Mapping land cover and estimating forest structure using satellite imagery and coarse resolution lidar in the Virgin Islands

Todd A. Kennaway,[a] Eileen H. Helmer,[b] Michael A. Lefsky,[c] Thomas A. Brandeis,[d] Kirk R. Sherrill[c]

[a]Department of Forest, Rangeland and Watershed Stewardship, Colorado State University, Fort Collins, CO 80523-1472, USA
kennaway@cnr.colostate.edu

[b]International Institute of Tropical Forestry, USDA Forest Service, Jardín Botánico Sur, 1201 Calle Ceiba, Río Piedras, PR 00926-1119, USA

[c]Center for Ecological Analysis of Lidar, Department of Natural Resources, Colorado State University, 131 Forestry Building, Fort Collins, CO 80523-1472, USA

[d]Southern Research Station, USDA Forest Service, 4700 Old Kingston Pike, Knoxville, TN 37919-5206

Abstract

Current information on land cover, forest type and forest structure for the Virgin Islands is critical to land managers and researchers for accurate forest inventory and ecological monitoring. In this study, we use cloud free image mosaics of panchromatic sharpened Landsat ETM+ images and decision tree classification software to map land cover and forest type for the Virgin Islands, illustrating a low cost, repeatable mapping approach. Also, we test if coarse-resolution discrete lidar data that are often collected in conjunction with digital orthophotos are useful for mapping forest structural attributes. This approach addresses the factors that affect vegetation distribution and structure by testing if environmental variables can improve regression models of forest height and biomass derived from lidar data. The overall accuracy of the 29 forest and non-forest classes is 72%, while most the forest types are classified with greater than 70% accuracy. Due to the large point spacing of this lidar dataset, it is most appropriate for height measurements of dominant and co-dominant trees (R^2 = 72%) due to its inability to accurately represent forest understory. Above ground biomass per hectare is estimated by its direct relationship with plot canopy height (R^2 = 0.72%).

Keywords: Land cover, decision tree software, discrete lidar, forest structure, regression modeling, Virgin Islands.

1 INTRODUCTION

Information on land cover, forest type and forest structure for the Virgin Islands is limited to maps of ecological zones and photo interpreted land cover [1,2] and relatively few forest inventory plots. The lack of current land cover data and robust techniques for updating that data, and the sparseness of forest inventory plots relative to the number of different forest types, pose challenges to land managers and researchers in ecologically vulnerable subtropical environments. These challenges are made more acute in Caribbean environments, because the interaction between trade winds and steep topographic gradients cause forest types to change over short distances [3] and high rates of disturbance lead to variable forest structure. In this study, datasets of land cover and forest type are derived from satellite imagery with decision tree software, illustrating a low cost, repeatable approach for creating

Received 13 Jul 2008; accepted 3 Dec 2008; published 12 Dec 2008 [CCC: 19313195/2008/$25.00]

such data. Although decision tree classification is becoming common in remote sensing, only a few studies use decision trees for detailed forest mapping of subtropical islands [4-6].

This study also addresses the characterization of forest structure with airborne light detection and ranging (lidar) when inventory data are sparse. Lidar adds a third (z) dimension to the spatial description of forest types with accurate estimates of vegetation height and above ground biomass [7]. No prior research has addressed the quantification of lidar derived forest structure in the Virgin Islands. Data describing indices of forest structure such as height and biomass can provide important information such as indicators of forest age, species richness and habitat.

Forest distribution and structure in the Virgin Islands has been modified for hundreds of years by both natural and human caused disturbances, including hurricanes and human exploitation. Prior to European colonization, indigenous peoples such as the Carib first cleared forest for food, shelter, and boat building materials. In the 1600 and 1700's Danish and British settlers arrived and began converting forest to intensive agriculture that included coffee, sugar cane and tobacco [8,9,10,11]. As a result, most old growth forest was cleared and has recovered as fragmented secondary forests after the gradual abandonment of agriculture through to the early twentieth century. The forest clearing had lasting impacts on forest structure, ecosystem function and species composition, including the introduction and extinction of exotic and endemic species [12].

Increased pressure from urban development has led to additional forest clearing in the Lesser Antilles [5] and Puerto Rico. The spatial pattern of forest clearing is often influenced by proximity to existing urban areas, roads and topography [13,4]. Islands such as St. Thomas [14] and Tortola have also experienced urban growth at the expense of forested areas over the last decade. About 65% of St. John is protected by the US Park Service including much of its semi-deciduous (including semi-evergreen) and deciduous forests. However, the unprotected low elevation dry forests on that and other islands, which have been shown to be important habitat for many avian species, are considered endangered and susceptible to increasing developmental pressures [15].

The overall goal of this study is to develop an approach for characterizing the structure of varied subtropical island forest formations when available inventory data are relatively sparse. To accomplish this goal, we developed three main objectives. The first objective is to test an improvement to a previously developed approach for using Landsat image mosaics to map land-cover and forest types in persistently cloudy, complex tropical landscapes with decision tree classification software [4-6,16,17]. The improvement is that we use panchromatic-sharpened image mosaics to increase spatial resolution in the resulting maps. We also test whether the approach is applicable to a large area, the Virgin Islands, which includes many islands. The second objective is to test if the coarse-resolution (shot spacing of 2.76m) discrete lidar data that are often collected in conjunction with digital orthophotos are useful for mapping forest structural attributes, including height and biomass, over the steep environmental gradients present on the islands of St. John and St. Thomas. Also, this study tests whether integrating Landsat ETM+ satellite imagery and environmental variables with the lidar data can improve models of forest structural attributes. Whether the large point spacing of such coarse resolution lidar will be adequate to accurately sample and model forest structure parameters has not been tested. Also unknown is whether the range of physiognomic types found across these islands will complicate the estimates of forest height and biomass. Several studies have shown that large footprint scanning lidar accurately predicts forest structure, including canopy height, basal area and above ground biomass in

Douglas fir/western hemlock forests [18,19]. Other studies accurately model forest metrics with discrete lidar, focusing on small tracts of homogeneous forest stands and small foot-print sensors [20]. Accurate estimates of forest structural attributes using regressions have been successfully performed such as height [21,22], aboveground biomass [23,24], and crown diameter [25,26]. Lidar based biomass has been estimated for a variety of forest types including, but not limited to, temperate mixed deciduous coniferous [25,23,27], temperate deciduous [24] and tropical rain forests [28]. The third objective is to summarize forest structure of the predicted forest structural attributes for each mapped forest class. This step allows us to characterize forest height and biomass for different forest types on St. John and St. Thomas. The datasets generated in the project will support other studies in the Virgin Islands, including avian monitoring surveys and the Forest Stewardship Program [29].

2 METHODS

2.1 Study area

The US and British Virgin islands (18°20'N, 64°40'W) are a part of the Caribbean's Lesser Antilles and are composed of six major and 40+ minor islands and cays. The major islands in the US territory include St. Thomas, St. John and St Croix, while the main islands in the British territory include Tortola, Virgin Gorda and Anegada (Figure 1). The islands have a combined area of about 50,000 ha, with subdued to rugged topography and elevations ranging from just below sea level in some wetlands to over 500 m on the island of Tortola. The climate is mostly subtropical, with a hot and humid rainy season that extends from May to November and a dry season that is tempered by trade winds. The geology of the islands consists of alluvial, sedimentary, volcanic and limestone strata. Ecological zones on the islands include Subtropical Moist and Dry forest *sensu* Holdridge [1,30].

Lidar Study Area

St. John and St. Thomas were selected as the lidar study area based on the availability of lidar data coverage. The island of St. John (18º22'N, 64º40'W) and the island of St. Thomas (18º21'N, 64º55') are about 5,000 and 7,200 hectares in area, respectively, and consist of mountainous topography with elevations ranging from sea level to 387 m on St. John and 471 m on St. Thomas. The woody vegetation on both islands is similar to other islands in the Virgin Islands and includes both late and early stage successional forests.

In 1956, the US Park Service established the Virgin Islands National Park (VINP). Protecting about 65% of St. John, it includes the island's interior high elevation semi-evergreen and deciduous forests. The long standing reserve status has helped protect most of the island's forests from development, creating one of the largest contiguous expanses of forest in the Lesser Antilles. The VINP provides unique research opportunities to study the island's diverse ecology and establishes a template for monitoring mature successional forest structure. In contrast, the forests of St. Thomas which make up about 69% of the island has not received protection status and developmental pressures and impacts can be observed island-wide.

2.2 Landsat Imagery and Reference Data

A land-cover and forest type map for the US and British Virgin Islands was created by supervised classification of Landsat ETM+ imagery using decision tree analysis software. An image mosaic for about the year 2000 was created from Landsat scenes of various dates. The

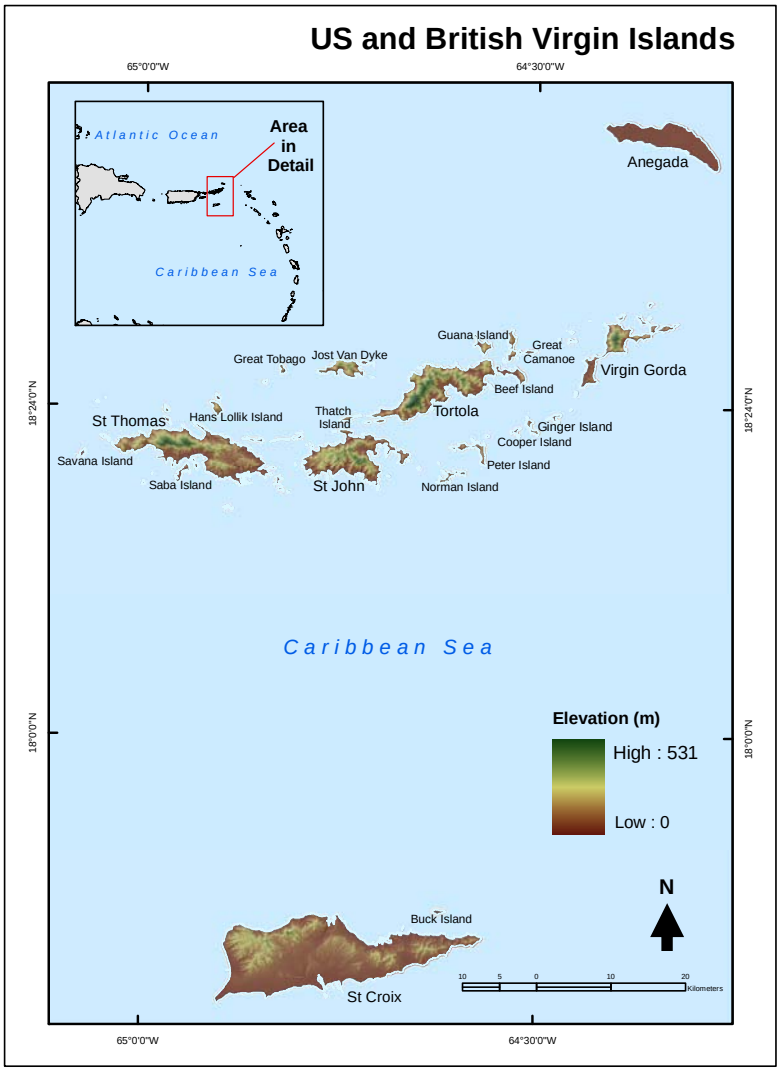

Fig. 1. Map of the study area.

reference scenes for the mosaic were World Reference System 2 Path/Row 004/047-048, both dated 27 Mar 00. The scenes used to fill cloud-masked or edge areas in Path/Row 004/047 were dated 02 Nov 01 (Path/Row 003/047), 17 Sep 99, 02 Aug 00, and 25 Jan 01. The scene used to fill cloud-masked areas in Path/Row 004/048, was dated 25 Jan 01. Cloud obstruction in the reference image was 20.9 % before and 5.3% after the cloud removal and mosaic process. The 30-m multispectral bands for each scene were first cloud-masked and then matched to the reference scene with regression tree normalization [31]. This technique models the relationship between co-located pixels from different image dates and estimates new image digital numbers (DNs) to fill in the cloud and cloud-shadow masked areas of the reference scene. In addition, the technique reduces atmospheric and phenological differences that occur with multi-date image mosaics [5]. Likewise, the 15-m panchromatic band for each scene was also matched to the reference panchromatic band with regression tree normalization models based only on the panchromatic bands. The matched panchromatic

image parts were then mosaicked, and the panchromatic mosaic was then used to pan-sharpen the 30-m mosaic of the multispectral bands. Principal components transformation was chosen to merge the native 30 m Landsat multispectral image bands (bands1-5, 7) for each scene with the 15 m panchromatic band (band 8). It was chosen based on 1) results from preliminary tests of other resolution merging methodologies (Brovey and Multiplicative) available in ERDAS Imagine, and 2) other studies [32-34] that have concluded that the principle components transformation method provides increased spatial resolution without degrading spectral discrimination.

Ancillary data were used to create an island-wide predictor variable dataset to assist in the classification of image pixels. Adding geographic data ancillary to satellite imagery improves classification of land cover and forest types by reducing spectral confusion among vegetation classes [35,36], including in Caribbean island landscapes [6]. Topographic variables derived from United States Geological Survey (USGS) 30 m digital elevation models (DEM) for the US territory and 90 m Shuttle Radar Topography Mission (SRTM) elevation datasets resampled to 30 m for the British territory included elevation, slope and aspect [37,38]. Climatic variables included mean annual precipitation and temperature [39]. Variables derived from USGS Digital Line Graphics (DLG) for the US islands and scanned topographic maps for the British islands at a scale of 1:24,000 that were registered to the image mosaic include distance to primary and secondary roads, distance to streams and ravines, and distance to coastlines [40]. The ancillary predictor data was spatially co-registered with the cloud free image mosaic and stacked with the Landsat ETM+ reflectance bands 1-5, 7, and two band indices, resulting in an 18 band image mosaic for the classification. The band indices included the Landsat ETM+ image bands to produce the normalized difference vegetation index (NDVI) and 4/5 band ratio, which are useful indictors of vegetation vigor and forest structure [41-43].

Field surveys in 2005 and consultation with experts enabled us to discern land-cover and forest type in reference imagery and the classification image mosaic. The reference imagery included 1 m IKONOS panchromatic sharpened imagery for the US and British Islands and additional 1m color digital ortho quarter quads (DOQQ) for the US islands. Land cover and forest type were then identified in the satellite imagery. Forest type was identifiable in both the reference imagery and the Landsat imagery by color, tone and texture as well as spatial indicators including aspect and elevation. Difficulties distinguishing forest type were encountered in areas that were transitional between semi-deciduous and seasonal evergreen forest. Field survey data proved useful in identifying these transitional areas in the reference imagery.

Training data for the image classification model was derived using the reference imagery and field data collected in the 2005 reconnaissance survey. About 25 to 250 multiple pixel polygons were distributed for each class throughout the extent of the study area. Training data samples collected over a large inter-island extent ensured thorough representation of each class and provided a full range of variability for the class. For example, there are often spectral variations in similar forest types where the image scenes were radiometrically matched in the cloud elimination procedure [44]. Target classes included sunlit and shadowed woody vegetation types, sunlit and shadowed green and senescent pasture, mangrove, wetland, and non-forested classes (Table 1).

We used the woody vegetation classification system designated at the formation level (Table 1) that [45] adapted for Landsat imagery classification from [46]. Areas with less than

Table 1. Classification schema definitions including forest and non-forest classes.

Forest, forest/shrub, woodland and shrubland (Dry and Moist)	***Forest is defined as lands with > 25% cover of tree that co-dominate with shrubs***
Drought Deciduous Woodland	25-60% woody canopy cover with understory of grasses and forbs affected by grazing
Drought Deciduous Young Forest and Forest Shrub	Young drought deciduous secondary forest with Leucaena leucocephala and Acacia muricata common
Drought Deciduous Xeric Coastal Shrubland with Succulents	Very dry drought deciduous shrubland dominated by succulents and exposed soil and rock
Deciduous, Evergreen Mixed Forest and Shrubland with Succulents	Deciduous, drought deciduous and evergreen forest and shrub species (succulents common)
Evergreen Coastal Shrubland	Shrubland with > 75% evergreen species including hemi-sclerophyllous and sclerophyllous species
Semi-Deciduous Forest and Forest and Forest Shrub (Includes Semi-Evergreen Forest)	Stands with 25-75% deciduous woody canopy species
Semi-Deciduous Gallery Forest	Stands with 25-75% deciduous woody canopy species located in drainages
Seasonal Evergreen Forest and Forest Shrub	Stands with ≥75% evergreen woody canopy species (may drop leaves during drought)
Seasonal Evergreen Young Forest and Forest Shrub	Young seasonal evergreen secondary forest
Seasonal Evergreen Gallery Forest	Stands with ≥75% evergreen woody canopy species located in drainages
Seasonal Evergreen Forest with Coconut Palm	Stands with ≥75% evergreen woody canopy species dominated by coconut palm
Forested Wetland	
Mangrove	Mangrove forest
Seasonally Flooded Woodland	Disturbed forested wetland with 25-60% woody canopy cover and seasonal flooding or soil saturation
Agricultural land, pasture, hay, abandoned agriculture, grass	
Herbaceous Agriculture (Cultivated Lands)	Intensive agriculture and cultivated lands where activity is recent
Pasture, Hay, Abandoned Agriculture or other Grassy Areas	Areas with < 25% woody vegetation cover / recreation fields
Golf Course	Golf course
Coastal Grassland	Coastal grassland with < 25% woody vegetation
Non-forested and wetland	
Emergent Wetland	Emergent wetland permanent
Dry Salt Flats	Dry salt flats including mud flats
Quarries	Active or inactive quarries
Coastal Sand and Rock	Coastal sand (beaches) and coastal rock outcrops
Interior Rock (Virgin Gorda boulder)	Virgin Gorda rock outcrops (boulders)
Urban or built-up land	
High-Medium Density Urban	Land with > 80% urban features such as buildings, roads and impervious surfaces
Low-Medium Density Urban	Land with as low as 10-15% urban features and other land-cover types such as pasture or forest

25% woody vegetation cover are classified as pasture and grasslands; these include natural grasslands, abandoned agriculture and grazed or ungrazed pasture. Subtropical drought deciduous forest is defined by the Federal Geographic Data Committee (FGDC) as having at least 75% deciduous woody species [47]. Semi-deciduous forest [5] includes stands with 25-75% deciduous woody canopy species and includes semi-evergreen forest. Drought deciduous woodland has forest and shrub with a canopy cover of 25-60 % and an understory of grasses and forbs dominated by grazing or fire. Young leguminous secondary forest and shrub formations consisting primarily of *Leucaena leucocephala* and *Acacia muricata* were identified at lower elevations where recent or ongoing disturbance has occurred. Seasonal evergreen forest consists of at least 75% evergreen woody canopy species that may drop leaves during drought. Deciduous, evergreen and mixed forest with succulents includes a matrix of deciduous, drought deciduous and evergreen forest and forest shrub species containing succulents including *Stenocereus peruvianus, Leuchtenbergia principis and Opuntia tricantha*. Drought deciduous xeric coastal shrubland with succulents consists of very dry drought deciduous shrubland dominated by succulents and exposed or rocky soil. Evergreen coastal shrubland consists of a least 75% of evergreen species such as hemi-sclerophyllous *Coccoloba uvifera* and may include other sclerophyllous coastal shrub species. Low density urban land includes land with as low as 10-15% urban features and may include a mix of other land-cover types such as pasture, or forest. High-to-medium density urban land has greater than 80% urban features including buildings, roads and impervious surfaces.

2.3 Classification and Image Interpretation

See5 software (www.rulequest.com), a data mining program using decision tree algorithms, was used to predict land-cover and forest type pixel values [48]. In the last several years, decision tree classification techniques have been applied to a wide range of classification problems and have proven to be valuable to the classification of remote sensing imagery due to their flexibility, simplicity and computational efficiency [16,17]. First, a 10 trial adaptive boosting option was employed to improve the overall accuracy and reduce error of the decision tree algorithm by combining many individual classifiers (decision trees) into a single combined classifier [39]. Second, the default global pruning option was used to reduce the likelihood of over fitting the tree to the training data. The pruning process removes parts of the decision tree with relatively high error rates [49].

Manual editing of confused classes was required to correct for residual confusion between urban, barren and pasture areas. Several areas of drought deciduous young forest and drought deciduous woodland that were spectrally confused with pasture and grass were also manually edited. In addition, herbaceous agriculture on St. Croix was manually delineated, which accounted for 279 ha, or 0.01% of the total area mapped. Coastal grassland was manually recoded from pasture, including on small cays. Manual recoding was also necessary to delineate some drought deciduous young forest mostly on the island of St. Croix, and some boundaries between semi-deciduous and seasonal evergreen forest located on St. Thomas, St. John and Tortola. Coastal sand and rock was recoded throughout the classification. Finally, pixels were manually recoded in several high elevation areas of St. Thomas representing low density urban and a few areas of semi-deciduous and seasonal evergreen gallery forest. Areas in the image mosaic that were cloudy in all available Landsat images were manually interpreted from 1 m IKONOS panchromatic sharpened imagery and 1m color DOQQs (about 0.34% of the total mapping area).

2.4 Classification Accuracy Assessment

A stratified random sample was used to create 50 validation points for each land-cover class.

Each accuracy assessment point was verified with high resolution reference imagery and assigned a land-cover class. The reference imagery included IKONOS 1 m panchromatic sharpened imagery and 1m color DOQQs. An error matrix was created for each mapped class to estimate the overall percentage of correctly classified pixels, statistics for producer and user's accuracy and the Kappa coefficient, which is an indicator of the accuracy of chance agreement between classes [50,51]. Producer's accuracy is the proportion of correctly classified accuracy assessment estimates and user's accuracy estimates the proportional assignment of pixels to a correct class [5].

2.5 Lidar Data Processing

Discrete lidar data were collected during January and February 2004 by 3001 Inc. under contract to the US Army Corps of Engineers, using a Leica Geosystems ALS 40 sensor [52]. Geodetic control was established by a static GPS network covering the lidar survey area of Puerto Rico and the US Virgin Islands. In addition to data collection, Real-Time Kinematic GPS surveys were conducted to establish a network of ground truth data for statistical comparisons with the lidar data. The results of the comparisons indicate a Vertical Root Mean Square Error (RMSEz) of 9.26 cm on level smooth surfaces [52]. The spatial extent of the data in this study includes St. John and St. Thomas and surrounding small islands and cays. The lidar data were collected in conjunction with digital photos and used to generate and improve digital elevation models (DEM) for orthorectification. The raw point cloud data were provided by the contractor in .xyz format with a 2.76 m shot spacing and consists of multiple return measurements including first, last and intermediate return values.

The data were filtered into ground (minimum elevation) and non-ground returns to create a bare earth DEM and forest height estimates. TIFFS (Toolbox for Lidar Data Filtering and Forest Studies) was used to process the lidar data for extracting a bare earth DEM [53]. The filtering method used by TIFFS to create the DEM used a progressive morphological filter for removing non-ground measurements from the lidar elevation data [54]. Morphological filtering composes operations based on set theory to extract non-ground features from an image. The two fundamental operations include dilation and erosion which are used to enlarge (dilate) or reduce (erode) the size of features in continuous surfaces. By gradually increasing the window size of the filter and using elevation thresholds, the measurements of non-ground features such as vegetation and buildings are eliminated while topographic data are preserved [54]. The morphological filter algorithm used in [55] incorporates the assumption that non-ground objects such as buildings exhibit abrupt elevation changes while topographic elevation is gradual and continuous. This method is adaptive to local terrain and is applicable to rugged topography [53]. Once non-ground features were removed, terrain points were extracted from the approximated surface and a 5 m DEM was interpolated. Finally, the DEM tiles were mosaicked to form a continuous surface of the study area.

Programs written in IDL (Interactive Data Language, ITTVIS, 2008) were used to create a canopy height dataset from the non-ground elevation data [56]. The corrected DEM provided a minimum elevation surface which was subtracted from non-ground elevation lidar data to estimate canopy height. The resulting output was a continuous multiband image mosaic of canopy height variables with a 30 m pixel resolution. The image bands were of canopy height statistics including: (1) height percentiles (P-tile) below which certain percentages of data fall; (2) shot return profiles (SRP) that quantify the number of shots returned from 5 m height bins; (3) quadratic mean height (QMH) the root mean square of canopy lidar point height [57]; (4) maximum height; and (5) mean height (Table 2). Other areas where clouds obstructed the land and lidar measurements were treated as "no data". The cloud obstructed

area was about 180 hectares or 1.4% of the study area and was confined mostly to the island of St. John. In addition, a water mask created from USGS 1:24,000 scale DLG data representing coastlines was used to mask ocean elevation measurements [40].

Table 2. USDA Forest Service, Forest Inventory and Analysis data, lidar and environmental variables.

FIA Plot Variables	**Units/Type**	**Definition**
Average Height of all Trees (AHT)	m	Average height of all trees
Average Height of Dominant/Co-dominant Trees (HDCD)	m	Average height of dominant and co-dominant trees
Above Ground Biomass per Hectare (AGBH)	Mg/ha	Above ground live biomass per hectare
Crown Volume (CV)	m^3	Sum of the volume of all crowns for trees with a d.b.h. ≥ 12.5 cm. Estimated as an ellipsoid using crown ratio, tree height and crown radius in two perpendicular directions (V= 4pi/3)abc
Lidar Variables		
Minimum Elevation	m	Minimum lidar surface elevation
Maximum Height	m	Maximum lidar canopy height
Mean Height	m	Mean lidar canopy height
Quadratic Mean Height (QMH)	m	Root mean square height of the lidar points
Height Percentiles (P-tile)	m	Height at which a certain percent of data fall below
Shot Return Profile (SRP)	m	Number of shots returned from 5 meter height bins
Environmental Variables		
Landsat Bands 1-5, 7	Integer	Landsat image reflectance bands
Landsat NDVI, 4/5 Ratio	Float	Normalized Difference Vegetation Index, Band 4/Band 5
Aspect	Degrees	Aspect expressed in degrees
Sine Aspect	Radians	Sine of aspect in radians
Cosine Aspect	Radians	Cosine of aspect in radians
Curvature	Integer	Slope geometry indicating convex or concave geometry
Slope	Degrees	Slope expressed in degrees
Slope Position	Integer	Ridge or valley of any point in landscape
Degrees from North	Degrees	Degrees from 0 (north)
Landsat Classification	Nominal	Land-cover and forest type thematic values
Precipitation	mm/year	Total annual precipitation
Temperature	mm/year	Mean annual temperature

2.6 Forest Parameter Modeling

Regression models of field estimated canopy height and biomass were developed from lidar estimates of canopy height and environmental variables. The plot data was collected by the USDA Forest Service Forest Inventory Analysis (FIA) Caribbean program on St. John and St. Thomas in 2004 [58] providing field based forest structural information (Table 2). Each plot consists of four 7.3 m radius circular subplots in which all woody vegetation with a diameter at breast height (DBH, measured at 1.37 m) of $\geq$ 12.5 cm is surveyed. A single microplot with a 2.1 m radius nested within each subplot is used to survey woody vegetation saplings with a DBH between 2.5 and 12.5 cm [59]. Summary statistics were generated at the plot and subplot level. For the plot level data, circular and square plot extraction schemes were used to extract the lidar mosaic data using a 90 m window. In addition to the lidar data, this study examined if other explanatory variables such as ancillary environmental data improved the regressions and fit of the models of structural data (Table 2). Ancillary data included Landsat reflectance bands (bands 1-5, 7) and band indices (NDVI and the ratio of bands 4 and 5), total annual precipitation and mean annual temperature (Helmer, Daly and Plume, unpublished data), land-cover and forest type, and elevation and elevation derivatives such as slope, aspect, and sine and cosine of aspect expressed in radians. Two other topographic indices were included: slope position, which calculates the extent that each point is similar to a ridge or valley position as values 0 through 100 [60], and standard curvature, which is a measure of slope geometry indicating convex or concave topography [61]. A 24 band image stack containing the lidar canopy height statistics and the environmental variables was assembled, and data were extracted at every FIA plot and subplot location. The sample size for the plot and subplot level data was 18 and 72 observations respectively.

We developed stepwise regression models in JMP software (www.sas.com) and used the resulting regression equations to create images of predicted values with a 30 m pixel resolution for average height of dominant and co-dominant trees (HDCD), and above ground biomass per hectare (AGBH) on the islands of St. John and St. Thomas [62] (Table 3). Forest structural statistics were then calculated using a zonal operation for the predicted maps of HDCD and AGBH and related to the areas of the forest types mapped in the land-cover classification for St. John and St. Thomas.

3 RESULTS

3.1 Land-cover and Forest Type Classification

The land-cover and forest type classification consisted of 29 classes (Figure 2 and Appendix A and B). After manual editing and interpretation of residual cloudy areas the overall accuracy was 72%. The Kappa coefficient of agreement was 0.76 ± 0.01, which indicates a significant agreement between the reference and map classifiers. The main sources of error were confusion between low density and high-medium density urban lands, and between low density urban lands and pasture. Most of the forest types were classified with greater than 70% accuracy (Appendix A). However, some confusion occurred between semi-deciduous forest and seasonal evergreen forest. Also, deciduous, evergreen and mixed forest and shrubland with succulents showed some confusion with semi-deciduous forest, drought deciduous young forest shrub and woodland classes.

Table 3. Stepwise regression equations for forest structural response variables using Forest Inventory Analysis fully forested plot level data for the islands of St. John and St. Thomas. The predicted models are significant at $p < 0.05$ and 18 observations. Superscript symbols indicate significance levels of overall model: ^{o} ($p \leq 0.0001$), $^{+}$ ($p \leq 0.001$), $^{\Delta}$ ($p \leq 0.05$)

Response variable	Explanatory data (lidar / environmental variables)[1]	Predicted model	RMSE	Adj. R^2
Average height of dominant/co-dominant trees (HDCD)	SRP 2, SRP 6, PRECIP^{o}	HDCD = -14.03+ 10.67(SRP2) + 24.10(SRP6) + 0.00015(PRECIP)	0.78	0.73
Above ground biomass per hectare (AGBH) mapped	SRP 5$^{\Delta}$	AGBH = 58.29 + 220.17(SRP5)	22.78	0.36
Above ground biomass per hectare (AGBH) mapped	Observed HDCD^{+}	AGBH = -22.92 + 14.27(Observed HDCD)	15.29	0.72[2]
Above ground biomass per hectare (AGBH) mapped	Predicted HDCD$^{\Delta}$	AGBH = -9.89 + 12.37(Predicted Formula HDCD)	22.69	0.37[2]
Average height of all trees (AHT)	MEANH, COSAP, LSB 1, 5, 8$^{\Delta}$	AHT = 3.14 + 0.35(MEANH) + 0.69(COSAP) - 0.13(LSB1) - 0.084(LSB5) + 0.48(LSB8)	0.75	0.59
Crown volume (CV)	SRP 2, SRP 6, COASP^{+}	CV = 87.44 – 646.612(SRP4) + 1959.04(SRP6) + 76.69(COASP)	51.48	0.77
Height Penetration Index (HPI)	ASPECT, SLOPE, SLPPOS, PRECIP$^{\Delta}$	HPI = -0.87 - 0.0006(ASPECT) – 0.005(SLOPE) + 0.002(SLPPOS) + 1.28e-5(PRECIP)	0.09	0.56

[1] Lidar variable definitions: SRP (Lidar Shot Return Profile), PRECIP (Average Annual Precipitation mm/yr), MEANH (Lidar Mean Canopy Height), COSAP (Cosine of Aspect in radians), LSB (Landsat TM Band (band 8 = 4/5 ratio)), SLOPE (Slope in degrees), ASPECT (Aspect in degrees), SLPPOS (Slope position).
[2] Linear equations are stated in R2

Confusion among these classes can be explained by similarities in deciduous and drought deciduous forest shrub and woodland species. Finally, some confusion occurred between pasture, drought deciduous woodland and drought deciduous young forest shrub. Pasture often exhibits confusion with drought deciduous forest types in Landsat classifications due its composition of up to 25% drought deciduous woody vegetation species [5]. The mapped combined closed woody vegetation (closed forest and associated classes) for the seven main islands and their associated islets and cays, was 34,175 ha, which encompassed about 67.7% of the total mapped island area (Appendix A). Pasture, hay, abandoned agriculture or other grassy areas was the second most abundant class representing 12.2% of the total mapped area. The total urban and developed land area was about 6,367 ha in 2000 (Appendix B).

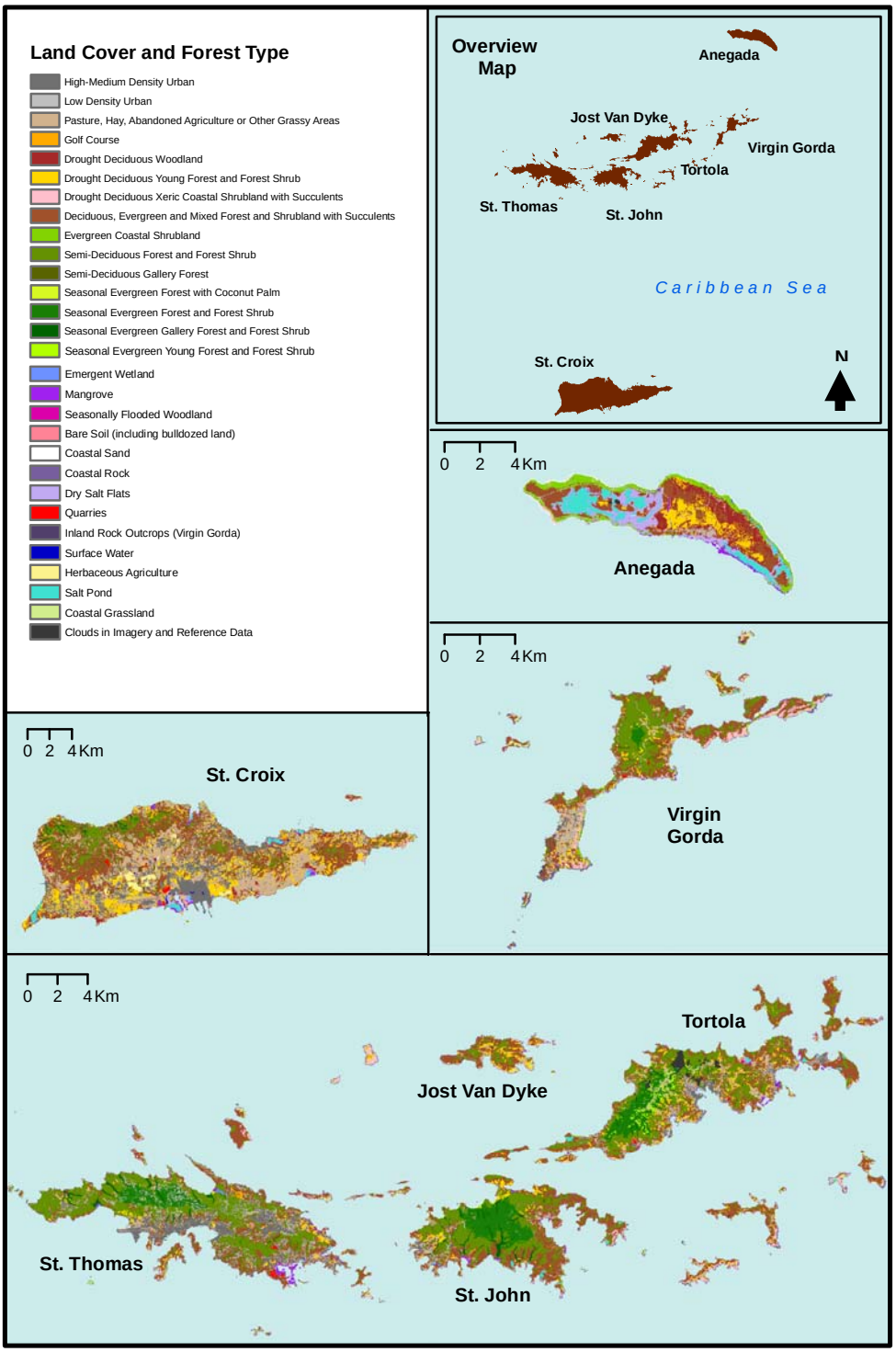

Fig. 2. Land cover and forest type classification for the United States and British Virgin Islands

3.2 Lidar Data Processing

The TIFFS algorithm efficiently filtered the large volume of point cloud data, although the processing required manual parameter tuning and multiple iterations to produce an adequate DEM. The mean shot spacing was one shot per 2.76 m. Mean canopy height ranged between 2.0 and 16.9 m with a mean and standard deviation of 5.4 and 2.2 m. The mean canopy height image with outliers removed shows a realistic representation of mature moist forest stands and short drought deciduous and xeric forest types. The minimum and maximum value of the elevation image before correction was -82.4 and 970.7 m with a mean and standard deviation of 14.0 and 48.3 m respectively. After correction using the TIFFS software, the minimum and maximum bare earth DEM values were -4.094 and 470.5 m with mean and standard deviation of 3.3 and 68.6 m. Negative values resulted from ocean shots that were not completely removed in the masking process due to minor edge matching differences in the water mask and the lidar data.

3.3 Estimating and Modeling Forest Structural Parameters

Models at both the plot and subplot levels that used only lidar indices were statistically significant but did not explain as much variance as other studies using lidar (adjusted R^2 for plot level HDCD = 0.49 and RMSE 1.07 m; adjusted R^2 value for subplot level HDCD = 0.18 and RMSE 1.79 m). Further data analysis indicated the existence of a regular pattern between the regression residuals, canopy closure and environmental conditions. To quantify these effects, we calculated a height penetration index as an indication of the distance that the lidar penetrated through the forest canopy. The height penetration index is defined as average height of dominant and co-dominant trees divided by lidar maximum height. Values of this index near 1.0 represent canopies where most returns are from dominant and co-dominant trees; values greater than 1.0 indicate higher level of penetration into the canopy. A stepwise regression of the height penetration index and related environmental variables show a strong dependence (adjusted $R^2 = 0.56$, Table 3). The environmental variables that were significant in the regression include aspect, slope, slope position and precipitation and are a major component of environmental gradients found in Caribbean landscapes that influence forest types and structure. Given the direct relationship between environmental variables and this characteristic of forest canopy structure, environmental variables were added as dependent variables.

Models of forest structural parameters that included environmental variables were significant at both the plot and subplot level, but plot level relationships explained more variance. For example, the adjusted R^2 value was 0.40 for the model of subplot HDCD but 0.73 for the plot level model. At the plot level, plot extraction using circular and square sample areas both provided statistically significant models, but the circular plot scheme explained more variance (adjusted R^2 for HDCD circular plot = 0.73; adjusted R^2 for HDCD square plot = 0.59).

The relationships between the lidar and environmental variables, and the inventory plot level measurements of HDCD are strong, but most of the environmental variables tested were insignificant. Regressions of HDCD with lidar shot return profile (SRP) variables and precipitation explain high levels of variance, with an adjusted R^2 of 0.73 (Table 3, Figure 3). The RMSE of 0.78 m for HDCD is 11.35% of the response mean of 6.87 m which is a relatively low error. The regression equation for HDCD is dependent on the explanatory variables shot return profile 2 (SRP2), shot return profile 6 (SRP6) and the environmental variable total annual precipitation resulting in the following multiple regression model with all explanatory variables significant at $p<0.05$ (Table 3):

$$HDCD = 14.03 + 10.66SRP2 + 24.10SRP6 + 0.00015PRECIP \quad (P < 0.0001) \qquad (1)$$

The predicted model for inventory HDCD explained a greater amount of variance and produced a more reasonable height map than did the equation for predicted Average Height of all Trees variable (adjusted $R^2 = 0.59$, Table 3). This suggests that the coarse resolution lidar is more sensitive to the height of the dominant and co-dominant trees than the average height of all trees and can be attributed to the inability of the lidar returns to account for the variability of the forest understory.

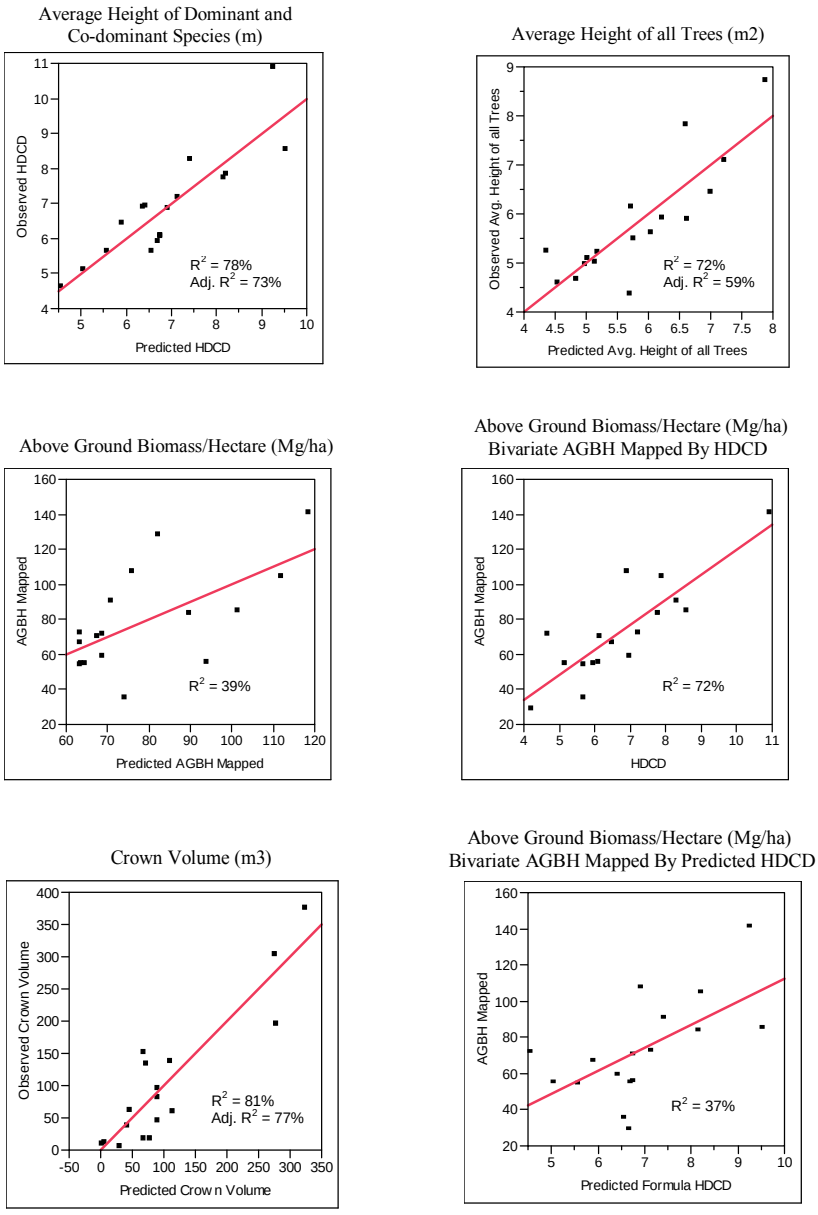

Fig. 3. Graphs of predicted versus observed forest structural variables

The relationship between the lidar data, environmental variables, and the inventory plot level measurements of AGBH explain only 36% of the variance. No environmental variables were significant and SRP 5 was the only significant explanatory variable in overall model with a p-value of 0.01 (Table 3). The model did not adequately map low biomass values in dry forest and forest shrub formations and did not consistently map biomass values in adjacent tall and short forested areas. Instead, an alternative approach that uses the strong relationship between AGBH and HDCD in the inventory plot dataset was applied. A bivariate fit between inventory HDCD and AGBH results in the following linear regression model explaining 72% of the variation with RMSE (15.29 m) (Table 3, Figure 4).

$$AGBH = -22.92 + 14.27HDCD \quad (P < 0.0001) \tag{2}$$

The map of the AGBH model shows the same trends as the predicted map of HDCD due to their direct linear dependence. Biomass estimates generally increase with an increase in elevation and canopy height (Figure 4). A few areas of predicted biomass located primarily in the island's high elevation ridge topography were estimated beyond the maximum range of the observations that form the regression model and exceeded values reasonable for subtropical moist forest in the region. A threshold was determined based on the largest biomass estimate (141.3 Mg/ha) from FIA plot data located in St. John. Predicted estimates exceeding the threshold were reclassified in the map to this maximum value. Larger biomass estimates are not reported on St. John, though dense mature moist forest stands in the drainages could show values exceeding the threshold. Modeled biomass values that exceed the range of reasonable observations occur on high elevation ridges; a location where biomass is unlikely to be near the FIA maximum. [63] summarized biomass estimates of Caribbean dry and moist forest from recent studies showing values for the Cinnamon Bay watershed measured after hurricane Hugo in 1989 to be about 131.5 tons/ha. Finally, [64] presents pre-hurricane Gilbert average biomass estimates for Rancho San Filipe, Mexico of 132 Mg/ha in forests similar to St. John, but consisting of drier forest types. In addition, a few low elevation areas consisting of mostly drought deciduous xeric coastal shrubland with succulents were modeled in the AGBH map with negative values due to the y intercept of the model at about 1 m. To correct this error, a minimum biomass estimate threshold was applied to the negatives values based on a minimum HDCD height of 1.2 m.

The relationship between the lidar data, environmental variables, and the inventory plot level measurements of Crown Volume explain 81% of the variance with RMSE (51.48 m). Although, crown volume was accurately estimated in the regression models, the predicted maps provided a poor representation of this variable throughout the study area showing large non-contiguous areas of negative and erroneous values at both high and low elevations and therefore not included as a final predicted map.

3.4 Forest Structural Summaries

Forest structural summaries derived from the predicted maps for HDCD and AGBH show height and biomass statistics for the islands of St. John and St. Thomas and include mapped drought deciduous, xeric and coastal classes located on the surrounding small islands and cays (Tables 4A and B). The mean height and biomass estimates are similar to FIA field estimates of moist forest types. Height and biomass estimates for seasonal evergreen, and semi-deciduous forest show a slight increase in the maximum and mean forest height and total biomass in St. John compared to St. Thomas in protected areas. This result may show the effect of the protected status of seasonal evergreen and semi-deciduous forest within the

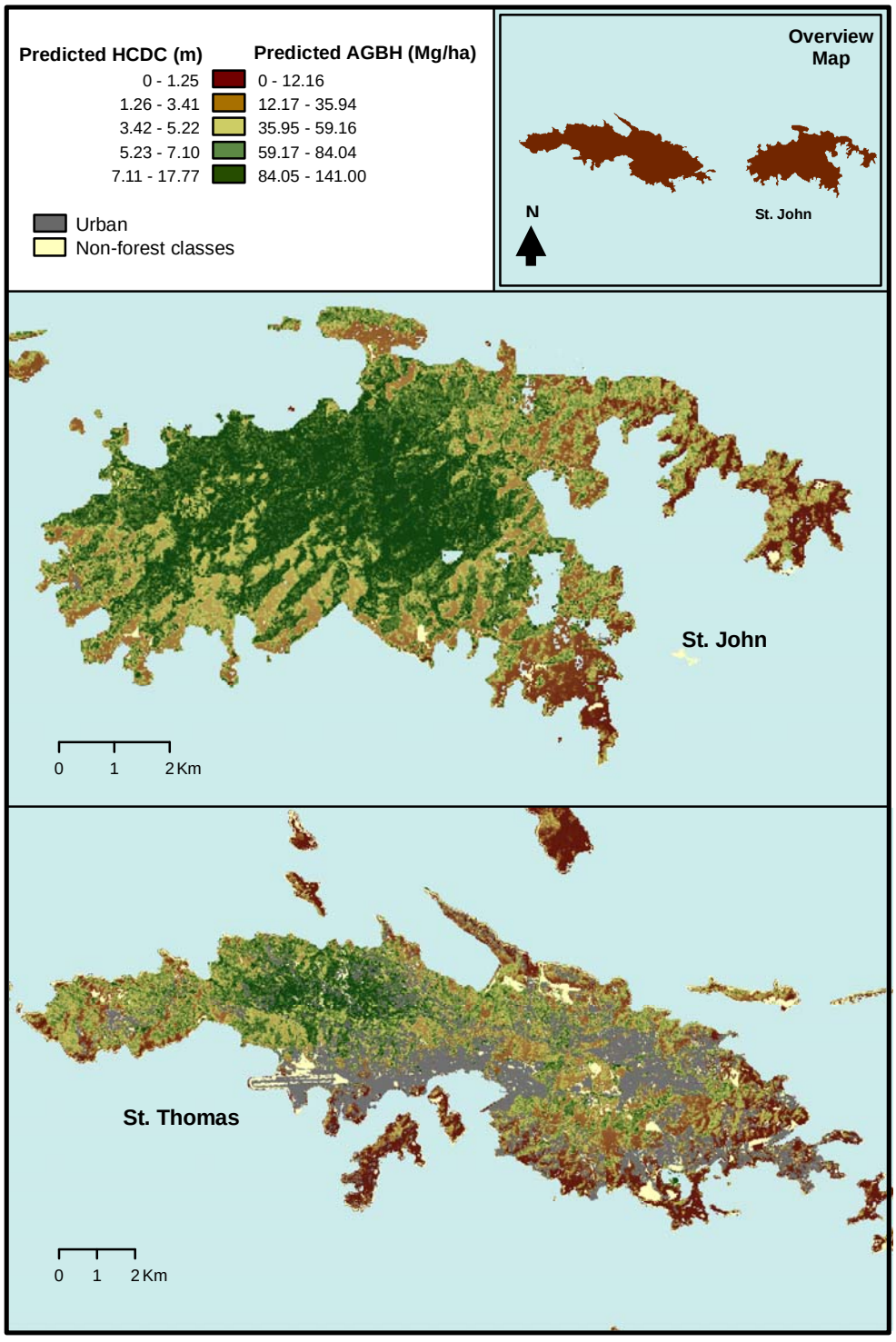

Fig. 4. Map of predicted average height of dominant and co-dominant trees (HDCD) and above ground biomass per hectare (AGBH) for the islands of St. John and St. Thomas.

VINP where these larger height values are identified. Drought deciduous xeric coastal shrubland and evergreen coastal shrubland has the smallest mean HDCD and AGBH estimates and show reasonable values for the short shrubland classes. Mean HDCD estimates for the largest dry forest formation, deciduous evergreen mixed forest and shrubland with succulents was 4.4 m for St. John and 3.0 m for St. Thomas with mean AGBH estimates of 39 Mg/ha and 21 Mg/ha respectively.

Table 4A. Forest structural summaries for average height of dominant and co-dominant trees for the mapped forest types on St. John and St. Thomas including surrounding small islets and cays. Slight discrepancies in forest type area between the predicted lidar maps and the land-cover classification are due to "no data" masks in the lidar data.

Average height of dominant and co-dominant trees (HDCD) (m)	**St. John**				**St. Thomas**			
Forest formation mapped	Min	Max	Mean	STD	Min	Max	Mean	STD
Drought Deciduous Woodland	-	-	-	-	1.3	6.9	3.4	1.8
Drought Deciduous Young Forest and Forest Shrub	1.9	11.0	4.3	1.5	0.1	12.2	3.1	1.7
Drought Deciduous Xeric Coastal Shrubland with Succulents	1.3	9.1	2.7	1.4	0.0	10.8	2.0	1.4
Deciduous, Evergreen Mixed Forest and Shrubland, with Succulents	1.3	13.9	4.4	1.6	0.0	12.2	3.0	1.6
Evergreen Coastal Shrubland	1.1	9.5	2.9	1.3	0.0	7.8	2.3	1.5
Semi-Deciduous Forest and Forest Shrub	1.5	15.6	6.2	1.4	0.6	13.6	4.8	1.4
Semi-Deciduous Gallery Forest	2.2	12.5	5.9	1.6	1.7	11.4	5.4	1.3
Seasonal Evergreen Forest and Forest Shrub	4.0	16.1	7.5	1.2	2.9	13.1	6.6	1.3
Seasonal Evergreen Young Forest and Forest Shrub	5.5	8.6	6.0	1.1	4.0	8.3	4.5	1.1
Seasonal Evergreen Gallery Forest	3.4	15.5	7.3	1.5	2.1	13.0	6.2	1.2
Seasonal Evergreen Forest with Coconut Palm	3.7	10.7	7.4	1.5	1.5	10.1	5.8	1.4
Mangrove	1.4	9.4	4.8	1.6	0.0	9.8	2.7	1.9
Seasonally Flooded Woodland	1.9	9.2	4.4	1.5	1.3	10.7	3.7	2.2

Table 4B. Forest structural summaries for above ground biomass per hectare for the mapped forest types on St. John and St. Thomas including surrounding small islets and cays. Slight discrepancies in forest type area between the predicted lidar maps and the land-cover classification are due to "no data" masks in the lidar data.

Above ground biomass per hectare (AGBH)	**St. John (Mg/ha)**				**St. Thomas (Mg/ha)**				**Both islands**
Forest formation mapped	Min	Max	Mean	STD	Min	Max	Mean	STD	Total (Gg)
Drought Deciduous Woodland	-	-	-	-	1	75	27	24	0
Drought Deciduous Young Forest and Forest Shrub	4	134	39	21	0	141	24	22	131
Drought Deciduous Xeric Coastal Shrubland with Succulents	0	108	17	18	0	132	11	16	34
Deciduous, Evergreen Mixed Forest and Shrubland, with Succulents	0	141	39	23	0	141	21	20	1,057
Evergreen Coastal Shrubland	0	112	19	18	0	88	14	18	22
Semi-Deciduous Forest and Forest Shrub	1	141	65	20	0	141	46	20	2,426
Semi-Deciduous Gallery Forest	8	141	61	22	1	140	54	19	65
Seasonal Evergreen Forest and Forest Shrub	34	141	84	17	18	141	71	19	1,088
Seasonal Evergreen Young Forest and Forest Shrub	55	100	77	15	34	95	62	15	4
Seasonal Evergreen Gallery Forest	25	141	81	20	7	141	66	18	262
Seasonal Evergreen Forest with Coconut Palm	30	130	81	21	1	121	59	20	14
Mangrove	0	111	45	23	0	118	19	25	43
Seasonally Flooded Woodland	4	108	40	21	1	129	31	31	5

4 DISCUSSION

4.1 Land-cover Classification Techniques

The use of decision tree classification techniques to map land cover and forest types in subtropical environments has been highly successful in past studies [17,4,5]. This study uses the techniques tested in previous work to create the first Landsat ETM+ land-cover and forest type classification for the US and British Virgin Islands. Unique to this study is that for the first time we created 15-m panchromatic-sharpened cloud-free mosaics and used them in a classification to simultaneously map land-cover and forest types over a large inter-island extent.

The classification approach provided several advantages over traditional techniques in this study area. First, the Principal Components transformation for creating 15 m panchromatic sharpened Landsat imagery was effective for increasing the spatial resolution of the Landsat multi-spectral bands. The resulting imagery enhanced the ability to discern sparsely vegetated surfaces, urban features in low density urban areas and linear features such as riparian corridors. However, difficulty in distinguishing areas of less than a few pixels, like very small forest patches, small clearings, or man-made structures in forest still proved difficult due to the moderate resolution of the panchromatic sharpened imagery.

The second advantage of this approach was the utility of the mostly cloud-free image mosaic for training data collection and classification. The technique devised by [6] provided a repeatable method for replacing the cloudy areas with images from different dates in which vegetation phenology is generally normalized to the base scene. This process makes training data collection in the cloud-filled areas easier by providing a single image mosaic with minor differences in image tone [5].

4.2 Estimating and Modeling Forest Structural Parameters

Stepwise regression was effective for estimating canopy height, as measured in FIA plot data, from lidar canopy height estimates and environmental variables. Testing different multiple regression models showed that the large point spacing of the coarse resolution lidar is most suitably modeled at the plot level, with a 90 m window circular plot extraction design. In addition, the large point spacing of the lidar data is most appropriate for height measurements of dominant and co-dominant trees, due to the inability to accurately measure forest understory. This limitation is the most likely reason why the models of above ground biomass per hectare were inaccurate. A significant portion of the biomass in Caribbean landscapes is located in the forest understory and contributes greatly to overall biomass estimations. The proportion of biomass contribution decreases with an increase in average tree height and is calculated to be about 94% based on a FIA plot with an HDCD of 4.1 m and 24% for a FIA plot with a HDCD of 10.9 m. This assumption is further supported by a stepwise regression using lidar and environmental variables to estimate field measurements of above ground biomass that don't include the biomass contribution measured on microplots (trees and shrubs with diameters < 12.5 cm). The regression to estimate biomass without the microplot contribution was significantly related to lidar SRP 1, SRP 6, QH and Landsat TM band 1, explaining 81% of the variance compared to only 36% of the variance for AGBH with scaled micro plot. This supports the conclusion that the coarse resolution of the lidar data is more sensitive to the larger dominant forest structure located in the upper canopy.

The regression equation for the "Height Penetration Index" and the importance of environmental variables for predicting forest structure suggests that canopy structure varies substantially with environmental conditions in this study area. Field observations and photo interpretation show how forest structure on these islands changes with aspect and elevation, which also influence solar input, wind, and climate. Because of these interactions, estimating/mapping forest structure in Caribbean environments may require stratification of field plots by forest type, which will increase the cost of field work for these studies by the number of strata. In this study, which used FIA data collected with systematic sampling, we did not have this variety of sites across the main environmental gradient. An alternate sampling scheme should allow the appropriateness of using environmental variables in this way.

The HDCD and AGBH models yield reasonable maps of those variables for St. John and St. Thomas across the range of elevations, showing a trend of increasing canopy height and biomass with an increase in elevation (Figure 4). Field observations and reference data confirm the mapped results, which show taller forest stands associated with higher elevations as well as slopes and drainages sheltered from prevailing winds and solar radiation. Canopy height and biomass estimates on windward south and southeast aspects are smaller than those on lee slopes. While most low lying areas with relatively short forest types are accurately predicted in the HDCD model, several short forest types estimated adjacent to coastal areas, including drought deciduous xeric shrubland with succulents and evergreen coastal shrubland, have scattered pixels that exceed reasonable heights. This result may be due to isolated large trees, land-cover classification errors, or lidar height errors from non-forest features. Also, a few of the large watershed drainages in the predicted maps show taller forest stands and larger biomass estimates than have been identified in FIA surveys and reference imagery, while some of the watershed drainages at lower elevations depict height estimates that may be too short (Figures 2 and 4). This discrepancy may be caused by high density forest structure and the inability of the lidar to penetrate the canopy to estimate accurate minimum elevation data or the lack of multiple returns for the lidar data in some areas. Flat urban features such as recreation fields, golf courses, and the St. Thomas airport runway as well as some pasture are predicted as zero height in the maps, although urban features such as buildings depict reasonable height estimates.

The forest structural summaries provide a quantitative overview of the average height and biomass of the forest types in the study area. Seasonal evergreen and semi-deciduous forests including gallery forests, which represent the dominant moist forests on St. John, and St. Thomas, had the largest mean height and biomass. Semi-deciduous forest on both islands accounted for the largest total combined biomass. These moist forest types contain the tallest forest stands and greatest biomass, providing an indicator of structure, species richness and habitat suitability, and highlighting their importance for protection status. Mean maximum height and biomass for seasonal evergreen forest and semi-deciduous forest are slightly larger on St. John compared to St. Thomas, especially for the seasonal evergreen forests that are well represented under protection status (Table 4A and B). In contrast, drought deciduous xeric coastal shrubland with succulents and evergreen coastal shrubland forest types depict the smallest mean height and biomass estimates. Low elevation deciduous evergreen mixed forest and shrubland with succulents represents the largest biomass estimate for the dry (drought deciduous) forest types. Incidentally, in the Caribbean, the unprotected low elevation dry forests represent the greatest danger of deforestation due to increasing developmental pressure and are considered critical habitat for many endemic vegetation species and important habitat for Neotropical migratory birds [65].

Prior to this study, applications for low density lidar data include topographic modeling, feature extraction and floodplain and coastal mapping [66,67]. For example, the Louisiana Statewide Lidar Project was initiated due to increased flood loss rates experienced by the FEMA National Flood Insurance Program and provides low cost lidar derived high resolution topographic data to update floodplain maps [67]. The project incorporates the Leica Geosystems ALS 40 lidar mapping system used in this study with 3 m point spacing and the subsequent acquisition of aerial photography to develop products pertaining to floodplain mapping at a regional and watershed scale. While the use of coarse resolution discrete lidar to map topographic features is known, this study shows new applications for this type of data providing a cost effective technique to map forest structure in a subtropical environment that can be applied to forestry applications in other settings.

5 CONCLUSION

Decision tree classification using cloud free image mosaics of panchromatic sharpened Landsat ETM+ images proved effective for mapping land cover and forest types in the Virgin Islands providing more detail than previous mapping efforts. The overall accuracy of the 29 forest and non-forest classes was 72%, with most of the forest types classified with greater than 70% accuracy. The mapped combined closed woody vegetation for the seven main islands and their associated islets and cays, was 34,175 ha, which encompassed about 67.7% of the total mapped island area.

The coarse resolution discrete lidar data accurately modeled forest structure including average height of dominant and co-dominant trees, average height of all trees and above ground biomass per hectare on the islands of St. John and St. Thomas, where forest types vary dramatically with topography and environmental factors. However, due to the large point spacing, the lidar data is more indicative of upper canopy dominant and co-dominant tree height as opposed to the average height of all trees. The data did not completely account for variability in the forest understory. This limitation prevents accurate biomass estimates using a stepwise regression approach due to lack of important understory contributions of woody vegetation. However, as this study concludes, biomass can be estimated by its direct relationship with canopy height.

This study shows that mapping forest height and biomass can be performed from coarse resolution discrete return lidar sensors in Caribbean landscapes. The resulting quantification of forest structure enables better characterization of the forest types from a passive optical image classification. In addition, this study shows that regression modeling of forest height and biomass can be performed using limited plot data that does not represent the complete range of height values found in the forest types. Another major advantage of this type of lidar data is the relatively low cost of data acquisition, because this type of data are often collected in conjunction with high resolution airborne digital photos. Additional studies are necessary to further test how field plot sampling design may improve the models. This may include stratification of field plots by forest type, accounting for a comprehensive range of heights for each environmental condition. However, this approach will increase the cost of field work based on the number of strata assigned.

Acknowledgements

This project was funded by the U.S. Forest Service (USFS) International Institute of Tropical Forestry (IITF) State and Private Forestry Program. The work is a contribution to the Caribbean Forest Inventory and Analysis (FIA) Program and Southern Research Station Forest Inventory Analysis Program and was completed in cooperation with the USFS Rocky Mountain Research Station (RMRS) and Colorado State University's Department of Forestry, Rangeland and Watershed Stewardship. Thanks also to Humfredo Marcano-Vega for assistance with identification of forest types. We also thank Dr. David Theobald for input and review, and expert advice.

Appendix A. Forest and non-forest areas including the User's and Producer's accuracies for the 2000 U.S. Virgin Islands land-cover and forest type classification. Area totals rounded to integer values (Area totals include small islands and cays associated with major islands).

Forest and non-forest classification name	User's Accuracy[1] (%)	Producer's Accuracy[1] (%)	St. Croix (ha)	St. John (ha)	St. Thomas (ha)	Total Area (ha)
Forest, forest/shrub, woodland and shrubland						
Drought Deciduous Woodland	63	81	409	0	1	554
Drought Deciduous Young Forest and Forest Shrub	70	54	2974	190	211	3499
Drought Deciduous Xeric Coastal Shrubland with Succulents	73	62	93	85	160	473
Deciduous, Evergreen Mixed Forest and Shrubland, with Succulents	84	60	6153	1558	1794	9649
Evergreen Coastal Shrubland	82	79	117	44	101	423
Semi-Deciduous Forest and Forest Shrub	77	65	1770	1584	2587	6083
Semi-Deciduous Gallery Forest	74	90	528	41	64	797
Seasonal Evergreen Forest	84	72	0	783	453	1392
Seasonal Evergreen Young Forest and Forest Shrub	62	88	0	1	6	157
Seasonal Evergreen Gallery Forest	76	85	128	175	147	611
Seasonal Evergreen Forest with Coconut Palm	78	95	1	11	8	193
Forested Wetland						
Mangrove	83	66	185	48	105	487
Seasonally Flooded Woodland	59	81	0	8	4	152
Urban or built-up Land						
High-Medium Density Urban	78	78	2747	80	1040	4023
Low-Medium Density Urban	68	67	399	219	798	1551
Agricultural Land, Pasture hay abandoned agriculture or other grassy areas						
Herbaceous Agriculture – Cultivated Lands[1]	86	100	275	0	0	461
Pasture hay abandoned agriculture or other grassy areas (i.e. soccer fields)	74	55	5173	54	261	5617
Golf Course	89	95	86	0	28	298
Coastal Grassland	70	83	53	17	39	262
Non-forested and Wetland						
Emergent Wetlands	66	89	12	1	17	185
Dry Salt Flats (Includes mud flats)	76	81	135	9	2	303
Quarries	94	98	62	0	28	282
Coastal Sand	79	89	104	38	35	345
Coastal Rock	61	87	61	96	242	547
Interior Rock (Virgin Gorda boulder)	74	89	0	0	0	163
Bare Soil (including bulldozed land)	78	89	37	5	16	225
Salt Pond	90	75	232	42	20	459
Surface Water	80	97	61	1	3	242

[1]User's and producer's accuracies include both US and British Virgin Islands

Appendix B. Forest and non-forest areas for the 2000 British Virgin Islands land-cover and forest type classification. Area totals rounded to integer values (Area totals include small islands and cays associated with major islands).

Forest and non-forest classification name	Anegada (ha)	Jost Van Dyke (ha)	NPCG[1] (ha)	Tortola (ha)	Virgin Gorda (ha)	Total Area (ha)
Forest, forest/shrub, woodland and shrubland						
Drought Deciduous Woodland	643	0	0	1	8	652
Drought Deciduous Young Forest and Forest Shrub	440	157	69	454	109	1229
Drought Deciduous Xeric Coastal Shrubland with Succulents	28	103	162	68	209	570
Deciduous, Evergreen Mixed Forest and Shrubland, with Succulents	841	422	464	2044	808	4579
Evergreen Coastal Shrubland	544	12	28	90	82	756
Semi-Deciduous Forest and Forest Shrub	0	156	30	1824	637	2647
Semi-Deciduous Gallery Forest	0	3	0	29	22	54
Seasonal Evergreen Forest	0	0	0	633	38	671
Seasonal Evergreen Young Forest and Forest Shrub	0	0	0	149	0	149
Seasonal Evergreen Gallery Forest	0	0	0	87	0	87
Seasonal Evergreen Forest with Coconut Palm	0	0	1	4	3	8
Forested Wetland						
Mangrove	91	4	2	68	10	175
Seasonally Flooded Woodland	0	0	0	10	0	10
Urban or built-up Land						0
High-Medium Density Urban	6	5	3	303	47	364
Low-Medium Density Urban	107	28	12	426	147	720
Agricultural Land, Pasture hay abandoned agriculture or other grassy areas						
Herbaceous Agriculture – Cultivated Lands[1]	0	0	0	5	0	5
Pasture hay abandoned agriculture or other grassy areas (i.e. soccer fields)	78	31	63	376	135	683
Golf Course	1	0	0	0	0	1
Coastal Grassland	7	35	47	22	40	151
Non-forested and Wetland						
Emergent Wetlands	0	0	0	3	0	3
Dry Salt Flats (Includes mud flats)	581	1	3	11	2	598
Quarries	0	0	0	11	4	15
Coastal Sand	98	10	18	36	25	187
Coastal Rock	9	67	165	128	107	476
Interior Rock (Virgin Gorda boulder)	0	0	1	0	20	21
Bare Soil (including bulldozed land)	0	0	1	16	3	20
Salt Pond	514	1	7	44	5	571
Surface Water	2	1	1	5	1	10

[1] NPCG (Norman, Peter, Cooper, Ginger islands)

References

[1] J. J. Ewel and J. L. Whitmore, "The Ecological life zones of Puerto Rico and the U.S. Virgin Islands," Forest Service research paper ITF-018, Institute of Tropical Forestry, Rio Piedras, Puerto Rico, pp. 72 (1973).

[2] The University of the Virgin Islands, Conservation Data Center, "Rapid Ecological Assessment," St. Thomas, VI. U.S.A., pp. 56 (2000).

[3] J. S Beard, "The natural vegetation of the Windward and Leeward Islands," *Geogr. Rev.***4**, pp 350 (1949).

[4] T. Kennaway and E. H. Helmer, "The forest types and ages cleared for land development in Puerto Rico", *GIScience and Remote Sensing***44** (4), 356-382 (2007) [doi:10.2747/1548-1603.44.4.356].

[5] E. H. Helmer, T. Kennaway, D. Pedreros, M. Clark, H. Marcano, L. Tieszen, T. Ruzycki, S. Schill and C. M. S. Carrington, "Land cover and forest formation distributions for St. Kitts, Nevis, St. Eustatius, Grenada and Barbados from decision tree classification of cloud-cleared satellite imagery," *Caribb. J. Sci.* **In Press**, (2008).

[6] E. H. Helmer and B. Ruefenacht, "A comparison of radiometric normalization methods when filling cloud gaps in Landsat imagery," *Can. J. Remote Sensing***33**(4), 325-340 (2007).

[7] M. A. Lefsky, W. B. Cohen, G. G. Parker, and D. J. Harding, "Lidar remote sensing for ecosystem studies," *Bioscience* **52**(1), 19-30 (2002) [doi:10.1641/0006-3568(2002)052[0019:LRSFES]2.0.CO;2].

[8] R. O. Woodbury and P. L. Weaver, "The vegetation of St. John and Hassel Island: U. S. Virgin Islands," U.S. DOI National Park Service, Georgia (1997).

[9] J. Rogozinski, *A brief history of the Caribbean. From the Arawak and Carib to the present, (revised edition)*, Penguin Group, New York, pp. 30 (2000).

[10] P. L. Weaver, "Estate Thomas Experimental Forest: St. Criox, U.S. Virgin Islands: Research history and potential', General technical Report IITF-30, USDA Forest Service International Institute of Tropical Forestry, San Juan, Puerto Rico, pp. 62 (2006).

[11] P. L. Weaver, "A summary of 20 years of forest monitoring in Cinnamon Bay watershed, St. John, U.S. Virgin Islands," General Technical Report IITF-34, USDA Forest Service International Institute of Tropical Forestry, San Juan, Puerto Rico pp. 47 (2006).

[12] S. N. Oswalt, T. J. Brandeis, and B. P. Dimick, "Phytosociology of vascular plants on an international biosphere reserve: Virgin Islands National Park, St. John, US Virgin Islands," *Caribb J. Sci.* **42**(1), 53-66 (2006).

[13] E. H. Helmer, "Forest conservation and land development in Puerto Rico," *Landscape Ecol.* **19**(1), 29-40 (2004) [doi:10.1023/B:LAND.0000018364.68514.fb].

[14] T. J. Brandeis and S. N. Oswalt, The status of the U.S. Virgin Islands' Forests, 2004, Resource Bulletin SRS-122, USDA Forest Service Southern Research Station, Asheville, NC (2007).

[15] P. G. Murphy, and A. Lugo, "Ecology of tropical dry forest", *Annu. Rev. Ecol. Syst.***17**, 67-88 (1986).

[16] M. A. Friedl, C. E. Brodley A. H. Strahler, "Maximizing land cover classification accuracies produced by decision trees at continental to global scales," *IEEE T. Geosci. Remote* **37**(2), 969-977 (1999).

[17] M. B. Carreiras, M. C. Pereira, and E. Shimabukuro, "Land-cover mapping in the Brazilian Amazon using SPOT-4 vegetation data and machine learning classification methods," *Photogramm. Eng. Remote Sensing* **72**(8), 1449-1464 (2006).
[18] M. A. Lefsky, W. B. Cohen, S. A. Acker, G. G. Parker, T. A. Spies, and D. J. Harding, "Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir Western Hemlock forests," *Remote Sensing of Environment***70**(3), 339-361 (1999) [doi:10.1016/S0034-4257(99)00052-8].
[19] J. E. Means, S. A. Acker, D. J. Harding, J. B. Blair, M. A. Lefsky, W. B. Cohen, M. E. Harmon, and W. A. McKee, "Use of large-footprint scanning lidar to estimate forest stand characteristics in the western Cascades of Oregon," *Remote Sens. Environ***. 67**(3), 298-308 (1999) [doi:10.1016/S0034-4257(98)00091-1].
[20] V. Thomas, P. Treitz, J. H. McCaughey and I. Morrison, "Mapping stand-level boreal forest biophysical variables for a mixedwood boreal forest using lidar: an examination of scanning density," Can. J. For. Res. **36**(1), 34-47 (2006) [doi:10.1139/x05-230].
[21] E. Næsset and T. Okland, "Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve," *Remote Sens. Environ*. **79**(1), 105-115 (2002).
[22] S. C. Popescu, R. H. Wynee, and R. F. Nelson, "Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window size," *Comput. Electron Agr*. **37**(1), 71-95 (2002) [doi:10.1016/S0168-1699(02)00121-7].
[23] K. S. Lim and P. M. Treitz, "Estimation of Above ground Forest Biomass from airborne Discrete Return Laser Scanner Data Using Canopy-based Quantile Estimators," *Scand J For Res.* **19**(6), 558-570 (2004) [doi: 10.1080/02827580410019490].
[24] G. Patenaude, R. A. Hill, R. Milne, D. L. A. Gaveau, B. B. J. Briggs and, T. P. Dawson, "Quantifying forest above ground carbon content using LiDAR remote sensing," *Remote Sens. Environ***. 93**(3), 368-380 (2004) [doi:10.1016/j.rse.2004.07.016].
[25] S. C. Popescu, R. H. Wynne, and R. F. Nelson, "Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass," *Caribb J. Sci*. **29**(5), 564-577 (2003).
[26] F. Morsdorf, E. Meier, B. Kötx, K. I. Itten, M. Bobbertin, and B. Allgöwer, "LIDAR-based geometric reconstruction of boreal type forest stands at single tree level for forest and wildland fire management," *Remote Sens. Environ***. 92**(3), 353-362 (2004) [doi:10.1016/j.rse.2004.05.013].
[27] N. Skowronski, K. Clark, R. Nelson, J. Hom, and M. Patterson, "Remotely sensed measurements of forest structure and fuel loads in the Pinelands of New Jersey," *Remote Sens. Environ***. 108**(2), 123-129 (2007) [doi:10.1016/j.rse.2006.09.032].
[28] R. Nelson, "Modeling Forest Canopy Heights: The Effects of Canopy Shape," *Remote Sens. Environ.* **60**(3), 311-326 (1997) [doi:10.1016/S0034-4257(96)00213-1].
[29] USDA Forest Service, "Forest Stewardship program (Puerto Rico and Virgin Islands region)," U. S. Department of Agriculture, Forest Service, International Institute of Tropical Forestry, Puerto Rico (2007).
[30] L. R. Holdridge, *Life zone ecology, (Revised edition)*, Tropical Science Center, San José, Costa Rica, pp. 206 (1967).
[31] E. H. Helmer and B. Ruefenacht, "Cloud-free satellite image mosaics with regression trees and histogram matching," Photogramm. Eng. Remote Sensing **71**(9), 1079-1089 (2005).

[32] ERDAS, Inc, "Leica Geosystems ERDAS Imagine version 8.7," 2007, http://gi.leica-geosystems.com/LGISub1x20x0.aspx, (10 November 2007).
[33] L. Fox, M Garrett, R. Heasty, and E. Torres, "Classifying wildlife habitat with pan-sharpened Landsat 7 imagery," *ISPRS Commission, I/FIEOS 2002 Conference Proceedings*, pp. 11 (2002).
[34] T. M. Lillesand, R. W. Kiefer, and J. W. Chipman, *Remote sensing and image interpretation (Fifth edition)*, John Wiley and Sons, New York, pp. 784 (2004).
[35] A. H. Strahler, "Stratification of natural vegetation for forest and rangeland inventory using Landsat imagery and collateral data," Int. J. Remote Sens. **2**(1), 15-41 (1981) [doi:10.1080/01431168108948338].
[36] A. K. Skidmore, "An expert system classifies eucalypt forest types using thematic mapper data and a digital terrain model," Photogramm. Eng. Remote Sensing **55**(10), 449-1464 (1989).
[37] D. M. Gesch, Oimoen, S. Greenlee, C. Nelson, M. Steuck, and D. Tyler, "The National Elevation Dataset", Photogramm. Eng. Remote Sensing **68**(1), 5-11 (2002).
[38] T. G. Farr and M. Kobrick, "Shuttle Radar Topography Mission produces a wealth of data," Amer. Geophys. Union Eos **81**, 583-585 (2000).
[39] United States Geological Survey, 1:24,000 digital line graphs, Reston, Virginia, 2007, http://edc.usgs.gov/products/map/dlg.html, (September 20, 2007).
[40] M. A. Friedl and C. E. Brodley, "Decision tree classification of land cover from remotely sensed data," Remote Sens. Environ. **61**(3), 399-409 (1997) [doi:10.1016/S0034-4257(97)00049-7].
[41] M. Fiorella and W. J. Ripple, "Determining successional stage of temperate coniferous forests with Landsat satellite data," Photogramm. Eng. Remote Sensing **59**(2), 239-246 (1993).
[42] E. H. Helmer, W. B. Cohen, and S. Brown, "Mapping montane forest successional stage and land use with multi-date Landsat imagery," International Int. J. Remote Sens. **21**(11), 2163-2183 (2000) [doi:10.1080/01431160050029495].
[43] R. A. Schowengerdt, *Remote sensing: models and methods for image processing (Second edition)*, Elsevier, New York, pp. 522 (1997).
[44] E. H. Helmer, O. Ramos, T. d. M. Lopez, M. Quiñones, and W. Diaz, "Mapping forest type and land cover of Puerto Rico, a component of the Caribbean biodiversity hotspot," Caribb J. Sci. **38**(3), 165-183 (2002).
[45] A. Areces-Mallea, A. S. Weakley, X. Li, R. G. Sayre, J. D. Parrish, C. V. Tipton, and T. Boucher, "A guide to Caribbean vegetation types: classification systems and descriptions," The Nature Conservancy, Washington, D.C. (1999).
[46] FGDC, "National Vegetation Classification Standard," Federal Geographic Data Committee, Vegetation Subcommittee. FGDC-STD-005, U.S. Geological Survey, Reston, Virginia (1997).
[47] RuleQuest Research, Data Mining Tools See5 and C5.0, 2007 http://www.rulequest.com/see5.htm, (November 20, 2007).
[48] J. R. Quinlan, *C4.5: programs for machine leaning*, Morgan Kaufmann, California (1993).
[49] R. G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data," Remote Sens. Environ. **37**(1), 35-46 (1991) [doi:10.1016/0034-4257(91)90048-B].
[50] S. V. Stehman and R. L. Czaplewski, "Design and analysis for thematic map accuracy assessment: fundamental principles," Remote Sens. Environ. **64**(3), 331-334 (1998) [doi:10.1016/S0034-4257(98)00010-8].
[51] 3001 Inc., Puerto Rico Ground Control Report, Contract No: DACW43-03-D-0504 (2004).

[52] Q. Chen, "Airborne lidar data processing and information extraction," Photogramm. Eng. Remote Sensing **73**, 109-112 (2007).
[53] K. Zhang, S. Chen, D. Whitman, M. Shyu, J. Yan, and C. Zhang, "A progressive morphological filter for removing nonground measurements from airborne lidar data," IEEE T. Geosci. Remote **41**(4), 872-882 (2003) [10.1109/TGRS.2003.810682].
[54] Q. Chen, P. Gong, D. Baldocchi, and G. Xie, "Filtering airborne laser scanning data with morphological methods", Photogramm. Eng. Remote Sensing **73**(2), 175-185 (2007).
[55] M. A. Lefsky, Interactive Data Language (IDL) algorithm for Lidar point cloud data, Colorado State University, Fort Collins, Colorado (2007).
[56] M. A. Lefsky, "Application of Lidar Remote sensing to the estimation of forest canopy and stand structure", PhD Thesis,University of Virginia (1997).
[57] T. J. Brandeis and S. N. Oswalt, "The status of U.S. Virgin Islands' forests, 2004," Resource Bulletin SRS-122, USDA Forest Service Southern Research Station, Ashville, NC (2007).
[58] USDA Forest Service, "Field Procedures for Puerto Rico and the Virgin Islands, Supplement C to SRS Regional Manual 1.56", USDA Forest Service, Southern Research Station, Forest Inventory and Analysis Research Work Unit, Knoxville, TN, 2002, http://srsfia2.fs.fed.us/data_acquisition/manual.shtml (June 25, 2008).
[59] D. Hatfield, Slope Position AML (ARC/INFO 7.2.1). USDA Forest Service, Pacific Northwest Region, Portland, Oregon (1999).
[60] B. Ruefenacht, M. V. Finco, M. D. Nelson, R. Czaplewski, E. H. Helmer, J. A. Blackard, G. R. Holden, A. J. Lister, D. Salajanu, D. Weyermann, K. Winterberger, "Conterminous U.S. and Alaska forest type mapping using forest inventory and analysis data," Photogramm. Eng. Remote Sensing **In Press**, (2008).
[61] SAS Institute Inc., "JMP Statistical Software", 2007, http://www.jmp.com.
[62] P. L. Weaver, "Forest productivity of the Cinnamon Bay watershed, St. John, U.S. Virgin Islands," Caribb. J.Sci. **32**, 89-98 (1996).
[63] D. F. Whigham, I. Olmsted, E. C. Cano, and M. E. Harmon, "The impact of Hurricane Gilbert on trees, litterfall, and woody debris in a dry tropical forest in the Northeastern Yucatan Peninsula," Biotropica **23**(4), 434-441 (2006) [doi:10.2307/2388263].
[64] P. G. Murphy and A. E. Lugo, *Dry Forests of Central America and the Caribbean*, pp. 9-34 in S.H. Bullock, H.A. Mooney and E. Medina, *seasonally dry tropical forests*, Cambridge University Press, New York, NY (1995).
[65] A. Wehr, and U. Lohr, "Airborne laser scanning-an introduction and overview," ISPRS J.Photogramm. **54**(2-3)**,** 68-82(1999) [doi:10.1016/S0924-2716(99)00011-8].
[66] R. Cunningham, D. Gisclair, and J. Craig, "The Louisiana Statewide LIDAR Project," (n.d.), ftp://ftp-fc.sc.egov.usda.gov/NCGC/products/elevation/la-lidar-project.pdf, (September 20, 2007).