophysics Laboratory.

Blair, J. B., Rabine, D. L., & Hofton, M. A. (1999). The laser vegetation imaging sensor: A medium-altitude, digitisation-only, airborne laser altimeter for mapping vegetation and topography. *ISPRS Journal of Photogrammetry and Remote Sensing*, *54*, 115–122.

Drake, J. (2001). *Estimation of tropical forest biomass*. Dissertation thesis, University of Maryland, College Park, MD.

Drake, J., Dubayah, R., Knox, R., Clark, D., & Blair, J. B. (2002). Sensitivity of large-footprint lidar to canopy structure and biomass in a neotropical rainforest. *Remote Sensing of Environment*, *81*(2–3), 378–392.

Finney, M. A. (1998). FARSITE: Fire area simulator — model development and evaluation. (RP-4): 1 —+.

Hudak, A. T., Lefsky, M. A., Cohen, W. B., & Berterretche, M. (2002). Integration of lidar and Landsat ETM plus data for estimating and mapping forest canopy height. *Remote Sensing of Environment*, *82*(2–3), 397–416.

Hunsaker, C. T., Boroski, B. B., & Steger, G. N. (2002). *Relations between canopy cover and occurrence and productivity of California spotted owls, predicting species occurrences, issues of accuracy and scale.* Covelo, CA: Island Press.

Hyde, P., Dubayah, R., Peterson, B., Blair, J. B., Hofton, M., Hunsaker, C., et al. (2005). Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems. *Remote Sensing of Environment*, *96*(3–4), 427–437.

Hyyppä, J., Hyyppä, H., Inkinen, M., & Engdahl, M. (1998). Verification of the potential of various remote sensing devices for forest inventory. *Proceedings of IEEE geosciences and remote sensing society* (pp. 1812–1814). Pasadena, CA: California Institute of Electrical and Electronics Engineers.

Lefsky, M. A. (1997). Application of lidar remote sensing to the estimation of forest canopy and stand structure. Dissertation thesis, University of Virginia.

Lefsky, M. A., Cohen, W. B., Acker, S. A., Parker, G. G., Spies, T. A., & Harding, D. (1999). Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. *Remote Sensing of Environment*, *70*, 339–361.

Lefsky, M. A., Cohen, W. B., & Spies, T. A. (2001). An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in Western Oregon. *Canadian Journal of Forest Research*, *31*, 78–87.

Lefsky, M. A., Harding, D., Cohen, W. B., Parker, G., & Shugart, H. H. (1999). Surface lidar remote sensing of basal area and biomass in deciduous forests of eastern Maryland, USA. *Remote Sensing of Environment*, *67*, 83–98.

Moghaddam, M., Dungan, J. L., & Acker, S. (2002). Forest variable estimation from fusion of SAR and multispectral optical data.. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(10), 2176–2187.

Nelson, R., Krabill, W., & Maclean, G. (1984). Determining forest canopy characteristics using airborne laser data. *Remote Sensing of Environment*, *15*, 201–212.

Nelson, R., Krabill, W., & Tonelli, J. (1988). Estimating forest biomass and volume using airborne laser data. *Remote Sensing of Environment*, *24*, 247–267.

Nilsson, M. (1996). Estimation of tree heights and stand volume using an airborne lidar system. *Remote Sensing of Environment*, *56*, 1–7.

North, M. P., Franklin, J. F., Carey, A. B., Forsman, E. D., & Hamer, T. (1999). Forest stand structure of the northern spotted owl's foraging habitat. *Forest Science*, *45*(4), 520–527.

Pu, R. L., & Gong, P. (2004). Wavelet transform applied to EO-1 hyperspectral data for forest LAI and crown closure mapping. *Remote Sensing of Environment*, *91*(2), 212–224.

Skole, D., & Tucker, C. (1993). Tropical deforestation and habitat fragmentation in the Amazon: Satellite data from 1978 to 1988. *Science*, *260*, 1905–1909.

Slatton, K. C., Crawford, M. M., & Evans, B. L. (2001). Fusing interferometric radar and laser altimeter data to estimate surface topography and vegetation heights. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(11), 2470–2482.

Treuhaft, R. N., & Cloude, S. R. (1999). The structure of oriented vegetation from polarimetric interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, *37*(5), 2620–2624.

Treuhaft, R. N., & Siqueira, P. R. (2000). Vertical structure of vegetated land surfaces from interferometric and polarimetric radar. *Radio Science*, *35*(1), 141–177.

Waddell, K., & Hiserote, B. (2003). *Technical documentation for the integrated database, version 1.0.*: USDA Forest Service Pacific Northwest Research Station.

Weishampel, J. F., Blair, J. B., Knox, R. G., Dubayah, R., & Clark, D. B. (2000). Volumetric lidar return patterns from an old-growth tropical rainforest canopy. *International Journal of Remote Sensing*, *21*(2), 409–415.

Wulder, M., Hall, R., Coops, N. C., & Franklin, S. (2004). High spatial resolution remotely sensed data for ecosystem characterization. *BioScience*, *54*(6), 511–521.

Wulder, M., & Seeman, D. (2003). Forest inventory height update through the integration of lidar data with segmented Landsat imagery. *Canadian Journal of Remote Sensing*, *29*(5), 536–543.