Surface Lidar Remote Sensing of Basal Area and Biomass in

Deciduous Forests of Eastern Maryland, USA

Michael A. Lefsky(1,4), D. Harding(2), W.B. Cohen (4)

G. Parker(3), H.H. Shugart(1)

(1) University of Virginia, Department of Environmental Science

Charlottesville, VA 22903

(2) Laboratory for Terrestrial Physics, NASA Goddard Space Flight

Center, Greenbelt, MD 20771

(3) Smithsonian Environmental Research Center,

P.O. Box 28, Edgewater MD 21037

(4) Current Address: USDA, Forest Service, PNW Research Station

Forestry Sciences Laboratory

3200 SW Jefferson Way, Corvallis, OR 97331 USA

running head: Lidar remote sensing of forests

figures: 6

tables: 3

galley and proof correspondence to:

Michael Lefsky
USDA, Forest Service, Pacific Northwest Research Station
Forestry Sciences Laboratory
3200 SW Jefferson Way
Corvallis, OR 97331 USA
telephone: (541) 758-7765
email: lefsky@fsl.orst.edu

ABSTRACT

A method of predicting two forest stand structure attributes, basal area and aboveground biomass, from measurements of forest vertical structure was developed and tested using field and remotely sensed canopy structure measurements. Coincident estimates of the vertical distribution of canopy surface area (the canopy height profile), and field-measured stand structure attributes were acquired for two datasets. The chronosequence dataset consists of 48 plots in stands distributed within 25 miles of Annapolis, MD, with canopy height profiles measured in the field using the optical-quadrat method. The stemmap dataset consists of 75 plots subsetted from a single 32 ha stem-mapped stand, with measurements of their canopy height profiles made using the SLICER (Scanning Lidar Imager of Canopies by Echo Recovery) instrument, an airborne surface lidar system. Four height indices, maximum, median, mean and quadratic mean canopy height (QMCH) were calculated from the canopy height profiles. Regressions between the indices and stand basal area and biomass were developed using the chronosequence dataset. The regression equations developed from the chronosequence dataset were then applied to height indices calculated from the remotely sensed canopy height profiles from the stemmap dataset, and the ability of the regression equations to predict the stemmap plot's stand structure attributes was then evaluated. The QMCH was found to explain the most variance in the chronosequence dataset's stand structure attributes, and to most accurately predict the values of

the same attributes in the stemmap dataset. For the chronosequence dataset, the QMCH predicted 70 % of variance in stand basal area, and 80 % of variance in aboveground biomass, and remained non-asymptotic with basal areas up to 50 m^2ha^{-1}, and aboveground biomass values up to 450 Mg*ha^{-1}. When applied to the stemmap dataset, the regression equations resulted in basal areas that were, on average, under-estimated by 2.1 m^2ha^{-1}, and biomass values were under-estimated by 16 Mg*ha^{-1}, and explained 37 % and 33 % of variance, respectively. Differences in the magnitude of the coefficients of determination were due to the wider range of stand conditions found in the chronosequence dataset; the standard deviation of residual values were lower in the stemmap dataset than on the chronosequence datasets. Stepwise multiple regression was performed to predict the two stand structure attributes using the canopy height profile data directly as independent variables, but they did not improve the accuracy of the estimates over the height index approach.

1.0 INTRODUCTION

   Most studies of forest development focus on what is commonly
referred to as "stand structure", the size and number of woody
stems per unit area, and related statistics (Miles 1979, Oliver
and Larson 1996). However, the forest canopy, "the collection of
all leaves, twigs, and branches formed from the combination of all
the crowns in the stand" (Maser 1989), is another functionally and
structurally critical component of the forest. The canopy is
responsible for the majority of material and energy exchanges with
the atmosphere, a critical habitat for forest biota, and a
controlling influence over the micro-climate of the forest
interior. Increasingly, species vertical position is recognized
as a major determinant of successional status (Wierman and Oliver
1979, Aber 1979, Bicknell 1982, Gulden and Lorimer 1985, Smith
1986, Oliver and Larson 1996), and therefore canopy structure, the
"organization in space and time, including the position, extent,
quantity, type, and connectivity, of the aboveground components of
vegetation" (Parker 1995), plays a dynamic role in forest
development.

   Studies of forest development have focused on the size and
number of stems because they are conveniently measured. Study of
forest canopies has been hindered by the difficulty of
characterizing canopy structure (Nadkarni and Parker 1994), and
various methods have been developed to do so from more easily
obtained measurements such as tree diameter distributions (Mawson

et. al. 1976). A new remote sensing device developed at NASA's Goddard Space Flight Center, SLICER (Scanning Lidar Imager of Canopies by Echo Recovery, Blair et al. 1994, Harding et al. 1994), is able to rapidly measure the vertical distribution of canopy surface area, through the integration of laser altimetry and surface lidar (light detection and ranging) techniques.

Laser altimetry is an established technology for obtaining accurate, high resolution measurements of surface elevations (Krabill et al. 1984, Bufton et. al. 1991). Laser altimetry is used to measure the distance between the sensor and the object sensed through the precise timing of the round-trip return time of the backscattered reflection of a short duration pulse of laser light. The first generation of laser altimeters for remote sensing of vegetation were designed to record the height to the first surface intercepted by the laser over a relatively small sampling area, or footprint, usually less than one meter in diameter (Arp et al. 1982, Schrier et al. 1984, 1985, Ritchie et al. 1993, Menenti and Ritchie 1994, Weltz et al. 1994). Returns from the top surface of the forest canopy were combined with subsequent measurements of distance to the forest floor, obtained through gaps in the forest canopy, to infer the height of the dominant trees. A more technically advanced version of this approach involves recording, for each individual small footprint, the distance to the first return from the upper surface of the vegetation, and to the last return from the ground surface. The distance between these two measurements is inferred to be the

vegetation height for each footprint. Measurements made using
these techniques have proved useful for predicting canopy height,
timber volume and forest biomass (Maclean and Krabill 1986, Nelson
et al. 1988a, 1988b, Naesset 1997a & 1997b), species type (Jensen
et al. 1987) , and percent canopy cover (Ritchie et al. 1993, and
Weltz et al. 1994).

The SLICER instrument is one of a new generation of systems
(Aldred and Bonnor 1985, Nilsson 1996) that augment traditional
first-return laser ranging with a surface lidar capability. In
surface lidar, the power of the entire return laser signal is
digitized, resulting in a waveform that records the vertical
distribution of the backscatter of laser illumination from all
canopy elements (foliar and woody) and the ground reflection, at
the wavelength of the transmitted pulse (1064 nm, in the near-
infrared). The use of relatively large footprints (5-15 m) is
designed for the recovery of returns from the top of the canopy
and the ground in the same waveform, while remaining small enough
to be sensitive to the contribution of individual crowns of
eastern deciduous species. Currently, the SLICER system has been
mounted on various aircraft platforms and flown over sites in a
range of footprint number and size configurations. Details of the
technical aspects of SLICER can be found in Blair et al. (1994)
and Harding et al. (1994).

Motivation for work relating forest attributes to lidar
sensed canopy structure has been enhanced by the announcement that
VCL, the Vegetation Canopy Lidar mission, has been funded by

NASA's Earth System Science Pathfinder (ESSP) program (Dubayah 1997). Scheduled to be launched in mid 2000, VCL will provide global coverage of surface LIDAR data similar to that used in this study, with transects of contiguous 25 meter footprints spaced every 2 km along the earth's surface.

1.1 Objectives

This current work is part of a larger effort to verify the ability of SLICER to accurately measure canopy height profiles (Lefsky 1997), to relate the canopy height profiles to simple stand structure attributes (this paper), and to relate changes in the canopy height profiles from a 300 year chronosequence to the processes of stand dynamics (Lefsky 1997). The aims of this paper are: 1) to determine if estimates of two stand structure attributes, basal area and aboveground biomass, can be made using indices derived from field-measured canopy height profiles, 2) to determine if regression equations developed from field-measured canopy height profile indices can accurately predict the same two stand structure attributes when applied to indices derived from canopy height profiles measured by SLICER, and 3) to determine the relative power of stepwise multiple regression using the elements of the CHP, and simple regression using height indices, to predict basal area and aboveground biomass.

2.0 METHODS AND MATERIALS

2.1 Overview

Two datasets, stemmap and chronosequence, each consisting of spatially coincident measurements of canopy and stand structure, were collected in the coastal plain of Maryland, USA. For both datasets, two stand structure attributes, basal area and aboveground biomass, were derived from field measurements of tree diameter at breast height. The canopy structure of the plots in both datasets was quantified using the canopy height profile measurement, the distribution of foliage as a function of height. The canopy height profiles of plots in the chronosequence data were measured in the field using the optical point-quadrat method of Aber (1979). The canopy height profiles of plots in the stemmap dataset were measured by the SLICER scanning lidar instrument using a processing algorithm based on the principles of the optical-quadrat method. Canopy height indices, including maximum, mean, median, and quadratic mean canopy height (QMCH) were calculated for plots from both datasets using their associated canopy height profile measurements. Regressions between canopy height indices and basal area and aboveground biomass were developed using the chronosequence dataset. These same regression equations were then applied to the remotely-sensed height indices from the stemmap dataset, and the resultant estimates of basal area and biomass were compared to those measured in the field.

2.2 Data Collection and Pre-processing

2.2.1 <u>Floristics</u>

Data used in this work were collected as part of a larger project describing the structural, floristic and environmental development of tulip-poplar stands. The tulip-poplar association is the most common upland forest association in the coastal plain and much of the piedmont of mid-Atlantic North America, from the Carolinas to New Jersey. Though variable in composition, tulip poplar (*Liriodendron tulipifera*) occurs at most stages of succession. Its life cycle begins with high populations of sweet gum (*Liquidambar styriciflua*) or tulip poplar following agricultural abandonment or timbering; these species persist for several decades. In mature stands, the canopy is composed of oaks, hickories, beech and some tulip poplar, with a diverse complement of mid- and sub-canopy species. Pines (*Pinus virginiana* and *P. taeda*) are rarely a major component in these forests.

2.2.2 Chronosequence dataset

The chronosequence dataset consists of 48 plot observations from stands dispersed within a 25 mile radius of the Smithsonian Environmental Research Center (SERC), located in Edgewater, Maryland, USA (Brown and Parker 1994). Stand structure information for the chronosequence dataset was collected using variable-sized plots that were scaled roughly to the maximum height of the canopy; the average plot was 20 m X 50 m. In each plot, the species and breast-height diameters of all living woody plants less than or equal to 2 cm in diameter were recorded. These data

were used to estimate basal area directly, and to estimate the aboveground woody biomass of each plot through the use of an allometric equation (Table 1). The equation used was that of Monk et al.(1970), which was developed in a forest of similar composition:

$$\log_{10}B = 1.9757 + 2.5371 \log_{10}DBH \qquad [Eq. 1]$$

where B is the biomass per stem, in grams, and DBH is the diameter at breast height, in centimeters. Total biomass per unit area for each plot was calculated as the total biomass of every measured stem, divided by the area of the plot. Monk et al(1970) did not report the error of their regression coefficients.

The canopy height profile (CHP) variable used to describe canopy structure in this dataset is a modification of MacArthur and Horn's(1969) foliage height profile (FHP) variable. While some investigators have measured height profiles directly, through stratified clipping (Fujimori 1971) or point quadrat techniques (Warren-Wilson 1958,1965, Miller 1967, Ford and Newbould 1971), these methods have largely been supplanted by the optical-quadrat method. Using this method, optical point quadrats are established and multiple observations of vertical distance to first leaf intersection are made using a camera equipped with a zoom telephoto lens. This distribution is used to estimate the cumulative percent cover of foliage as a function of height. The estimate of cover is transformed into the vertical distribution of foliage using a method that assumes that leaf angle is constant and that the horizontal distribution of leaves is random.

Using these assumptions the amount of foliage that results in the observed changes in cover can be calculated, using an equation derived from the Poisson distribution:

$$FHP_c(h) = -\ln(1-cover(h)) \qquad [Eq.\ 2],$$

where $FHP_c(h)$ is the cumulative one-sided leaf surface area (or LAI, Leaf Area Index), expressed as a fraction of projected ground area, above height h, and cover(h) is the fraction of sky not covered by foliage, above height h. The actual FHP is calculated from $FHP_c(h)$ by calculating the additional LAI at each height interval, with respect to that above it. The theory behind the original application of this technique is found in MacArthur and Horn (1969), and a validation of the method is presented in Aber (1979).

The FHP is the distribution of *foliage* surface area as a function of canopy height, from the ground to the top of the canopy. In contrast, the canopy height profile (CHP) is the surface area of *all* canopy material, foliar and woody, as a function of height. Combining foliar and non-foliar materials was necessary so that field and remotely sensed canopy height profiles could be compared. This is because the single-wavelength SLICER system cannot distinguish between various sources (bark, foliage, soil) of backscattered illumination. In order to measure the CHP in the field, the distribution of the height to the first intersection of any canopy structure type is recorded, rather than only intersections of foliage. Either canopy or foliage height profiles can be calculated as relative, (with the total vector

11

scaled to one) or absolute (with the total vector scaled to the total leaf or plant area index of the canopy). In this work relative canopy height profiles are used exclusively.

2.2.3 <u>Stemmap Dataset</u>

The stemmap dataset combines field observations of forest stand structure with coincident remotely-sensed observations of canopy structure. The measurements of basal area and biomass for this dataset came from an existing 32 ha stand at SERC in which every stem greater than 20 cm dbh has been mapped. The SLICER instrument was flown over the stand in September of 1995, in a five-beam cross-track configuration. The SLICER footprints were georeferenced by combining the ranging data with laser pointing and aircraft position data, obtained by a Inertial Navigation System and a kinematic Global Position System trajectory, respectively. The stemmap and SLICER transects were then registered to a digital orthophoto quadrangle (Maryland DNR, 1991), which was re-projected to the UTM projection (Figure 1a). The stemmap was geolocated by matching the roads as recorded on the photo to the areas without trees within the stemmap, which are associated with the roads. After conversion of the georeferenced SLICER data to the UTM projection, a systematic offset of unknown source was noted between forest edges in the orthophoto and as expressed by the SLICER canopy height profile. In order to ensure proper registration between the stemmap and the SLICER footprints,

the SLICER data was translated to match forest edges in the orthophotos. On this basis, the error in the relative positions of the stemmap and SLICER transects should be reduced to less than 15m.

To calculate basal area and biomass from the stemmap dataset, the geo-referenced tree and SLICER waveform data were processed using programs written in IDL (Interactive Data Language, Research Systems Inc, Boulder Colorado). For this study, the transect of remotely sensed data was five laser footprints wide, with each footprint nominally 10 meters in diameter, and nominally spaced at 10 meter intervals in along- and cross-track dimensions. Data from the two outer footprint positions in the transect were discarded due to anomalous height measurements. The anomalous height measurements are thought to be due to low instrument signal-to-noise caused by misalignment between the footprint crosstrack scan pattern and the outer edges of the instruments receiver field-of-view. Three by three blocks of SLICER footprints were selected from the central three of the five cross track footprints (See Figure 1b); each 3 x 3 block was considered to be a single plot. Of a possible 104 samples within the vicinity of the stem-mapped stand, 75 were selected for analysis. The remaining plots were eliminated due to their proximity to either the edge of the stem-mapped area, a clearing, or roads within the stand, because they overlapped with other plots, or were in the vicinity of an instrumented tower within the stemmap.

A mask was generated for each 3 x 3 block of SLICER waveforms (Figure 1b), to determine which stems fell within the area sensed by the instrument. The mask was fit to the outermost positions of the four corner waveforms, which were calculated as occurring 7.07 m from the center point of each corner waveform. All stems within the mask were extracted and the total basal area and biomass (as calculated using Eq 1) of those stems was divided by the area of the mask, in hectares(See Figure 1b).

Canopy height profiles for each plot in the stemmap dataset were calculated using the plot's 9 SLICER waveforms. Validation of the SLICER system and the processing software's ability to remotely sense canopy height profiles can be found in Lefsky (1997). Briefly, we hypothesized that the power of the backscattered laser illumination is subject to the same process of occlusion observed in the field measurements of height to first intersection, and modified the MacArthur-Horn method to apply this approach to the SLICER return energy waveforms. The most critical step in the modification of the MacArthur-Horn routine was the separation of the portion of the waveform returned from the ground surface from the balance of the waveform (Figure 2a). The ratio of the power of the "ground return" to the total signal power is inversely proportional to the total canopy cover, but to estimate canopy cover, the ratio must be adjusted to account for differences in ground and canopy reflectance at 1064nm. We did this by assuming that the ratio of canopy and ground reflectance is approximately 2:1. The total horizontal canopy cover at each

height increment can then be calculated, which allows the use of the MacArthur-Horn equation (Figure 2b). The processing that implements this algorithm was tested using four dissimilar stands at SERC. A two-sample, uneven sample size, Chi-square goodness-of-fit analysis was performed to determine if there were statistically significant differences between field and SLICER derived canopy height profiles. The canopy height profiles measured in the field and from SLICER were statistically indistinguishable (Lefsky 1997). In this work, an interactive version of the waveform processing algorithm was used to improve the identification of the position of the ground return.

2.2.4 <u>Comparison of Stemmap and Chronosequence Stand Structure Attributes</u>

Stand structure attributes for the chronosequence dataset were measured considering all stems greater than 2 cm. The stemmap, in contrast, was mapped using a minimum diameter of 20 cm. To allow comparison of the attributes measured using the 20 cm and 2 cm limits, we generated a set of 20 cm limit structure attributes for the chronosequence dataset, so that we had both 2 and 20 cm limit stand structure attributes for that dataset. Using this data, we estimate that stems between 2 and 20 cm DBH account for 13% of total basal area, and 8% of aboveground biomass. Regressions with the 20 cm structure attributes as independent variables, and the 2 cm structure attributes as dependent

variables was performed. The resulting regression equations were then evaluated for their suitability in estimating 2 cm limit basal area and biomass from the 20 cm limit data; ie., to determine if there was variability in the 2 cm limit measurements that was not explained by the 20 cm limit measurements, specifically in the range of conditions over which we want to apply these equations.

2.3 Canopy Height Indices

To relate the field and remotely-sensed canopy height profile measurements to the stand structure attributes, we reduced the vector information in the CHP to four height indices. We chose to use the following height indices: maximum canopy height, mean canopy height, median canopy height, and QMCH. In this discussion, the canopy height profile is treated as a vector of one meter high elements, with the value for each element equal to the fraction of the total profile in the height range of that element. For example, the first element in the CHP vector represents the fraction of total canopy surface area between 0 and 1 meters above the ground.

Maximum canopy height is calculated as the height of the highest canopy height profile element that has a value greater than zero. Median canopy height is calculated as the height of the highest element below which no more than 50% of the total canopy height profile is distributed. Mean canopy height is calculated as

the summation of the product of the canopy height profile and the height of each element. The QMCH is defined as:

$$QMCH = \sqrt{\sum_{1}^{max.height} CHP[i]*i^2} \qquad [Eq\ 2],$$

where CHP[i] is the fraction of total foliage at height "I".

Regressions between coincident field and SLICER measurements of the maximum, median and quadratic mean canopy height have previously been performed using a dataset of 12 plots in two eastern deciduous forests, as described in Lefsky (1997), where each of the plots had both field and SLICER measurements of canopy structure. Among the twelve plots were four from the chronosequence dataset. Analysis at that time indicated that SLICER-measured indices of height were closely correlated with those measured in the field (field vs. SLICER height, R^2=76%, median height, R^2=68%, quadratic mean canopy height, R^2=78%, e.g. Figure 3). In that work, a positive bias was noted in the equations relating the SLICER measurement of height indices to field measurements. Subsequent re-analysis of that dataset indicates that intercepts of those equations are not significantly different from zero, and their slopes are not significantly different from 1.0. Therefore, no corrections were applied to the SLICER measured indices of canopy structure in the current work. To evaluate the relationships among the four height indices, we plotted them against each other and calculated their r^2 values, using data from the chronosequence dataset.

2.4 Linear Regression

    Linear regression was used to develop equations relating
height indices to basal area and biomass (Objective 1). Four
equations were developed for both basal area and aboveground
biomass, one using each of the four height indices, for a total of
eight equations. Data for the development of the equations came
from the chronosequence dataset, which has field-measured height
indices. Regression between the height indices and aboveground
biomass indicated that there was a consistent, positive
correlation between the independent values and the variance of the
residuals, and therefore aboveground biomass values were
transformed using a square-root. To assess the relative
explanatory power of each height index, the r^2 between each height
index and both stand structure attributes in the chronosequence
dataset was calculated, as well as the standard deviation of the
residuals. For the square-root transformed aboveground biomass,
the r^2 reported is for the transformed variable, all other
statistics were calculated using the back-transformed predictions.

    The resulting equations were then applied to the SLICER-
measured height indices from each plot of the stemmap dataset
(Objective 2), to produce predicted stand structure attributes for
that dataset. The applicability of the regression equations to the
stemmap dataset was evaluated in two ways. For each equation we
calculated the residual between the predicted and observed
(stemmap) stand structure attributes, and the residuals' mean and

standard deviation. A final check on the applicability of the chronosequence derived equations to the stemmap data was performed using regression between the values predicted for the stemmap plots using each of the equations, and those observed in the stemmap dataset. The coefficients of these regressions were then tested to determine if they differed significantly from those expected if the actual relationship was the identity equation:

$$\text{Predicted} = B_0 + B_1 * \text{Observed},$$

where $B_0=0$ and $B_1=1$. Preliminary results indicated that the r^2 values between the predicted and observed stand attributes for the stemmap dataset were smaller than those which were obtained for the regressions between the height indices and stand attributes made using the chronosequence dataset. One difference between the two datasets is the narrower range of conditions, of both stand attributes and height indices, in the stemmap dataset. In order to determine if this was a factor in the lower r^2 values, correlation coefficients between each of height indices and each of the stand attributes were calculated, for four datasets. These datasets were: all plots from the chronosequence dataset, all chronosequence plots less than 30 meters tall, all chronosequence plots greater than or equal to 30 meters tall, and all stemmap plots greater than or equal to 30 meters, which included all the stemmap plots.

2.5 Stepwise multiple regression

Stepwise multiple regression has been proposed as a method to predict basal area and stem volume from vertical canopy profile measurements (Hyyppa and Pulliainen, 1994). Stepwise multiple regressions were performed to check if more variance in the stand structure attributes was explained by linear combinations of the canopy height profile elements than by the indices derived from it. These regressions were conducted using the elements of the canopy height profile, aggregated to 11 four-meter resolution bins, as the independent variables. For example, the first independent variable was the fraction of the CHP between 0 and 3 meters above the ground.

3.0 RESULTS

3.1 Overview

Regression analysis between 20 cm and 2 cm stand attributes indicated that basal area and aboveground biomass measurements made using the 20 cm limit could accurately predict 2 cm attributes (Figure 4a & 4b). Therefore, stand structure attributes for the plots from the stemmap dataset, which had been measured using a 20 cm limit, were transformed using the resulting regression equations, so that they could be compared directly to attributes from plots in the chronosequence dataset. Height indices calculated from each plot's CHP were highly correlated with each other, but maximum canopy height was least well correlated with the other indices (Figure 5). High correlations between all of the height indices and both stand structure attributes were observed as results of the regression analysis using data from the chronosequence dataset (Figure 6). We then calculated estimated stand attributes for the stemmap dataset, using the equations from this chronosequence regression analysis and height indices from the stemmap dataset (Figure 6). While correlation coefficients between the observed and these new predicted estimates of stand attributes were lower than the correlation coefficients obtained when developing the original equations, other measures of the predictive power of these equations, such as the standard deviation of residuals, were similar (Tables 2,3). The QMCH index was found to be the most reliable predictor of basal area and biomass when results from the

stemmap and chronosequence datasets were considered. Equations developed using stepwise multiple regressions explained as much variance as those resulting from simple regression using the height indices, but their predictions of the stand attributes of the stemmap plots were not as accurate.

3.2 Prediction of 2 cm Stand Structure Attributes from 20 cm Stand Structure Attributes

Using the chronosequence dataset, strong linear relationships were found between basal area and biomass measured with the 2 cm limit and the same attributes measured with the 20 cm limit. For both basal area and aboveground biomass (Figure 4a&b) the relationships consist of a range of low values in which the two variables are weakly correlated, and a range of higher values in which the two variables are very strongly related. The range of values found for each stand attribute from the stemmap data (as measured using the 20 cm limit) are indicated on the graphs, and indicate the range over which we want to use the relationship. Within the range of stand structure attributes observed in the stemmap dataset, the relationship between then 2 and 20 cm limit indices remains highly correlated and linear. Regression equations for each attribute were developed using only the data that fell within the range where the relationship between the 20 cm and 2 cm measurements were well correlated. The equations developed are:

$\text{Basal Area}_2 = 12.503 + 0.809 * \text{Basal}_{20}$, r2=93%, P<0.0001, and,

$\text{Biomass}_2 = 38.016 + 0.934 * \text{Biomass}_{20}$, r2=99%, P<0.0001.

These equations have been used to estimate 2 cm limit stand structure attributes for the stemmap plots. All subsequent analyses are for 2 cm limit attributes.

3.3 Relationships Among Canopy Height Indices

High levels of correlation were found between each of the four canopy height indices, as illustrated in Figure 5, but there was some variability . Maximum canopy height has the lowest correlation to the other variables, especially for taller stands. Median canopy height is highly correlated to the two mean height indices. The two mean height indices, the mean canopy height and the QMCH, are the most highly correlated of the variables.

3.4 Relationship of Canopy Height Indices to Basal Area and Biomass

For the chronosequence dataset, all four height indices considered in this work were highly correlated with both basal area and aboveground biomass, with r^2 values between 60% and 80% (Table 2a). In all cases, the correlation between the height indices and aboveground biomass was higher than the corresponding correlation between the height indices and basal area. The standard deviations of the residuals resulting from each regression are lowest when the r^2 values are highest. Of the four equations predicting basal area, the QMCH has the largest r^2 value and the smallest standard deviation of residuals. Of the four equations predicting aboveground biomass, maximum canopy height

and the QMCH both explain 80% of variance, but maximum canopy height has a smaller standard deviation of residuals. The absolute differences in these two indicators (r^2 and the standard deviation of residuals) between the height indices are small in magnitude, and the differences in r^2 are non-significant statistically.

The ability of each of the regression equations to predict the basal area and biomass of plots in the stemmap dataset was evaluated in several ways. Scatterplots of predicted and observed basal area and biomass are presented in Figure 6. For each stand attribute, the two best equations were selected for inclusion in Figure 6 on the basis of the goodness-of-fit statistics presented in Table 2. Examination of the figures indicate that the variability of the stemmap dataset plots is similar to that of the chronosequence dataset. However, the r^2 values of the predicted vs observed regression equations are much lower than those of the regression equations predicting stand attributes from height indices. This result must be viewed within the context of the wider range of conditions observed in the chronosequence plots. For a constant number of data-points distributed around a linear relationship with a specified standard deviation, the amount of variance explained by the linear relationship declines with decreasing range of the independent variable.

This effect is demonstrated by Table 3, which documents the coefficients of determination between each of height indices and the two stand structure attributes for the four dataset defined in the methods. The high determination coefficients obtained using

the entire chronosequence dataset are maintained in the subset of plots shorter than 30m tall. In every case but one (QMCH and Basal Area), determination coefficients drop when chronosequence plots equal to or taller than 30m are considered. When these determination coefficients are compared to those calculated using the stemmap dataset differences still exist, but they are much smaller than the differences obtained when the stemmap plots are compared to the total chronosequence dataset.

Another index of the strength of the relationships include the mean and standard deviation of residuals. The mean residual indicates the total accuracy of all the predictions made with an equation, while the standard deviation of residuals indicates the ability of each equation to predict individual values. The mean residual values for the basal area equations resulted in a error between -9.9 and 7.2 m^2*ha^{-1} , or between -25 % and +19 % of the mean basal area for the stemmap plots. The mean residual values for the aboveground biomass equations resulted in a error between -130.1 and 80.2 $Mg*ha^{-1}$, or between -34 % and 21 % of the mean biomass for the stemmap plots. Of the four equations predicting basal area and biomass, mean canopy height and the QMCH have the mean residual values with the lowest magnitudes.

The standard deviation of residual values for the basal area equations ranged between 4.4 and 6.5 m^2*ha^{-1}, and were in each case lower than those observed in the original regressions. Although the differences between the equations are small in magnitude, they

25

do represent meaningful differences when compared to the standard deviation of the dependent variable, basal area, which is 5.5 $m^2 \ast ha^{-1}$. The equation that uses the median canopy height has residuals who standard deviation is larger that 5.5 $m^2 \ast ha^{-1}$, due to its extremely poor fit. The reduction of the standard deviation from 5.5 to 5.3 $m^2 \ast ha^{-1}$ represents a 7 % reduction in variance, while the reduction from 5.5 to 4.4 represents a 37% reduction. Similarly, the standard deviation of the stemmap biomass values is 56.4 $Mg \ast ha^{-1}$. The equations using maximum and median canopy height have residuals who standard deviation is larger than 56.4 $Mg \ast ha^{-1}$, also due to their poor fit. The reduction of the standard deviation of biomass residuals from 56.4 to 50.9 $Mg \ast ha^{-1}$ represents a 20% reduction in variance, while the reduction of the standard deviation from 56.4 to 46.7 $Mg \ast ha^{-1}$ represents a 33 % reduction in variance. Of the four equations predicting basal area, the QMCH and the mean canopy height had the lowest standard deviation of residuals. The QMCH and the mean canopy height also had the lowest standard deviation of residuals of the four equations predicting aboveground biomass.

Of the predicted vs. observed regression for the four equations predicting basal area, both the maximum and median canopy height had a slope and intercept significantly different from 1 and 0, respectively. The other equations had non-significant p-values of similar magnitude. Of the four equations predicting biomass, the equations using maximum, median and mean

canopy height had either one or both coefficients that were significantly different from those expected.

3.5 Stepwise Multiple Regression

Results from the stepwise multiple regression analysis are presented in Table 2B. The equations predicting basal area and biomass make their predictions from the same three variables, the fraction of the total profile between 16 and 20 meters above the ground, between 28 and 32 meters above the ground, and between 36 and 40 meters above the ground. The three slope coefficients appear to increase exponentially as a function of height. The equation predicting basal area explains nearly as much variance as the most correlated height index (69% vs. 70% for the QMCH), and has a lower standard deviation of residuals then any height index (7.7 vs. 7.8 m^2*ha^{-1} for QMCH). The equation predicting aboveground biomass explains slightly more variance than as the most correlated height indices (81% vs. 80% for maximum canopy height and QMCH), and also has a lower standard deviation of residuals then any height index (61.4 vs. 75.4 $Mg*ha^{-1}$).

The equations derived from the stepwise multiple regression were then applied to the stemmap dataset (Table 2B). Both equations have residual statistics that were near the best of the height index equations. The predicted vs. observed regressions for both basal area and biomass had intercepts that were significantly different from what would be expected, assuming the identity

relationship (See Figure 6), and the equation predicting basal area had a slope that was also significantly different.

4.0 DISCUSSION

4.1 Relationship of Canopy Height Indices to Basal Area and Biomass

The development of equations relating height indices to basal area and biomass indicated that, although there were some differences in the predictive ability of the height indices, those differences were small, and statistically non-significant. Nevertheless, the canopy structure information summarized in the median, mean, and quadratic mean canopy height indices did improve their estimates of stand basal area, albeit non-significantly, relative to the maximum canopy height. Maximum canopy height was as good or better than the other variables at predicting aboveground biomass. This reflects a difference in what the two attributes (basal area and biomass) represent, and in how they are calculated. Basal area, as the name suggests, is a two-dimensional measurement, and increases, on a per stem basis, as a function of DBH squared. Biomass is three-dimensional, the product of wood density and stem volume, and stem volume is a function of the product of stem basal area and height. This means that, on a per stem basis, biomass must increase as a function of DBH to a power greater than 2, as in Eq. 1. In practice, this means that, as a function of stem diameter, stem biomass increases more steeply

than stem basal area. As an example, the ratio of the basal areas of stems 95 cm and 2 cm in diameter is 2060:1, while the ratio of the biomass of those two stems is 16000:1, an eight-fold increase. As a result, the relative contribution of large stems to aboveground biomass is greater than their relative contribution to basal area. Height indices that mostly reflect the height of the largest trees in a stand, such as maximum canopy height and the QMCH, should be most highly correlated with biomass. Conversely, basal area is more sensitive to the number and size of smaller stems. Indices which take into account the average position of foliage, such as the median, mean and quadratic mean canopy height, should be more highly correlated with basal area. This is because these indices represent the average height of all trees, not just the largest ones. The QMCH index includes information about the distribution of tree heights, unlike maximum canopy height, but weights the importance of the taller tree heights, unlike mean or median canopy height. This may explain why it is highly correlated with both stand structure attributes. The relative contribution of large stems to biomass and basal area may also explain why the correlation coefficients between the height indices and biomass are larger than the correlations between the height indices and basal area.

The quadratic mean canopy height is an index developed during the course of this work, and has no literature supporting it. Its development was suggested by the quadratic mean diameter used in

forest mensuration-- the diameter of the tree with the average basal area. The QMCH, as defined earlier (Eq. 3), is the square root of the summation of the product of the canopy height profile and each element's squared height. Niklas (1994) reports that for a dataset of angiosperm "champion" trees, the relationship between tree height and diameter is

$$H = 19.1D^{0.474} ,$$

transforming to predict diameter results in

$$D = (H/19.1)^{2.1}.$$

Note the similarity of the exponent of the transformed equation to the square power used to weight the elements of the canopy height profile, and the similarity of the exponent of the first equation to the square root used to transform the weighted canopy height profile (Eq 3). This suggests the canopy height profile is being weighted by a factor that is proportional to the diameter required to support it, and conversely its average is transformed, by the square root, to a variable that is proportional to height.

The application of regression equations, developed using the chronosequence dataset, to the stemmap dataset indicated that there were relevant differences in the equations suitability. The QMCH and mean canopy height were the best predictors of basal area for plots in the stemmap, but the QMCH is marginally superior in all but one aspect (mean residual) of regression quality. For the prediction of biomass, the QMCH and maximum canopy height are very

similar in terms of regression quality, but when applied to the
stemmap data the equation using mean canopy height has a higher
standard deviation of residuals, and the slope and intercept of
the predicted-vs.-observed regression line for the mean canopy
height is significantly different from the identity line. While
the r^2 values associated with the stemmap dataset were lower than
those associated with the chronosequence dataset, this is due to
differences in the range of conditions found in each dataset, not
the strength of relationships relating height indices to stand
structure attributes.

4.2 Stepwise Multiple Regression

The objective of the stepwise multiple regression analysis
was to see if the individual elements of the canopy height profile
could improve the predictions of stand structure attributes, as
compared with simple regressions using height indices. The
stepwise multiple regressions explained as much variance as linear
regression with canopy height indices, but the resulting equations
were less applicable to the stemmap dataset than the best height
index, the QMCH. Working with a dataset of canopy height profiles
for two sites with differing composition, Lefsky (1997) suggests
that the height index approach is qualitatively preferable,
because it is probably less site specific than predictions made
directly from the canopy height profile. That the percentage of
variance explained by each approach is similar is indicative that

the height index approach is probably explaining as much variance as can be explained.

4.3 Height Index Approach

The height index approach used in this work is similar to that previously used in the analysis of first-return laser altimeters. Many approaches to the analysis of that data have followed the work of MacLean (1982), who showed that the area between a line following the height of the canopy, and another following the ground surface, is closely and linearly related to the natural logarithm of stand volume. The area between these two lines measures the average height of the upper surface of the canopy. Large footprint surface lidar systems, such as the one used in this paper, do not provide a high resolution record of this measurement. However, within the large footprint waveform, the distribution of vertical surfaces is recorded. A weighted height index performs the same function as the canopy height trace-- except that it integrates the height distribution of the entire canopy, not just it's outer surface, as in Naesset (1997a). The canopy height profile transformation further serves to correct the vertical distribution of returned power to reflect the power available for return from the canopy at each successive level through the canopy. The fact that a height index (the QMCH) that is weighted towards the top of the canopy does better than one that isn't suggests that the height of the upper canopy surface may still be an important index for predicting stand attributes.

The value of the coefficient of determination between the stand structure attributes and height indices for both datasets are consistent with those reported in the laser altimetry literature (Maclean 1982, Nelson et al. 1988a, Nelson et al. 1988b, Nilsson 1996, Hyyppa and Hallikainen 1996, Naesset 1997a). A determination coefficient (r^2) of 61 % has been reported for the prediction of basal area (Hyyppa and Hallikainen 1996), while values between 53 % and 92 % have been reported for the prediction of stem volume and biomass. We calculated an adjusted r^2 of 70 % for basal area and an adjusted r^2 of 80 % of biomass for the chronosequence generated regression equations. While these equations were developed using field estimates of canopy structure, we have shown that they are applicable to the prediction of stand structure attributes from canopy height indices measured using the SLICER scanning lidar system. While the r^2 values of the predictions of basal area and biomass of plots from the stemmap dataset were lower than those obtained with the chronosequence dataset, the standard deviation of the residuals for both datasets are nearly equal. The higher r^2 values for equations developed with the chronosequence data reflects its wider and more uniform distribution of conditions.

We have seen that the regressions relating the field measured QMCH index to basal area and biomass are applicable to remotely sensed height indices, for the range of stand conditions observed in the stemmap dataset. Can this conclusion be applied to the whole range of conditions found in the chronosequence dataset? If

the error in the remotely sensed QMCH estimate is constant throughout the range of canopy structure conditions, we can. At present, no direct evidence is available to answer this question. Some steps in the processing of the raw waveform data, such as the delineation of the ground return, could tend to introduce a constant error into the estimates of the canopy height profile, which will have an larger proportional effect on shorter stands, but such an effect has not yet been identified. If the ability to remotely sense the canopy height profile is constant, it is logical to conclude that the overall strength and the coefficients of the relationship between remotely-sensed height and field-collected stand structure attributes is the same as that found between optical-quadrat method height indices and stand structure indices.

When considered along with the results of Lefsky (1997) which concludes that field and SLICER sensed canopy height profiles were statistically indistinguishable, this work supports the premise that SLICER and field collected profiles are directly comparable. If the success in validating SLICER is extended to other forest types, it would provide an unprecedented level of flexibility in developing remote sensing applications using surface lidar techniques. Whereas conventional optical and radar remote sensing platforms do not have conveniently measured field analogues for their measurements, existing relationships between maximum canopy height and forest ecosystem structure and function can be applied directly to surface lidar remote sensing. For those forest types

where the MacArthur-Horn technique can be applied, field estimates of the canopy height profile can offer "proof of concept" support to new analyses, without the difficulty of obtaining laser altimetry and geo-locating the laser footprints in the field, although the 15 m error in footprint position did not seem to overly effect this current work. Current work to establish a laser altimetry profile measurement capability for use in the field will increase the desirability of this approach.

5.0 CONCLUSIONS

Indices measuring the vertical distribution of canopy structure are highly correlated with stand basal area and aboveground biomass. Relationships developed using field measured canopy height profiles were found to be applicable, in varying degrees, to remotely sensed canopy height profiles. The quadratic mean canopy height (QMCH) was the height index which was both highly correlated to both basal area and aboveground biomass, and had the best overall predictions of the stand attributes of the remotely sensed dataset. We find that reasonable because the QMCH is a weighted average of the canopy height profile, where the weights are proportional to the amount of woody structure required to support foliage at each height. Stepwise multiple regression of basal area and biomass using the canopy height profile vector as independent variables did increase the power of the field-measured

regression equations, but were not as applicable to the remotely

sensed dataset as was the QMCH.

5.0 LITERATURE CITED

Aber, J.D. 1979. Foliage-height profiles and succession in a Northern Hardwood forest. Ecology. 60:18-23.

Aldred, A. H., and G. M. Bonnor. 1985. Application of airborne lasers to forest surveys, Petawawa National Forestry Institute, Canadian Forestry Service, Information Report PI-X-51, 61p.

Arp, J.D., Griesbach J.C, and Burns, J.P. 1982.Mapping in tropical forests: a new approach using laser APR. PERS 48(1): 91-100.

Bicknell, S.H. 1982. Development of canopy stratification during early succession in northern hardwoods. For. Ecol. Manage. 4:41-51.

Blair, J.B., Coyle, D.B., Bufton, J.L., and Harding, D.J. 1994. Optimization of an airborne laser altimeter for remote sensing of vegetation and tree canopies. Proceedings of IGARSS'94.

Brown, M.J., and Parker, G.G. 1994. Canopy light transmittance in a chronosequence of mixed-species deciduous forests. Can. J. For. Res. 24:1694-1703.

Bufton, J.L., Garvin, J.B., Cavanaugh, J.F, Ramos-Izquierdo, L., Clem, T.D., and Krabill, W.B., 1991. Airborne lidar for profiling of surface topography. Optical Engin. 30: 72-78.

Dubayah, R., J. B. Blair, J. L. Bufton, D. B. Clark, J. JaJa, R. Knox, S. B. Luthcke, S. Prince, and J. Weishample. 1997. The Vegetation Canopy Lidar Mission. Pages 100-112. Land Satellite Information in the Next Decade II: Sources and Applications. ASPRS, Washington D.C.

Ford E.D. and Newbold, P.J. 1971. The leaf canopy of a coppiced deciduous woodland. J Ecol 59:842-862

Fujimori, T. 1971. Analysis of forest canopy on the basis of a *Tsuga heterophylla* forest. Japanese J. of Eco. 21:134-139

Gulden, J.M. and Lorimer, C.G. 1985. Crown differentiation in even-aged Northern Hardwood forest of the great lakes region. For. Ecol. Mgmt. 10:65-86.

Harding, D.J., Blair, J.B., Garvin, J.G., and Lawrence, W.T. 1994. Laser altimeter waveform measurement of vegetation canopy structure. Proceedings of IGARSS'94.

Hyyppa, J and Pulliainen, J. 1994. Inventory by compartments using radar-derived stand profiles. IGARSS Conference 1: 108-110

Hyyppa, J. and Hallikainen, M. 1996. Applicability of airborne profile radar to forest inventory. Remote Sens. Environ. 57:39-57

Jensen, J. R., E. J. Christensen, and R. Sharitz. 1987. Non-tidal wetland mapping in South Carolina Using Airborne Multi-Spectral Scanner Data. Remote Sensing of Environment 16: 1-12.

Krabill W.B., Collins, J.G., Link, L.E., Swift, R.N., and Butler, M.L. 1984. Airborne Laser topographic mapping results PERS 50(6):685-694

Lefsky, M.A. 1997. Application of lidar remote sensing to the estimation of forest canopy and stand structure. Ph.D. Dissertation, Department of Environmental Science, University of Virginia, Charlottesville, Virginia.

MacArthur, R.H. and Horn, H.S.   1969. Foliage profile by vertical measurements. Ecology 50:802-804.

Maclean, G.A. 1982. Timber volume estimation using cross-sectional photogrammetric and densitometric methods. Master's thesis, University of Wisonsin, Madison, WA

Maclean, G.A. and Krabill, W.B. 1986. Gross merchantable timber volume estimation using an airborne lidar system. Can. J. Remote Sens. 12:7-8

Maryland Department of Natural Resources. 1991 Digital Orthophoto Quarter Quads (DOQ), Annapolis, Maryland, USA

Maser, C. 1989. Forest primeval - the natural history of an ancient forest. Sierra Club Books, San Francisco 282pp

Mawson, J.C., Thomas, J.W., DeGraaf, R.M. 1976. Program HTVOL: the determination of tree crown volume by layers. USDA Forest Service Research Paper NE 354.

Menenti, M and Ritchie, J.C. 1994. Estimation of effective aerodynamic roughness of Walnut Gulch watershed with laser altimetry measurements. Water Resour. Res 30:1329-1337

Miles. J. 1979. Vegetation Dynamics. Chapman & Hall, London.

Miller, P.C. 1967. Tests of solar radiation models in three forest canopies. Ecology 50:802-804

Monk, C.D., Child G.I. and Nicholson, S.A. 1970. Biomass, litter and leaf surface area estimates of an oak-hickory forest. Oikos, 21:134-141

Nadkarni, N. and Parker. G.G. 1994. A profile of forest canopy science and scientists- who we are, what we want to know and obstacles we face: results of an international survey. Selbyana 15:38-50

Naesset, E. 1997a. Estimating timber volume of forest stands using airborne laser scanner data. Remote Sensing of Environment 61:246-253

Naesset, E. 1997b. Determination of mean tree height of forest stands using airborne lidar scammer data. ISPRS Journal of Photogrammetry and Remote Sensing, 52:49-56

Nelson, R., Krabill, W. and Tonelli, J. 1988a. Estimating forest biomass and volume using airborne laser data. Remote Sensing Environ. 24:247-267

Nelson, R., Swift R., and Krabill, W. 1988b. Using airborne lasers to estimate forest canopy and stand characteristics. J. For. 86:31-38

Niklas, K.J. 1996 Plant Allometry: The Scaling of Form and Process. University of Chicago Press, 1996 395pp

Nilsson, M. 1996. Estimation of tree heights and stand volume using an airborne lidar system. Remote Sens. Environ. 56:1-7 (1996)

Oliver, C.D, and Larson, B.C. 1996. Forest Stand Dynamics. John Wiley & Sons, New York. 435pp

Parker, G.G., O'Neill, B.C., and Higman, D. 1989. Vertical profile and canopy organization in a mixed deciduous forest. Vegetatio 89: 1-12

Parker, G. G. 1995. Structure and microclimate of forest canopies. in M. Lowman and N. Nadkarni, eds. Forest Canopies - A Review of Research on a Biological Frontier. Academic Press, San Diego.

Ritchie, J.J, Evans, D.L., Jacobs, D., Everitt, J.H., Weltz, M.A. 1993. Measuring canopy structure with an airborne laser altimeter. Transactions of the ASAE. Vol 36(4)1235-1238

Schreier, J., Logheed, L.,Gibson, J.R. and Russell, J. (1984) Calibration an airborne laser profiling system. PERS 50(11):1591-1598

Schreier, H., Loughead, J., Tucker, C., and Leckie D. (1985) Automated measurements of terrain reflection and height variations using an airborne infrared system. Int. J. Rem. Sens 6(1):1-1-103

Smith, D.M. 1986. The Practice of Silviculture. 8th ed. John Wiley and Sons, New York.

Warren-Wilson, W.J 1958. Analysis of the spatial distribution of foliage by two-dimensional point-quadrats. New Phytologist, 59, 92-101

Warren-Wilson, W.J. 1965. Stand Structure and light penetration. I. Analysis by point quadrats. Journal of Applied Ecology 2,383-390

Weltz, M.A., Ritchie, J.C., and Fox, H.D. 1994 Comparison of laser and field measurements of vegetation height and canopy cover, Water Resources Research, 30, 1311-1319

Wierman, C.A. and Oliver, C.D. 1979. Crown stratification by species in even-aged mixed stands of Douglas-fir -- western hemlock. Can. J. For. Res. 9:1-9.

FIGURE LEGENDS

1. A) Illustration of the SLICER footprints overlain on the SERC stemmap, and B), detail of the sampling geometry.

2. Steps in the transformation of the lidar waveform into an estimate of the vertical distribution of canopy surface area, the canopy height profile.

3. Regressions between field and SLICER measured quadratic mean canopy height (QMCH) from Lefsky (1997). Field QMCH = 2.002 + 0.798 * SLICER QMCH, R^2=0.78.

4 A) Comparison of basal area measured using a 2 and 20 cm minimum DBH limit calculated with data from the chronosequence dataset, and the observed range of basal areas (calculated with the 20 cm minimum DBH limit) in the stemmap dataset, and B) the same for aboveground biomass.

5. Scatterplots comparing each of the four height indices from the chronosequence dataset to each other. N=48.

6. Scatterplots of predicted vs. observed basal area (a-c) and biomass (d-f), for the chronosequence and stemmap datasets, as predicted from the two best height index regression equations, and by stepwise multiple regression. Dashed line indicates the

identity line, solid line indicates the predicted vs. observed

regression for the stemmap dataset.

Table 1. Stand attributes for the chronosequence and stemmap datasets.

Dataset:	Chronosequence	Stemmap
Stem Data Source	Forty-eight plot observations.	75 plots subsetted from a 32 ha stemmap.
Canopy Height Profile Source	Optical quadrat method.	SLICER waveforms
Number of Plots	48	75
Mean Basal Area(m^2*ha^{-1})	36.1	37.5[1]
Mean Aboveground Biomass (Mg*ha^{-1})	235.9	239.0[1]
Maximum Canopy Height (m)		
Maximum	40.0	44.0
Mean	26.5	36.2
Minimum	4.0	30.0
Mean QMCH (m)	14.4	18.1

Notes: 1) Basal area and Aboveground Biomass reported for the stemmap dataset are for predicted 2 cm stand structure attributes.

Table 2. Regression equations for linear(A) and stepwise multiple regressions(B). Columns 1 through 4 record the name of the dependent variable, the regression equation for each height index, and the r^2 and standard deviation of residuals for equations developed using data from the chronosequence dataset (n=48). Columns 5 and 6 record the mean and standard deviation of residuals which result from the application of each equation to height indices from the stemmap dataset (n=75). Columns 7 through 9 record the r^2, and coefficients of regression between predicted and observed basal area and biomass from the stemmap dataset, and the significance of the difference between the observed "predicted-observed" regression, and identity.

A. Results of linear regressions

Dependent Variable	Equation	r^2	Stdev[1] of Resid.	Mean Residual	Stdev[1] of Residuals	r^2	B_0 $P(B_0=0)$	B_1 $P(b_1=1)$
	Original Regressions using field data from the Chronosequence dataset(n=48)			Application of Regressions to Stemmap Data (n=75)				
				Residual Statistics		Predicted vs. Observed Values for the Stemmap Dataset (Observed=B_0+B_1Predicted)		
Basal Area (m2/Ha)	7.84 + 1.07 * Maximum Canopy Height	60% p<0.0001	8.9	-9.9	5.3	16%	9.3 p=0.02	0.60 p<0.0001
	9.80 + 2.13 * Median Canopy Height	66% p<0.0001	8.3	7.2	6.5	3%	31.2 p<0.0001	0.22 p<0.0001
	6.34 + 2.30 * Mean Canopy Height	65% p<0.0001	8.3	1.9	4.7	28%	3.4 p=0.36	0.96 p=0.68
	6.05 + 2.08 * QMCH	70% p<0.0001	7.8	-2.1	4.4	37%	-4.8 p=0.17	1.1 p=0.47
Biomass (Mg/Ha)	(2.77 + 0.44 * Maximum Canopy Height)^2	80% p<0.0001	73.9	-130.1	62.1	20%	87.8 p=0.001	0.41 p<0.0001
	(4.78 + 0.79 * Median Canopy Height)^2	70% p<0.0001	91.6	80.2	68.1	0%	220.7 p<0.0001	0.12 p<0.0001
	(3.16 + 0.88 * Mean Canopy Height)^2	73% p<0.0001	89.0	28.4	50.8	21%	80.8 p=0.002	0.75 p=0.019
	(2.90 + 0.80 * QMCH)^2	80% p<0.0001	75.1	-16.5	46.7	33%	28.1 p=0.22	0.83 p=0.07

1. Stdev = Standard Deviation

B. Results of Stepwise Multiple Regression

Dependent Variable	Equation	r^2	Stdev[1] of Residual	Mean Residual	Stdev[1] of Residuals	r^2	B_0 $P(B_0=0)$	B_1 $P(b_1=1)$
	Original Regressions using field data from the Chronosequence dataset(n=48)			Application of Regressions to Stemmap Data (n=75)				
				Residual Statistics		Predicted vs.Observed Values for the Stemmap Dataset (Observed=B_0+B_1Predicted)		
Basal Area (m2/Ha)	21.5+(49.1*CHP[16:20]) +(156.2*CHP[28:32]) +(244.8*CHP[36:40])	69% p<0.0001	7.7	6.0	4.5	35%	10.6 p=0.004	0.854 p=0.123
Biomass (Mg/Ha)	91.1+(403.2*CHP[16:20]) +(1597.0*CHP[28:32]) +(4109.8*CHP[36:40])	81% p<0.0001	61.4	50.2	45.8	36%	86.3 p<0.0001	0.81 p=0.004

Table 3. Comparison of coefficients of determination for four datasets; all chronosequence plots, chronosequence plots less than 30 m tall, chronosequence plots greater than or equal to 30 m tall, and all stemmap plots greater than or equal to 30 m tall. Canopy height is abbreviated by CH.

Dependent Variable	Independent Variable	Chrono-sequence All Plots	Chrono-sequence Plots < 30m	Chrono-sequence Plots => 30m	Stemmap Plots =>30m
Basal Area	Maximum CH	60%	47%	24%	16%
	Median CH	66%	44%	40%	3%
	Mean CH	65%	40%	35%	28%
	QMCH	70%	44%	46%	37%
Biomass	Maximum CH	80%	72%	39%	20%
	Median CH	70%	67%	18%	0%
	Mean CH	73%	63%	15%	21%
	QMCH	80%	69%	29%	33%
Number of Plots		48	24	24	75

A.

Location of SLICER waveform

Location of trees > 20 cm DBH

UTM Y Coordinate

4305800

4305600

4305400

4305200

UTM X Coordinate

364200   364400   364600   364800

B.

Individual SLICER waveform, 10 m footprint.

Stemmap sample mask
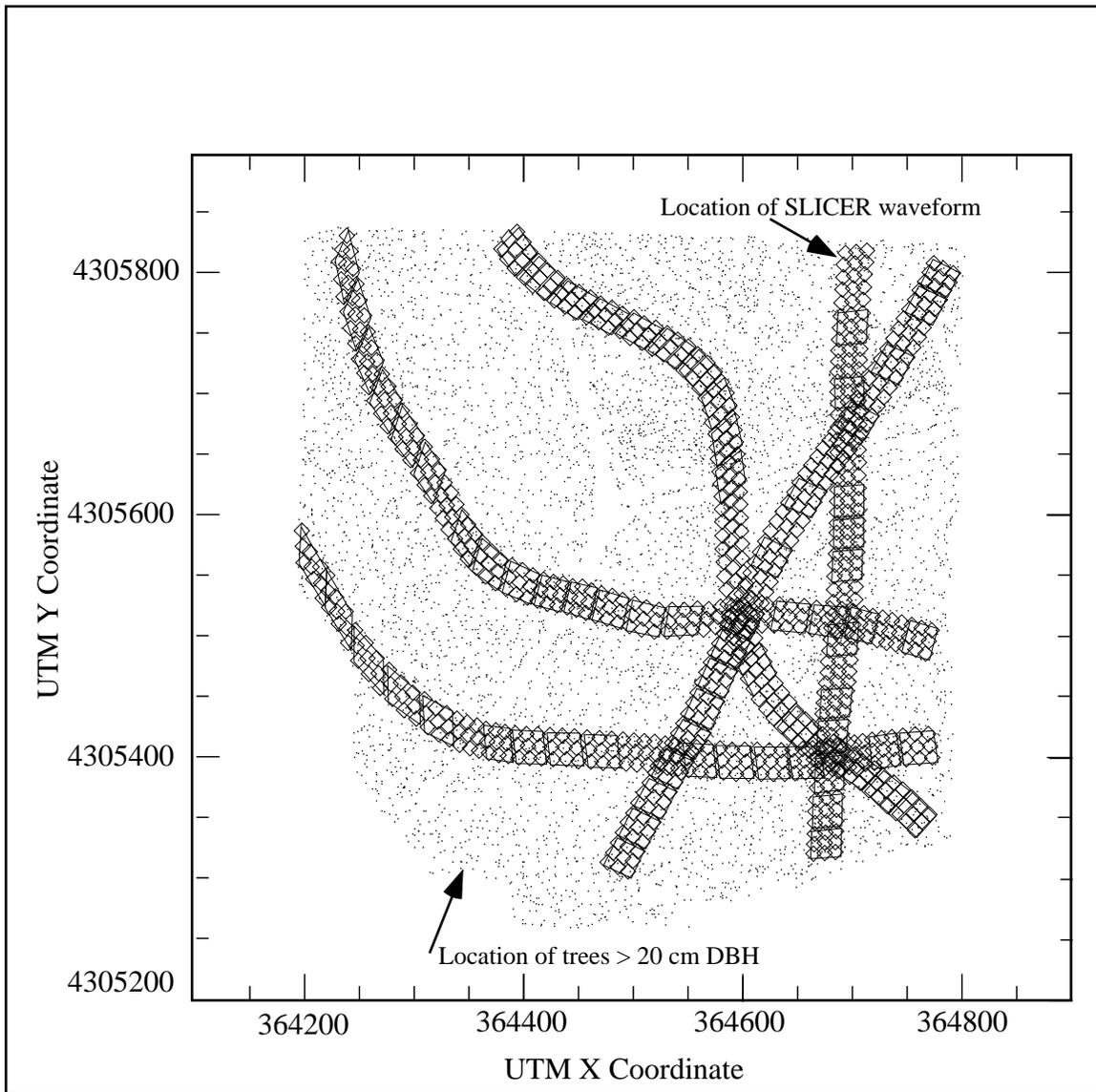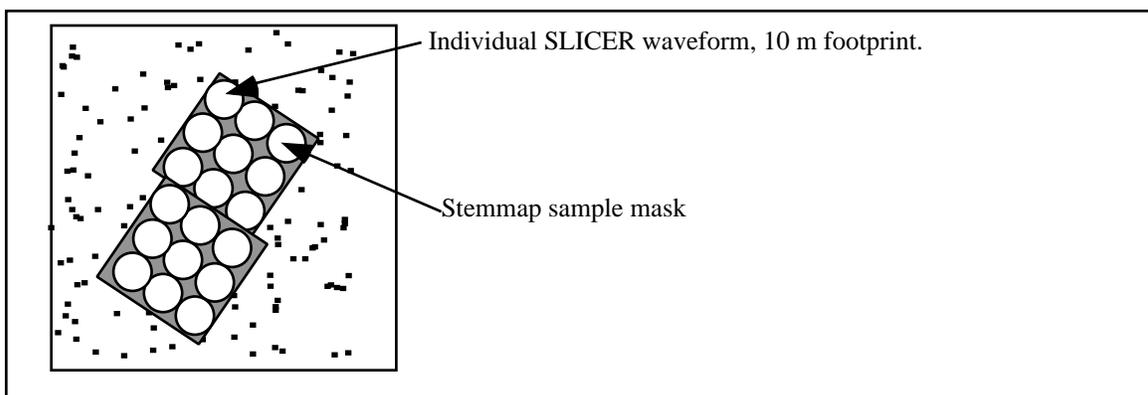
A. Ground Return Processing

1. The first step in processing the lidar waveform is the identification of the peak of the ground return, which is assumed to be the mean elevation of the forest floor.
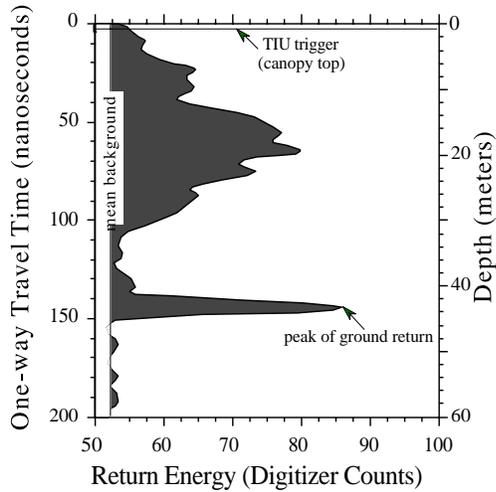
2. The posterior half of the ground return is defined as the total signal between peak and height at which the power of the signal falls below background noise.
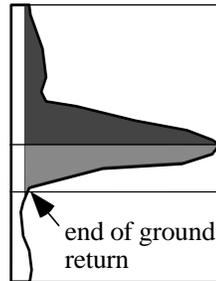
3. The posterior half of the ground return is copied and flipped vertically to define the anterior half of the ground return. Power greater than the level established by the ground return is assumed to be understory canopy.

TIU trigger (canopy top)

mean background

peak of ground return

One-way Travel Time (nanoseconds)

Depth (meters)

Return Energy (Digitizer Counts)

end of ground return

beginning of ground return

understory canopy
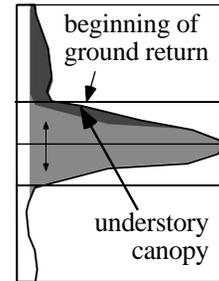
B. Canopy Height Profile Calculation

Normalized Cumulative Return Energy

Canopy Height

Raw Canopy Cover

Return Energy (Digitizer Counts)

Height Above Forest Floor (m)

Cumulative fractional canopy power return

Cumulative Canopy Cover

Fractional Cover

Cumulative CHP
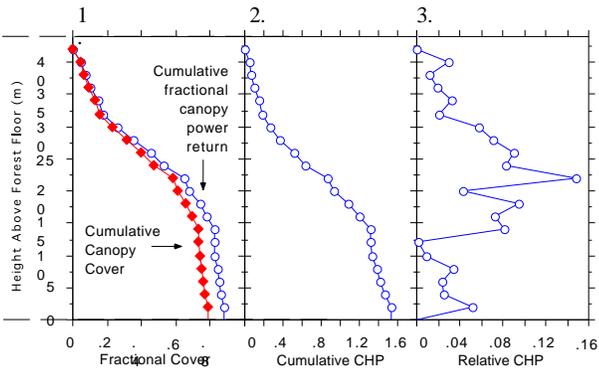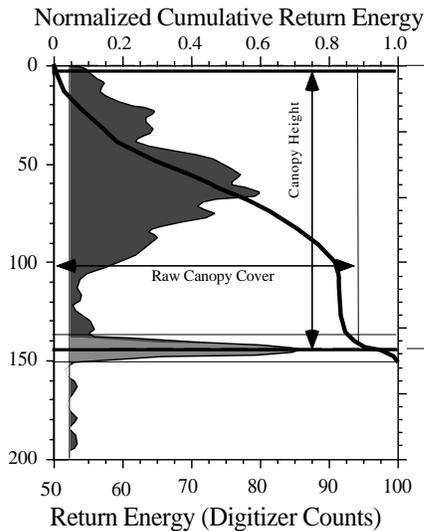
Relative CHP

1. After the ground return is delineated from the remainder of the waveform the fraction of total power returned from the canopy can be calculated. Canopy cover can then be estimated by correcting the canopy power for the relative reflectivities of foliage and soil.

2. The cumulative cover fraction can then be transformed using the MacArthur-Horn (1969)

Basal area (m^2*ha^{-1}) calculated with 2 cm minimum DBH limit vs. Basal area (m^2*ha^{-1}) calculated with 20 cm minimum DBH Limit. Observed range of independent conditions.

Biomass ($Mg*ha^{-1}$) calculated with 2 cm minimum DBH limit vs. Biomass ($Mg*ha^{-1}$) calculated with 20 cm minimum DBH Limit. Observed range of independent conditions.