THE USE OF AIRBORNE LASER ALTIMETRY TO ESTIMATE TROPICAL FOREST BASAL AREA, VOLUME, AND BIOMASS

by

Ross Nelson

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Forestry

APPROVED:

RG Oderwald

R.G. Oderwald, Chairman

T. G. Gregoire

H. E. Burkhart

M. R. Reynolds

P. V. Bolstad

December 20, 1994
Blacksburg, Virginia

The Use of Airborne Laser Altimetry to Estimate Tropical Forest Basal Area, Volume, and Biomass

by

Ross Nelson

Committee Chairman: Dr. Richard Oderwald

Forestry

ABSTRACT

Airborne laser profiling is used to estimate tropical forest basal area, volume, and biomass. A procedure is developed and tested to divorce the laser and ground data collection efforts. The procedure involves 1) the collection of airborne laser data over the area of interest; 2) the collection of ground data in areas of similar forest cover type (in most cases in the same study area, but not necessarily coincident with the laser transects); 3) computer simulation of the forest canopy based on spatial and forest mensuration data available from the ground survey; 4) development of regression relationships between laser and ground measurements based on the simulated forest canopy; and 5) the use of these regressions with the airborne laser data using simple random sampling or line intercept sampling techniques. Variants of this simulation procedure were tested using three distinct data sets acquired in and over the tropical forests of Costa Rica. On two of the three study sites, airborne laser estimates of basal area, volume, and biomass grossly misrepresented ground estimates of same. On the third study site, where the widest ground transect samples were acquired, airborne and ground estimates agree within 10%. Basal area, volume, and biomass prediction inaccuracies in the first two study areas are tied directly to disagreement between simulated laser estimates and their airborne measurements of average canopy height, height variability, and canopy density.

A number of sampling issues were investigated in order to define a large area inventory procedure which utilizes airborne laser data in conjunction with a limited ground sampling effort divorced spatially and temporally from the airborne laser sampling phase. The following results were noted in the analyses of the three study areas. 1) Four ground transect segment lengths were considered: 25, 50, 75, and 100m.

The 25m segment length introduces a level of variability which may severely degrade prediction accuracy in these Costa Rican primary tropical forests. This effect is more pronounced as transect width decreases. A minimum transect length which mitigates significant mean square errors by capturing a representative spatial sample of the primary tropical forest is on the order of 50m. 2) Gaps between airborne laser segments on the order of 15m mitigate the effects of spatial autocorrelation. 3) The decision to transform or not to transform the dependent variable (eg., biomass) is by far the most important factor of those considered in this experiment. The natural log transformation of the dependent variable increases prediction error, and error increases dramatically at the shorter segment lengths. The most accurate models are simple linear models with forced zero intercept and an untransformed dependent variable. 4) Results do not suggest any apparent, consistent advantage to the use of parametric or nonparametric regression techniques; either is appropriate. 5) General linear models are developed to predict basal area, volume, and biomass using airborne laser height metrics. Laser metrics include average canopy height, all pulses ($\bar{h}_a$), average canopy height, canopy hits ($\bar{h}_c$), and the coefficients of variation of these terms (c_a and c_c). Coefficients of determination, where calculated for comparable models with intercepts, range from 0.4 to 0.6. 6) Line Intercept Sampling techniques may be used to estimate $\bar{h}_a$, $\bar{h}_c$, canopy profile area (p), canopy density (g), canopy area, and canopy volume directly using airborne laser data without the need to identify individual tree crowns in the airborne laser data. Efforts to extend these metrics to estimates of basal area, volume, or biomass by attempting to relate canopy area to basal area and canopy volume to merchantable woody volume or above-ground dry biomass were unsuccessful. 7) The assumption of canopy shape needed to develop the simulated airborne laser measurements has a significant impact on the accuracy of the basal area, volume, and biomass estimates. Estimates of volume or biomass developed assuming an elliptical crown, parabolic crown, and conical crown are respectively 5%, 12%, and 23% larger than estimates developed based on spherical cross-sectional assumptions. Based on this body of research, airborne laser and ground sampling procedures are proposed for use for reconnaissance level surveys of inaccessible, forested regions.

ACKNOWLEDGEMENTS

Most of the data utilized in the course of this study were collected by others and generously shared. Dr. Armond Joyce (Stennis Space Center) and Dr. Steve Sader (University of Maine, Orono) provided me with the airborne laser data, the Tileran and LaSelva-5m ground transect data, and the allometric equations for volume and biomass. Drs. Diane and Milton Lieberman gave me their 20m-wide LaSelva ground transect data. Thanks go to the staff at the LaSelva Biological Research Station, and to Drs. Deborah and David Clark in particular, for their support of my 1993 field efforts. This unfunded research could not have been done without the support and generosity of these people. I appreciate it.

Also:

To my parents, Margaret and William Nelson, who thought and continue to think that all education is important and worth pursuing...

To NASA and, indirectly, to the American taxpayer, who gave me the time and financial means to disappear for two years to take the classes I needed....

To the professors and students at VPI, to the people of Blacksburg, and to the people out there in Giles County on the New River and along Spruce Run whose friendliness and openness gave me much pleasure and made my two years in the area some of the happiest and most contented of my life...

I thank y'all.

TABLE OF CONTENTS

LIST OF FIGURES

page

LIST OF TABLES

page

page

I. INTRODUCTION

Airborne laser profiling data may be used to remotely measure tree height. Pulses from a laser transmitter carried aboard an aircraft are directed toward the ground to collect ranging data from aircraft to the top of the canopy, and in some instances, from aircraft to ground. The forward motion of the aircraft in conjunction with the sequential pulses of the laser produce a height transect through the vegetation canopy. Mathematical relationships have been established between the sequential laser height measurements and ground and/or photo measurements of forest canopy height, canopy density (Nelson et al., 1984; Aldred and Bonner, 1985), timber volume, and forest biomass (Maclean and Krabill, 1986; Nelson et al., 1988a). These regression relationships can, in turn, be used in conjunction with airborne LIDAR (LIght Detection And Ranging) data acquired over a particular study area in order to estimate certain biometric characteristics of that area (Nelson et al., 1988b).

Previous studies which have related laser height measures to various forest canopy attributes have depended on the accurate location of the path of the laser profile on the ground. Once specific sections of the laser transect are located on the ground, regression relationships may be developed which estimate the forest characteristic(s) of interest (e.g., ground-measured tree height, total tree biomass, bole volume) as a function of laser metrics related to the canopy height. Laser canopy height data acquired throughout the entire study area serves as the independent variable in regressions which predict height, biomass, or volume.

This need to accurately locate laser transects on the ground limits the use of the airborne laser data to those areas with adequate ground control. Accurate ground control requires features which can be precisely located on the ground and in the laser transect (e.g., roads, edges of lakes, streams, buildings, forest/nonforest or plantation edges). Those areas most appropriate for laser ranging, inaccessible areas, tropical forests, typically have few ground control features. The tropical forests especially are characterized by extensive runs of unbroken canopy. The primary purpose of this research is to develop a laser sampling procedure which may be used to estimate various forest characteristics over regions with no consistent ground control. Necessarily then, this procedure must separate the laser and ground sampling phases.

The overall objective of this research is to develop robust sampling and analysis techniques to facilitate the use of airborne laser profiling for forest biomass and volume assessment when no consistent ground control is available. The specific subobjectives which follow provide a general outline of the study approach.

1. Develop and implement an algorithm which can be used to formulate regression equations used to predict basal area, volume, and biomass using airborne laser data without depending on laser line-ground line colocation. The algorithm involves the following:

- Develop a computer program which utilizes ground reference information to simulate the height characteristics of forest canopies. This program utilizes fixed-area plot data, including tree location, tree bole and canopy size information, to construct a three dimensional digital representation of the plot. The program handles different forest canopy shapes and ground plot sizes.

- Develop a computer program which simulates laser transects across the computer model of forest canopy heights. The program accumulates simulated laser measurements of forest canopy height and pairs these simulated laser measurements with corresponding ground metrics of interest, eg., woody bole volume, biomass, tree height, basal area.

- Process these laser-ground measurement files using simple linear and multiple linear regression techniques in order to develop predictive equations.

- Process the airborne laser data using these predictive equations in order to estimate basal area, volume, and biomass. Assess the accuracy of the laser estimates of forest volume and biomass by comparing the laser results to comparable ground measurements.

2. Investigate the effects of various regression models and regression development techniques on the strength of the prediction models and on the accuracies of the predicted basal area, volume, and biomass. Look at the issues of ground sampling and airborne laser sampling segment length and

ground transect width to try to describe a sampling methodology appropriate to the tropical forests of Costa Rica.

3. Investigate the use of line intercept sampling techniques to develop per-area estimates of basal area and woody volume or biomass. Compare LIS approaches with the more traditional regression approaches to determine if LIS techniques are a suitable vehicle for basal area, volume, and biomass estimation.

4. Using the forest canopy height simulator and the laser transect simulator, quantify the effects of different tree canopy shapes on simulated laser-ground relationships.

5. Discuss and describe the information which needs to be collected on the ground and the procedures which should be used to collect these ground data, in order to reliably simulate forest canopies in a tropical environment. Describe a specific procedure to sample tropical forests, both from the air using lasers and on the ground.

The list of research objectives encapsulates a procedure which is used to develop estimates of tree basal area, volume, and biomass for three different data sets on two different study sites in Costa Rica.

II. LITERATURE REVIEW

Laser ranging measurements are basically measurements of time. Knowing the speed of light, one can calculate the distance associated with the time difference between pulse generation and pulse return. Pulse timers now are reaching into the picosecond range (10^{-12} seconds) but many of the airborne ranging systems currently operating use timers able to discern differences of one nanosecond (10^{-9} seconds, equivalent to a vertical height resolution of approximately 15 cm). Some systems, such as the NASA Airborne Oceanographic LIDAR (AOL) used in this study, have the capability to monitor not only the primary return from the target, but also significant secondary returns from a single laser pulse. If the target is a tree or multistory forest canopy, a portion of the initial pulse may continue through the canopy to lower canopy layers and ultimately to the ground. These secondary returns may be picked up by instrumentation aboard the aircraft so that, in some instances, the capability exists to discern tree height from a single laser pulse (Figure 1).

Most aircraft laser systems acquire profiling data; sequential pulses are directed out the bottom of the aircraft and the forward motion of the aircraft establishes a series of height measurements along a line (Figure 2). A combination of the laser cycling time and the aircraft ground speed determines the pulse interval on the ground. Post-mission processing may be done to identify and "connect" ground hits; interpolation between ground shots provides a canopy height estimate for every pulse along the transect.

II.A. Airborne Lasers for Bathymetry and Terrain Mapping

Airborne laser metric systems were originally developed in the 1960s and 1970s to facilitate terrain mapping and bathymetry studies. Hickman and Hogg (1969) were the first to demonstrate the capabilities of a laser pulse to penetrate through water for bathymetric analyses. Their system made multiple measurements on individual laser pulses which permitted calculations of distances between the aircraft, the water surface, and the ocean bottom, yielding water depth. Hoge et al. (1980) used the NASA LIDAR to measure water depths in the Delmarva area. Link and Collins (1981) describe many of the early terrain-bathymetry mappers, including an early variant of the laser mapper used to collect the airborne LIDAR data in this study. They also note that the greatest hindrance in the use of airborne laser data then was keeping track of the location of the laser transect on the ground (or ocean bottom). Similar problems

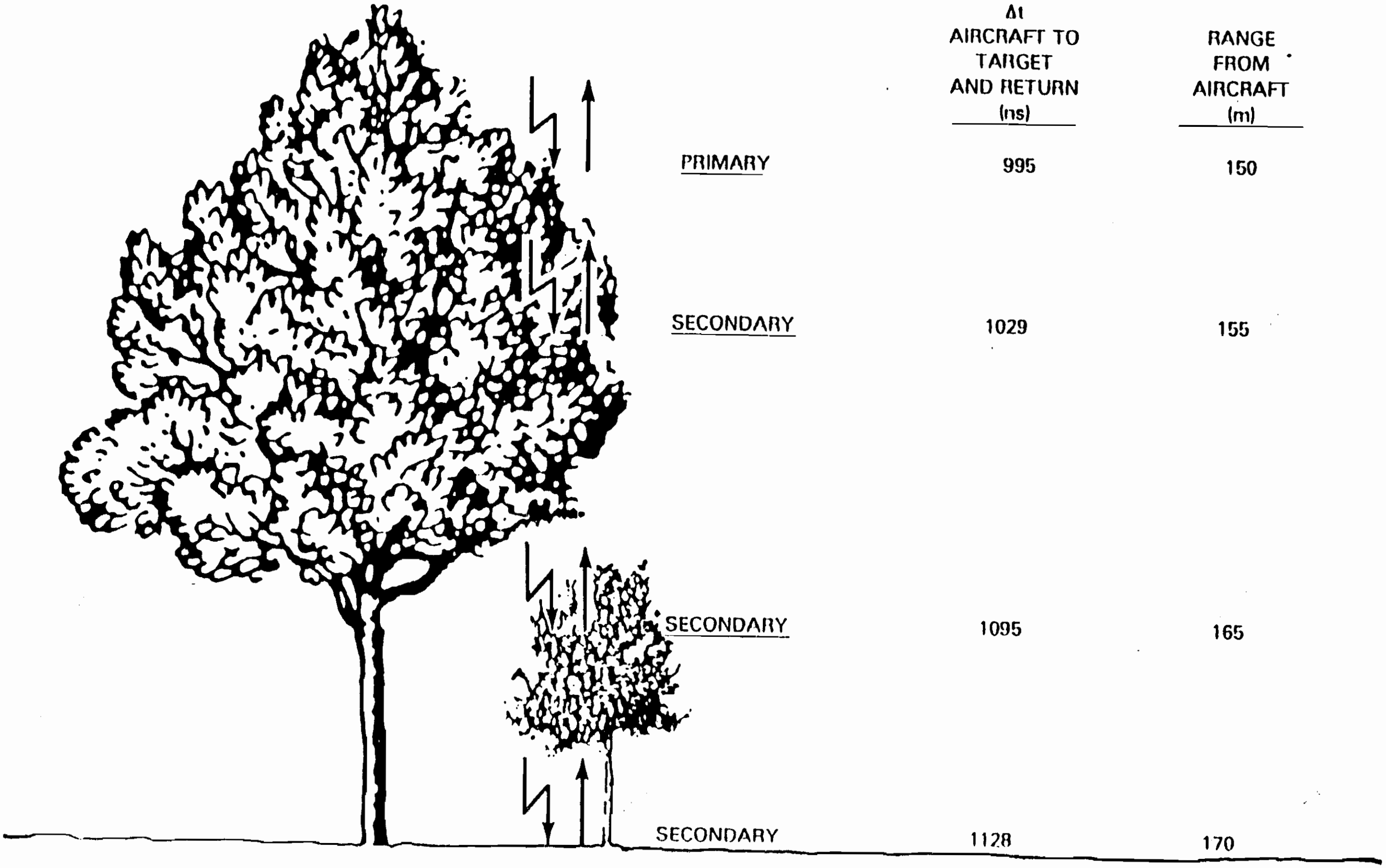


Figure 1. A schematic illustrating primary and secondary pulse returns from a single laser pulse.

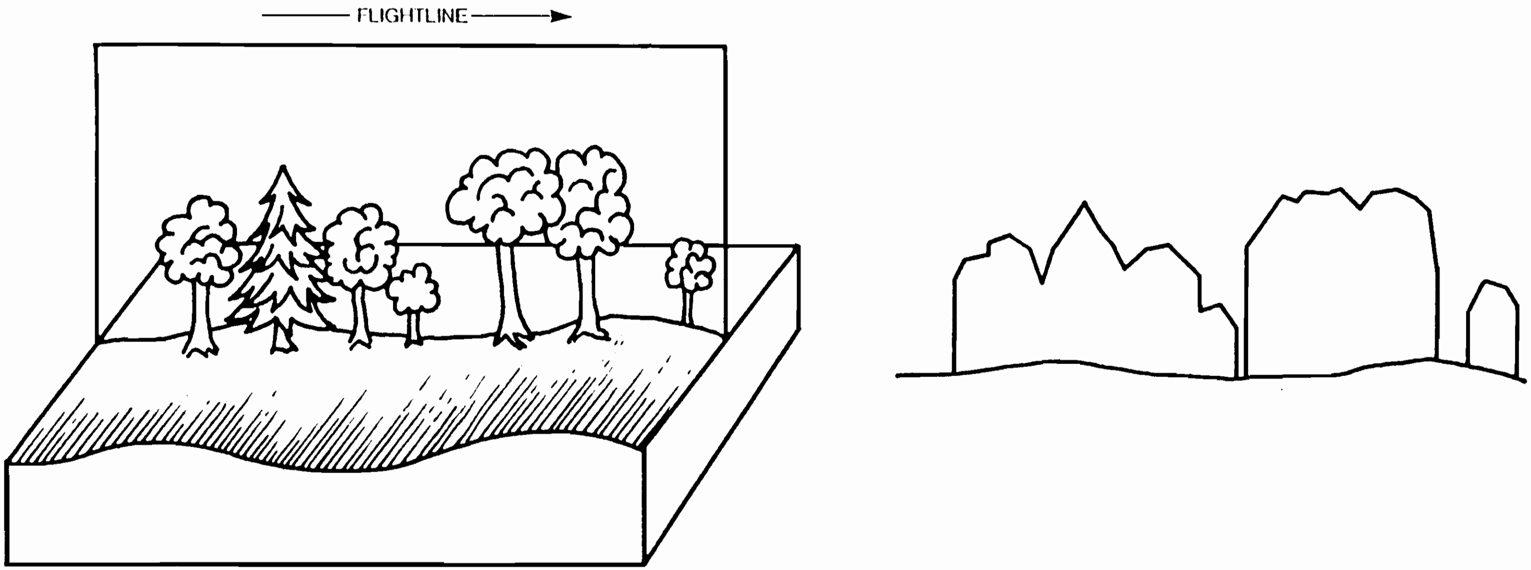


Figure 2. A vertical cross-section of a forest canopy measured by sequential laser pulses along a flightline (Nelson et al., 1984).

constrain the use of airborne LIDAR for forest mensuration even today, especially in closed canopy situations. The laser-ground registration problem, in fact, provides the impetus for this research.

Secondary return tracking, developed for bathymetric applications, also enables terrain mappers to "see" through vegetation to the land surface to facilitate topographic measurements. The vegetation in these terrain studies was a semitransparent source of noise. For instance, Krabill et al. (1984) found that airborne laser topographic measurements of a watershed near Memphis, Tennessee agreed with ground reference measurements within 12-27 cm over open ground, but agreement dropped to 50 cm root mean square (RMS) in forested areas.

Other laser researchers depend only on intermittent pulse penetration through micro-openings in the canopy to find ground. Arp et al. (1982) demonstrated the penetration capabilities of an airborne laser system in the tropics. They used a continuous wave (CW) helium-neon laser to topographically map a proposed reservoir site in Venezuela. The spot size at the target was 20 cm, and the authors presented vegetation profiles (accurate to plus or minus 3 m) depicting ground and canopy traces, suggesting an ability to adequately penetrate dense tropical greenery. Schreier et al. (1984) used a near-infrared gallium-arsenide laser profiling system to obtain terrain data over sites in southern Ontario. They concluded that the laser terrain mapper produced "height measurements on the same order of accuracy as conventional methods" - mean differences on the order of 20-30 cm. In addition they found that vegetation canopy heights could be readily measured. Jackson and Richie (1988) and Richie and Jackson (1989) report on the uses of a laser terrain mapper for identifying gully erosion. They found that gullies with depths as shallow as 10-30 cm and 50 cm wide could be identified using a moving average filter.

II.B. Airborne Lasers for Vegetation Assessment

Those characteristics which make airborne laser systems useful for mapping extensive sections of terrain also make them useful for assessment of the vegetation directly. These characteristics include high sampling rates, extensive areal coverage, ability to penetrate and/or "see" beneath a vegetation canopy, and accurate ranging measurements. In this context, vegetation is no longer a source of noise but the primary focus.

The foundations of laser forest mensuration lie with photogrammetric techniques developed to assess tree height, volume, and biomass. Aerial stand volume tables have been developed to estimate forest volume based on average canopy height, crown diameter, tree counts per unit area, and/or crown closure (Paine, 1977; Avery and Burkhart, 1983). Spurr (1960) introduced a variant by relating canopy profile area to stand volume. The canopy profile area (CPA) is that area between a trace of the top of the forest canopy and the ground over a given distance. Spurr measured CPA photogrammetrically by selecting transect endpoints and then sampling canopy heights along the transect. Smith (1969) developed correlations between photogrammetrically estimated CPA, film density values, and ground-measured timber volume using a Kelsh plotter to estimate tree heights and a spot densitometer to estimate film darkness. Maclean (1982) and Maclean and Martin (1984) continued in the photographic vein, investigating further the relationships explored by Smith. They found that the natural logarithm of timber volume is linearly related to the corrected CPA. This corrected CPA was calculated by subtracting the area of micro-openings (a function of the film density readings) from the uncorrected CPA. They also found that stratification by species significantly improved the strength of the ln(volume) - CPA relationship. A few years after this work, an airborne laser system was used to obtain CPA estimates previously retrieved photogrammetrically.

Numerous investigators in the U.S., Canada, and the U.S.S.R. have participated in airborne laser research for forest mensuration. Much of the Russian work revolves around a researcher named V.I. Solodukhin who began investigating the theoretical aspects of laser profiling in the mid 1970s and moved into airborne laser reconnaissance for forest mensuration in the early 1980s. Solodukhin et al. (1977a) used a ground-based laser rangefinder to draw tree crown profiles of felled trees, noting good agreement between laser results and tape-measured results. They concluded that a "laser profileograph" having a range of up to 300 m could be used to remotely sense tree crowns from aircraft. Solodukhin et al. (1977b) discussed the theoretical mathematics of randomly transecting tree crowns to predict tree height and canopy diameter. They found that the difference between the mean laser canopy height (the mean of numerous random transects across a given canopy) and the actual height depended on canopy shape and is predictable if the canopy shape is known. They considered five different tree canopy shapes, including horizontal and vertical ellipsoids, spheroids, paraboloids, and vertical cylinders. Differences between mean laser heights and the actual tree heights were greatest for the conically-shaped crowns, lowest for the vertical cylinders. They analyzed simulated laser profiling data of spruce, pine, dwarf-pine, and larch stands and found that mean laser canopy height underestimated ground-measured heights by 26% in spruce stands, 8-10% in the other species. The more coniform the canopy, the greater the laser underestimate of true height. Solodukhin et al. (1979) report the first Soviet trials with an airborne laser profiler; in these trials, the laser ranging measurements were graphically recorded on a photo-paper strip. By 1985, though still recording data graphically, Solodukhin et al. (1985) related profilogram measurements to ground-measured crown and stand heights and densities. Stolyanov and Solodukhin (1987) report the move to magnetic recording of the laser profileograms. The overall impression of the Russian work is one where the researchers develop a strong theoretical basis for the expected responses of the airborne laser, but are constrained by computer and instrumentation development.

U.S. and Canadian investigations have compared laser and ground or photo estimates of canopy height and canopy density. Nelson et al. (1984) analyzed airborne laser data acquired over an oak-hickory forest in south-central Pennsylvania. They found that mean laser estimates of canopy height slightly underestimated photogrammetric measurements, differing on the average by 60 cm. They also found that the presence or absence of secondary returns beneath a forest canopy were related to canopy density. Laser metrics were used to explain 50-60% of the variation in photointerpreted canopy closure estimates. Research by Schreier et al. (1984) demonstrated the accuracy of the laser tree height measurements. They found the mean height difference between laser and photographic height estimates were 24 cm, and 95% of all laser heights were within 1.8 m of heights determined photogrammetrically. Schreier et al. (1985) used laser height, the strength of the laser return, and the variability of the strength of return to develop an automated discriminator to differentiate conifer and broadleaf forests near Chalk River, Ontario. They concluded that 1) lasers could be used to accurately measure tree heights; 2) laser reflection measurements may be used to differentiate coniferous and broadleaved trees; and 3) multiple wavelength transmitters could be used to provide a multispectral component for automated vegetation identification. Aldred and Bonner (1985) found that laser heights underestimated ground or photogrammetric estimates of tree height, and concluded that laser height measures should generally be within plus or minus 2 m of the reference value two-thirds of the time. They also found that 20% crown cover density classes were correctly identified 62% of the time using airborne laser data, and laser estimates of canopy closure were within one 20% class 89% of the time. Aldred and Bonner provided additional information on the effects of laser spot size and canopy morphology. Jensen et al. (1987) presented a different application of laser height data by combining airborne multispectral scanner (MSS) and airborne LIDAR data. They found, on a wetland study area in South Carolina, that in-situ measurements of tree height were highly correlated (R-square value of 0.988) with LIDAR heights. Using the MSS data, they produced a vegetation classification map of the study area. The laser transect data was used to characterize the mean height of each vegetation class, with the end result being a three dimensional vegetation map of the entire study area. Currie et al. (1989) integrated a laser rangefinder and a multispectral video system in order to develop a system which could generate a three-dimensional digital representation of the forest canopy. They found that 1) on the average, tree heights measured using the laser profiler were within 0.5 m of ground-measured tree heights; and 2) crown closure was more easily determined in single-species coniferous stands because of the difficulty differentiating ground cover from deciduous crowns.

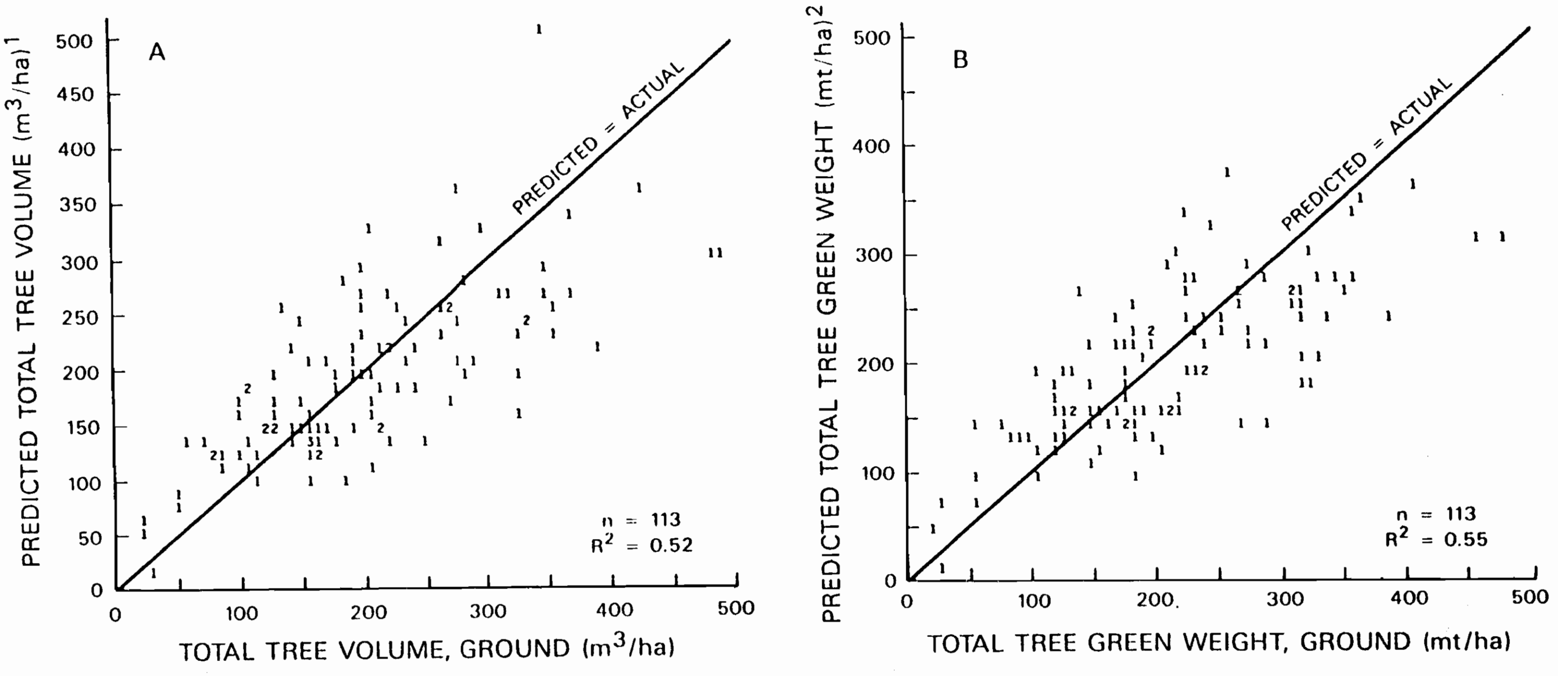
Maclean and Krabill (1986) and Nelson et al. (1988a,b) investigated the utility of airborne LIDAR for forest volume and biomass assessment. Maclean and Krabill's work on the Delmarva penninsula builds directly upon volume-canopy height relationships established photogrammetrically by Maclean and Martin (1984). In their 1986 study, Maclean and Krabill showed that laser-based variants of CPA could be used to predict gross merchantable timber volume. They found that a stratification by species, effectively a conifer-hardwood breakdown, improved the strength of these mathematical relationships. Nelson et al. (1988a) conducted a regression study to predict total tree volume and above-ground woody biomass using airborne laser height metrics. They found, using data acquired over the southern pine forests of southwestern Georgia, that the models explained between 53-65% of the volume or biomass variance. The best model follows:

$$\ln(v_t) = a_1(\bar{h}_a) + a_2 \ln(\bar{h}_a) + a_0$$

where v_t = total tree green weight;
$\bar{h}_a$ = average canopy height (proportional to CPA);
and a_0, a_1, a_2 are the linear regression coefficients.

Figure 3 illustrates the relationship between ground-based estimates of volume or biomass and the associated predicted values for the 113 - 20 meter plots considered. The researchers concluded that, with limitations, laser profiling data may be used to predict timber volume and biomass. "Mean volume and biomass values - ground vs. laser - are within 3% when 38 test plots were checked." However site-specific variability indicates that a laser profiler may fall short as an accurate, site-specific mapping tool. Nelson et al. (1988b) applied these regression equations to the laser transect data collected over the entire 6000 ha study area to develop tractwide estimates of forest volume and biomass. They used semi-variogram techniques to estimate the distance between independent laser pulses (i.e., pulses which are not autocorrelated). A single laser height observation every 11m along a transect was used to estimate volume and biomass. Their results showed that laser-based estimates of biomass and volume are precise; comparable sections of the flightlines yielded forest estimates within 3-6%. The laser-ground comparisons yielded differences of 7-8% although the authors note that the proximity of the laser and ground samples may result in an underestimate of the true difference.

Hyyppa et al. (1992) placed a ranging scatterometer (radar) on a helicopter in an attempt to determine tree heights and volume on a pine-spruce forest in Finland. (See Hallikainen et al. (1990) for an overview; Hyyppa (1993) provides comprehensive equipment descriptions and results.) The scatterometer collects ranging data to the top of the forest canopy and to the ground and, like an airborne laser profiler, it may be used to recover the canopy profile. They compared the microwave-measured forest canopy profile to ground measurements obtained on 20x20 meter plots spaced along the flight transects. Hyyppa et al. (1992) used the microwave ranging data to retrieve 1) total volume per hectare; 2) dominant height; and 3) mean height. They found that the correlation coefficient between ground-measured and radar-measured mean and dominant tree height was 0.99 with a slope of approx. 1.0 for both relationships. The results suggest that microwave height measurements can be used to directly, accurately measure ground attributes. Ground-measured and radar-estimated total volume were also highly correlated



volume: $\ln(v_t) = 0.03470(\bar{h}_o) + 0.44810[\ln(\bar{h}_o)] + 3.61016$

biomass: $\ln(v_t) = 0.02595(\bar{h}_o) + 0.51176[\ln(\bar{h}_o)] + 10.54605$

where v_t = total tree volume (m^3) or total tree green weight (kg)
$\bar{h}_o$ = average canopy height

Figure 3. Relationship between predicted versus ground-measured total tree merchantable volume and total tree green weight for 113 - 20m plots located in the southern pine forests of southwestern Georgia (Nelson et al., 1988a).

(r=0.86), though the radar seemed to overestimate ground-measured volume in the high-volume stands. Hyyppa (1993, personal communication) has compared radar and laser estimates of tree height and stem volume and noted that the radar results are more accurate. He cautions that his predictive radar models work best in open stands since the models require that individual tree crowns be identified and measured. Work by Hyyppa and others at the Helsinki University of Technology suggests that the microwave region may be more useful for characterizing forest canopy attributes because of improved canopy penetration capabilities.

The studies reviewed above require ground observations coincident with the remotely sensed observations in order to develop the regression equations used for prediction. Fufilling the requirement for coincident data is difficult for two reasons. First, it is difficult to locate the laser transect on the ground, and the difficulty and locational inaccuracy increases as the frequency of ground control decreases. Second, laser and ground observations are not directly comparable even if the registration is perfect. The laser transect is a two dimensional slice through the forest canopy, while the associated ground plots, to this point, have been fixed area or polyareal plot samples. Although most of the research to date has relied on 1) ground transect-laser transect registration, and 2) areal plot sampling on the ground, the linear nature of fixed-wing, airborne laser data suggests that line intercept sampling techniques shoud be investigated to relate ground and aircraft measurements.

II.C. Line Intercept Sampling

Although the literature is muddy on this issue, Pielou (1983) and Warren (1990) take care to draw a distinction between Line Intercept Sampling and Line Intersect Sampling (LIS is used as an acronym for both). The two sampling procedures are similar in that items are selected for sampling only if they are transected by a sample line randomly placed. The two differ in their treatment of an individual object and that object's length of intersection with the sample line. Line Intercept Sampling involves measurement of the length of interception of the transect line with the object. The length (or sum of the lengths) of interception is used to estimate areal cover. Line Intersect Sampling, on the other hand, is a method of unequal probability sampling where the measured attribute of an object (eg., volume) is weighted by its length of intersection. Line Intercept Sampling is utilized in this study for two reasons. First, at this point, it is not possible to reliably differentiate individual tree crowns in the airborne laser data, a prerequisite to the use of Line Intersect Sampling. Second, even if crowns could be differentiated, metrics of interest (eg., basal area, volume, biomass) could not be measured on the individual trees unless the airborne laser transect is exactly located on the ground and the intersected trees sampled, thereby nullifying any benefit to be derived from remotely sensing the forest canopy. Both approaches, however, are related, the nomenclature is still murky and not always respected, and their development is intertwined. LIS history and theory is reviewed below because it provides one statistical framework for compiling biometric measurements on a forest canopy. LIS is considered in this study as an alternative to more traditional approaches for estimating basal area, volume, or biomass.

Canfield (1941) was one of the first to use a line intercept method for assessing vegetation. The procedure involves running a straight, randomly oriented transect across a portion of the study area and measuring characteristics of interest on the intercepted vegetation. Canfield used the technique to measure rangeland grass and shrub densities and species composition. These figures were calculated as percentages, i.e., the length of line intercepting vegetation versus the total length of the line. Volume estimates required laying out a "belt", a transect of finite width, within which clippings were made. Though the precursors for a line interception procedure were presented, volumetric assessment depended on a fixed-area plot survey.

Warren and Olsen (1964) were the first to develop and report on a procedure whereby measurements made along a line could be used directly for volume estimation. In searching for an efficient means of sampling logging residue, they surmised that the number of pieces intersected along a randomly oriented transect (line) of known length should be related to the amount of logging waste. Their formula for volume follows:

$$V = \frac{660(a)(n)}{(I)(B)}$$

where V = volume per acre,
a = an estimate of volume/piece, derived by limited preliminary sampling,
n = number of pieces intersected by the line,
I = a factor depending on piece orientation relative to the sampling line,
and B = line length, in chains.

Using their approach, then, logging slash volume could be determined quickly (once "a" and "I" were determined via preliminary sampling) by counting the number of line intersections along a given transect length. The authors noted that "a" would have to be recalculated if the logging practices changed, and "I" would vary by transect if the population being sampled exhibited a preferred orientation. If such an orientation exists, as might occur in skyline yarding, then the sampling team should locate their transects at right angles to the long axis of the oriented materials. At this point, it should be noted that preferred orientations may be present in forests flown by an airborne laser. Such phenomena may occur at local or regional scales. At the local scale, the presence of plantations, straight drainage features; or roads might require cross-transection. At a regional scale, linear colonization effects (as found in Rondonia, Acre, and Mato Grosso, Brazil), topographic effects (e.g., the Ridge and Valley of the Appalachians), and hydrologic features such as large rivers - the Amazon and its primary tributaries - may require that laser transects be laid out at systematic angles.

Van Wagner (1968) generalized Warren and Olsen's findings to the measurement of wood volume on the ground in any form. He provides the theoretical mathematics to prove that volume can be calculated based on piece diameter only. He also presents an error analysis showing how error is reduced as the number of systematically-oriented transects increases. De Vries (1973) presents the generalized overview and theory behind Line Intersect Sampling for estimation of slash volume. De Vries (1974) reports formulas for two and three stage sampling using Line Intersect Sampling techniques at the primary stage. Such multistage techniques would be useful where specific characteristics associated with individual intersected pieces are of interest.

Pickford and Hazard (1978) report the results of a comprehensive simulation study of Line Intersect Sampling for logging slash volume assessment. They found that the sample size necessary for a given precision is inversely proportional to piece density, which confirms findings by Warren and Olsen (1964). Variance reduction, then, depends not on the length of the line, rather, it hinges on the number of pieces intersected. In terms of laser forest mensuration, this finding suggests that shorter total laser line lengths are needed in closed canopy situations; an open canopy would require a longer total line length. The observation, though important from a ground sampling perspective, is somewhat moot with airborne laser sampling since 1) an aircraft LIDAR experiment can quickly overwhelm the researcher with transect data; and 2) individual tree canopies are not currently considered in airborne laser data analysis of closed canopies.

Meeuwig and Budy (1981) applied Line Intersect Sampling techniques to pinyon-juniper forests in the southwest to estimate biomass and/or volumes of different tree components (cordwood, slash, foliage). They suggest that point sampling <u>or</u> LIS techniques be used to assess pinyon-juniper stands. The choice is made depending on local tree morphology. The low growing characteristics of the juniper often make point sampling difficult; providing the impetus to use LIS. LIS samples with probability proportional to crown diameter; tally trees are identified as those where the line intersects any portion of the tree canopy, thereby circumventing the need to occularly estimate tree dbh. Their formula for calculating per-area biomass is straightforward and, like Warren and Olsen (1964), depends on preliminary samples to establish the relationship between crown diameter and biomass.

$$B = (k/L)\sum_{i=1}^{n}\left(\frac{b}{c}\right)_i$$

where B = biomass in kg/ha,
k = 10,000 m^2/ha,
L = transect length in m,
n = number of sample trees intercepted along transect of length L,
and $\left(\frac{b}{c}\right)_i$ = ratio of biomass to crown diameter (kg/m) for tree i, as estimated by regression.

Others involved in illustrating and extending the applications of Line Intersect Sampling include Berisford et al. (1985), who used the multistage approach reported by De Vries (1974) to assess bark beetle populations in logging slash. Hansen (1985) used Line Intersect Sampling techniques in conjunction with aerial photography to estimate wooded strip resources in Kansas. He used a ratio-of-means estimator to relate the number of line intersections to wooded strip length on a subset of townships in Kansas. He then applied this estimator to line intersect counts for all townships throughout the state. Hazard and Pickford (1986) continued the logging slash simulation work first reported by Pickford and Hazard (1978). They simulated nonrandom slash patterns and repeatedly sampled these populations using different Line Intersect Sampling procedures (e.g., varying line orientations, numbers of lines). They found that, in nonrandom populations, 3 lines 60 degrees apart and randomly oriented issued from each point of a systematic grid provided the maximum accuracy.

Ritchie et al. (1990) demonstrated the use of airborne lasers and Line Intercept Sampling techniques to estimate shrub canopy closure on a rangeland in south Texas. They compared airborne laser estimates of canopy closure to ground estimates developed using line intercept sampling techniques. They found, using their pencil-beam system (footprint size, 15 cm) that 1) field and laser measurements of canopy density were highly correlated, $R^2 = 0.96$; however 2) differences between the estimates were significant. The laser data underestimates ground-measured densities by and average of three percent. They conclude that "...airborne laser measurements of land surface features can provide estimates of canopy cover for large areas of rangelands."

The findings of these studies provide some guidance for assessing potentially nonrandom populations using an airborne laser. First, transect lines should be flown at right angles or, if possible, at 60 degree angles, to circumvent problems associated with predictable spatial patterns that might be associated with topography, drainage, or man's activities. Second, expect to fly longer transects and to expand (spatially) the ground reference data collection as the forest opens up. In general, the acquisition of airborne laser data will not be a limiting factor, however canopy density should be given consideration when the ground segment is sampled. Third, if Line Intercept Sampling techniques are used to assess forest biomass, volume, basal area, or some other forest metric not directly related to canopy density or tree height, then researchers processing airborne laser data must rely on regression relationships to infer these metrics.

Hansen (1985) concisely summarizes the two general approaches available under the heading of Line Intersect Sampling. The first, reported by De Vries (1973, 1986) uses the inverse of the probability of selection to weight each item intersected. In the case of "needles" (eg., logging slash), the probability of selection is a function of the length of the item. In the case of "circles", the probability of selection is a function of the diameter (eg., crown diameter). This length must be measured for each item intersected unless length is the only measurement of interest and variance does not have to be calculated. In airborne laser sampling, the analyst cannot measure the true length of an intersected item, i.e., the true length of a log or the actual maximum canopy diameter. The analyst can, at best, only determine that an item has been intersected and, possibly, measure the length of that intersection. Therefore, the analyst is forced to use regression estimators to infer the relationship between the apparent canopy diameter (that canopy length actually intercepted by the laser profiler) and the actual canopy diameter (as measured on the ground). The inferred canopy diameter, then, could be used as the weighting factor (i.e., the reciprocal of the probability of intersection) in the first of the two general approaches discussed by Hansen (1985). The second approach, which in fact Hansen (1985) uses in his study, involves the development of the linear relationship between the number of intersections along the transect and the characteristic of interest, in his study, wooded strip length. This second general approach has the advantage that variance estimates can be made without measuring or inferring the "needle length" or diameter of each item intersected. However, this approach has the distinct drawbacks that 1) the number of intersections should be related to the ground metric of interest (i.e., volume or biomass in this laser study); and 2) individual tree crowns would have to be reliably identified in the airborne laser data in order to employ this LIS approach. There is little reason to believe that any sort of coherent regression relationship could be developed to support (1), and attempts to automate (2) have proven difficult and unreliable due to indistinct crown junctures in a closed forest canopy.

The LIS studies reviewed thus far have dealt with items where the important characteristic is the length of the item. The length of a stick of wood drives the calculation of the probability of selection - the longer the randomly oriented stick, the more likely that item is to be transected by a line. Sequential airborne laser pulses intersect tree crowns which are basically circles; the probability of intersecting a circle is a function of the diameter of that circle. The Line Intersect Sampling calculations for "needles" vs "circles" differ, and are presented in De Vries (1973); De Vries and Van Eijnsberger (1973) present calculations necessary to handle arbitrary shapes. The relevant aspects of both articles are summarized in De Vries (1986), a reference that may be more accessible to U.S. researchers.

Kaiser (1983) reviews Line Intercept Sampling procedures which do not require or rely on the identification and enumeration of number of objects intersected, unless number of items is the measurement of interest. Given 1) a population of particles P_i, i=1,N of arbitrary shapes, eg., tree crowns, in region R which has area A; 2) a randomly located transect in R of fixed length L with orientation θ; 3) that each P_i has area a_i, and canopy volume v_i; 4) that d_i and c_i are equal to the length of intersection and area of intersection, respectively, of particle P_i with the line in R which contains the transect; and 5) t_i is a Bernoulli variable, $t_i = 0$ if P_i not intersected, $t_i = 1$ if P_i intersected (see Figure 4). In a forest, then, a_i is the area covered by the vertical projection of the tree crown and v_i is the canopy volume, or that volume included in the space between the top of tree i's canopy and the vertical projection of canopy area on the ground. If tree i is vertically profiled using an airborne laser, d_i is the vertical projection of the length of transect intersection and c_i is the area of the vertical planar surface between the laser trace in the canopy and the laser trace on the ground. The ground measurements a and v can be estimated using the laser metrics c and d as follows:

$$\frac{\sum_{i=1}^{N} a_i}{A} = 1/L \sum_{i=1}^{N} t_i \; d_i ,$$

$$\frac{\sum_{i=1}^{N} v_i}{A} = 1/L \sum_{i=1}^{N} t_i \; c_i$$

As pointed out by Kaiser (1983), "only the length of intersection ($\sum d_i$) needs to be measured to obtain an unbiased estimate of $\sum a_i$/A" and "only the area of intersection ($\sum c_i$) needs to be measured to obtain an unbiased estimator of $\sum v_i$/A". These equations are simple and intuitive. The $\sum a_i$/A is really just a

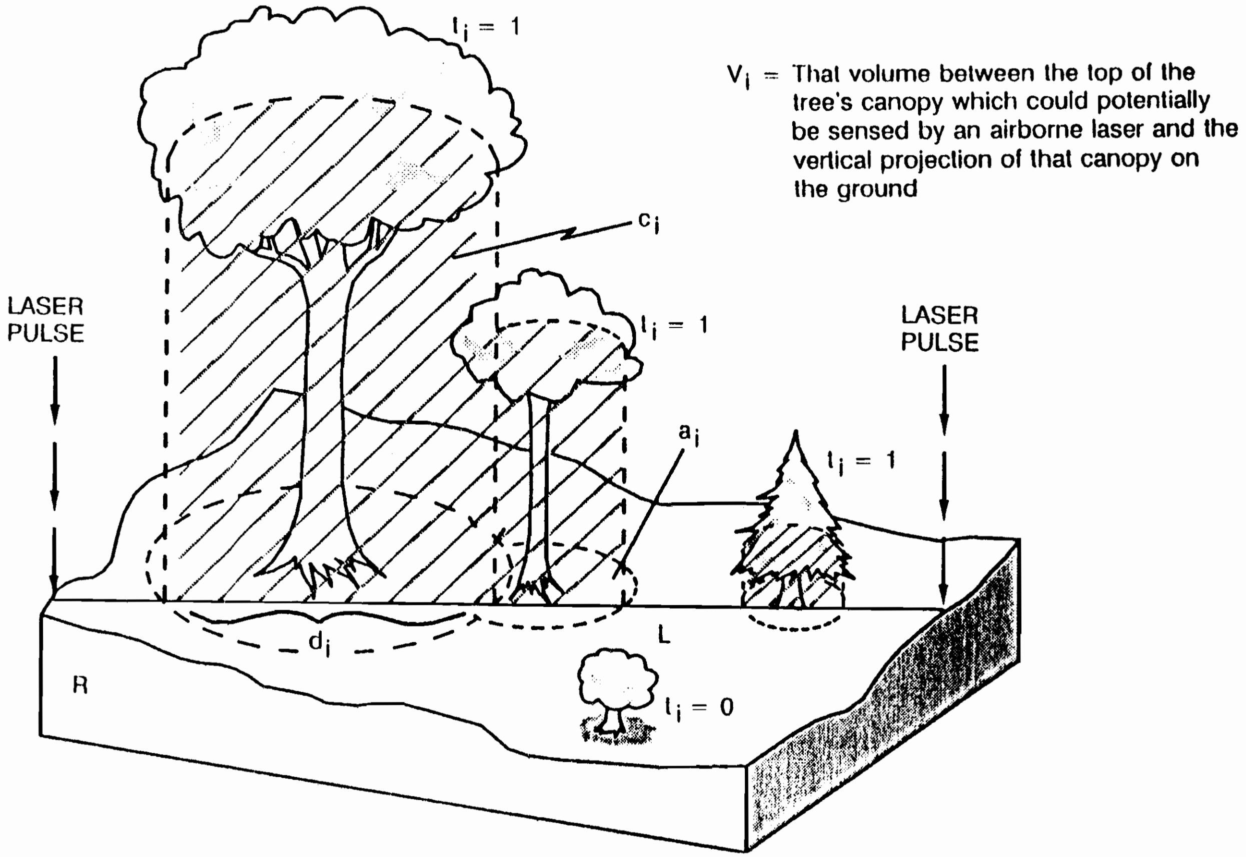


Figure 4. Airborne laser measurements made in the context of Line Intercept Sampling. c_i is tree i's cross-sectional area (canopy profile area) between the laser trace in the canopy and the vertical projection of that trace on the ground. a_i is the crown area as viewed from above. d_i is the horizontal length of canopy intersection. L is the total length of the laser line in region R, and t_i is a Bernoulli variable denoting transect intersection with a tree crown.

measure of canopy closure, with 0 equivalent to no canopy and 1.0 equivalent to complete closure. It seems reasonable that a ratio of areas, canopy vs entire area, would be equivalent to a ratio of line lengths, canopy versus the entire line. Less intuitive, perhaps, is the fact that a ratio of volume per unit area, which might be viewed as an average canopy height, could be estimated by dividing cross-sectional canopy area by total line length. The Kaiser Line Intercept procedures were investigated in this study to see if the statistical approaches could be used to estimate basal area and forest volume or biomass. The specific formulas as they apply to the airborne laser mensuration problem are detailed in the following sections.

III. PROCEDURE

The overall objective of this research is to develop a procedure which can be used to predict basal area, merchantable volume, and total above-ground dry biomass of tropical forests using airborne laser data. A key component of the research is the development of a procedure which does not rely on airborne laser - ground transect colocation. The research, then, includes two distinct components, i.e., 1) the development of methods to estimate ground measurements of interest (eg., biomass) using airborne laser data and ground data which are not necessarily coincident; and 2) the development of a statistical framework within which these methods may be used in order to produce valid estimates of the ground measurements of interest.

Section III.A provides details concerning the characteristics of the study area and ground and laser data collection procedures. Included are descriptions of the three ground reference data sets and the two airborne laser data sets used to develop and test the computational and statistical procedures.

Section III.B outlines the methods used to produce estimates of basal area, volume, and biomass based on airborne laser profiling data. A forest canopy height simulator constructs the height structure at the top of the forest canopy based on fixed-area ground plot measurements. This two-dimensional height array is randomly transected to simulate an airborne laser profile. The result is a concatenation of ground measurements and simulated airborne laser measurements which may be used to develop regression equations relating the two. These prediction equations are applied to the airborne laser data to develop areal estimates and variance estimates for the ground measurements of interest.

Section III.C deals with issues of statistical/procedural methodology. This section outlines the approach taken to try to find the best way to develop predictive equations and to analyze the airborne laser data in order to effect the most accurate, stable estimates of basal area, volume, and biomass. This portion may be broken into three distinct sections. Section III.C.1 describes the use of a simple linear regression procedure with a single independent laser variable, average canopy height, to predict basal area, volume, and biomass. An exponential model ($b = \exp(a_0 + a_1\bar{h}_a)$), the utility of which has been documented by previous researchers, is one of numerous simple linear regressions considered in order to investigate the mechanics of developing predictive equations used to process the airborne laser data. The effects of four factors - regression approach, ground transect segment length, linear model, and airborne laser segment length - are assessed to quantify their effects on estimation of basal area, volume, and biomass. All possible combinations of the different factor levels are applied to a simple linear model using average height as the independent variable. The results are assessed to identify empirically 1) those factors which significantly affect prediction accuracy; and 2) those combinations of treatments which produce the most accurate estimates of basal area, volume, and biomass. Section III.C.2 extends the results of the previous section to multiple independent laser variables in order to determine the predictive utility of additional laser metrics. Section III.C.3 considers line intercept sampling (LIS) in order to compare and contrast this sampling procedure with the more traditional approaches described in III.C.1 and 2.

Section III.D is a sensitivity analysis to document the effects of changing canopy shape on prediction equations and on airborne laser estimates of basal area, volume, and biomass. This portion of the study uses the appropriate procedural and statistical methods identified above to quantify the effects of changing canopy shape. This research extends seminal work reported by Nelson and Yaya (1990).

Figure 5 summarizes the overall procedure used to develop estimates of ground metrics of interest (eg., basal area, biomass) using airborne laser data. This analysis procedure was tested using three different ground transect data sets, one near the town of Tileran, Costa Rica, two on the LaSelva Biological Station near Puerto Viejo, Costa Rica.

III.A. Study Areas and Data Collection

Two airborne laser data sets and two ground transect data sets were collected in the mid 1980s, one air-ground set near Tileran, the second air-ground set at LaSelva (Sader, 1987). At the time of the collections, the ground segments were thought to be coincident with sections of the airborne laser transects. Subsequent analyses by Sader and Joyce (1989, personal communication) called the accuracy of these colocations into question to the extent that specific sections of the airborne laser data could not be registered with the ground segments. A third ground data set was collected independently of the laser investigation in the late 1980s on LaSelva. By chance this ground transect data set is in the neighborhood of a portion of the LaSelva airborne laser data.

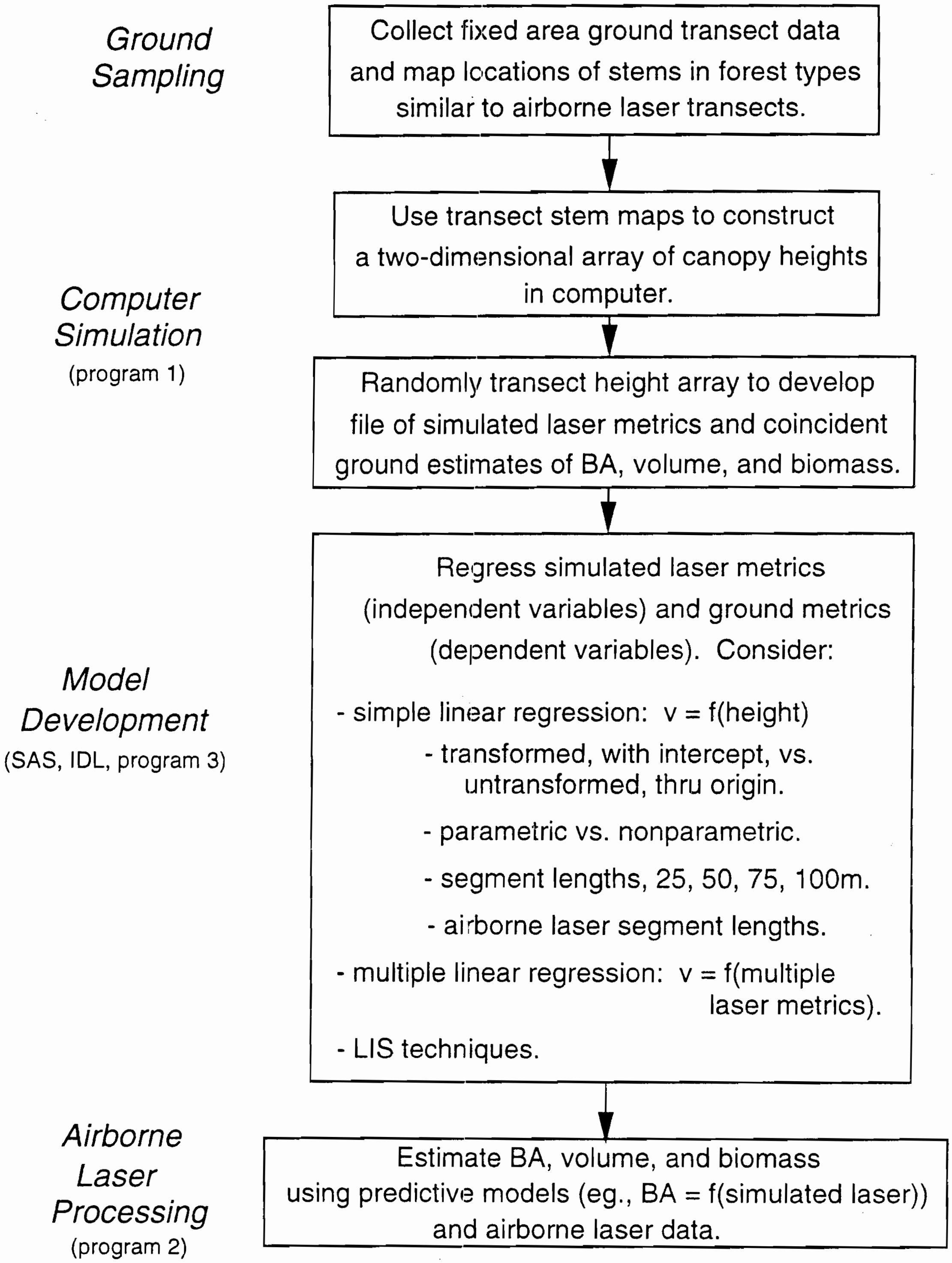


Figure 5. A flowchart of the general procedure used to develop airborne laser-based estimates of basal area, volume, and biomass.

III.A.1. Airborne Laser Data Collection

Two study areas in Costa Rica were flown by the NASA P-3a Airborne Oceanographic LIDAR configured with a Nd:YAG laser pulsing at 400 hz. The nominal aircraft ground speed was 100 m/s, resulting in a nominal along-track sampling frequency of 0.25 m (Sader, 1987). Nominal flying height was 400-500 m above terrain; laser beam divergence was set to 5 milliradians (mr), providing a spot size at the target of 2.0 - 2.5 m. The field of view (FOV) of the detector was a 5 x 5 mr square, providing an actual FOV at the ground of 2.0 x 2.0 m up to 2.5 x 2.5 m. The FOV was boresighted with the transmitted laser pulse. A waveform digitizer was used to record secondary laser pulse returns. These waveform data are processed post-flight to locate ground beneath the forest canopy. The waveform digitizer is described by Hoge et al. (1980), Nelson et al. (1984), and Krabill et al. (1984).

The P-3a also had onboard two photographic systems to record aircraft progress across terrain. A 35 mm half-frame, true color camera acquired imagery coincident with the laser transects. A large format (23 x 23 cm, 9 x 9 in) T-11 near-infrared research camera was used to obtain synoptic coverage of the study areas. Foresters used these photographs to locate, as best they could, selected laser transects on the ground for purposes of ground reference data collection.

The two study sites flown in Costa Rica in October, 1984 were located 1) near the town of Puerto Viejo de Sarapiqui at the LaSelva Biological Station on the east side of the divide, and 2) near the town of Tileran on the west side of the continental divide (see Figure 6). On the morning of October 19, 1984, laser profiling data were acquired over the La Selva Biological Station under scattered clouds. LaSelva is on the Carribbean side of the continental divide, and it has no distinct, predictable dry season (David Clark, 1993, personal communication). La Selva hosts tropical wet and tropical premontaine wet forest (Holdridge et al., 1971); the majority of the 700 ha biological reserve is in primary forest. (Note: Since this mission was flown, La Selva has acquired a substantial additional holding. The Biological Station now comprises 1333 ha, with 961 ha in primary forest - recent information note entitled Estacion Biologica La Selva, no date, no author.) Laser data were acquired along two flightlines oriented approx. NW-SE. Each flightline incorporated 2-3 km of laser data within the Station borders (see Figure 7). Figure 8 depicts selected portions of the airborne laser data acquired along transects on the LaSelva Biological Research Station. Figure 9 depicts height frequency diagrams for each of the two flightlines.

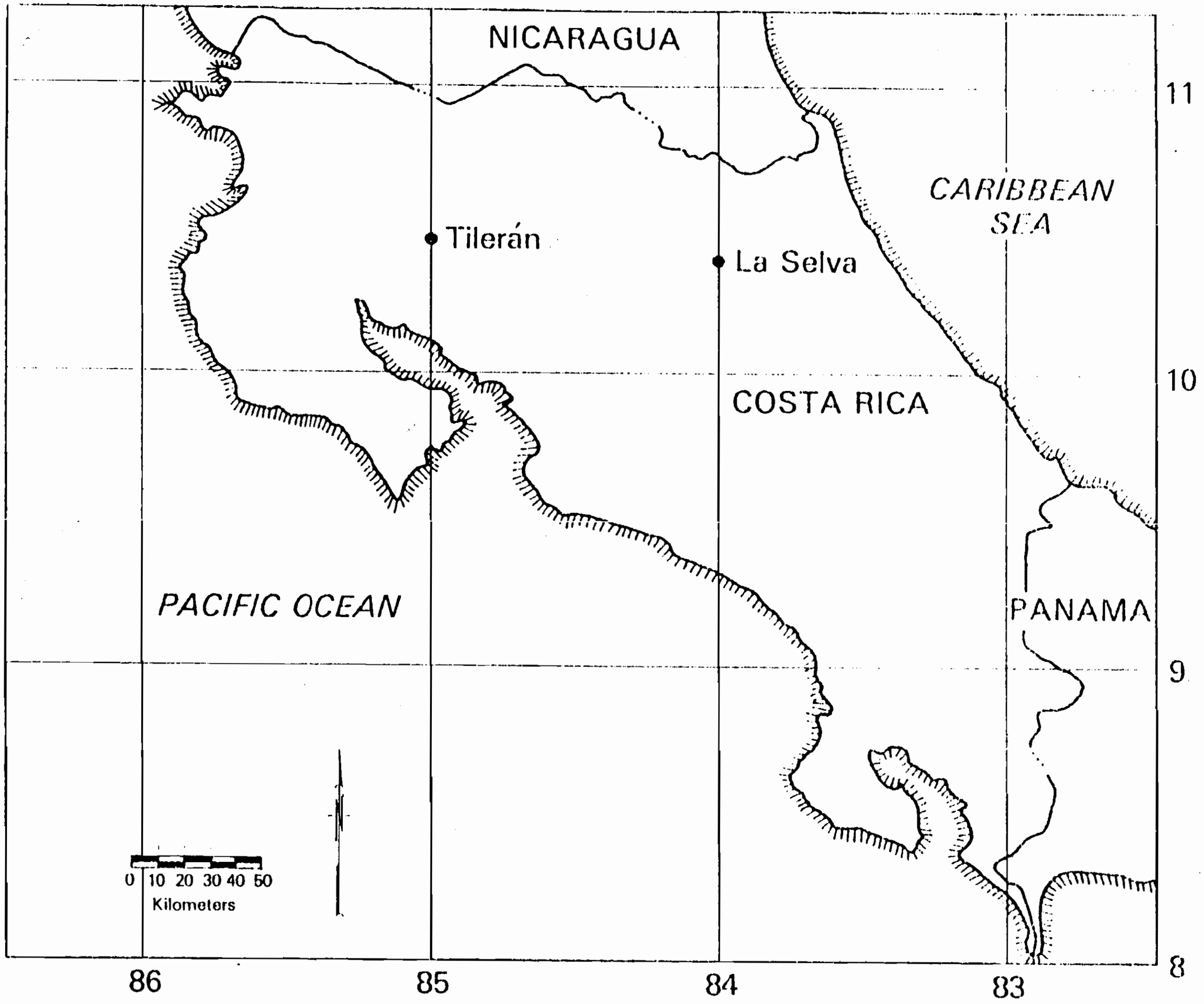


Figure 6. Location of the study areas in Costa Rica.

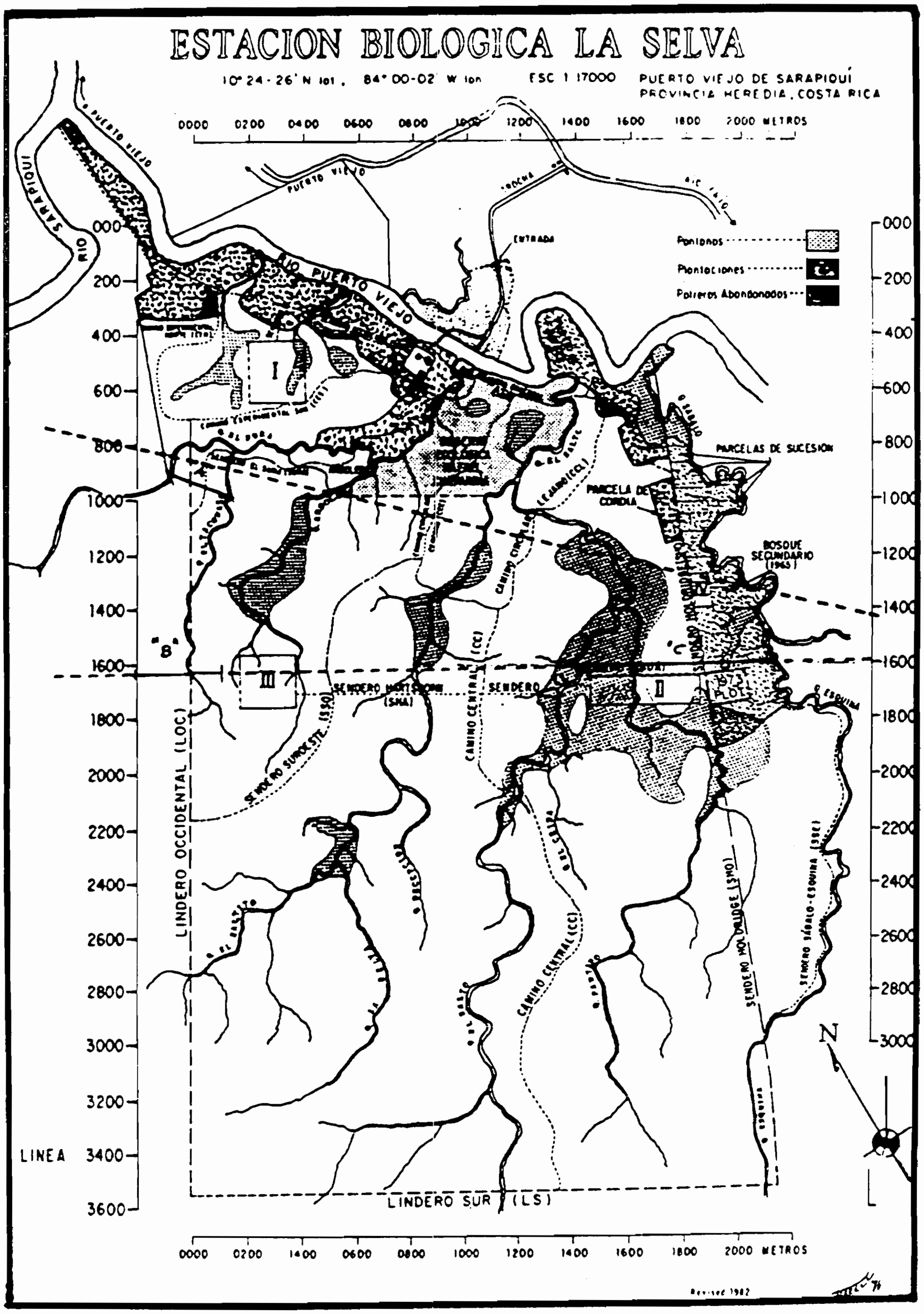


Figure 7. Map showing the two flightlines flown over the LaSelva Biological Research Station near Puerto Viejo, Costa Rica.

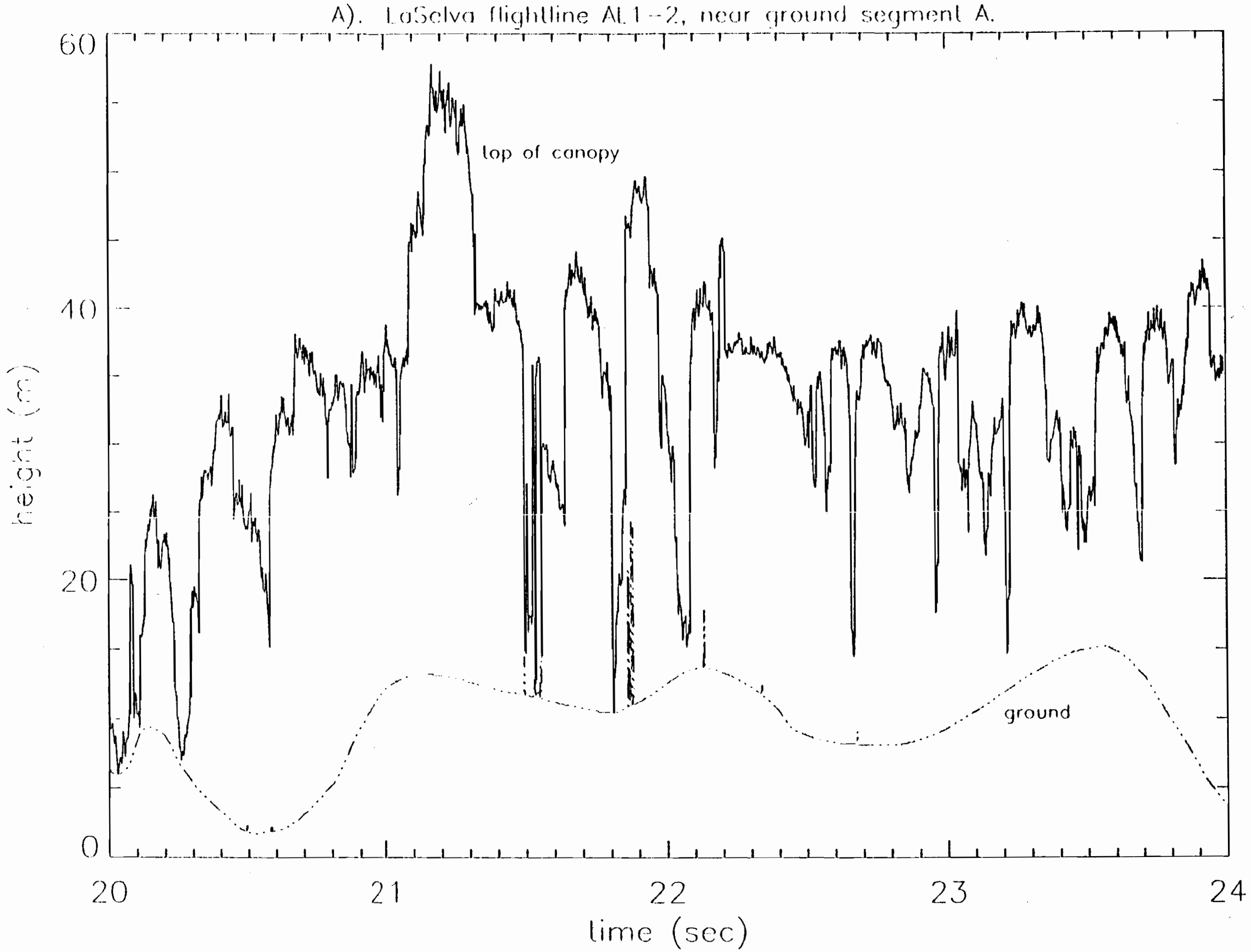
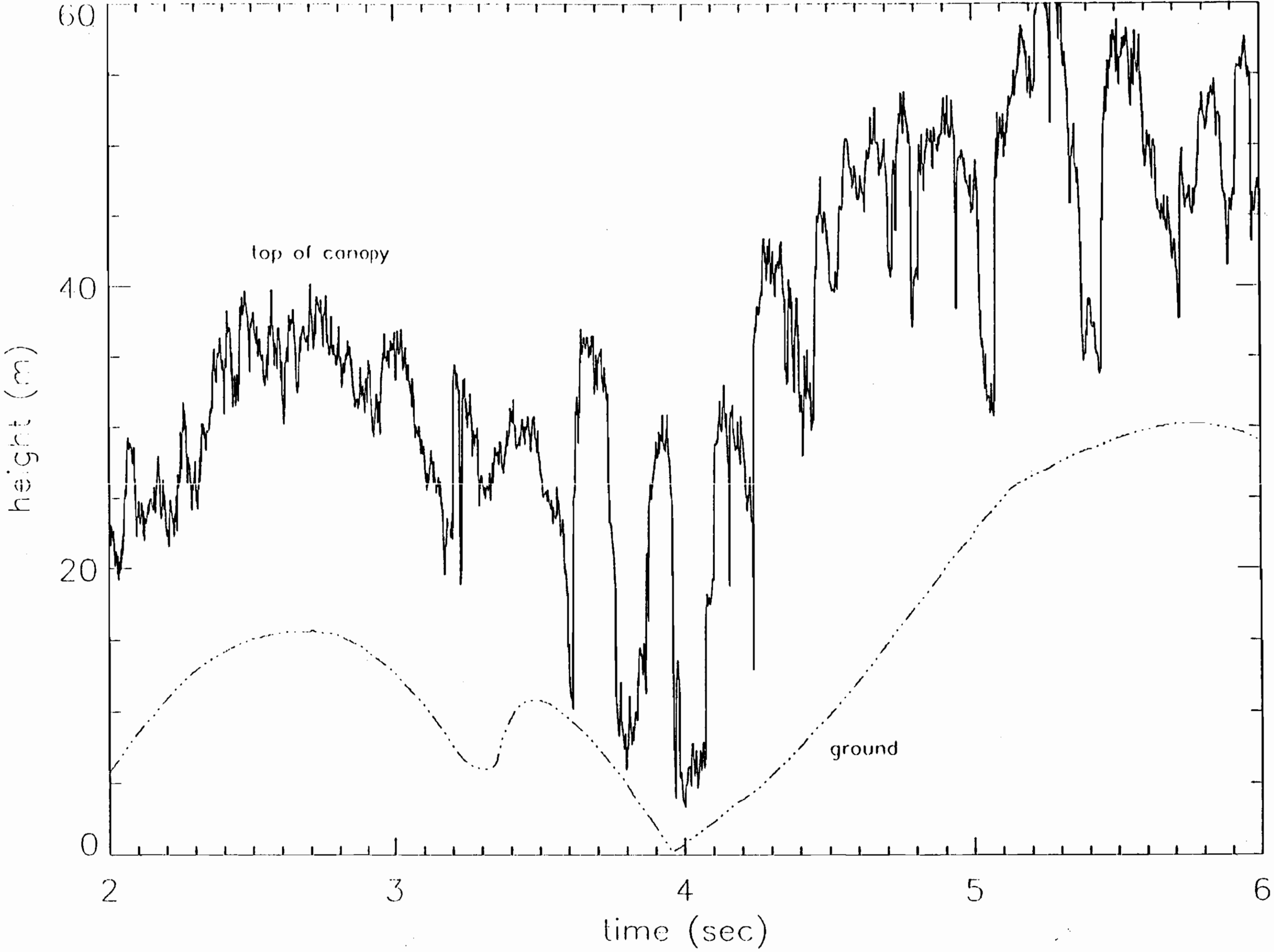
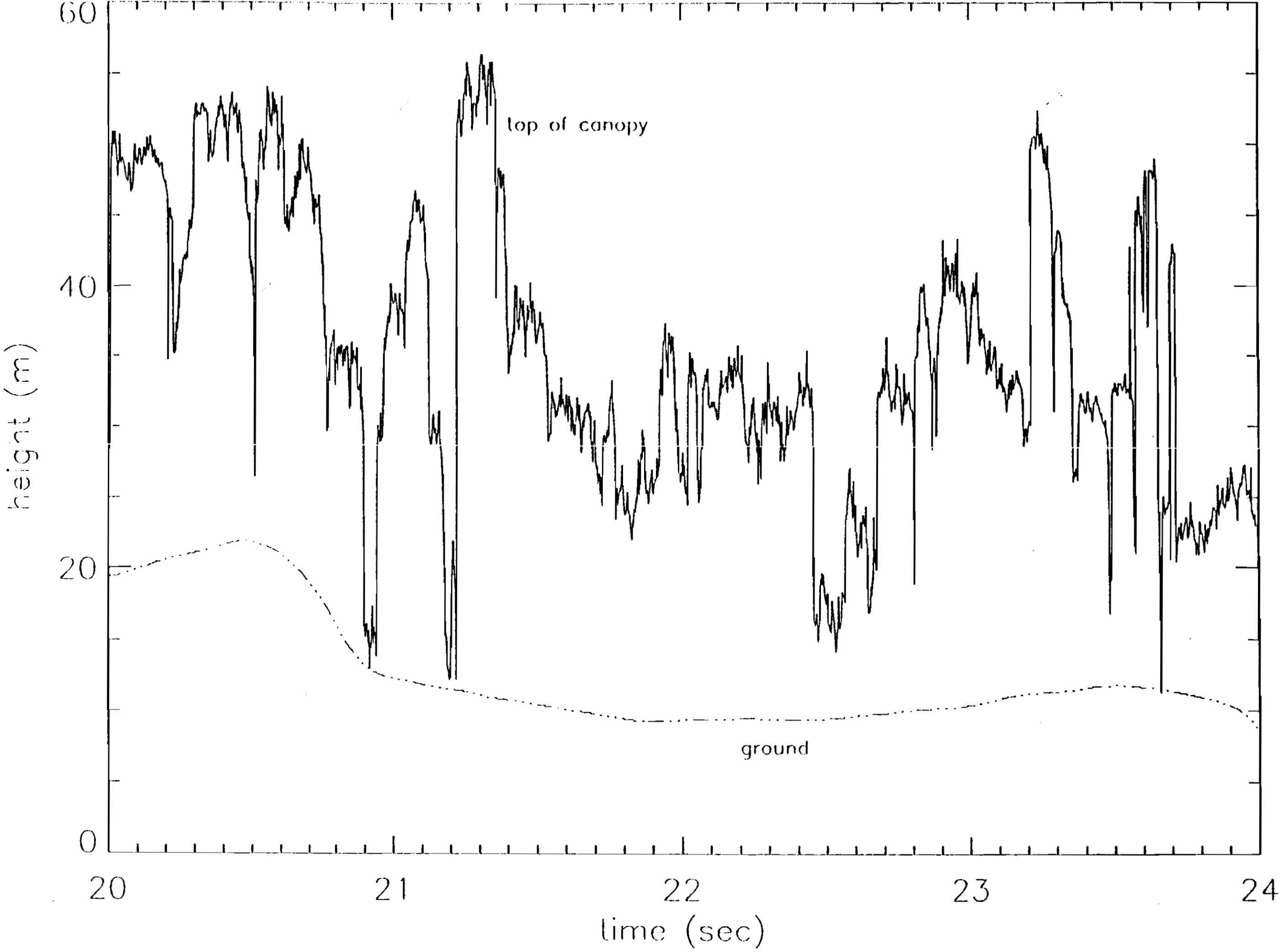


Figure 8. Airborne laser data acquired at LaSelva A) along the NW portion of flightline AL1-2, the northernmost of the two lines, near Seg. A, B) along the NW end of AL1-1, near the 20m transects and Seg. B, and C) along the SE end of AL1-1, near Seg. C.



B). LaSelva flightline AL1-1, near ground segment B and the 20m ground transects.

C). LaSelva flightline AL1-1, near ground segment C.



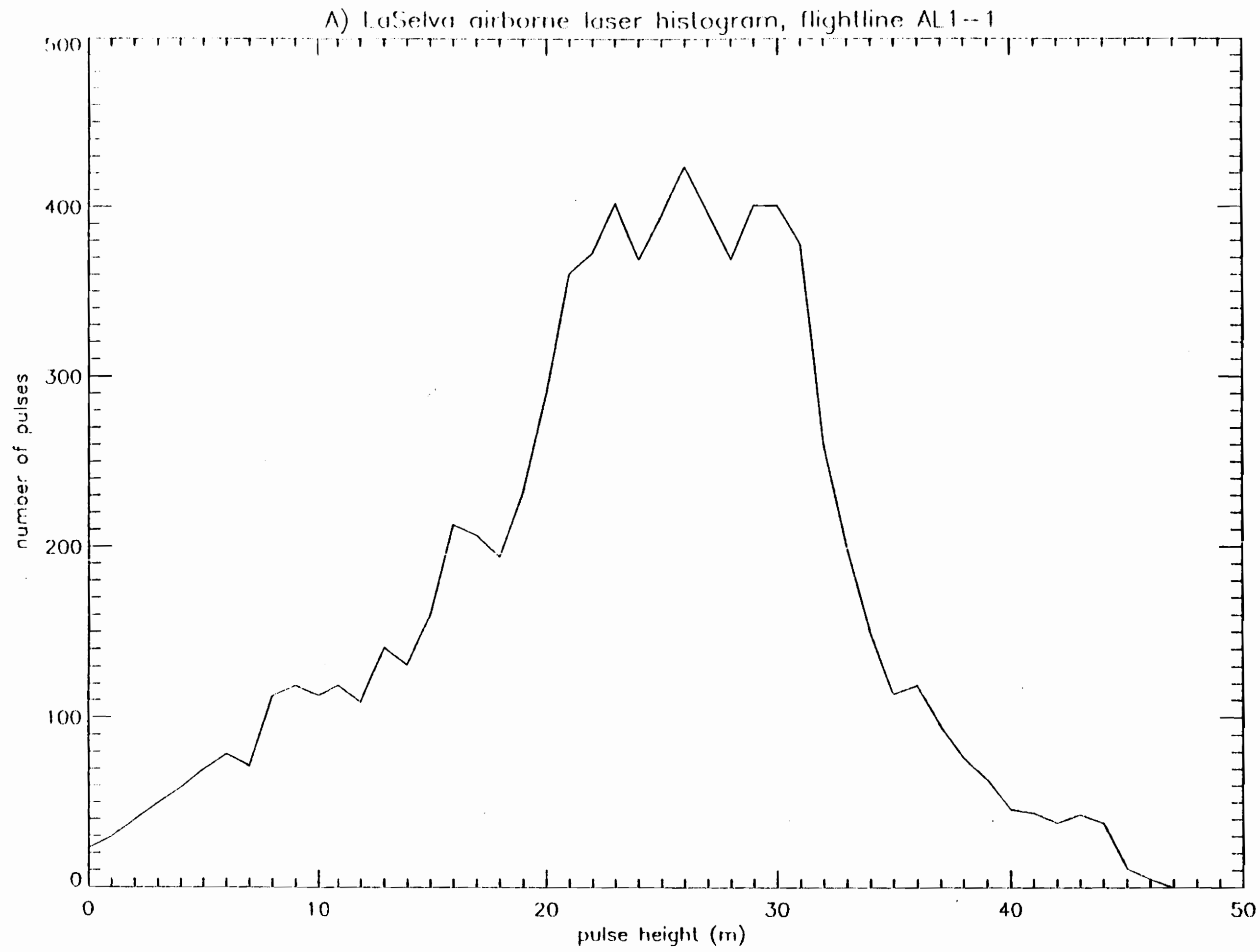
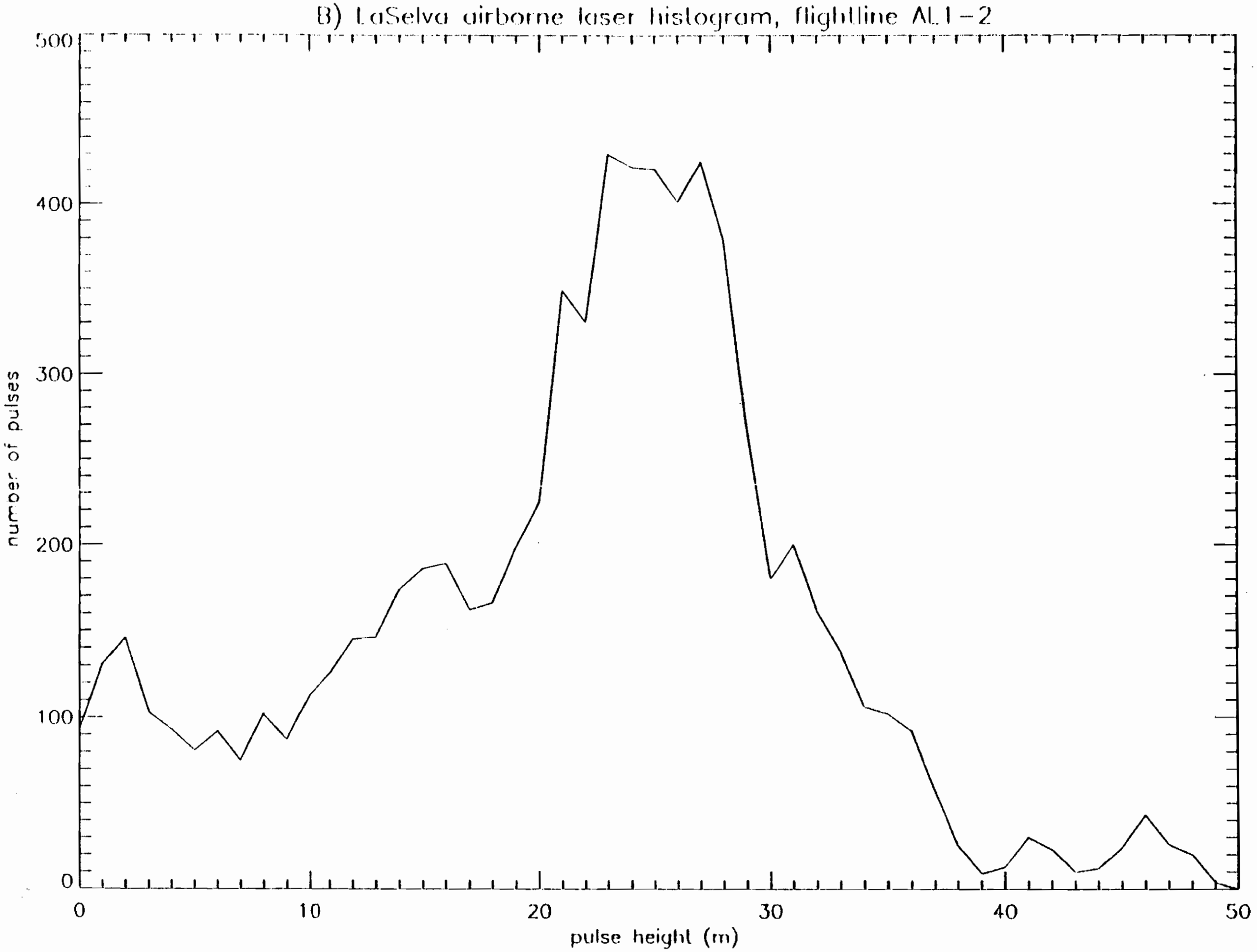

A) LaSelva airborne laser histogram, flightline AL1--1

Figure 9. Height frequency distributions of the airborne laser data for the two flightlines flown over the LaSelva Biological Research Station. A) LaSelva airborne laser histogram, fl. AL1-1 (CRT2F1, 3.609-24.000 seconds). B) LaSelva airborne laser histogram, fl. AL1-2 (5.592-24.829 seconds).



B) LaSelva airborne laser histogram, flightline AL1-2

The Tileran site was flown on the afternoon of October 19, 1984 under clear skies. The forests in the Tileran area are categorized as tropical moist forest (Holdridge et al., 1971). The area around Tileran consists of small, relatively open, patchy forest stands which have been disturbed by selective logging, however, there are remnant primary forest trees. The site is on the Pacific side of the continental divide, as such it has a distinct dry season (Armond Joyce, NASA/NSTL, 1989, personal communication). Six roughly parallel flightlines, each less than one kilometer in length and spaced approximately 200 m apart perpendicular to the slope, were acquired (see Figure 10). The ground transect was located relative to flightline 1/3; only that flightline is considered in the course of this study. Figure 11 illustrates the height characteristics of the Tileran forest along flightline 1/3, as measured by the airborne laser. Figure 12 reports the height frequency distributions of the Tileran airborne laser flightline (1/3).

Once the airborne LIDAR measurements were collected, the data were processed to identify the ground trace beneath the forest canopy. This ground trace was located based on laser pulses which travel directly to ground and also on pulses which exhibit significant secondary returns from the ground. The ground trace was digitized, thereby permitting the calculation of a canopy height for each laser pulse.

III.A.2 Ground Data Collection

The ground data collection varied between sites, but in both instances the foresters conducting the ground survey used the 35 mm and T-11 photography in an attempt to locate ground transects coincident with the actual laser trace on the ground. Lack of ground control, even in these relatively accessible forests, made this coregistration job difficult. In Tileran, a starting point for transect 1/3 was located in the photography and on the ground and a 500m compass line was established. No point other than the start point was used for line location. All trees whose boles were located within 5 meters of this line were tallied (i.e., a 10m wide transect) in May 1985. Tree species at Tileran were not noted, evidently because there are too many individual species and the occurance of any particular species is too infrequent. At LaSelva, five distinct ground transects were located in the general vicinity of the laser flightlines. Two - one kilometer ground transects, two - 300m transects, and one 400m transect, each 5m wide, were collected in January and February, 1985. All trees greater than 10 cm dbh were tallied if their bole centroids fell within 2.5 m of the ground transect line. Tree species were noted at La Selva because, atypically, five species make up

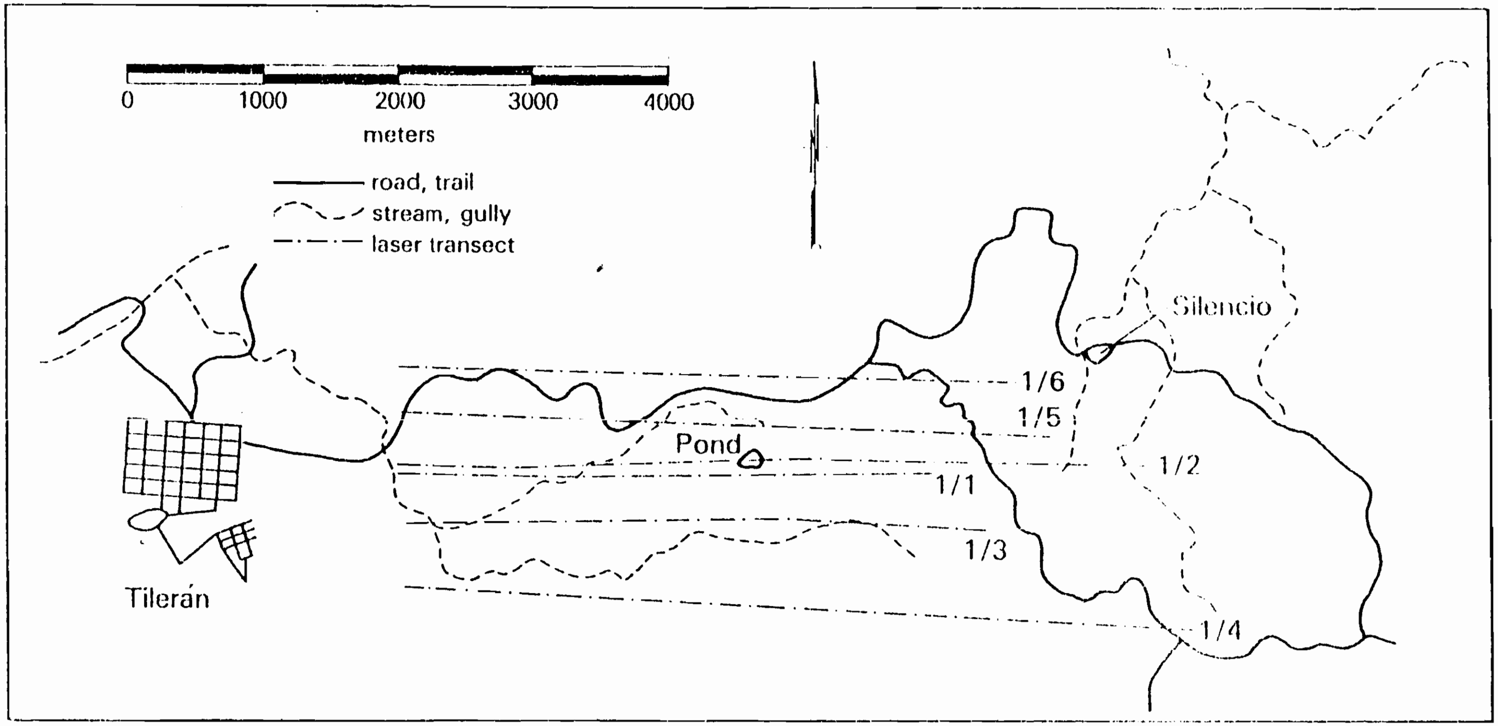


Figure 10. Map showing the six flightlines flown east of the town of Tileran. One of six lines, 1/3, is considered in this study.

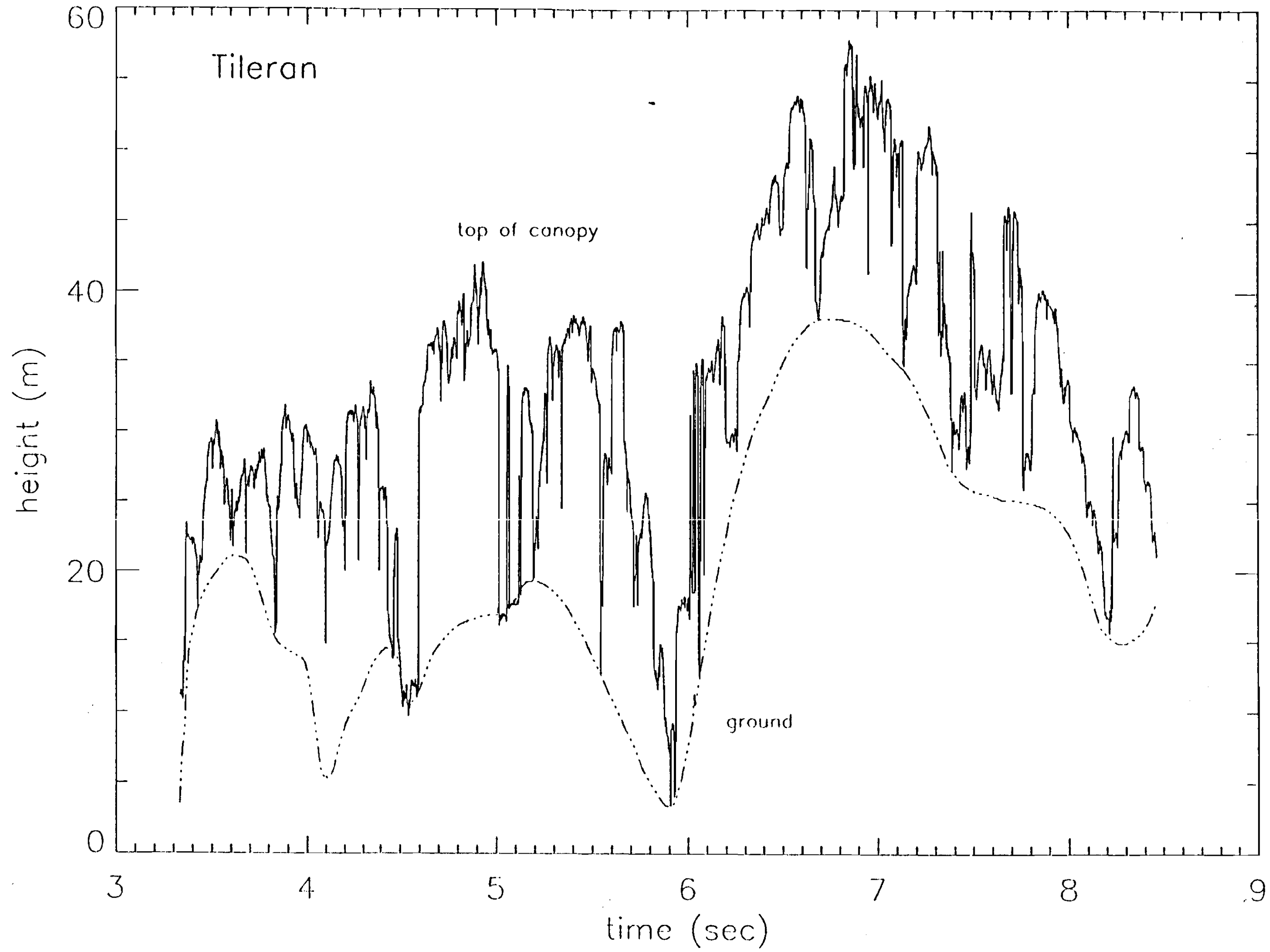


Figure 11. Airborne laser data acquired near Tileran, flightline 1/3.

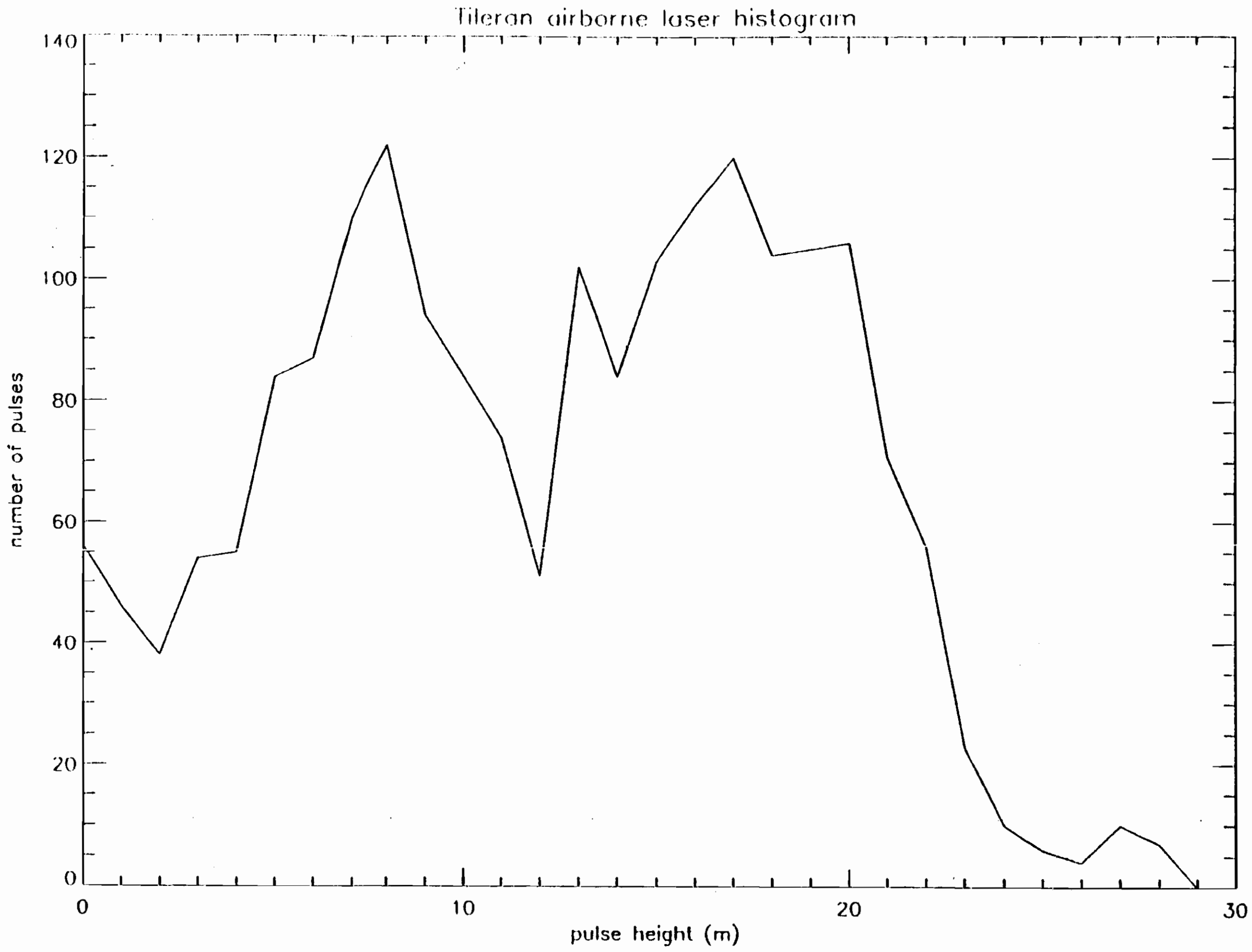


Figure 12. Height frequency distribution of the airborne laser data for flightline 1/3 flown near Tileran.

the bulk of the standing woody biomass; one species, Pentaclethra macroloba, comprises 20-30% of the total forest volume. The other four species or species groups in the majority at La Selva are Carapa guianensis , Virola spp., Vochysia ferruginea , and Tetragastris foramensis. The ground data collected on each tree sampled on the Tileran and LaSelva ground transects are reported in Table 1.

Table 1. Ground data collected on each tree within a transect at Tileran and LaSelva.

variable	measurement accuracy(m)	comment
species		LaSelva only
dbh	0.01	occasional measurements to 0.001m for trees<0.15m dbh-LaSelva only
total height	1	
height to first branch	1	
crown diameter	1	occasional measurements to 0.5m for trees with small crowns - LaSelva only
distance along transect	0.1	
distance perpendicular to transect	0.05	

Tree volume or tree biomass estimates were calculated for each tree tallied along the transects using the information collected in the ground survey. Merchantable volume was calculated for each of the 265 trees sampled near Tileran. A non species-specific volume equation published by Lojan (1966) was used.

$$\log(v) = 2.03986 \log(d_b) + 0.779 \log(h_b) - 4.07682$$

where v = merchantable volume, in cubic meters,
d_b = dbh in centimeters,
and h_b = tree height to first major branch, in meters.

The dbh, total height, and bole volume histograms are presented in Figure 13 for the 265 trees sampled along the 500m Tileran ground transect.

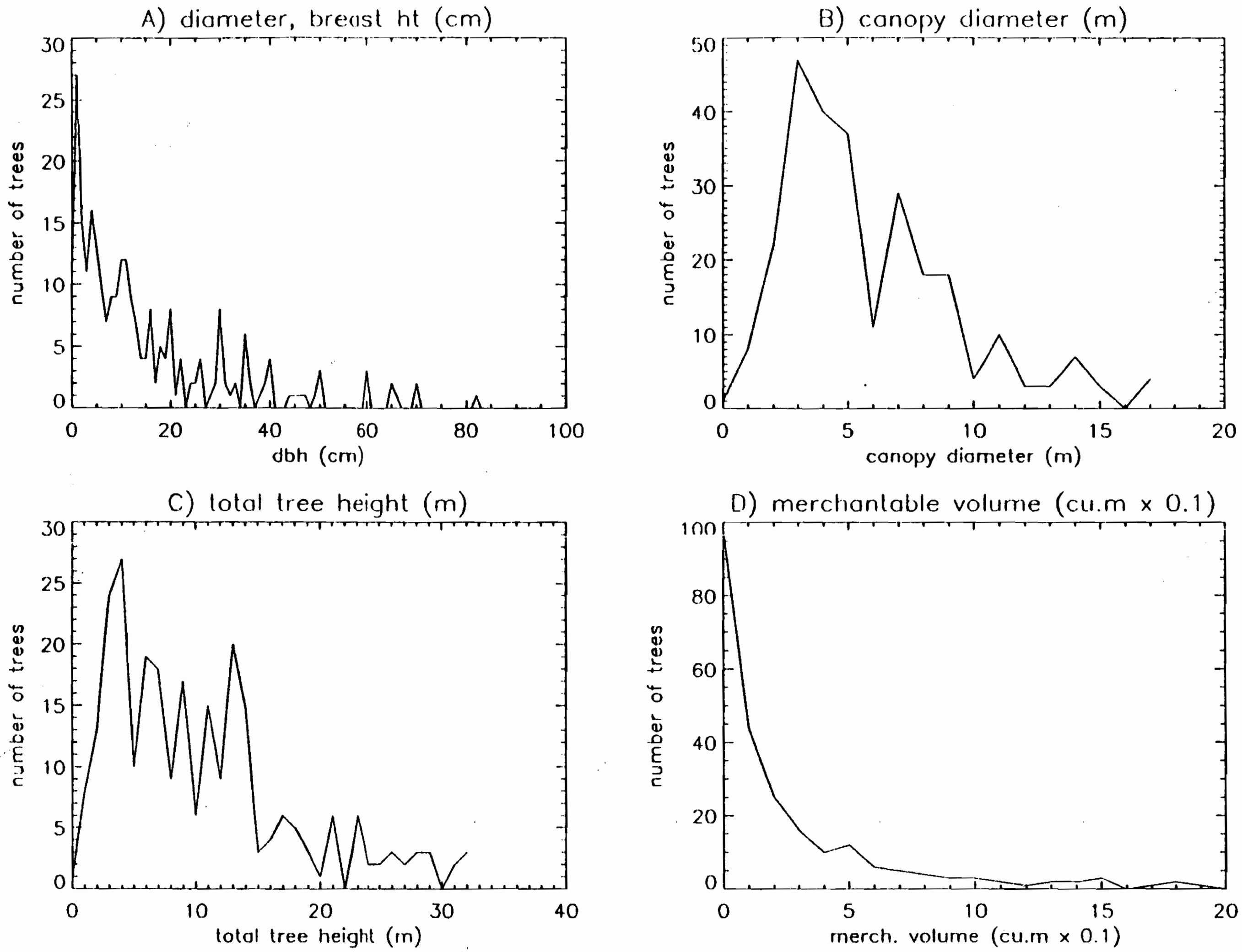


Figure 13. Frequency distributions of A) tree diameter breast height; B) canopy diameter; C) total tree height; and D) bole volume for the 265 trees found along the 10m x 500m Tileran ground transect.

Approximately 8 km SE of the LaSelva Biological Station, in tropical wet forest, 96 trees were destructively sampled in order to derive dry weight measures for the total above-ground portions of the tree. The dry weights of the foliage, branches, and commercial stem were measured on each tree. These weight measures were logarithmically related to tree diameter and total height for the five tree species listed above. All other species were grouped together to produce a sixth set of generic biomass equations. The general form of the biomass equations follow:

$$\begin{bmatrix} \text{total tree dry weight (kg)} \\ \text{commercial stem dry weight} \\ \text{branch dry weight, or} \\ \text{foliage dry weight} \end{bmatrix} = a_0 + a_1 \ln(d_b^2 \, h_t)$$

where a_0 and a_1 are regression coefficients, and h_t is total tree height.[1]

Most R-square values for these relationships exceeded 0.90; all exceeded 0.76. These biomass equations were used to calculate ground estimates of the components of total tree dry weight for all trees sampled along the ground. Figure 14 depicts the dbh, total height, and total above-ground dry weight biomass frequency distributions for all of the trees sampled on all five transects.

A second LaSelva ground data set was acquired from Milton and Diane Lieberman, researchers at the University of North Dakota. They've mapped stems on two transects, 20m x 150m and 20m x 200m, located in the general vicinity of Plot III on the west side of the primary forest reserve. These plots are, by chance, located in the general vicinity of the northwestern end of the southernmost laser flightline acquired over LaSelva. The equations noted above were used to calculate total above-ground dry biomass for each of the trees mapped on these two subtransects. Though the amount of area incorporated in these two 20m wide transects is approximately one-ninth of the area encompassed by the 5m wide data, the wider transects permit at least a cursory look at the effects of transect sampling width on canopy simulation and estimation of basal area and biomass. Figure 15 illustrates the dbh, total tree height, and total above-ground dry biomass for all of the trees sampled along the two 20m wide transects.

---

[1] Source: Personal communication from Edgar Ortiz Malavasi, Director, Departamento de Ingenieria Forestal, Cartago, Costa Rica, to Dr. Armond Joyce, Principal Investigator, NASA/NSTL, dated March 24, 1987.

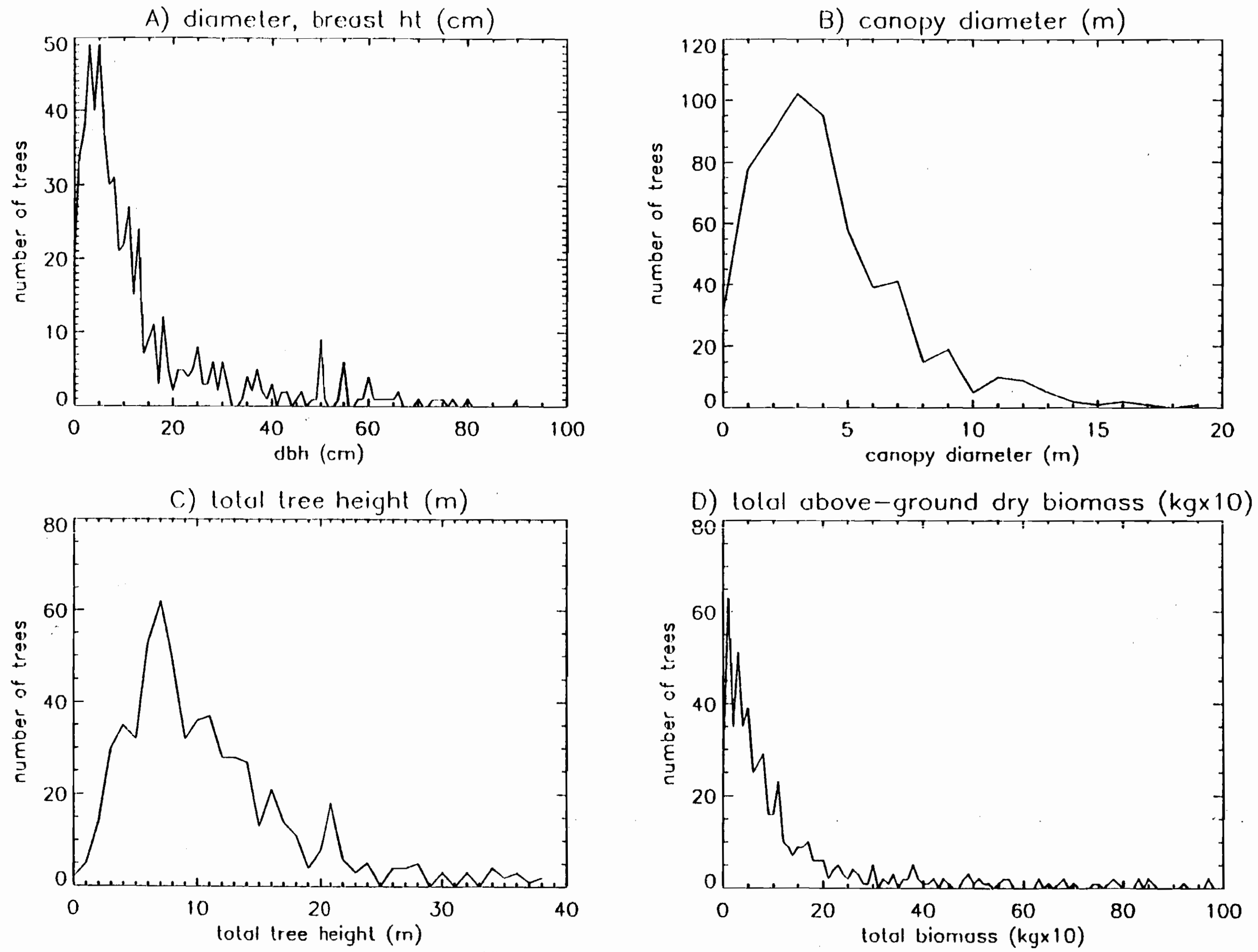


Figure 14. Frequency distributions of A) tree diameter breast height; B) canopy diameter; C) total tree height; and D) total above-ground dry biomass for the 604 trees found along the 3 km of 5m wide LaSelva transects (five transects totalling 3000m).

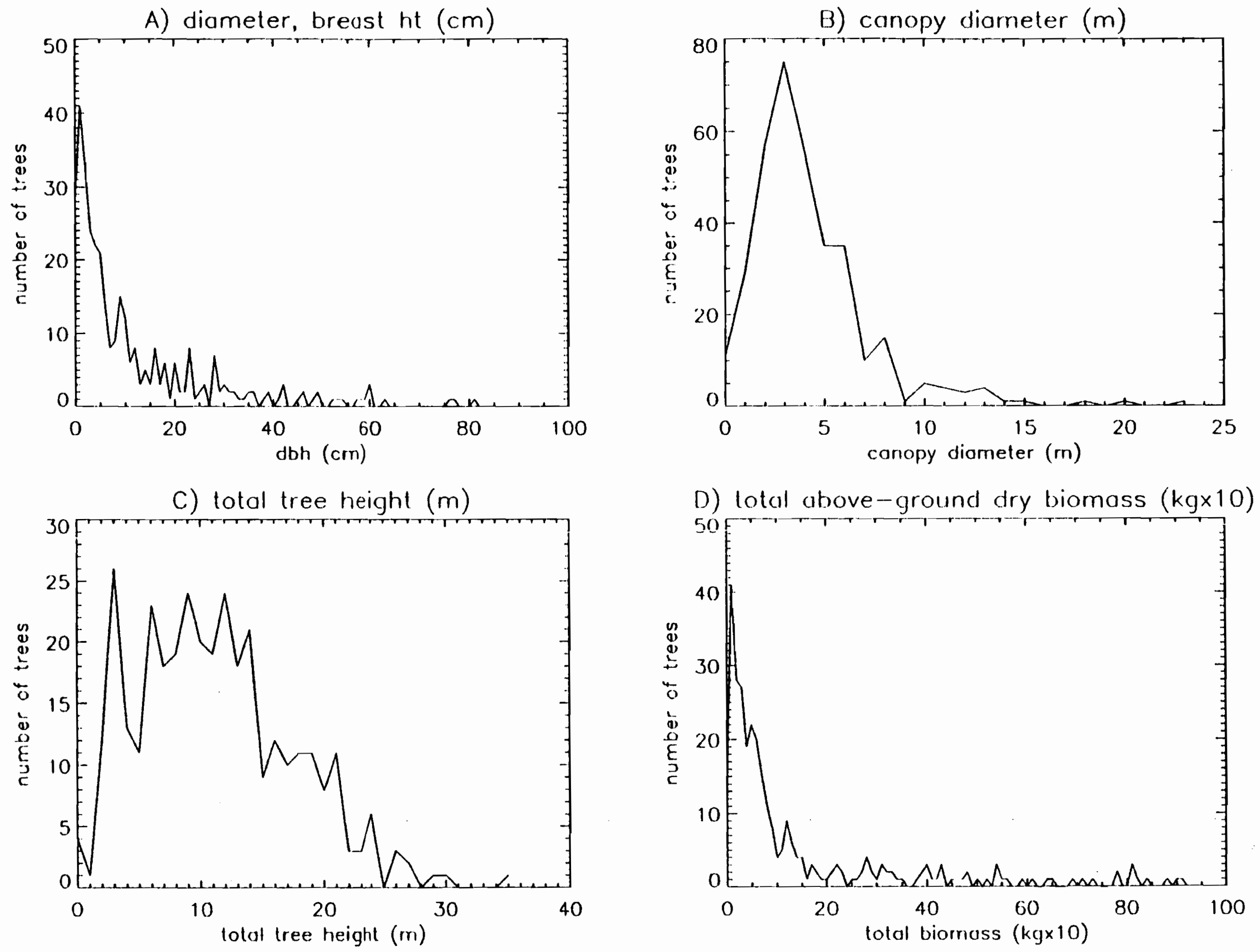


Figure 15. Frequency distributions of A) tree diameter breast height; B) canopy diameter; C) total tree height; and D) total above-ground dry biomass for the 345 trees found along the 350m of 20m wide LaSelva transects (two transects totalling 350m).

In all, then, three different ground reference data sets are considered, one near Tileran and two on the LaSelva Biological Research Station. Though these datasets are not spatially coincident with the airborne laser data, sections of the airborne laser flightlines in the general proximity of the ground transects may be identified. Table 2 lists the ground reference data set and its associated (though not coincident) airborne laser data. The airborne laser segment is reported in seconds, where the aircraft is traveling approximately 100 m/s. Airborne laser lengths may differ appreciably from the ground segment lengths. Exact laser endpoints were difficult to determine due to lack of local ground control; best guesses were made based on topographic and forest boundary clues in the airborne laser data. Given minimal control, these endpoints should be considered inexact.

Table 2. A listing of ground reference data and the associated airborne laser data. Though not spatially coincident, the ground reference data and the airborne laser data were most likely acquired in/over similar forest types and, on the average, may be comparable.

Ground Ref. Data Set	Dimensions	Airborne Laser Data File	Flightline	Airborne Laser		
				Laser Time(sec)		approx. Length
				Begin	End	
Tileran	10x500m	CRT1F2	1/3	3.333	8.465	513m
LaSelva-5m	5x3000m (5 transects)	CRT2C[2]	1/6,1R/6	3.609	34.243	3963m
Seg.A	5x300m	CRT2F2	1R/6	21.2	24.2	300m
Seg.B	5x300m	CRT2F1	1/6	0.85	3.85	300m
Seg.C	5x400m	CRT2F1	1/6	20.0	24.0	400m
AL1-1	5x1000m	CRT2F1	1/6	7.3	16.8	950m
AL1-2	5x1000m	CRT2F2	1R/6	10.8	19.2	840m
LaSelva-20m	20x350m (2 transects)	CRT2F1	1/6	0.860	9.300	844m
Tr1	20x200m	location wrt airborne laser not known				
Tr2	20x150m	location wrt airborne laser not known				

The ground transect information, along with calculated estimates of volume (Tileran) and biomass (LaSelva) served as input for canopy height simulation. The mapped stand information is used to generate a two-dimensional digital array of canopy height values which may be used subsequently to simulate airborne laser profiling transects. Metrics measured along the simulated laser transects are paired with

---

[2] CRT2C is a concatenation of CRT2F1 and CRT2F2. CRT2C permits the two flightlines to be processed as one by the airborne laser processing program.

ground measurements used to produce that transect. A simulated laser-ground data set is compiled for use in subsequent statistical analyses. Listings of the ground transect data are available from this author.

III.B. Computational Procedures - Forest Simulator

This research emphasizes the development and testing of mensurational procedures which utilize airborne LIDAR data to estimate tree volume and biomass without the reliance on laser-ground colocation. The colocation problem is addressed by simulating the height characteristics of the forest canopy in the computer using mapped stand data. Once the two dimensional array of canopy heights is constructed, the array is randomly transected and the simulated laser metrics (eg., canopy density, average canopy height, average transect height) and the associated ground reference information (eg., seen, unseen, and total basal area and volume or biomass) are tallied and reported.[3] These laser and ground metrics may be subsequently processed using various regression procedures. The simulated laser measurements serve as the independent variables in those statistical procedures and various ground metrics serve as the dependent variable. Once the regression relationships (ground=f(laser)) are established, the equations may be applied to the actual airborne laser data in order to estimate basal area and volume or biomass per hectare.

The ground transect data need not be collected coincidently with the airborne laser data. Ground data may be collected in areas removed in time and space from those areas covered by the airborne laser. An underlying assumption of these simulation analyses is, however, that the forest cover type characteristics of the ground transects are similar to those forests transected by the airborne laser. The forests sampled on the ground and by the airborne laser should be similar in their structural characteristics - overstory/understory relationships, stem densities, approx. age, height, phenology, species composition. Also, the range of forest conditions sensed by the laser should be reflected in the ground sample data so that these conditions can be simulated in the computer.

The forest canopy simulator consists of three major procedural subsections: 1) the canopy height simulator; 2) the laser transect simulator; and 3) a ground reference evaluation section. The forest canopy

---

[3] Seen and unseen refer to a tree's canopy position with respect to an airborne laser. If any portion of a particular tree's canopy can be potentially sensed by an airborne laser, in other words, if any portion of a tree's canopy is visible when the forest is viewed vertically from above, then that tree is "seen". The terms roughly correspond to overstory and understory.

simulator makes use of mapped transect data collected on the ground in forest similar to that flown by the airborne laser. Output consists of 1) the two-dimensional transect height array; 2) a similarly dimensioned tree identification array; 3) a data file consisting of simulated laser metrics and associated ground reference measurements (for subsequent use in statistical analyses); and 4) a formatted report of the simulated laser and ground metrics. The particular capabilities of each section of the forest canopy simulator are described below. The computer code, written in Fortran-77, is available from this author.

III.B.1. The Canopy Height Simulator

The purpose of the canopy height subsection of the overall simulation program is to produce a two-dimensional representation of forest canopy heights given mapped tree data acquired along a ground transect. In the ground sampling phase, trees are tallied along a long narrow swath; positional information is as important as biometric data for each tree since the height characteristics of the sample swath(s) will be reconstructed in the computer. The ground measurements, both tree size and tree location, are used to reconstruct the canopy swath in the computer. The program reads tree attribute and location data to develop a two-dimensional canopy height array. The two-dimensional array represents the length and width of the ground transect; each element (pixel) in the height array represents that area on the ground associated with the laser spatial resolution. For instance, if, as in this study, the airborne laser was pulsing at 400 hz and the aircraft ground speed was 100 m/sec, then the canopy height array could be dimensioned to reflect the 0.25m along-track horizontal sampling resolution of the laser. If tree measurements were acquired along a ground transect 300m x 5m in a covertype similar to one sensed by the airborne laser, then the resultant canopy height array would be dimensioned to 1200 samples (300m at 0.25m/sample) by 20 rows (5m at 0.25m/sample). As currently configured, the analyst can control those items listed in Table 3 under the canopy height simulator.

In order to run the simulation program, a ground data file which contains canopy shape information for each tree tallied must be available or a canopy shape must be assumed. The program can currently handle four canopy shapes - conic, parabolic, elliptic, or spheric. The following ground data are used to calculate canopy heights for a particular tree: 1) total height; 2) height to first branch; and 3) crown diameter. If the canopy shape is known or assumed, the equations in Table 4 may be used to generate a

Table 3. Items which may be set by the analyst prior to running the simulation program. The overall simulation program is divided into three distinct sections: 1) the canopy height simulator; 2) the laser transect simulator; and 3) the ground reference evaluation section.

1. Canopy Height Simulator:

- length and width of the ground transect.
- laser spatial resolution (in same units as length and width).
- canopy shape:
  - -conic
  - -parabolic
  - -elliptic
  - -spheric
  - -random assignment of 4 shapes.
- adjacency - trees "grow" into transect based on assumption that transect data is representative of trees growing on adjacent tracts.
- standard deviation of within-canopy heights (a canopy roughness factor).
- thinning - analyst can randomly remove any number of trees.

2. Laser Transect Simulator:

- number of randomly-oriented simulated laser lines which transect the long axis of the height array.
- minimum tree height: height below which canopy is defined as shrub, i.e., a non-tree canopy hit.
- segment length and gap length: each random, simulated laser transect may be subsectioned into equidistant segments with gaps between segments.

3. Ground Reference Evaluation: no analyst controls.

Table 4. Equations used to calculate tree canopy heights at different (x,y) locations in an individual tree crown based on total tree height, height to first branch, and crown diameter (see also Nelson and Yaya, 1990, for these equations). All variable units of length, including x and y, are identical.

variables:	(x,y):	The grid location of the point under consideration. Distance along- and across-transect from the ground transect origin.
	z_i:	Height of the crown at a particular x,y location. If solution to equation is negative, z_i is set to zero.
	h_t:	total tree height.
	h_b:	height to first branch.
	r_c:	crown radius, one-half of the crown diameter.

Conic:

$$z_i = h_t - \left(\frac{h_t - h_b}{r_c}\right)\sqrt{(x - r_c)^2 + (y - r_c)^2}$$

Parabolic:

$$z_i = h_t - \left[\left(\frac{h_t - h_b}{r_c}\right)\left((x - r_c)^2 + (y - r_c)^2\right)\right]$$

Elliptic:

$$z_i = \sqrt{(h_t - h_b)^2\left[1 - \frac{(x - r_c)^2}{r_c^2} - \frac{(y - r_c)^2}{r_c^2}\right]} + h_b$$

Spheric:

$$z_i = (h_t - r_c) + \sqrt{r_c^2 - (x - r_c)^2 - (y - r_c)^2}$$

tree's canopy height array. Conical, parabolical, and elliptical canopies are assumed to extend symmetrically down from the top of the tree to the height of the first branch. The spherical tree canopy extends symmetrically down from the top of the tree until the canopy diameter limit is reached; the height to the first branch is not considered. Nelson and Yaya (1990) illustrate (see Figure 16) the canopy construction procedure and report the effects of canopy shape on simulated laser measurements.

Adjacency effects can also be controlled in the canopy height simulator. Adjacency addresses the problem of tree canopies which might hang over the ground transect but whose boles are outside the swath width and hence are not tallied. The problem is addressed by assuming that tree sizes and locations are identical in the similarly sized plots which abutt the primary transect (see Figure 17). This assumption takes a "cookie cutter" approach to the question of the location and size of trees which might surround the ground transect being simulated in the computer. The net effect of this assumption is that trees which grow out of the transect on, for instance, one side of the plot will "grow" back into the plot on the side opposite. Trees whose boles nestle in a corner will "grow" into a transect in the three remaining corners. Transects simulated using adjacency, then, are more dense and, on the average, taller than transects simulated without adjacency. Adjacency is used in all simulations.

A within-canopy variability measure is used to give texture to the canopy surface of each tree. The random variation of heights within individual tree canopies is estimated from airborne laser data by drawing a smooth curve through the tree top and measuring the variation around that smooth curve. The variation associated with the surface roughness of an individual tree or collection of trees is assumed to be normally distributed.

III.B.2. The Laser Transect Simulator

A laser profile across the top of a forest canopy is simulated using the two-dimensional array of canopy heights. Additionally, a given simulated laser transect can be divided into equally-sized pieces, or segments, with gaps optionally left between these segments. The simulated laser metrics which are reported for each segment are listed in Table 5 and are explained below.

An airborne laser, or a simulation of an airborne laser, can measure only a limited number of forest canopy characteristics. Specifically, a pencil-beam laser system can measure canopy density (canopy hits

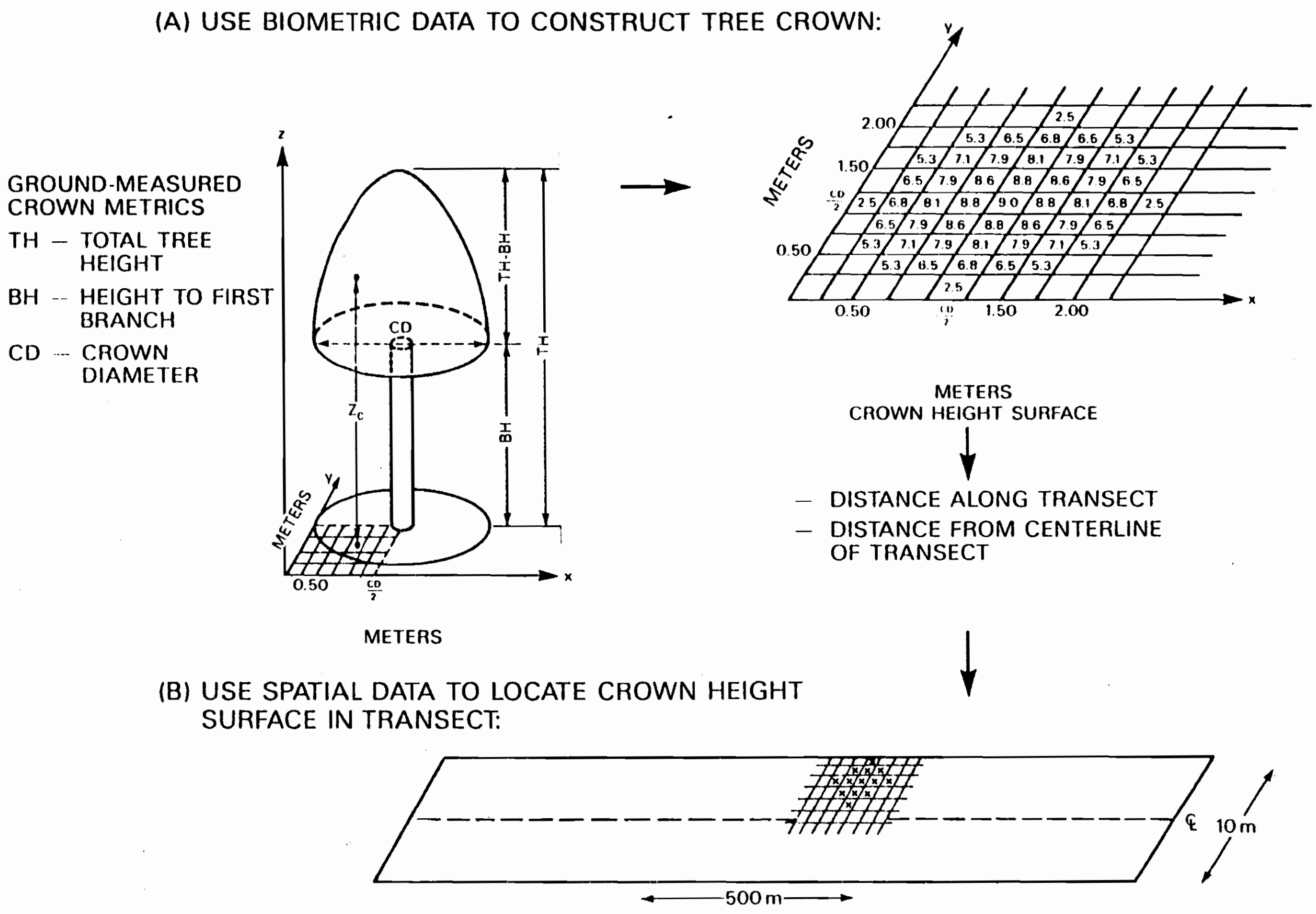


Figure 16. A schematic of the two-step procedure used to construct a canopy height array. A) Individual tree mensuration data are used to simulate a single tree crown. B) The simulated crown height values are inserted into the transect using measurements of distance along and perpendicular to the transect centerline (Nelson and Yaya, 1990).

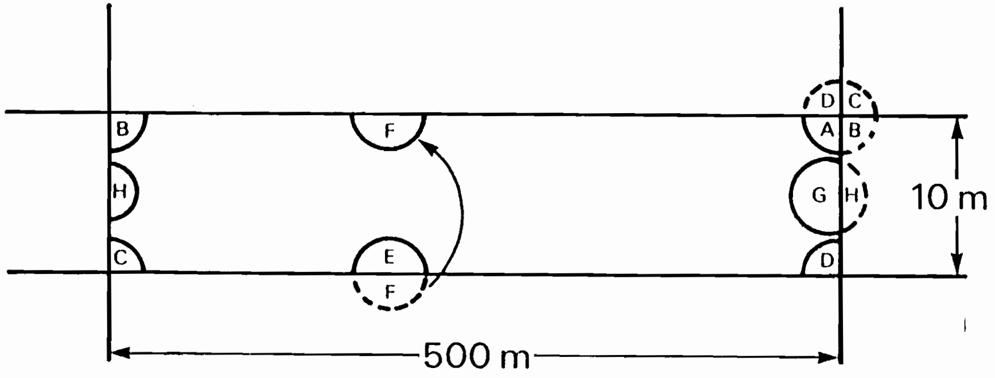


Figure 17. A schematic illustrating the effects of enabling adjacency on trees whose crowns overhang the borders of the fixed-area transect. Tree A will "grow into" the plot at locations B, C, and D. Trees E and G will be represented in the plot at locations F and H respectively.

Table 5 A listing of simulated laser and ground metrics produced by the simulation procedure.

Simulated Laser Metrics: for a particular segment

- canopy closure (%), g
- average canopy height, all pulses (m), $\bar{h}_a$
- average canopy height, canopy hits only[12] (m), $\bar{h}_c$
- coefficient of variation of average canopy height, all pulses, c_a
- coefficient of variation of average canopy height, canopy hits, c_c
- weighted sum of deviations, all pulses, w_a
- weighted sum of deviations, canopy hits, w_c
- standard deviation of canopy heights weighted by average height, all pulses, s_w

Ground Reference Metrics: per hectare estimates for a particular segment

- crown area (m^2/ha)
- canopy volume (m^3/ha)
- number of trees, seen
- number of trees, unseen
- basal area, seen[13] (m^2/ha), b_s
- basal area, unseen (m^2/ha), b_u
- volume or biomass, seen (m^3/ha or kg/ha), v_s
- volume or biomass, unseen (m^3/ha or kg/ha), v_u

Line Intercept Sample Ratios:

- seen basal area per ha/seen crown area per ha, r_b
- unseen basal area per ha/seen basal area per ha, s_b
- total basal area per ha/seen crown area per ha, t_b
- seen volume or biomass per ha/seen canopy volume per ha, r_v
- unseen volume or biomass per ha/seen volume or biomass per ha, s_v
- total volume or biomass per ha/seen canopy volume per ha, t_v

---

[12] g, $\bar{h}_a$, $\bar{h}_c$, and p (canopy profile area) are interrelated. For a given transect or segment:

$$\bar{h}_a = \left(\frac{g}{100}\right)\bar{h}_c = \left(\frac{p}{\text{segment length}}\right)$$

[13] b_t represents total basal area per hectare and is calculated by summing b_s and b_u. Likewise, v_t represents total merchantable volume or total above-ground dry biomass and is calculated by summing v_s and v_u.

versus ground hits, not leaf area index) and canopy heights. Canopy heights in conjunction with distance along the laser transect provides measures of canopy profile area which, if equally sized segments are considered, are directly related to average canopy height as calculated using all pulses. Height variation in this study is characterized by the coefficient of variation and by a weighted sum of deviations from the mean height response. Density, height, and height variation are used to predict ground reference attributes.

Simulated laser data may be easily utilized to count the number of distinct tree crowns intercepted. If individual tree crowns can be automatically identified in airborne laser data, then lasers could be used to estimate stem counts per unit area. Simulation work with respect to stem counts per hectare was not conducted in the course of this study. Significant difficulties are involved with accurately and reliably identifying tree crowns in the actual airborne data using an automated crown-finding algorithm.

Five different measures of laser height variability along a given segment were calculated in an attempt to mathematically characterize stand structure. Mean canopy height provides only a portion of the overall stand picture; the variability associated with a mean height estimate may be used to infer tree size and therefore may be helpful in inferring woody biomass. For instance, a tall, relatively open stand may have the same mean height (or canopy profile area) as a short, densely stocked forest. The tall, open stand, which may well support significantly more woody stem biomass per unit area, will likewise have a larger laser height variance. Variance, in this case, is used as a descriptor for canopy structure. As listed in Table 5, the five variance measures include 1) coefficient of variation of all pulses along a segment; 2) coefficient of variation of pulses intercepting the forest canopy only; 3) a weighted sum of laser height deviations of all pulses; 4) a weighted sum of laser height deviations of pulses intercepting the forest canopy only; and 5) the standard deviation of canopy pulse heights weighted by the average height of all pulses.

The weighted sums of laser height deviations (3 and 4, directly above) take the following form:

$$w_a = \sum_{i=1}^{n_a} \left(\frac{h_i}{\bar{h}_a} \right) (h_i - \bar{h}_a)$$

$$w_c = \sum_{i=1}^{n_c} \left(\frac{h_i}{\bar{h}_c} \right) (h_i - \bar{h}_c)$$

where n_a = total number of pulses in segment;
n_c = number of pulses which hit tree canopy;
h_i = height of pulse i;
$\bar{h}_a$ = average height of all pulses in a segment;
$\bar{h}_c$ = average of canopy hits in a segment.

w_a will be small in those situations where $h_i \approx \bar{h}_a$, i.e., dense, even-aged stands, and become larger as the stand opens up. This metric's tendency to increase as patchiness increases is more pronounced when all pulses along a segment are considered. w_c will be small in those situations where $h_i \approx \bar{h}_c$, regardless of patchiness or canopy density.

The final variance metric considered was developed with three basic generalities in mind (exceptions to these generalities can be easily found): 1) the taller the mean height, the higher the biomass; 2) dense stands have greater biomass than open stands for a given height and height variance; and 3) a single super-emergent may account for the majority of the woody biomass on a given plot, and a large height variance connotes an all-aged canopy structure which in turn connotes high biomass. If these premises hold, then the following laser metric may explain significant variation in ground-measured biomass:

$$s_w = (\bar{h}_c)\ (g)\ (s_c) = (\bar{h}_a)\ (s_c)$$

(see Table 5 for definition of terms). This metric is expected to increase as the canopy grows taller, more dense, and more varied.

All of these metrics were derived using ground transect data and the canopy height simulator to produce a simulated, two-dimensional, forest canopy height array. This canopy height array is randomly transected to simulate airborne laser flightlines. These simulated laser metrics are collected along with the ground measures reported in the following section.

III.B.3. Ground Reference Evaluation

Ground attributes such as basal area, volume, or biomass must be compiled from the ground data which used to generate simulated canopy transects. In addition, ground reference values such as average canopy height and canopy density must be determined for the simulated segments. These ground reference measurements are coupled with the simulated laser measurements in order to develop the laser-ground regressions needed to process airborne laser profiling data. Further, seen/unseen relationships must be identified and maintained, since the simulated laser metrics reflect only the characteristics of those trees that can be sensed by an airborne laser, while ground reference data incorporates both seen and unseen trees. The ground reference data used to produce the simulated canopy height array (used to develop simulated laser metrics) is analyzed in this section of the simulation. The objective is to produce a file of ground measurements and coincident, simulated laser measurements for regression analysis. The ground metrics are listed in Table 5, and are developed by summing the appropriate attributes of all of the ground reference trees measured along a transect or transect segment.

III.C. Establishing Ground Measure-Laser Metric Relationships

Predictive equations relating airborne laser metrics to ground measures of interest are established using simple linear and multiple linear regression techniques. This section describes the variables and equation forms considered in the regression analyses. The overall approach involves a three step process to identify optimal or near-optimal (in terms of prediction accuracy) single variables or sets of variables to predict basal area, volume, or biomass. First, simple linear regression techniques using average canopy height ($\bar{h}_a$) as the independent variable are employed to assess the effects of four factors which may influence prediction accuarcy. Second, using those factor levels which ranked highly in the SLR analyses, multiple linear regression techniques are employed to identify optimal variable combinations. Third, again considering the SLR results, LIS techniques are tested and compared to SLR and MLR results. Each of the three steps is discussed below.

The first of the three steps utilizes a simple linear regression model with average canopy height, $\bar{h}_a$, as the independent variable in order to assess the effects of four factors which may affect prediction accuracy. Previous researchers, Maclean and Krabill (1986), Nelson et al. (1988a,b), have utilized $\bar{h}_a$ to explain significant variability in simple and multiple linear regressions involving ln(volume or biomass) as the dependent variable. The four factors assessed in this first step include 1) parametric versus nonparametric regression techniques; 2) ground transect segment length; 3) airborne laser segment length; and 4) the use of a natural logarithmic transformation (to control variance) versus no data transformation. Those factor levels found to improve (or be least detrimental to) prediction performance as measured by the accuracy of the estimates are utilized in steps two and three.

Multiple linear models are considered in the second step using eight laser metrics as potential independent variables. Variable selection procedures are applied to each of the three ground-laser data sets to identify optimal variable subsets. Art (i.e., subjective judgement) takes over from science in this section of the analysis when variable subsets common to all three ground-laser data sets are identified in order to make the resultant "optimal" variable combinations more generic. Also, it is assumed that the factors investigated using SLR techniques behave similarly in the MLR context.

The third step also employs MLR techniques to see if the LIS ratios needed to estimate basal area, volume, or biomass can be accurately predicted using some subsets of the eight laser metrics. Those factor levels found optimal in the SLR analysis are applied here. The detailed procedures associated with each of these three steps are reported below. All procedures are run with adjacency enabled and assuming that all canopies are elliptically shaped.

III.C.1. Simple Linear Regression with One Independent Laser Variable

Initial efforts to develop models to predict basal area, volume, and biomass center on assessing variants of a basic model commonly employed in previous airborne laser studies:

$$\ln(b) = a_0 + a'_1 \ (p) = a_0 + a_1 \ (\bar{h}_a)$$

where b = basal area, volume, or biomass;
_p = canopy profile area along a 1 km flight transect, m^2;
$\bar{h}_a$ = average canopy height, all pulses, m;
and a_0, a'_1, and a_1 are SLR coefficients such that $a'_1 = 0.001a_1$.

Variants of this simple linear model are assessed to see if performance, in terms of model fit, mean square error, and/or prediction accuracy, can be improved. The effects of the following four factors are considered:

- Factor 1 - regression procedure: 2 levels, parametric versus nonparametric simple linear regression techniques. Preliminary scatterplots of height versus volume or biomass illustrate the presence of potentially highly influential observations. A single large tree (wrt dbh) in a particular segment of a ground transect produce such outliers. A nonparametric SLR procedure is utilized in an attempt to mitigate the impact of any highly influential observations (HIOs) which may be present.

- Factor 2 - ground transect segment length: 4 levels, 25m, 50m, 75m, and 100m. Selection of an appropriate segment length is a trade-off between two competing factors. The shorter the segment length, the larger the sample size per unit field effort. The larger sample sizes facilitate regression studies. The shorter the segment length, however, the higher the variation in the simulated laser and associated ground metrics. Segment length is constrained by canopy diameter. Segments should not be any smaller than the average canopy diameter along a transect; ideally segments should not be any smaller than the maximum canopy diameter along any ground transect. This constraint is imposed to avoid situations where the simulated laser and ground metric attributes of a particular segment are driven by the attributes of a single tree. In Tileran, the maximum canopy diameter is 18m; at LaSelva, the maximum canopy diameter recorded along those transects associated with the laser overflight (the 5m wide transects) is 20m. The maximum canopy diameter recorded along the 20m wide transect near plot III at LaSelva is 25m. The following segment lengths are analyzed 1) to determine the effects of changing segment length on the accuracy of the basal area and biomass estimates; 2) to illustrate the stability of this assessment procedure; and 3) to try to determine an optimal segment length for analysis purposes. Segment lengths of 25m, 50m, 75m, and 100m are considered to determine if shorter lengths lend any instability to the strength of the prediction equations or to estimates of basal area, volume, or biomass.

- Factor 3 - airborne laser segment length: 3 levels, 1) entire airborne flightline; 2) an airborne segment length equal to the ground transect segment length used to generate the predictive equations with a 15m gap between segments; and 3) and airborne segment length equal to the ground transect segment length without a gap between segments. The shorter the laser segment, the higher the likelihood that single, tall trees with large crown diameters will produce an unusually high estimate of forest canopy height. Though basal area, volume, and biomass variances cannot be calculated when the flightline is used in its entirety, the use of the entire flightline eliminates the possibility of overestimates due to height variability seen at the shorter segment lengths. A 15m gap is introduced to mitigate the effects of spatial autocorrelation.[4] The presence of a 15m gap may reduce autocorrelation by separating sequential segments by the width of an average tree crown. Within-stand autocorrelation effects will still be present, but the gap should preclude artificial variance reductions due to height measurements taken within the same tree's canopy.

- Factor 4 - model used: 2 levels, logarithmic simple linear model, with intercept ($\ln(b) = a_0 + a_1(\bar{h}_a)$) versus no transformation with the line forced through the origin ($b = a_1(\bar{h}_a)$). The natural log model serves as a scientific standard. The untransformed model with no intercept (suggested by scatterplots of height with basal area, volume, biomass) is considered to see if it is inherently more accurate, stable. Prediction accuracies might be improved by fitting an untransformed model in order to avoid bias in the detransformed estimates. Transforming and detransforming linear relationships result in a bias which is the same as that bias encountered when a sample median is used to estimate the population mean in a skewed univariate data set (Miller, 1984). The systematic bias introduced by detransforming a model employing the natural log tranformation is downward. The bias can be partially counteracted by applying a multiplicative adjustment factor to the detransformed estimate. The adjustment factor is the exponentiated value of 0.5 $\hat{\sigma}^2$, where $\hat{\sigma}^2$ is the least squares estimate of

---

[4] In Tileran, for instance, the average canopy diameter is 6.85m ± 3.48m (one standard deviation). Using Tchebysheff's Theorum to conservatively estimate the appropriate lag,

$$P(\text{canopy diameter} \geq 6.85 + 2(3.48)) < \frac{1}{2^2},$$

or

$$P(\text{canopy diameter} \geq 13.82m) < 0.25.$$

Hence, a 14m lag between adjacent transect segments in Tileran would be appropriate 1) to remove most within-crown autocorrelation effects; and 2) to increase sample size while maintaining transect independence. In Tileran, then, a 14m lag will mitigate autocorrelation effects due to adjacent measurements made in the same crown. At LaSelva, where the mean canopy diameter is 5.14m ± 3.04m based on the 5m wide transect data, and 6.02m ± 3.04m based on 20m transect data, the appropriate separation between segments would be approximately 12m.

the error mean square of the linearized model (Miller, 1984). Consistent underestimation problems with the transformed models would suggest the use of this correction factor.

All possible combinations of the four factors are applied to the three study areas in order to identify those factor level combinations which yield results closest to the ground reference values. Forty eight treatment combinations are considered on each of the Tileran and LaSelva-5m data sets (2 regression procedures x 4 ground segment lengths x 3 airborne segment lengths x 2 models). Twenty four treatment levels are considered on the LaSelva-20m data set due to the reduced number of ground segment lengths available. The two larger ground transect lengths (75m and 100m) did not generate sample sizes large enough to develop reliable regression estimates. The comprehensive listing of results are reviewed 1) to empirically identify those factors which have the greatest effects on prediction accuracy; and 2) to identify specific factor levels which produce the most accurate results.

III.C.2. Multiple Independent Laser Variables

The treatment levels considered above correspond to different methods of developing regression equations. In other words, any particular treatment combination is, in reality, a specific approach which is used to formulate predictive equations for the program which processes the airborne laser data. Factor levels identified as useful in the simple linear regression analyses reported above are next applied to situations involving multiple independent variables. The eight simulated laser metrics listed in Table 5 are considered to determine if 1) different, individual laser metrics are better predictors of basal area, volume, and biomass than average canopy height, all pulses; and 2) two or more simulated laser metrics explain significant additional variation in basal area, etc., with the result that prediction accuracy is improved.

As in the canopy simulation, adjacency is used, and all individual tree canopy shapes are assumed to be elliptical. Within-canopy variance is set to 0.117 m^2 and is assumed to be normally distributed (see sections III.B.2 and IV.A for the origin of this number). Using the predictive equations developed for each of the treatment level combinations, estimates of basal area, volume, or biomass are generated by processing airborne laser transect data. In the program which analyzes the airborne laser data, the minimum tree height is consistently set at 5.0m. Any laser pulse which measures a canopy height less than

5.0m is considered a shrub or ground (<0.5m) hit, and that height is not accumulated and does not contribute to the basal area, volume, or biomass of that segment.

The simple linear regression results will identify a particular ground segment length and airborne laser segment length (with or without gap) which may be employed in conjunction with a single laser variable, $\bar{h}_a$, to predict basal area, volume, and biomass. These SLR results provide a framework for the multivariate research. Parametric multivariate techniques are employed to identify those simulated laser metrics most useful for predicting the ground metrics of interest. Only the specific ground segment length and airborne laser segment length identified as optimal in the SLR investigation are considered in the multiple linear regression analyses. Only parametric multiple linear regression procedures are considered due to lack of ready access to nonparametric, multivariate linear procedures.[5] As in the simple linear regression investigation above, two generic models are considered: 1) dependent variable (eg., total merchantable volume) transformed using the natural logarithm with a linear model with intercept; and 2) dependent variable not transformed, with the linear model forced through the origin (i.e., no intercept term).

Statistical Analysis System software (SAS, 1989), including procedure PLOT, UNIVARIATE, CHART, and REG are used to describe and build models which relate simulated laser to ground metrics. Plotting routines are used to identify curvilinear or logarithmic relationships, or to identify problems with unequal variances at different levels of the independent laser variables. The SAS procedure REG is used to investigate strength of fit and to calculate regression coefficients for all possible subsets of the eight simulated laser metrics listed in Table 5. Coefficients of determination, Mallows Cp statistic, and mean square error are considered to identify a particular subset of laser variables which can be used to predict seen and total basal area, volume, and biomass across the three study areas. Though quantitative analyses are involved, the variable selection process is subjective since different variables and different numbers of variables present themselves as "optimal subsets" on the different study areas. An attempt was made

---

[5] Nonparametric multiple linear regression procedures are not available on the commercial statistical packages commonly used today. SAS, BMDP, SPSS, IMSL, and IDL (a graphics package with appreciable nonparametric analysis capability) do not provide nonparametric MLR capabilities. The only nonparametric multiple linear regression software known to this researcher is available in the Nonparametric Statistical System (NPSS) package available through Dr. Pirie, Statistics Department, at Virginia Tech, Blacksburg, Virginia.

by this investigator to identify variables which explained significant variation in a particular dependent variable on all three study sites so as to make the models more generic. This lack of study site specificity, this attempt to make models generic, means that most all of the models considered are not optimal in terms of maximizing R^2, or minimizing Cp or MSE. The strength of data fit, and most likely, the predictive capability of the models so identified are compromised in an attempt to derive laser variables which might be commonly employed, or at least considered, in subsequent investigations in other tropical forests.

Once the generic models are constructed, specific regression coefficients can be calculated using the SAS procedure REG. Collinearity diagnostics, i.e., variance inflation factors (VIFs), condition indices, eigenvalues, and variance proportions, are generated and checked to insure that the models are not overfit. In situations where significant collinearity problems exist - i.e., VIFs greater than 10, condition indices greater than 30 (SAS condition index), eigenvalue(s) close to zero, two or more variance proportions near one - the offending laser metrics are dropped from the model. (Note: Rules of thumb for collinearity diagnostics are based on communications with Dr. Myers, VPI Statistics Dept.)

Regressions are formulated for seen and total basal area, volume, and biomass for each of the three study areas. These equations are used in the airborne laser data processing program to generate estimates. The different MLR results are compared to discern the effects of model choice, and the MLR results are compared with SLR results to see if additional, independent laser variables significantly improve model fit and predictive capability.

III.C.3. Line Intercept Sampling

The airborne laser transect may be viewed as a line sample, and the airborne laser height measurements may be processed, with certain assumptions, using Line Intercept Sampling (LIS) techniques. Kaiser (1983) reports the techniques which may be used to relate laser measurements taken along a profiling line, i.e., length of canopy intersection, canopy profile area, to ground metrics of interest, i.e., canopy area and canopy volume, respectively. Canopy volume is that volume between the top of the canopy and the ground, which equals the product of average canopy height and stand area. Canopy area and canopy volume, in turn, may be directly related to basal area and woody biomass, respectively, <u>if</u> a stable, predictable relationship can be found between canopy area and basal area and between canopy volume and woody biomass, as follows:

For a single transect:

$$\hat{b}_s = \hat{r}_b \;\; (10{,}000 m^2/ha) \;\; \left(\frac{1}{L}\right) \sum_{i=1}^{n} y_i \qquad (1)$$

$$\hat{v}_s = \hat{r}_v \;\; (10{,}000 m^2/ha) \;\; \left(\frac{1}{L}\right) \sum_{i=1}^{n} p_i \qquad (2)$$

where $\hat{r}_b$ = the ratio of basal area to canopy area, as estimated by
regression relating simulated laser to ground measurements,
$\hat{r}_v$ = the ratio of woody volume or biomass to crown volume,
as estimated by regression,
y_i = length of intersect across tree crown i,
p_i = canopy profile area of tree crown i,
and n = number of individual tree crowns intersected along the
laser transect.

For k multiple transects, DeVries (1986) suggests weighting individual transect estimates by transect length, as follows:

$$\hat{b}_s = \frac{\left[\sum_{j=1}^{k} L_j \left(\hat{r}_b (10{,}000 m^2/ha) \frac{1}{L_j} \sum_{i=1}^{n_j} y_{ij}\right)\right]}{\sum_{j=1}^{k} L_j} = \frac{\left[\hat{r}_b (10{,}000) \sum_{j=1}^{k} \sum_{i=1}^{n_j} y_{ij}\right]}{\sum_{j=1}^{k} L_j} \qquad (3)$$

$$\hat{v}_s = \frac{\left[\sum_{j=1}^{k} L_j \left(\hat{r}_v (10{,}000 m^2/ha) \frac{1}{L_j} \sum_{i=1}^{n_j} p_{ij}\right)\right]}{\sum_{j=1}^{k} L_j} = \frac{\left[\hat{r}_v (10{,}000) \sum_{j=1}^{k} \sum_{i=1}^{n_j} p_{ij}\right]}{\sum_{j=1}^{k} L_j} \qquad (4)$$

where y_{ij} = length of intersect across tree crown i, on transect j,
p_{ij} = canopy profile area of tree crown i, on transect j,
k = number of laser transects, and
and n_j = number of individual tree crowns intersected
along laser transect j.

Individual tree crowns do not have to be differentiated in order to estimate 1) length of crown intersected (for basal area estimation), or 2) canopy profile area (for volume or biomass estimation). Note that $\frac{1}{L}\sum_{i=1}^{n} y_i$ (from equation 1) is the length of the transect intersecting the forest canopy divided by the total length of the transect, which is a number between zero and one denoting canopy density. Likewise for biomass, $\frac{1}{L}\sum_{i=1}^{n} p_i$ (from equation 2) is the sum of the canopy areas divided by the total length of the transect, which simplifies to the average canopy height over the length of the transect. Individual tree crowns do not have to be identified and measured; the <u>sum</u> of all tree crown measurements along a given transect are paramount. Going one step further, it can be seen that the success of the use of the LIS procedure for predicting basal area and volume or biomass relys on 1) the strength of the relationship between canopy density and basal area and between average canopy height and woody volume or biomass; and 2) the predictability of the laser-ground ratios, r_b, r_v.

The airborne laser senses the seen (or, roughly speaking, the overstory) component of the forest canopy directly. The basal area, volume, and biomass components of the unseen portion of the forest canopy must also be enumerated; two approaches are tested. The first approach is consistent with that taken in the SLR and MLR analyses. The unseen component may be measured by developing predictive equations for seen and total basal area, volume, or biomass, and then subtracting. Under this scenario, then, additional equations may be written patterned on equations 1 - 4:

$$\hat{b}_t = \hat{t}_b(10{,}000 m^2/ha)\frac{1}{L}\sum_{i=1}^{n} y_i \qquad (5)$$

$$\hat{v}_t = \hat{t}_v(10{,}000 m^2/ha)\frac{1}{L}\sum_{i=1}^{n} p_i \qquad (6)$$

The multiple transect versions are configured similarly to equations 3 and 4. The unseen components are then deduced by subtracting the seen component from the total:

$$\hat{b}_u = \hat{b}_t - \hat{b}_s \qquad (7)$$

$$\hat{v}_u = \hat{v}_t - \hat{v}_s \qquad (8)$$

A second approach involves developing estimates of unseen basal area, volume, and biomass as a function of seen values. Implicit in this approach is the assumption that understory/overstory ratios are stable, or at least predictable, on a given study area. Seen values for basal area, volume, and biomass are calculated as in equations 1 and 2, or 3 and 4 if multiple transects are considered. Unseen values, then, are calculated as follows:

$$\hat{b}_u = \hat{s}_b \, \hat{b}_s \qquad (9)$$

$$\hat{v}_u = \hat{s}_v \, \hat{v}_s \qquad (10)$$

where s_b is the ratio of unseen basal area to seen basal area,

and s_v is the ratio of unseen volume (or biomass) to seen volume.

The seen and unseen estimates may be summed to estimate total basal area and total biomass per hectare.

$$\hat{b}_t = \hat{b}_s + \hat{b}_u \qquad (11)$$

$$\hat{v}_t = \hat{v}_s + \hat{v}_u \qquad (12)$$

Variance calculations rely on multiple, independent transect estimates. Variance cannot be calculated based on a single transect because the term depends on the relative locations of tree crowns along the transect (Kaiser, 1983). Observed variance, then, is calculated as a weighted sum of squares (DeVries, 1986), as follows:

$$\hat{\mathrm{var}}(\hat{b}_s) = \frac{\sum_{j=1}^{k} L_j(\hat{b}_{sj} - \hat{b}_s)^2}{(\sum_{j=1}^{k} L_j)\ (k-1)} \qquad (13)$$

where $\hat{b}_{s_j}$ is the estimate of seen basal area per hectare for transect j, and $\hat{b}_s$ is the estimate across all transects.

The sample variances for b_u, b_t, and v_s, v_u, and v_t are similarly calculated. These calculations assume independent laser transects, ones that are randomly or systematically located across tracts which exhibit unpatterned locations. Concerns with independence arise when attempts are made to increase sample size (i.e., k) by subsectioning transects.

The utility of the LIS procedure rests entirely on the stability or predictability of r_b, r_v, s_b, s_v, t_b, and t_v. Preliminary work showed that, in fact, these LIS metrics are not stable; they change with changing canopy condition. Figure 18 illustrates the variability of the different volume or biomass LIS metrics across different canopy density and canopy height conditions. As with the MLR analyses discussed previously, variable selection procedures are employed to identify those simulated laser metrics which explain the variability of r_b, r_v, s_b, s_v, t_b, and t_v. If these LIS metrics can be predicted using some combination of the eight simulated laser metrics, then estimates of basal area, volume, and biomass will be generated using the LIS equations reported above. These results will be compared with the SLR and MLR results to identify the airborne laser processing approach which produces 1) the most robust prediction equations - eg., high R^2, low MSE); and 2) the most accurate estimates of basal area, volume, and biomass.

III.D. Effects of Canopy Shape

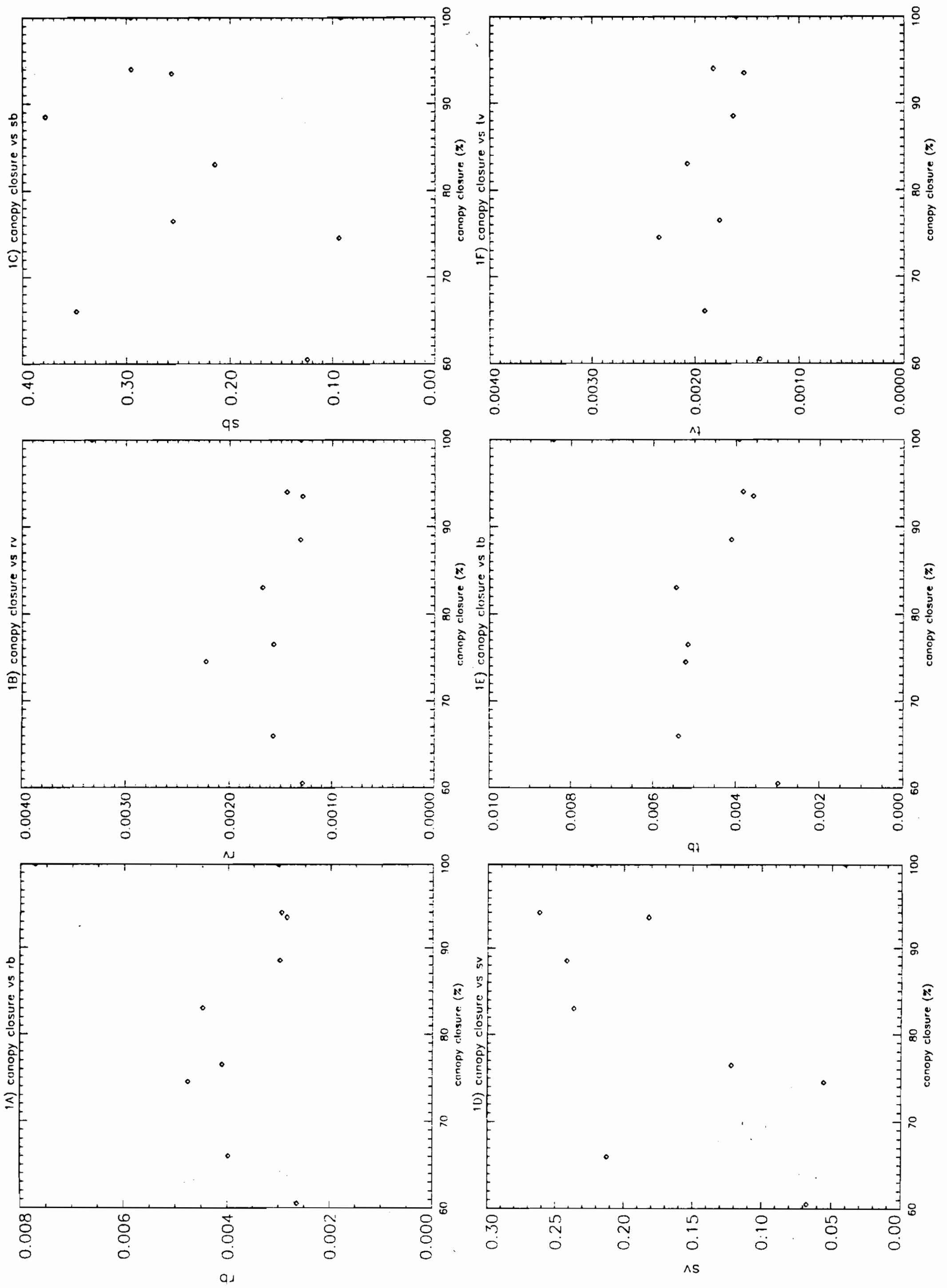
Simulations of forest canopy structure involve assumptions of canopy shape unless tree shapes are noted in the field. They were not noted in this study. The SLR, MLR, and LIS analyses are conducted assuming that all tree canopies are elliptically shaped. The canopy shape assumed in the course of the simulation can have a marked effect on subsequent regressions (Nelson and Yaya, 1990) since shape will affect the average canopy height and height variance. In general, the more conical the crown, the greater the discrepancy between laser measurements of canopy height and ground-measured total tree height. The canopy shape assumption, then, may have a significant effect on the accuracy of predicted basal area and biomass. A simple example is depicted in Figure 19 (from Nelson and Yaya, 1990, their Figure 2) where the natural log of timber volume is linearly related to total tree height. If the regression coefficients are developed assuming that the predominant canopy shape is conical, and then these regression equations are applied to airborne laser data flown over a forest where the actual canopy shape is, in general, more

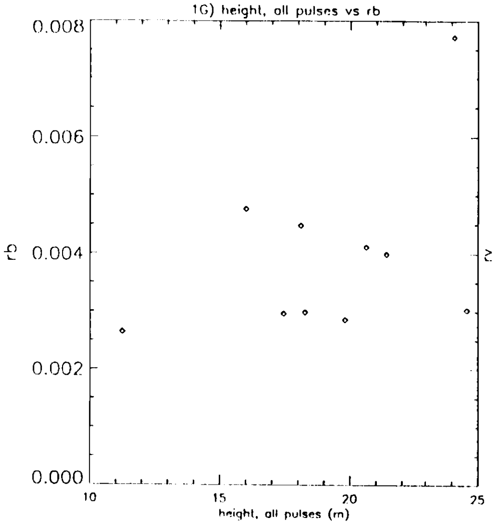
Figure 18. Scatterplots of LIS metrics r_b, r_v, s_b, s_v, t_b, and t_v as a function of canopy density or total tree height for the 1) Tileran, 2) LaSelva-5m, and 3) LaSelva-20m data sets (see Table 5 for variable definitions).

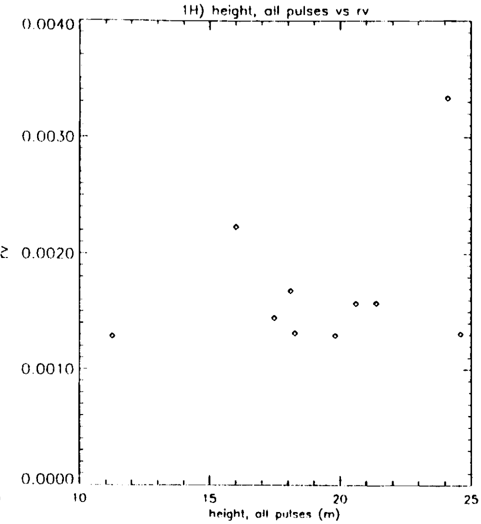
1A) Tileran, canopy closure vs. r_b.
1B) Tileran, canopy closure vs. r_v.
1C) Tileran, canopy closure vs. s_b.
1D) Tileran, canopy closure vs. s_v.
1E) Tileran, canopy closure vs. t_b.
1F) Tileran, canopy closure vs. t_v.
1G) Tileran, height, all pulses vs. r_b.
1H) Tileran, height, all pulses vs. r_v.
1I) Tileran, height, all pulses vs. s_b.
1J) Tileran, height, all pulses vs. s_v.
1K) Tileran, height, all pulses vs. t_b.
1L) Tileran, height, all pulses vs. t_v.

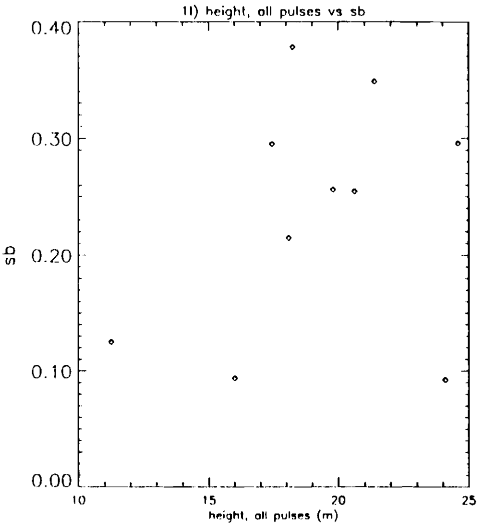
2A) LaSelva-5m, canopy closure vs. r_b.
2B) LaSelva-5m, canopy closure vs. r_v.
2C) LaSelva-5m, canopy closure vs. s_b.
2D) LaSelva-5m, canopy closure vs. s_v.
2E) LaSelva-5m, canopy closure vs. t_b.
2F) LaSelva-5m, canopy closure vs. t_v.
2G) LaSelva-5m, height, all pulses vs. r_b.
2H) LaSelva-5m, height, all pulses vs. r_v.
2I) LaSelva-5m, height, all pulses vs. s_b.
2J) LaSelva-5m, height, all pulses vs. s_v.
2K) LaSelva-5m, height, all pulses vs. t_b.
2L) LaSelva-5m, height, all pulses vs. t_v.

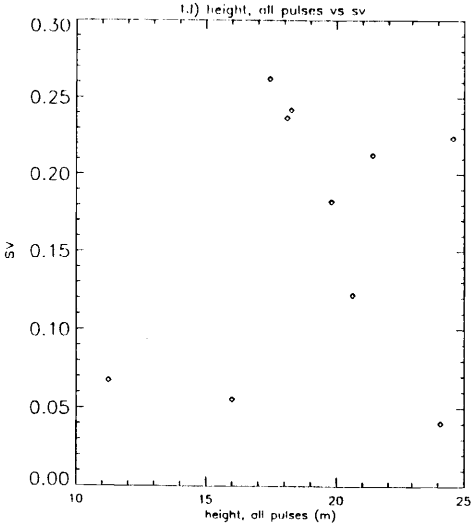
3A) LaSelva-20m, canopy closure vs. r_b.
3B) LaSelva-20m, canopy closure vs. r_v.
3C) LaSelva-20m, canopy closure vs. s_b.
3D) LaSelva-20m, canopy closure vs. s_v.
3E) LaSelva-20m, canopy closure vs. t_b.
3F) LaSelva-20m, canopy closure vs. t_v.
3G) LaSelva-20m, height, all pulses vs. r_b.
3H) LaSelva-20m, height, all pulses vs. r_v.
3I) LaSelva-20m, height, all pulses vs. s_b.
3J) LaSelva-20m, height, all pulses vs. s_v.
3K) LaSelva-20m, height, all pulses vs. t_b.
3L) LaSelva-20m, height, all pulses vs. t_v.

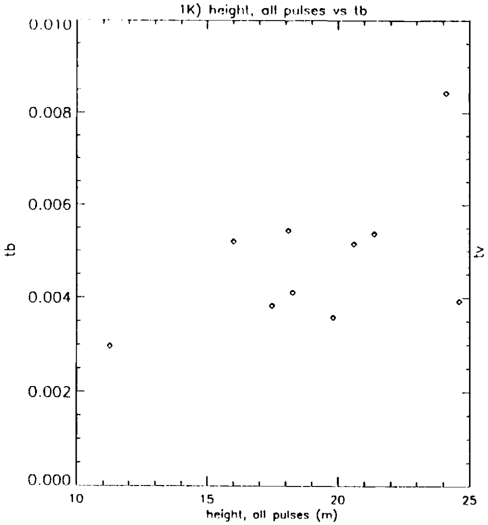

1A) canopy closure vs rb
rb
0.008
0.006
0.004
0.002
0.000
60
70
80
90
100
canopy closure (%)
1B) canopy closure vs rv
rv
0.0040
0.0030
0.0020
0.0010
0.0000
canopy closure (%)
1C) canopy closure vs sb
sb
0.40
0.30
0.20
0.10
0.00
canopy closure (%)
1D) canopy closure vs sv
sv
0.30
0.25
0.20
0.15
0.10
0.05
0.00
canopy closure (%)
1E) canopy closure vs tb
tb
0.010
0.008
0.006
0.004
0.002
0.000
canopy closure (%)
1F) canopy closure vs tv
tv
0.0040
0.0030
0.0020
0.0010
0.0000
canopy closure (%)

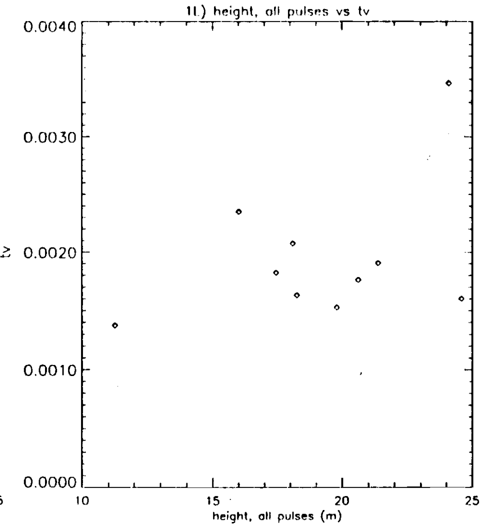

1G) height, all pulses vs rb

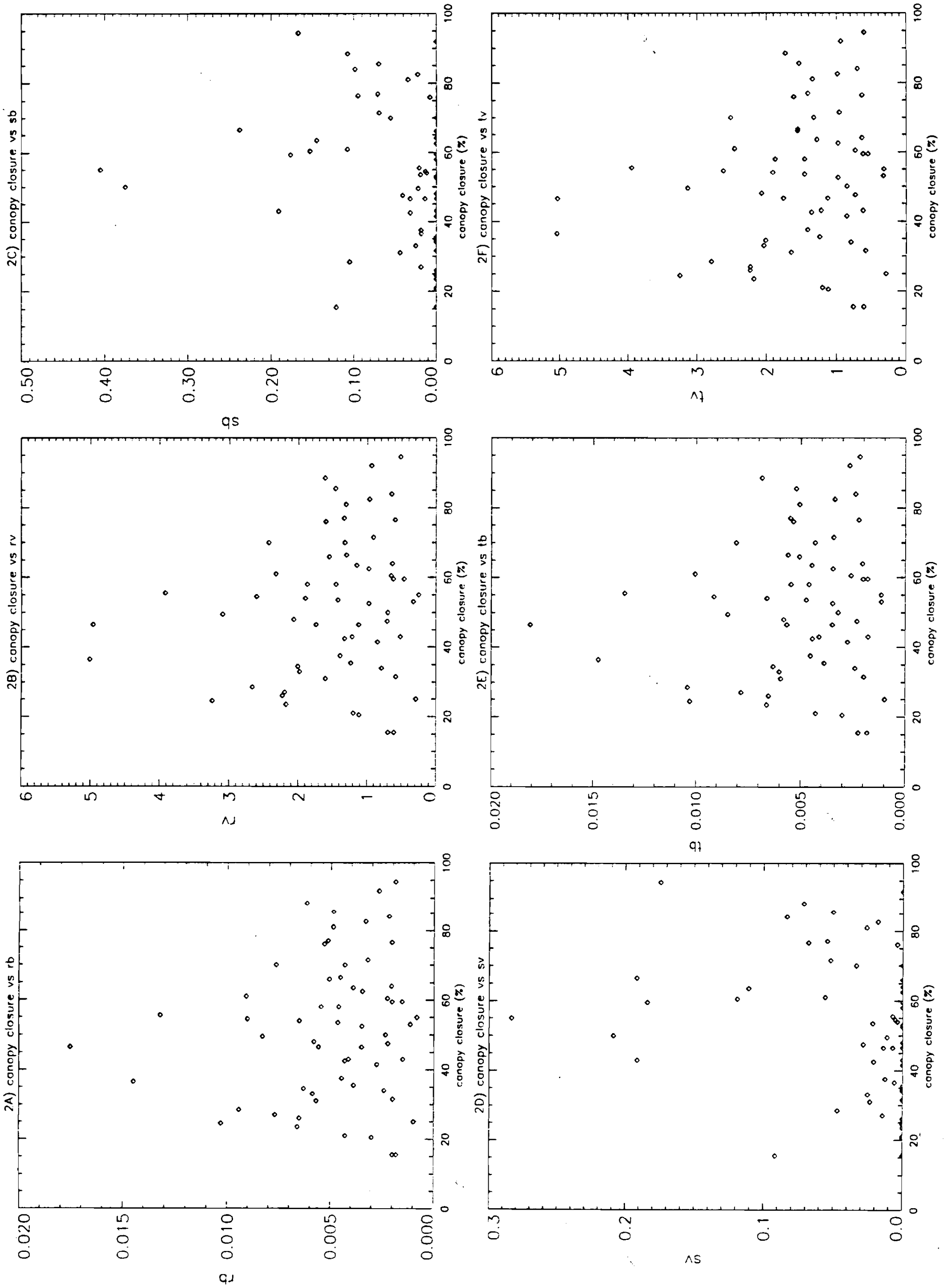

1H) height, all pulses vs rv

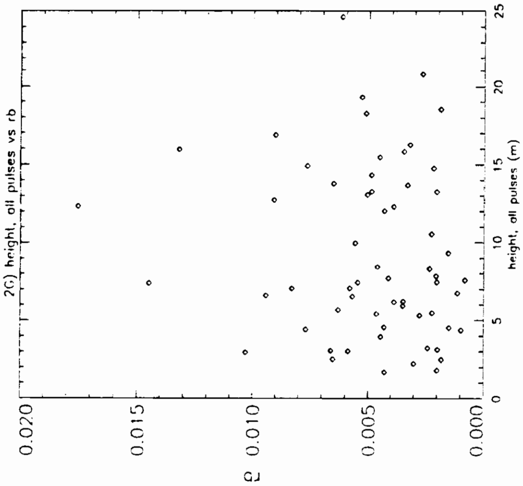

1I) height, all pulses vs sb

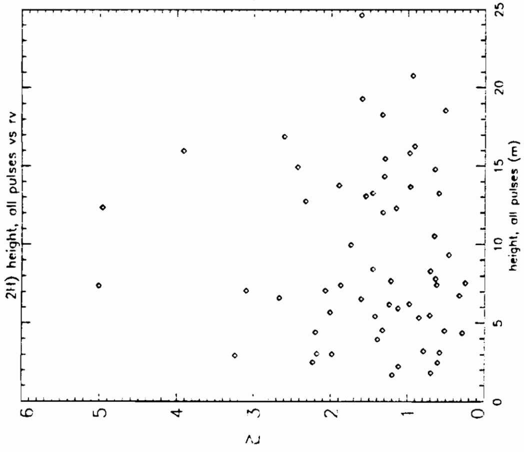

1J) height, all pulses vs sv

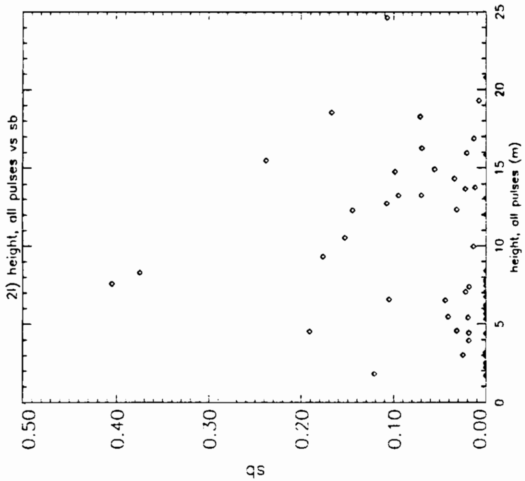

1K) height, all pulses vs tb

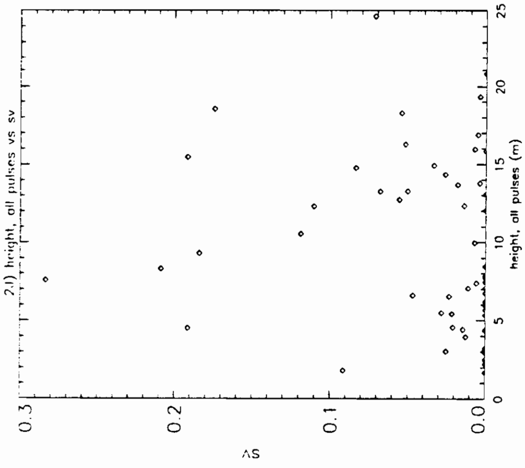

1L) height, all pulses vs tv

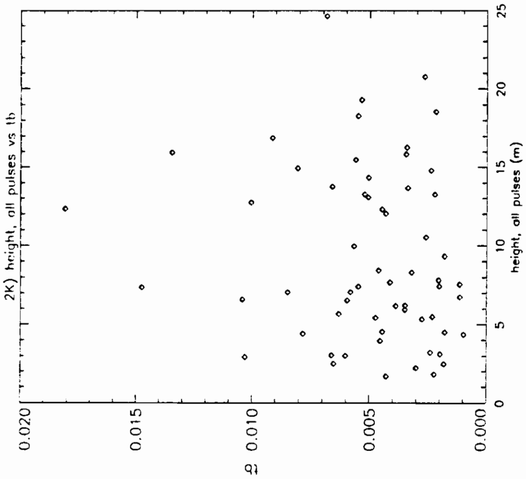

2A) canopy closure vs rb
rb
0.020
0.015
0.010
0.005
0.000
canopy closure (%)
0
20
40
60
80
100
2B) canopy closure vs rv
rv
6
5
4
3
2
1
0
canopy closure (%)
2C) canopy closure vs sb
sb
0.50
0.40
0.30
0.20
0.10
0.00
canopy closure (%)
2D) canopy closure vs sv
sv
0.3
0.2
0.1
0.0
canopy closure (%)
2E) canopy closure vs tb
tb
0.020
0.015
0.010
0.005
0.000
canopy closure (%)
2F) canopy closure vs tv
tv
6
5
4
3
2
1
0
canopy closure (%)

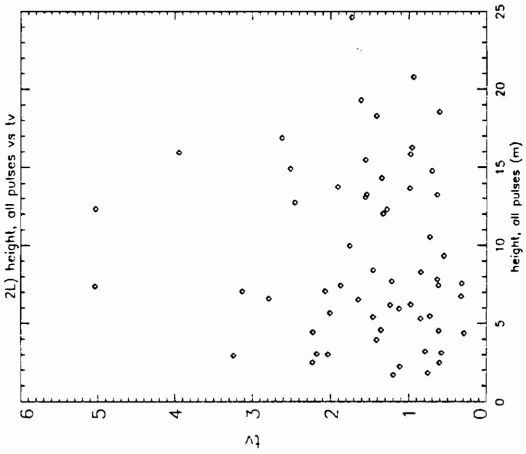

2G) height, all pulses vs rb
rb
0.020
0.015
0.010
0.005
0.000
0
5
10
15
20
25
height, all pulses (m)

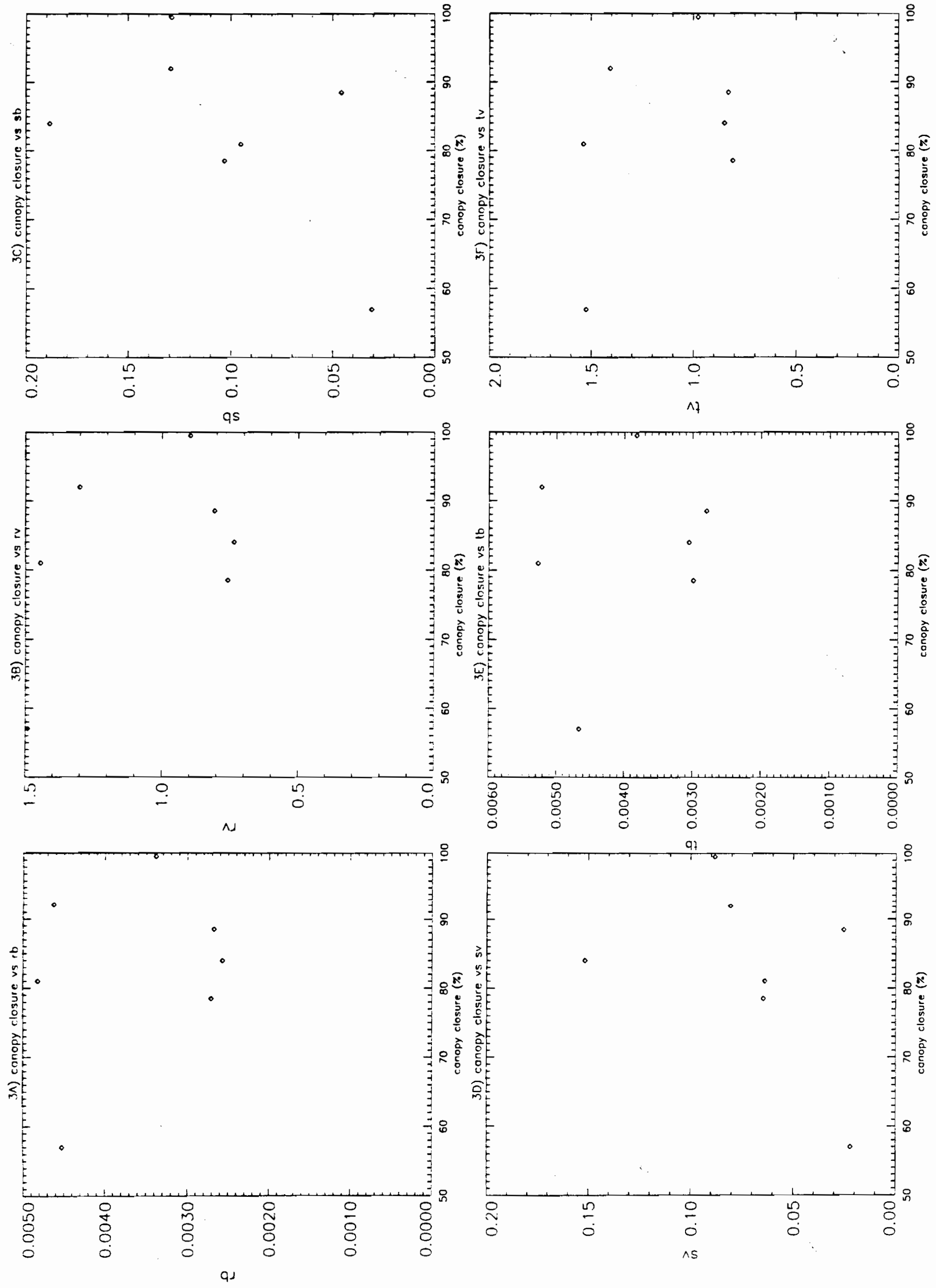

2H) height, all pulses vs rv
rv
6
5
4
3
2
1
0
0
5
10
15
20
25
height, all pulses (m)

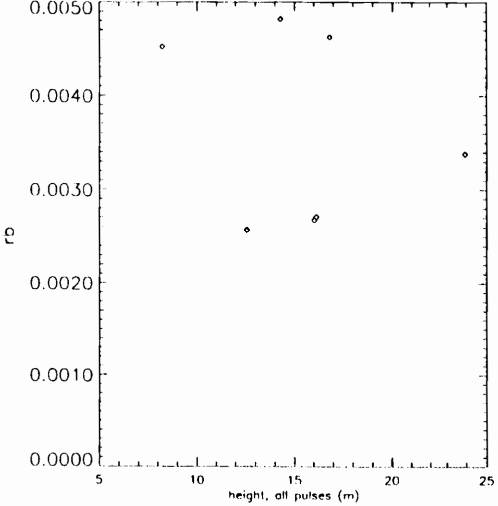

2I) height, all pulses vs sb
sb
0.50
0.40
0.30
0.20
0.10
0.00
0
5
10
15
20
25
height, all pulses (m)


2J) height, all pulses vs sv
sv
0.3
0.2
0.1
0.0
0
5
10
15
20
25
height, all pulses (m)

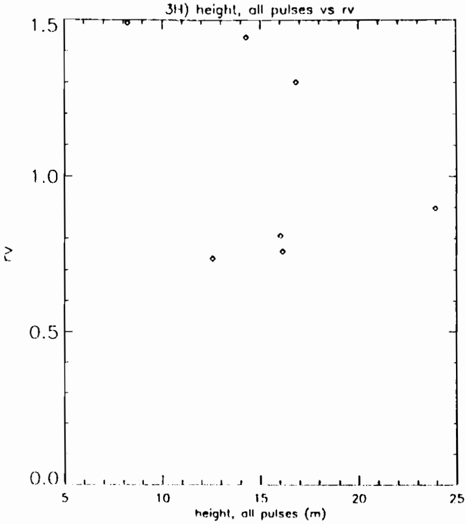

2K) height, all pulses vs tb
tb
0.020
0.015
0.010
0.005
0.000
0
5
10
15
20
25
height, all pulses (m)


2L) height, all pulses vs tv
tv
6
5
4
3
2
1
0
0
5
10
15
20
25
height, all pulses (m)

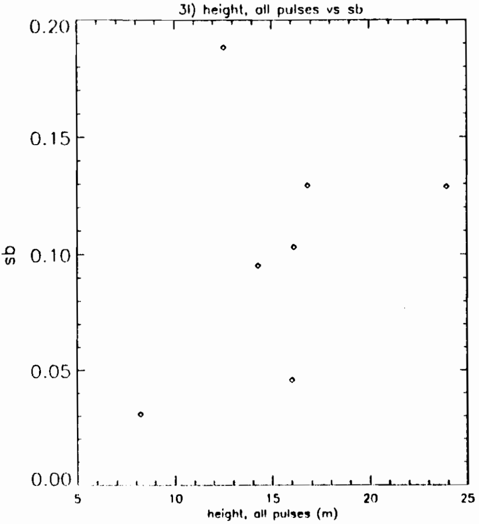

3A) canopy closure vs rb
rb
0.0050
0.0040
0.0030
0.0020
0.0010
0.0000
3B) canopy closure vs rv
rv
1.5
1.0
0.5
0.0
3C) canopy closure vs sb
sb
0.20
0.15
0.10
0.05
0.00
3D) canopy closure vs sv
sv
0.20
0.15
0.10
0.05
0.00
3E) canopy closure vs tb
tb
0.0060
0.0050
0.0040
0.0030
0.0020
0.0010
0.0000
3F) canopy closure vs tv
tv
2.0
1.5
1.0
0.5
0.0
50
60
70
80
90
100
canopy closure (%)



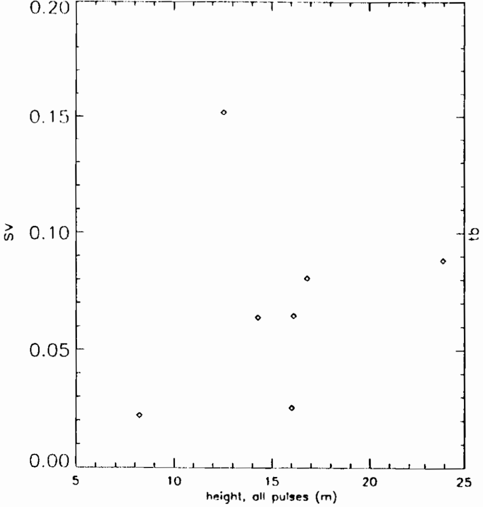
3G) height, all pulses vs rb



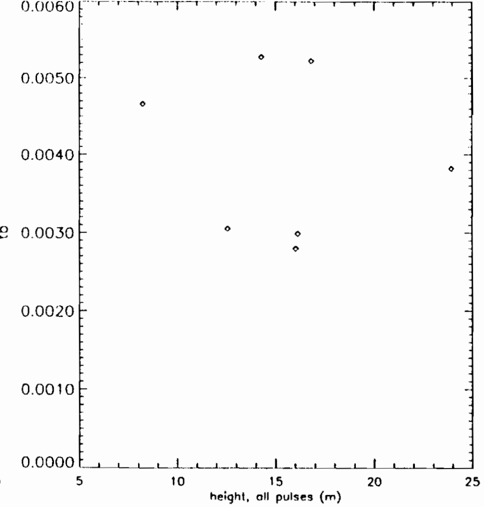
3H) height, all pulses vs rv



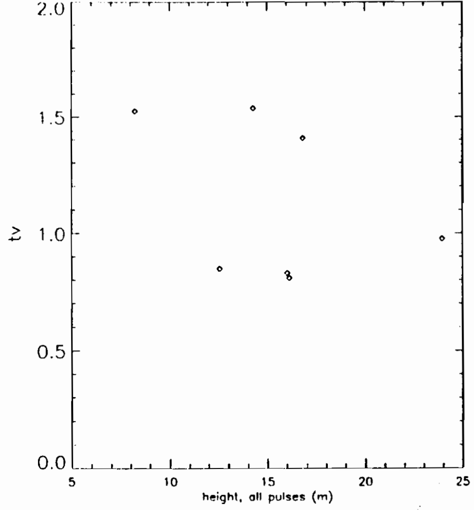
3I) height, all pulses vs sb



3J) height, all pulses vs sv



3K) height, all pulses vs tb



3L) height, all pulses vs tv

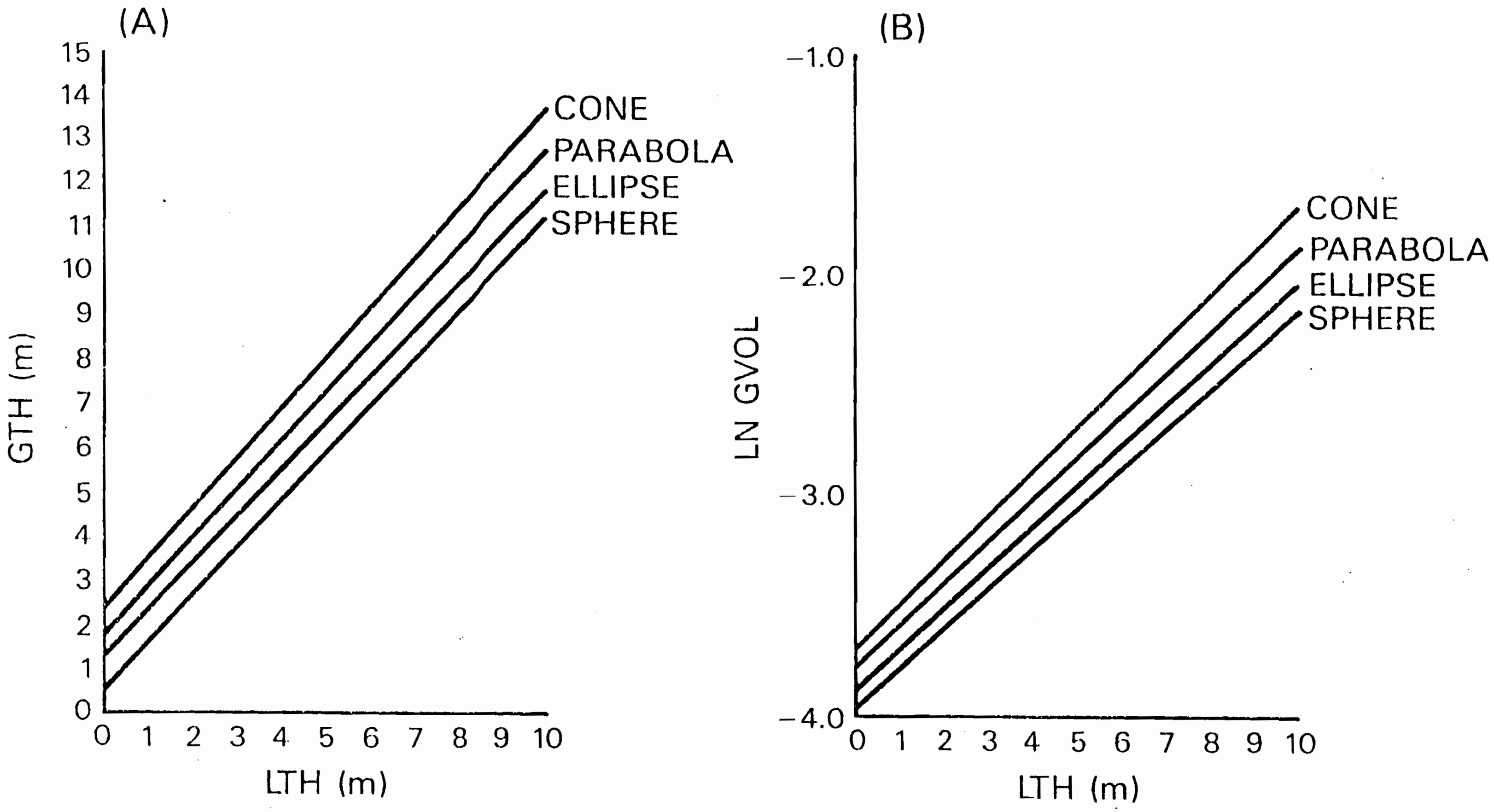


Figure 19. The relationship between laser-measured tree height and A) ground-measured tree height; and B) the ln(merchantable timber volume) for four canopy shapes (from Nelson and Yaya (1990), their Figure 2).

elliptical, then the predicted volume would overestimate actual volume on the ground. A sensitivity analysis is conducted on each of the airborne flightlines in order to quantify the effects of assumed canopy shape on predicted basal area, volume, and biomass. Separate sets of regressions are developed for each of the four shapes, conical, parabolic, elliptical, spherical, and also for a random selection of these four shapes. Of interest here is the extent to which basal area, volume, and biomass estimates diverge as a function of canopy shape for a particular airborne laser flightline. This section of research is meant to highlight and illustrate the importance of canopy shape in the simulation process.

IV. RESULTS

In order to develop estimates of basal area and biomass or volume using airborne laser data, certain parameters must be set 1) in the computer simulator; and 2) in the program which analyzes the airborne laser data. Once these parameters are set, estimates of basal area and biomass or volume require that the analyst:

1. assume a particular canopy shape (if shapes have not been noted in the field);
2. specify a segment, or sample transect length. This segment length is used a) to formulate the predictive regressions, and b) to segment the airborne laser data.

Initialization of the two parameters is discussed directly below. Issues dealing with the effects of numerous factors, including optimal segment length and canopy shape, are addressed in subsequent sections.

IV.A. Initialization of Parameters

Two parameters had to be set prior to initiating the analyses. The two parameters, 1) within-canopy height variation and 2) minimum tree height, are measures which may be estimated from airborne laser canopy profiles and from field observation, respectively. These are numbers which are site-specific and which may need to be changed from study to study.

Within-canopy height variation is some variance measure of the departure of a particular tree canopy from a smooth curve. The canopy height simulator, which constructs the topography of the top of the forest canopy based on mapped stand data, assumes a smooth canopy shape for each tree. The within-canopy height variation measure is used to generate canopy roughness. The number which must be provided to the canopy height simulator is an estimate of the standard deviation of the random height variation within an individual tree canopy. For lack of any evidence to the contrary, a normal distribution is assumed. Airborne laser profiles acquired over Tileran and LaSelva were plotted and measurements were taken on these profiles to develop a generic estimate of within-canopy variability. In order to estimate variability, the analyst identifies individual tree canopies on a laser plot, and then draws, free-form, close-fitting smooth curves to describe their general shape. Measurements are approximate and somewhat subjective, but they are made in order to generate a good-faith estimate of a relatively minor component in overall canopy height variability. Note that between-tree variation and within-tree height variation due to canopy curvature are the two larger components of the tree height variability and these are accounted for in the mapped stand data. Table 6 reports the height variation noted within-canopy in airborne laser data acquired over LaSelva and Tileran. Figure 20 illustrates the effects of inclusion of this random roughness factor in canopy construction.

Table 6. Within-canopy height variation noted in airborne laser data acquired over LaSelva and Tileran.

study area	airborne laser flightline	n	average distance(m)	standard deviation(m)
LaSelva	CRT1F1	15	1.04	0.28
	CRT2F1	17	0.60	0.29
	CRT2F2	15	0.77	0.34
	All LaSelva lines	47	0.80	0.35
Tileran	CRT1F2	12	0.61	0.32
	CRT1F3	14	0.76	0.33
	CRT1F4	12	0.60	0.33
	All Tileran lines	38	0.66	0.33
All Obs., LaSelva and Tileran		85	0.74	0.34

A second parameter, minimum tree height, must be supplied to the computer program which manipulates the airborne laser data to develop basal area, volume, or biomass estimates. This second parameter, derived from field observation, is an estimate of the general height of the shrub layer beneath the forest canopy. The minimum tree height parameter is used to make the forest canopy height measurements developed via simulation and those height measurements acquired by airborne laser ranging comparable. This lower bound (i.e., the minimum tree height) is necessitated by the fact that the mapped stand data, which inventories stems greater than 10 cm, essentially ignores a ubiquitous shrub layer. As a result, the simulated canopy is smooth and flat with canopy heights equal to zero in those areas where there are openings between trees. The minimum tree height is used to zero out height measurements where the airborne laser encounters holes in the overstory canopy.

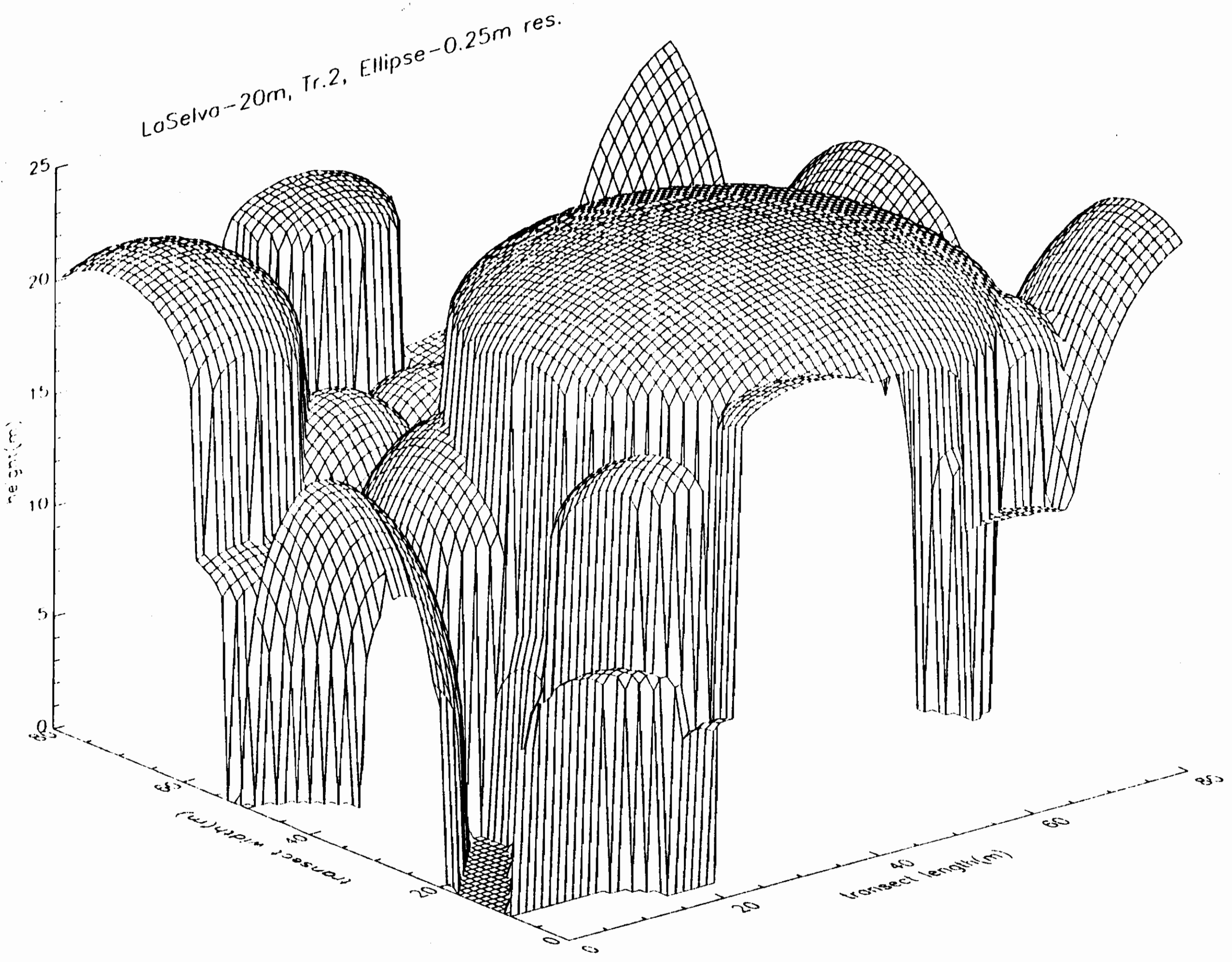
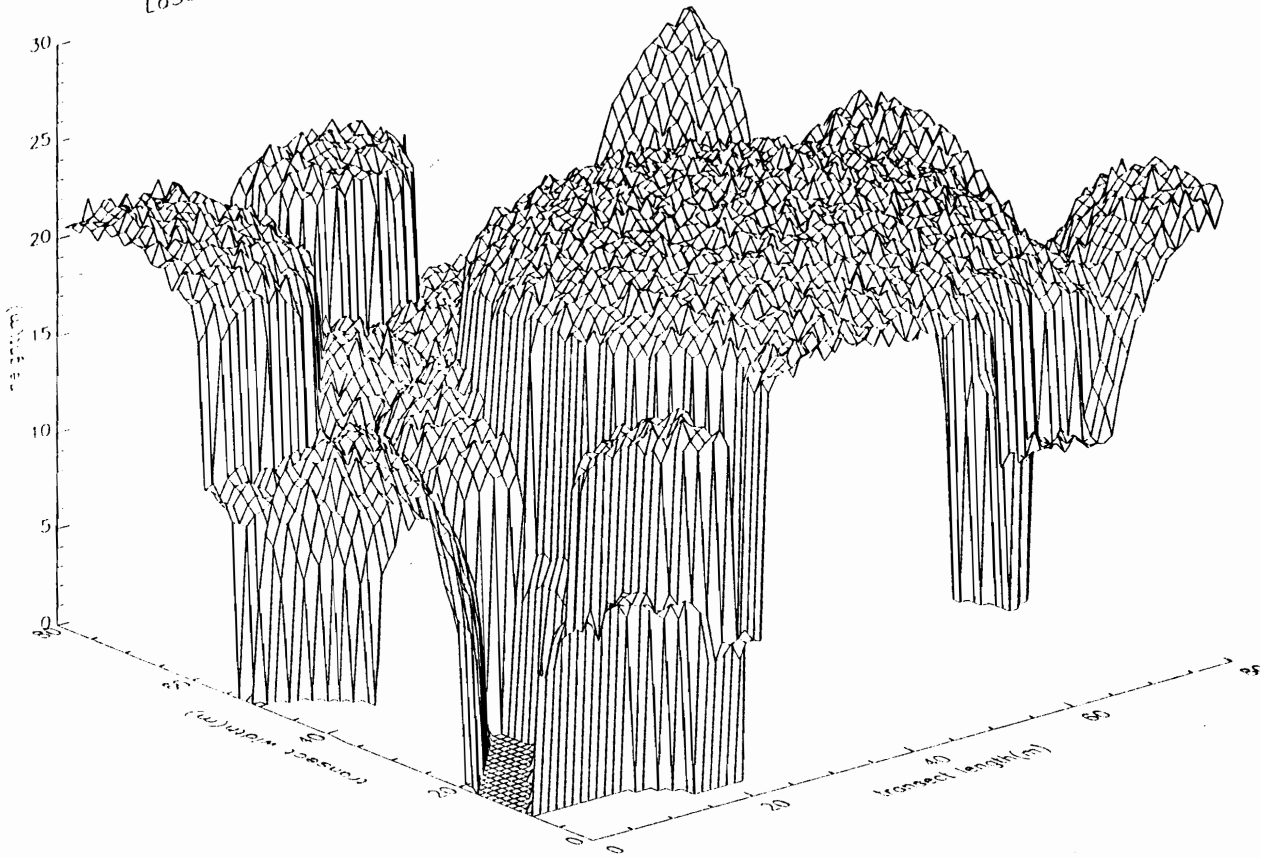


Figure 20. A section of the LaSelva-20m ground transect (65-85m along transect) simulated without (A) and with (B) canopy roughness enabled. Height is in meters; X and Y dimensions are 20m x 20m. An elliptical canopy shape is assumed, and adjacency is enabled.

LaSelva-20m, Tr.2, Ellipse 0.25m res.



Fieldwork done at LaSelva (44 points) indicates that a shrub layer, typically large ferns and small palms, extends 3-4m above the forest floor. No field observations were made at Tileran by this investigator. As a result of these field observations, the minimum tree height was set at 5m for the analyses of the LaSelva and Tileran data sets. This parameter dictates that any airborne laser pulse where canopy height is found to be between 0.5m and 5.0m will be considered a ranging measurement to the shrub layer. Any pulse with a ranging measurement of canopy height less than 0.5m is considered a ground hit. As such, airborne laser height measurements less than 5m are not included in height, canopy profile, or height variance calculations.

IV.B. Simple Linear Regression Analyses

Simple linear regressions using the average canopy height of all pulses along a segment ($\bar{h}_a$) as the independent variable were run to test the effects of four factors which might affect the accuracy of airborne laser estimates of basal area, volume, and biomass. Those four factors include: 1) regression procedure - parametric versus nonparametric regression techniques; 2) ground segment length - four lengths considered, 25, 50, 75, and 100 meters; 3) airborne laser segment length - entire flightline versus ground segment length with a 15m gap versus the ground segment length without a gap; and 4) simple logarithmic model, with intercept versus a simple linear model (no transformation) where the regression is forced through the origin. The purpose of this portion of the research is to use simple models to identify those treatment combinations which produce accurate results.

Scatterplots of total basal area and total merchantable volume (Tileran) or total above-ground biomass (LaSelva-5m and LaSelva-20m) versus average canopy height are presented in Figure 21. Regressions involving untransformed data, whether parametric or nonparametric, are forced through the origin. Such a constraint appears to be reasonable in the scatterplots and is based on the principle that a forest canopy with an average height of zero should not have measurable (or negative, depending on the intercept) basal area, volume, or biomass. As can be seen in the scatterplots, an intercept term is needed when the logarithmic model is considered.

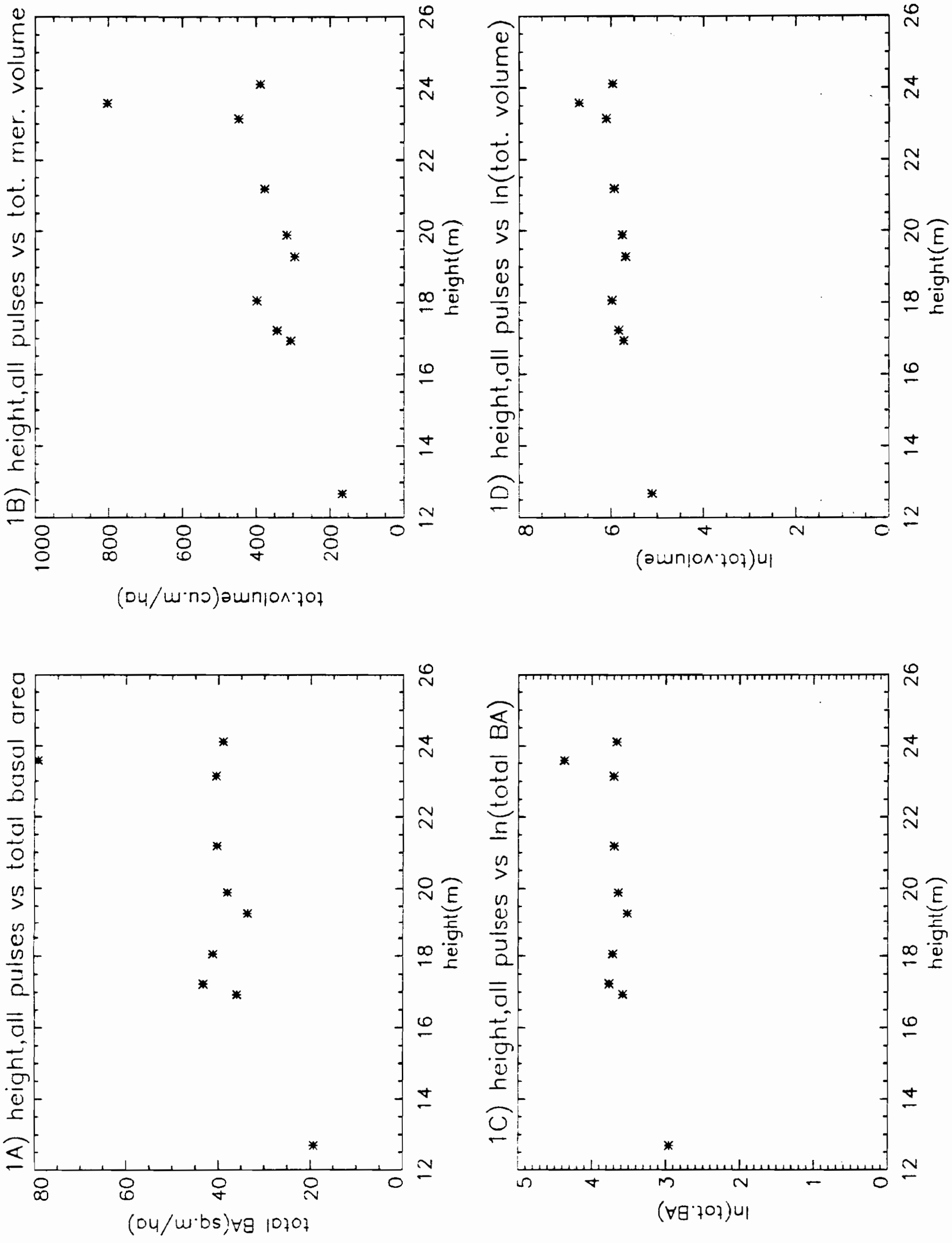
SAS procedure REG was run to calculate slopes and (where applicable) intercepts for those treatment combinations involving parametric regression. Nonparametric simple linear regression code was

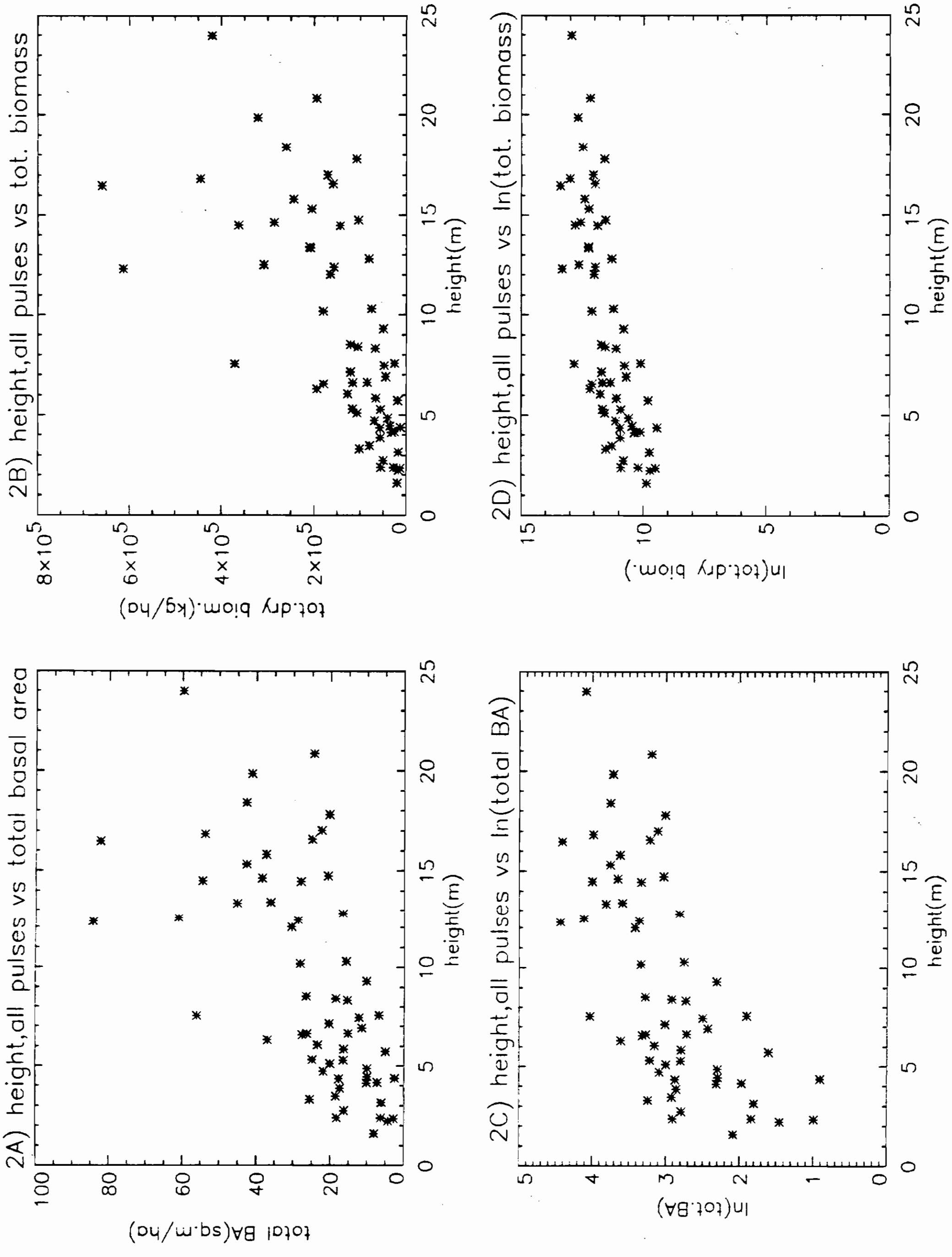
Figure 21. Scatterplots developed using the 1) Tileran, 2) LaSelva-5m, and 3) LaSelva-20m ground data sets and the canopy simulation program. Only 50m segment lengths are considered.

1A) Tileran, height, all pulses versus total basal area.
1B) Tileran, height, all pulses versus total merchantable volume.
1C) Tileran, height, all pulses versus ln(total basal area).
1D) Tileran, height, all pulses versus ln(total merchantable volume).

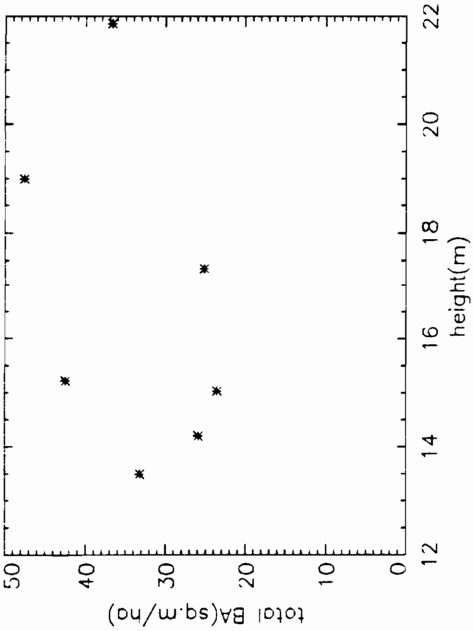
2A) LaSelva-5m, height, all pulses versus total basal area.
2B) LaSelva-5m, height, all pulses versus total above-ground dry biomass.
2C) LaSelva-5m, height, all pulses versus ln(total basal area).
2D) LaSelva-5m, height, all pulses versus ln(total above-ground dry biomass).

3A) LaSelva-20m, height, all pulses versus total basal area.
3B) LaSelva-20m, height, all pulses versus total above-ground dry biomass.
3C) LaSelva-20m, height, all pulses versus ln(total basal area).
3D) LaSelva-20m, height, all pulses versus ln(total above-ground dry biomass).

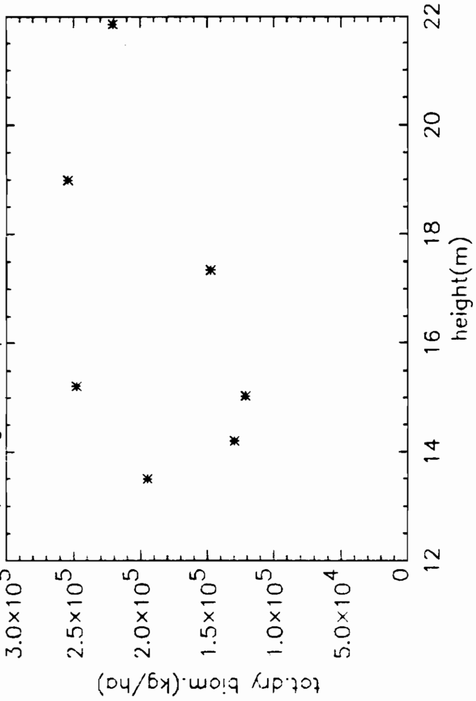

1A) height,all pulses vs total basal area
total BA(sq.m/ha)
height(m)
1B) height,all pulses vs tot. mer. volume
tot.volume(cu.m/ha)
height(m)
1C) height,all pulses vs ln(total BA)
ln(tot.BA)
height(m)
1D) height,all pulses vs ln(tot. volume)
ln(tot.volume)
height(m)
12
14
16
18
20
22
24
26
0
20
40
60
80
200
400
600
800
1000
1
2
3
4
5
6
8


2A) height,all pulses vs total basal area
total BA(sq.m/ha)
100
80
60
40
20
0
0
5
10
15
20
25
height(m)
2B) height,all pulses vs tot. biomass
tot.dry biom.(kg/ha)
8×10^5
6×10^5
4×10^5
2×10^5
0
0
5
10
15
20
25
height(m)
2C) height,all pulses vs ln(total BA)
ln(tot.BA)
5
4
3
2
1
0
0
5
10
15
20
25
height(m)
2D) height,all pulses vs ln(tot. biomass)
ln(tot.dry biom.)
15
10
5
0
0
5
10
15
20
25
height(m)

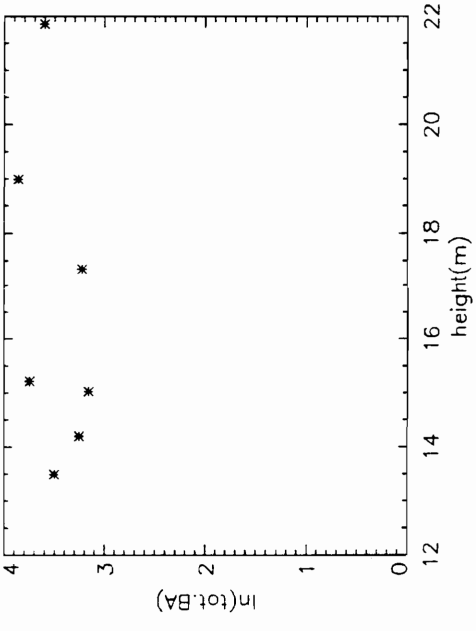
3A) height,all pulses vs total basal area



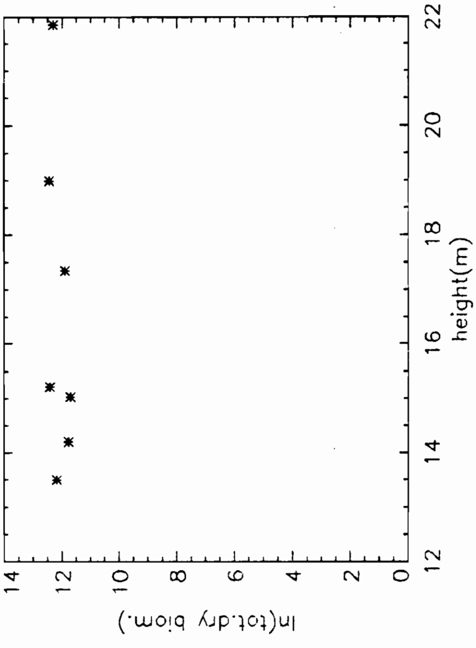
3B) height,all pulses vs tot. biomass



3C) height,all pulses vs ln(total BA)



3D) height,all pulses vs ln(tot. biomass)



written to estimate the corresponding nonparametric coefficients. In those regressions without intercepts, a simple slope estimator was employed where one slope was calculated for each observation, these candidates were ranked and the median was selected. Based on communications with Dr. Pirie, a Hodges-Lehmann (H-L) one sample estimator (Hollander and Wolfe, 1973) was also tested to calculate these slopes. The H-L estimator is reportedly more efficient than the simple slope estimator since the former ranks n(n+1)/2 averaged observations, whereas the latter considers only n unaveraged slopes. However, consideration of the H-L estimator was dropped when it was found that the H-L slope estimates in regressions without intercepts were consistently higher than those calculated using the simple slope estimator. Ten data sets were analyzed to compute both the H-L slope estimate and the simple nonparametric slope estimates. In 9 of 10 data sets, H-L slope estimates were larger than the simple estimates. As can be seen in the scatterplots, potentially highly influential observations tend to be much larger than the general population due to the effects of the occasional large tree in short segments. The H-L one sample estimator apears to be biased upwards in the presence of large HIOs since the H-L estimator involves more slope averaging. The nonparametric simple linear regression program, then, employed the simple slope estimator to develop all nonparametric regressions.

Table 7 lists the coefficients of determination associated with the parametric simple linear regressions (with intercept) developed for the logarithmic models. R^2 values are not reported for the nonparametric analyses or for the parametric analyses forced through the origin since, in these situations, the R^2 values may be unbounded at the lower or upper ends, are not comparable, and/or may be misleadingly inflated.[6] Table 7 indicates that, across most segment lengths and study areas (LaSelva-5m,

---

[6] R^2, in the context of least squares estimation with an intercept, "represents the proportion of variation in the response data that is explained by the model" (Myers, 1990, pg. 37), and this value is bounded by zero and one. R^2 are not reported from the results of the nonparametric analyses since the total sum of squares cannot be partitioned into sum of squares regression and sum of squares residual. An R^2-like measurement may be computed by calculating a corrected sum of squares regression ($\sum(\hat{y}_i - \bar{y})^2$) and dividing by the corrected sum of squares total ($\sum(y_i - \bar{y})^2$). This number, with a lower bound of 0, may be larger than 1 and is not comparable to a true R^2 due to the lack of additivity of the corrected sums of squares. R^2 values likewise are not reported for those parametric analyses where the regression is forced through the origin. In this situation, the R^2 analog for the zero intercept case, $R^2_{(0)}$, deals with uncorrected sums of squares with sum of squares regression equal to $\sum\hat{y}_i^2$ and sum of squares total equal to $\sum y_i^2$ (see Myers, 1990, pg. 39 for derivation). The ratio of these two sums of squares results in an $R^2_{(0)}$ measurement which tends to be larger than the R^2 value in the intercept case, regardless of any improvement in quality of fit. Various alternatives to $R^2_{(0)}$, where various corrected sums of squares are considered, have problems with their lower or upper bounds (less than zero or greater than one), depending on the method of computation.

25m segment length being the exception), merchantable volume and total above-ground biomass are better predicted by simulated laser metrics than basal area. No discernable pattern is noted across study areas with respect to segment length.

Table 8 lists the different mean square errors (MSEs) associated with the various predictive equations considered. Mean square error, in this table, is calculated as follows:

$$E = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}$$

where y_i = observed dependent variable (eg., total basal area);
$\hat{y}_i$ = predicted dependent variable;
and n = total number of observations used to develop the regression.

The total basal area, total merchantable volume, and total above-ground biomass data presented in Table 8 is depicted in Figure 22. The effects of decreasing segment length are obvious in the table and figure. In general, as segment length decreases, MSEs increase, with the largest rate of increase noted between the 25m and 50m segment lengths. The LaSelva-5m ln(total biomass) figure (22.2.c) is particularly noteworthy since the parametric and nonparametric MSEs increase approximately 1300 and 1500 percent respectively as segment length shortens from 50 to 25m. This remarkable increase is due to the presence of eight 25m segments without trees; the relatively large difference between the intercept and the origin inordinately drive up the MSE. The large increases noted as segment length decreases from 50m to 25m, across all three study areas, both regression techniques (parametric versus nonparametric), and both models (natural log transform versus no transform), suggest that segment lengths greater than 25m should be utilized in these tropical forests. Based on the results presented in Table 8 and in Figure 22, of the four segment lengths considered in this study, a 50m segment is optimal from the standpoints of mitigating MSE and field effort.

Subsequent work involves the estimation of seen, unseen, and total basal area, volume, and biomass using airborne laser data and prediction equations developed using simulated forest canopy data. Prior to presenting these data, however, four tables are presented below in order to provide a context for the airborne laser data analysis. The tables document various ground reference, simulated ground reference,

Table 7. Coefficients of determination for logarithmic models, parametric analyses.

Study Site	Segment Length	R-Square			
		ln(BA,seen)	ln(BA,total)	ln(V,seen)	ln(V,total)
Tileran	100 m	0.673	0.778	0.868	0.900
	75 m	0.310	0.441	0.524	0.601
	50 m	0.358	0.510	0.556	0.647
	25 m	0.527	0.645	0.589	0.666
LaSelva-5m	100 m	0.261	0.303	0.406	0.439
	75 m	0.239	0.284	0.379	0.412
	50 m	0.389	0.422	0.522	0.548
	25 m	0.497	0.527	0.330	0.335
LaSelva-20m	50 m	0.115	0.160	0.169	0.199
	25 m	0.343	0.466	0.392	0.480

Table 8. Mean square errors for logarithmic and untransformed models, parametric and nonparametric analyses.

A. Logarithmic Model (with intercept): ln(V) = a(h)+b

Study Site	Segment Length	Parametric Analysis				Nonparametric Analysis			
		ln(BA,seen)	ln(BA,total)	ln(V,seen)	ln(V,total)	ln(BA,seen)	ln(BA,total)	ln(V,seen)	ln(V,total)
Tileran	100 m	0.027	0.015	0.015	0.010	0.028	0.016	0.015	0.011
	75 m	0.078	0.053	0.061	0.044	0.083	0.056	0.062	0.045
	50 m	0.074	0.051	0.067	0.049	0.089	0.068	0.074	0.049
	25 m	0.154	0.117	0.241	0.191	0.158	0.125	0.260	0.226
LaSelva-5m	100 m	0.281	0.242	0.381	0.341	0.284	0.245	0.382	0.341
	75 m	0.296	0.275	0.374	0.352	0.297	0.275	0.375	0.353
	50 m	0.377	0.347	0.468	0.433	0.379	0.349	0.469	0.434
	25 m	0.668	0.629	6.073	6.062	0.670	0.630	6.653	6.651
LaSelva-20m	50 m	0.056	0.054	0.070	0.066	0.057	0.055	0.075	0.066
	25 m	0.216	0.167	0.270	0.220	0.217	0.168	0.270	0.221

B. Untransformed model (thru origin): V = a(h)

Study Site	Segment Length	Parametric Analysis				Nonparametric Analysis			
		BA, seen	BA, total	V, seen	V, total	BA, seen	BA, total	V, seen	V, total
Tileran	100 m	46	38	5533	5229	69	42	6691	6170
	75 m	132	126	14444	12856	163	158	19334	16052
	50 m	139	128	16274	14611	167	144	18337	16213
	25 m	325	294	42619	38135	344	308	44734	41937
LaSelva-5m	100 m	199	202	8.548E+09	8.482E+09	200	203	9.201E+09	9.295E+09
	75 m	193	196	8.780E+09	8.744E+09	200	199	9.136E+09	9.290E+09
	50 m	206	212	1.098E+10	1.095E+10	208	212	1.174E+10	1.185E+10
	25 m	320	319	1.762E+10	1.739E+10	330	329	2.194E+10	2.238E+10
LaSelva-20m	50 m	58	66	2.068E+09	2.216E+09	78	75	2.502E+09	2.589E+09
	25 m	235	223	8.847E+09	8.436E+09	248	250	1.004E+10	9.229E+09

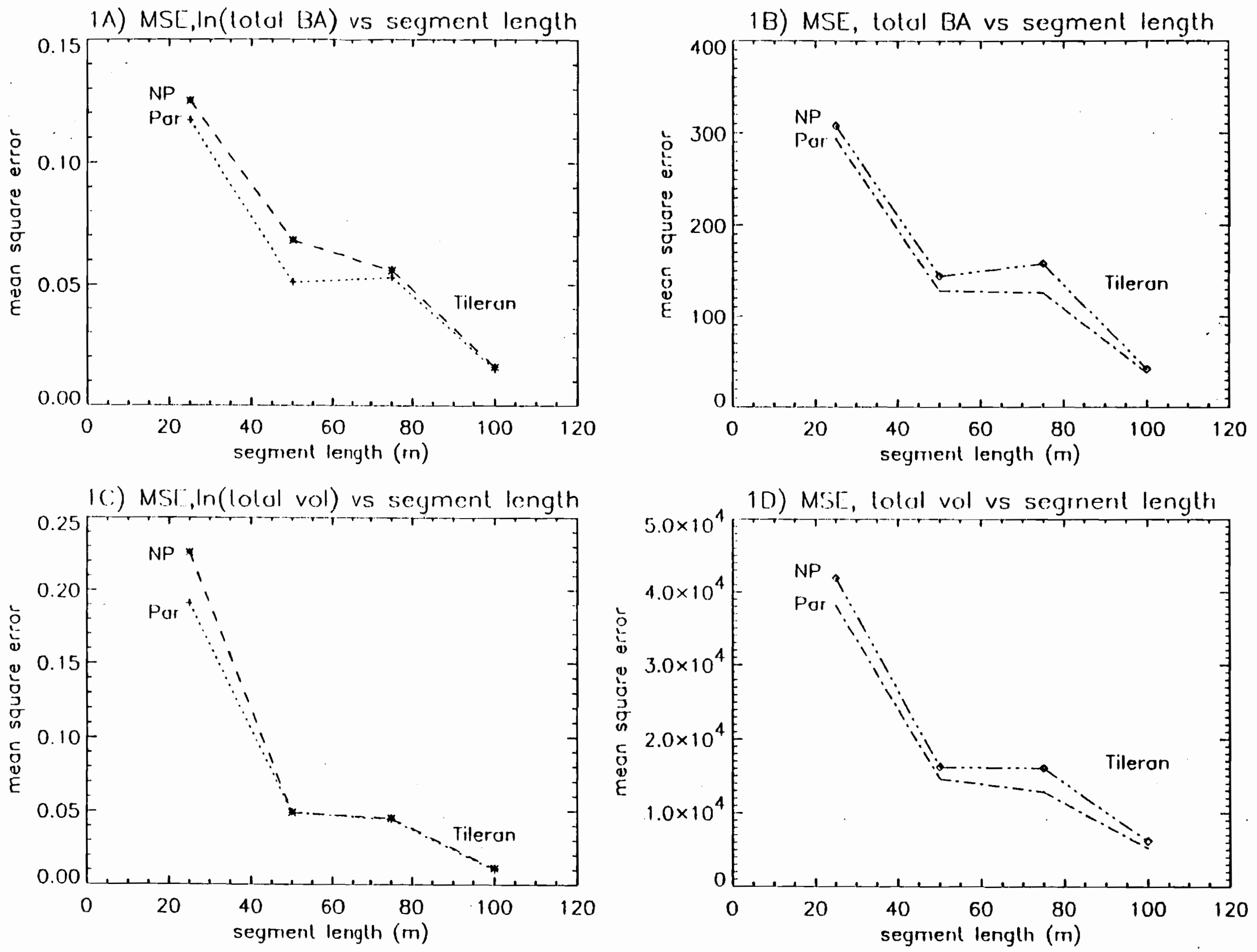
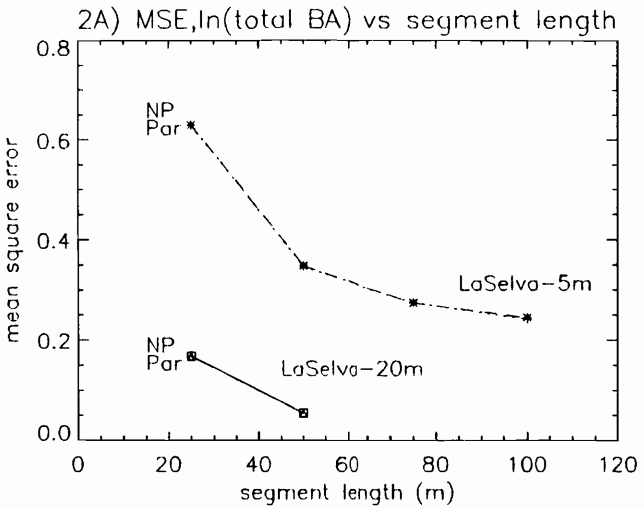
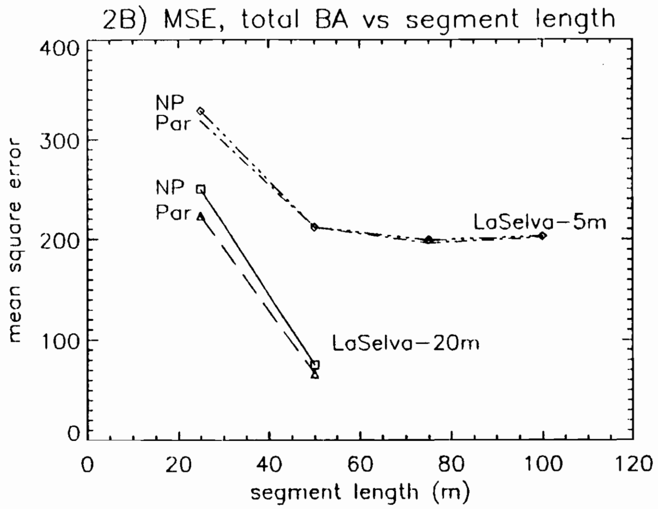
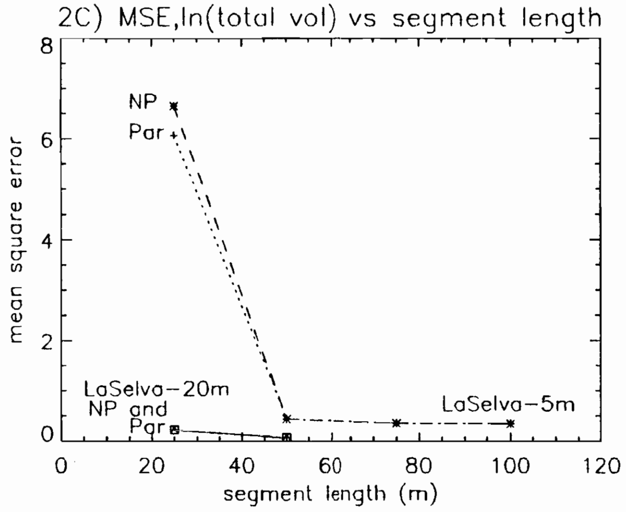
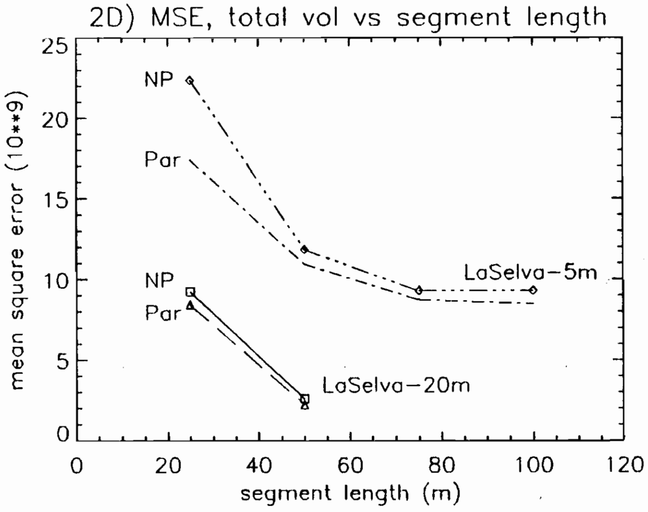


Figure 22. Mean square error for transformed and untransformed basal area, volume, and biomass versus segment length for Tileran, LaSelva-5m, and LaSelva-20m, for both parametric and nonparametric methods.


2A) MSE,ln(total BA) vs segment length


2B) MSE, total BA vs segment length


2C) MSE,ln(total vol) vs segment length


2D) MSE, total vol vs segment length

simulated laser, and airborne laser attributes. Table 9 reports the ground reference attributes for the three study areas.

Table 9. Ground reference summaries for Tileran, LaSelva-5m, and LaSelva-20m ground transect data sets.

Ground Reference	Tileran		LaSelva-5m		LaSelva-20m	
	mean	sd	mean	sd	mean	sd
Per-Tree:						
total tree ht.(m)	15.319	7.390	14.899	7.005	15.386	6.363
dbh (cm)	25.245	18.175	24.088	17.727	23.248	18.347
canopy diameter(m)	6.847	3.484	5.142	3.036	6.201	3.038
basal area(m^2)	0.079	0.191	0.070	0.149	0.069	0.171
merch.vol.(m^3)	0.741	2.138				
tot.dry biomass(kg)			402.318	1133.404	385.448	1049.405
n	265		604		345	
Per-Hectare:						
stems/ha	530.000		402.667		492.857	
basal area/ha	42.096		28.272		33.913	
merch.vol.(m^3/ha)	392.727					
tot.dry biomass(kg/ha)			162,000.1		189,970.6	
Sampling Particulars:						
total plot size(m x m)	10 x 500		5 x 3000		20 x 350	
total plot size(ha)	0.5		1.5		0.7	
No. ind. transects	1		5		2	

The LaSelva-5m results and, separately, the LaSelva-20m results were calculated by concatenating the individual, disjoint, ground transects into two large "supertransects", one for the five-5m wide units and one for the two-20m wide units. This method of calculation effectively weights each individual transect by number of stems and not by area, the thought being that the best estimates of canopy attributes will arise from a single large "superplot". Table 10 lists the simulated canopy results, i.e., the results of 1) building the forest canopy height structure in a two dimensional array, with the two dimensions representing the length and width of a ground transect; 2) determining the seen, unseen (roughly the overstory, understory) breakdown; and 3) running 30 random transects along the long axis of this height array to simulate laser transects and compiling the simulated laser responses. The summary results for the LaSelva-5m and LaSelva-20m transect sets were compiled from separate analysis of the individual ground reference transects, weighted by the lengths of the transects (as per DeVries, 1986, pgs. 255-256). The results from the simulation studies involving the individual transects (which were, in fact, used to calculate

Table 10. Simulated canopy results based on the ground transect data of Tileran, LaSelva-5m, and LaSelva-20m; summary of all transects. The LaSelva-5m and LaSelva-20m simulated transect means are calculated by weighting the individual transects by length, as per DeVries (1986). The individual transect measurements which are used to calculate these means are reported in the table which follows this one.

Simulated Transect Measurements (based on 30 random transects):

	Tileran		LaSelva-5m		LaSelva-20m	
	mean	sd	mean	sd	mean	sd
no. trees intersected (1 km fl)	167.733	9.274	149.922		220.191	
avg. height, all pulses(m)	19.854	0.512	9.941		16.830	
avg. height, canopy hits(m)	22.909	0.219	17.834		19.847	
coef. variation, all pulses	0.523	0.035	1.102		0.541	
coef. variation, canopy hits	0.322	0.013	0.443		0.309	
canopy closure(%)	86.665	2.154	55.116		84.738	
n	30		30		30	

Mapped Stand Evaluation:

	Tileran			LaSelva-5m			LaSelva-20m		
	seen	unseen	total	seen	unseen	total	seen	unseen	total
no. stems/ha	282	248	530	346.033	56.633	402.667	347.142	146.143	493.286
basal area/ha(m^2/ha)	34.62	7.48	42.10	26.871	1.399	28.271	30.767	3.141	33.914
merch.vol/ha(m^3/ha)	342.00	50.73	392.72						
total dry biomass(kg/ha)				156,679.6	5,317.6	161,997.2	177,470.9	12,498.9	189,969.8

the Table 10 LaSelva results) are presented in Table 11. The results presented in Tables 10 and 11 are based solely on the simulations of the ground reference transect data; the only assumptions required to generate these results follow:

1. Canopy shapes are circular when viewed from above.
2. The canopy shape is elliptic when viewed from the side, with one end of the major axis defined by the total tree height, the crown diameter defining the length of the minor axis, and the half-ellipsoid terminating at the height of the first branch.
3. Canopy roughness is distributed N(0,0.117 m^2).

The airborne laser data measurements comparable to those listed in Tables 10 and 11 are provided in Table 12. Remember that these airborne measurements were acquired in the same forest cover type, but not necessarily in close proximity to the ground transects. A few observations can be made immediately by comparing Tables 9, 10, and 12.

1. Table 9 suggests that the three study areas are similar in terms of individual tree height characteristics, including total tree height, dbh, and canopy diameter. However, higher stem counts in Tileran push per hectare estimates of basal area higher on this western site. Simulated tree canopy data (Table 10) illustrate that approximately half of the trees at Tileran are suppressed (i.e., unseen from above).
2. Simulated laser transects document study area dissimilarities. Suprisingly, the LaSelva-5m transect data stands as the oddball, relatively open (55% canopy closure), with few unseen stems. The low canopy density and low overall height connote a stand where the larger overstory trees are underrepresented.
3. A comparison of average heights, canopy densities, and coefficients of variation (CV) in Tables 10 and 12 document large disparities between ground measurements and airborne laser measurements. In Tileran, the airborne laser measured a canopy two-thirds the height of the simulated stand, with similar canopy closures and CVs. Conversely, the airborne laser measuring the forest canopy in the general vicinity of LaSelva-5m ground transects encountered a canopy taller and much denser (55% vs 95% canopy closure) than simulated measures. The LaSelva-5m ground - laser height differences are illustrated by comparing Figure 9A,B with Figure 14C. LaSelva-20m airborne and simulated laser measures more closely agree; heights and canopy density vary only 10-20 percent.

The differences between the simulated and airborne laser measurements are troubling. Which data sets are correct (i.e., most accurately describe actual ground conditions) and what are the reasons behind the discrepancies? Certainly, natural variation in tree heights and forest structure would account for some of the differences. However, considering the LaSelva-5m transect data, 3 km of 5m wide mapped stand data were collected and compared to approximately 4 km of airborne laser data. Intuitively, enough data were collected on the ground and from the air to adequately represent forest canopy conditions at LaSelva and to mitigate the effects of natural variation. The 5m wide transect data characterizes a forest 10m tall, with a canopy closure of 55%. Comparable airborne measurements describe a forest 23m tall and 95%

Table 11. Simulated canopy results based on the ground transect data of LaSelva-5m and LaSelva-20m; individual transect summaries are reported.

	Simulated Transect Measurements (n = 30 transects)							Mapped Stand Evaluation (per hectare)		
site	no.trees intersected	avg.ht. all pulses (m)	avg.ht. canopy hits (m)	coef.var. all pulses	coef.var. canopy hits	canopy closure (%)	crown position	no.stems /ha (#/ha)	basal area/ha (m^2/ha)	above-grnd dry biom. (kg/ha)
LaSelva-5m:										
Seg.A	131.778	15.721	22.347	0.796	0.386	70.353	seen	247	24.20	166,751.7
							unseen	53	1.10	4,274.1
							total	300	25.30	171,025.8
Seg.B	199.667	10.767	16.758	0.805	0.242	64.255	seen	440	12.28	60,825.6
							unseen	140	2.12	9,020.2
							total	580	14.40	69,845.8
Seg.C	133.917	12.130	21.462	0.981	0.330	56.523	seen	345	27.65	177,966.9
							unseen	75	2.23	11,198.4
							total	420	29.88	189,165.4
AL1-1	171.167	9.768	16.988	1.050	0.456	57.507	seen	352	23.81	130,433.6
							unseen	36	0.87	3,075.5
							total	388	24.68	133,509.1
AL1-2	125.600	7.258	16.199	1.384	0.553	44.848	seen	342	34.80	200,145.5
							unseen	46	1.47	4,409.6
							total	388	36.27	204,555.1
LaSelva-20m:										
Tr1	252.889	15.996	18.829	0.540	0.311	84.917	seen	350	34.47	194,550.1
							unseen	177	4.45	17,167.0
							total	527	38.92	211,717.1
Tr2	195.667	17.455	20.610	0.542	0.308	84.604	seen	345	27.99	164,661.6
							unseen	123	2.16	8,997.9
							total	468	30.16	173,659.4

Table 12. Airborne laser measurement summaries for those flightline segments thought to be closest to the ground transects. All measurements in this table are averages, either over an entire flightline or over a 50m segment of a flightline. No gaps were left between the 50m segments. The range of those 50m segment averages are reported as a method of exemplifying the measurement variation induced by subdividing the airborne laser transects. The Tileran airborne measurements are compiled from file CRT1F2, 3.333 to 8.465 seconds, approximately 500m of flightline. The LaSelva-5m airborne measurements are compiled from file CRT2C, 3.609 to 43.243 seconds, approximately 4km of flightline. The LaSelva-20m airborne measurements are compiled from file CRT2F1, 0.86 to 9.3 seconds, approximately 850m of flightline.

	Tileran			LaSelva-5m			LaSelva-20m		
		50m segment			50m segment			50m segment	
	entire fl.	low	high	entire fl.	low	high	entire fl.	low	high
height, all pulses(m)	12.185	5.780	18.858	22.901	5.525	36.243	20.548	5.750	31.823
height, canopy hits(m)	14.263	7.786	19.051	24.128	12.455	42.225	21.954	9.154	31.823
coef.var., all pulses	0.582	0.231	0.768	0.421	0.086	1.332	0.446	0.095	0.936
coef.var., canopy hits	0.379	0.193	0.426	0.342	0.086	0.506	0.349	0.095	0.538
canopy closure(%)	85.425	72.589	98.985	94.912	42.132	100.0	93.598	62.814	100.0

closed. (Note: This 95% canopy density figure was developed assuming that canopy heights less than 5.0m were laser pulses, not to trees, but to shrubs or ground.) A trip was made to the LaSelva Biological Research Station near Puerto Viejo, Costa Rica in September, 1993, nine years after the laser data were collected and eight years after collection of the ground transect data. Point measurements were made along lines roughly parallel to and in the proximity of the airborne laser profiles and 5m ground transects. At 25m intervals along a particular compass bearing, canopy density measurements were made using a spherical densiometer, and spatial and biometric data were collected on any tree whose canopy was intersected by a line vertically projected from the sample point. At a given point, for each tree whose crown was vertically intercepted, the following measurements were taken: 1) crown position (dominant, codominant, intermediate, suppressed); 2) distance along and perpendicular to the sample transect (i.e., the compass bearing); 3) total tree height; 4) dbh; and 5) shrub height in the general vicinity of the sample point. A total of 56 points were visited, half in the general vicinity of the northernmost flightline (AL1-2), half along the southern flightline (AL1-1). The results are reported in Table 13.

A number of items may be noted in Table 13:

1. The average perpendicular distance from the sample transect to a dominant or codominant tree (i.e., the tree whose canopy would intercept a given airborne laser pulse) is over 3.6m. A ± 2.5*m* wide ground sample transect will miss many of the trees which would interact with an airborne laser and which account for the canopy density over any particular point in the forest. It may be noted in field observations that 37 of 60 dominiant or codominant trees were greater than 2.5m from the sample transect line. Then, at LaSelva, a 5m wide transect will miss 62% of the trees which an airborne laser would sense if it traversed the same line.

2. Canopy densities ranged from 83.3% to 98.4%, with a mean of 94.5%. Only 4 of 56 points exhibited crown closures less than 90%, as measured using a spherical densiometer. The 1993 average canopy density as measured using a spherical densiometer agrees within one percent with the average canopy densities reported by the airborne laser at LaSelva. The simulated value developed using the ground transect data, 55%, is unrealistic.

3. The tallest tree found in this 56 point sample was 48.2m as measured using a Suunto hypsometer; the average dominant/codominant tree height was 32.6m. The average height of all trees tallied on these points 21.8m. This average height agrees closely with the comparable airborne laser measures, within 1.5m. The simulated height values based on the 5m wide transect data are again unrealistic. The sampled values are less than half of the 1993 ground measurements.

The LaSelva canopy, a tropical wet <u>primary</u> (read undisturbed) forest, appears to be all-aged, multi-level. If one assumes little overall change over the 10 year period between the mid-80's sampling work and the 1993 junket, then it is clear that the LaSelva-5m transect data is not useful for canopy simulation studies. The ground sampling methodology misses too many of the trees important for simu-

Table 13. Summary of point measurement data collected during a 1993 field trip to the LaSelva Biological Research Station. D/CD stands for dominant/codominant trees.

	AL1-1			AL1-2			All Points		
	mean	stan.dev.	n	mean	stan.dev.	n	mean	stan.dev.	n
avg. perpendicular distance from D/CD trees to transect (m)	3.76	2.44	31	3.58	2.76	29	3.67	2.58	60
avg.height D/CD trees (m)	31.96	6.34	31	33.26	8.58	29	32.58	7.47	60
avg.canopy density (%)	94.79	3.49	28	94.15	2.74	28	94.47	3.13	56
avg.height,all trees on pts.(m)	23.21	11.03	63	20.66	12.34	74	21.83	11.78	137
largest tree (m)	44.41			48.19			48.19		
smallest D/CD tree (m)	22.13			20.02			20.02		
smallest tree (m)	4.57			4.57			4.57		

lating canopy height structure. The fixed-area plot information is useful for developing areal estimates of the ground metrics of interest (eg., stems, basal area, and volume or biomass per hectare), but it does not accurately represent the primary tropical forest spatially. For comparative purposes however, the 5m transect data was subsequently utilized throughout the investigation for simulation studies. In other words, although the 5m wide transect data do not adequately characterize ground conditions sensed by an airborne laser, the data may be used to discern such things as effects of segment length and canopy shape on estimated basal area and biomass. Subsequent results using the 5m data should be evaluated only while considering that the 5m wide data set does not accurately characterize the spatial distribution of tree heights. Unfortunately, time in the field was constrained; a similar study could not be conducted at Tileran. The cause of the discrepancy between laser and ground height measurements at Tileran is unclear.

Lieberman et al. (1990) provide independent estimates of stem density, basal area, and total above-ground biomass per hectare for 3 - 4 hectare permanent plots established at LaSelva in 1969. These plots were remeasured in 1982; at that time stem density (greater than 10 cm dbh) was 446.0 boles/ha, basal area was 27.82 m^2/ha, and total above-ground dry biomass was 248.7 metric tons per hectare. They describe a stand with a high rate of turnover; one where the stand half life (that period of time within which 50% of the individuals tallied at a particular time are expected to die) is on the order of 30-34 years. However periodic inventories indicate an "overall stability in stand structure in spite of rapid turnover and high rates of mortality." Comparable values for data collected in 1985 on 5m wide transects (see Table 10) are 402.7 stems/ha, 28.27 m^2/ha, and 170.0 mt/ha. Comparable values for data collected in the late 1980's on 20m wide LaSelva transects (see Table 10) are 493.3 stems/ha, 33.91 m^2/ha, and 190.0 mt/ha. Though basal areas are comparable or higher in both transect data sets, biomass estimates are notably lower than the Lieberman et al. (1990) figure. The difference is due, in part, to the biomass equations used to calculate total above-ground dry biomass. Lieberman et al. (1990) used the following equation to generate per tree estimates:

$$v = 0.89b - 136$$

where v = total above-ground dry biomass in kg;

and b = basal area, in cm^2 (Lieberman and Lieberman, 1990);

The following generic biomass equation was applied to the majority of the trees in the LaSelva transect data:

$$\ln v = 0.94949 \quad \ln(d_b^2 h_t) - 3.36815$$

where d_b = dbh in cm;
and h_t = total tree height, in m (Malavassi, 1987, personal communication with Dr. A. Joyce).

The Malavassi equation used to generate biomass estimates in this study tends to underestimate dry metric tonnage of all trees except the smallest relative to the Lieberman and Lieberman (1990) equation. This finding does not explain any portion of discrepancies noted between the ground reference data and airborne laser estimates. However, this intercomparison does suggest that 1) there may be appreciable variation in tropical forest biomass estimates due solely to the allometric equations used to estimate biomass; and 2) the single comparison done in this study suggests that the LaSelva-5m and LaSelva-20m biomass estimates may be low relative to actual ground conditions as reported by Lieberman and Lieberman (1990).

Table 14 reports estimates of basal area, merchantable volume, and above-ground dry biomass developed using predictive equations which use average canopy height, all pulses ($\bar{h}_a$) as the sole independent variable. To produce the results in Table 14, airborne laser measurements of $\bar{h}_a$ were used in equations developed using simulated laser measurements of a forest canopy simulated from ground reference transect data. The table is lengthy, but the comprehensive results are provided in order 1) to compare and contrast parametric versus nonparametric results; 2) to compare transformed and untransformed model results; 3) to discern the effects of changing segment length; and 4) to see if there is any consistent difference between estimates developed from the entire flightline versus airborne flightline segments, with and without gaps. The purpose of this portion of the research is to identify the treatment combination(s) which produce(s) the most accurate estimates of basal area, volume, and biomass. All treatment combinations utilize the same single, independent variable, $\bar{h}_a$, whose predictive utility is well-founded in the scientific literature.

Table 14. Comprehensive listing of seen, unseen, and total basal area, merchantable volume, and total above-ground dry biomass results for Tileran, LaSelva-5m, and the LaSelva-20m data sets. These estimates are generated using airborne laser data and various equations which all employ average canopy height, all pulses ($\bar{h}_a$) as the single independent airborne laser variable.

A) Tileran, 1) parametric estimates, a) natural log model (with intercept), for 1) 100m, 2) 75m, 3) 50m, and 4) 25m segments.
A) Tileran, 1) parametric estimates, b) no transformation, SLR through origin, for 1) 100m, 2) 75m, 3) 50m, and 4) 25m segments.
A) Tileran, 2) nonparametric estimates, a) natural log model (with intercept), for 1) 100m, 2) 75m, 3) 50m, and 4) 25m segments.
A) Tileran, 2) nonparametric estimates, b) no transformation, SLR through origin, for 1) 100m, 2) 75m, 3) 50m, and 4) 25m segments.
B) LaSelva-5m, 1) parametric estimates, a) natural log model (with intercept), for 1) 100m, 2) 75m, 3) 50m, and 4) 25m segments.
B) LaSelva-5m, 1) parametric estimates, b) no transformation, SLR through origin, for 1) 100m, 2) 75m, 3) 50m, and 4) 25m segments.
B) LaSelva-5m, 2) nonparametric estimates, a) natural log model (with intercept), for 1) 100m, 2) 75m, 3) 50m, and 4) 25m segments.
B) LaSelva-5m, 2) nonparametric estimates, b) no transformation, SLR through origin, for 1) 100m, 2) 75m, 3) 50m, and 4) 25m segments.
C) LaSelva-20m, 1) parametric estimates, a) natural log model (with intercept), for 1) 100m, 2) 75m, 3) 50m, and 4) 25m segments.
C) LaSelva-20m, 1) parametric estimates, b) no transformation, SLR through origin, for 1) 100m, 2) 75m, 3) 50m, and 4) 25m segments.
C) LaSelva-20m, 2) nonparametric estimates, a) natural log model (with intercept), for 1) 100m, 2) 75m, 3) 50m, and 4) 25m segments.
C) LaSelva-20m, 2) nonparametric estimates, b) no transformation, SLR through origin, for 1) 100m, 2) 75m, 3) 50m, and 4) 25m segments.

For each of the treatment combinations listed above, the airborne laser data is analyzed 1) as an entire flightline; 2) broken into segments which are the same length as the ground transect data, with a 15m gap between segments; and 3) broken into segments which are the same length as the ground transect data without a gap between segments.

A. Tileran:

1. Parametric:

a. transformation (with intercept): model: ln(V)=a(h)+b

1. 100m segments

		Airborne Laser						
	Simulated	100m equation						
	Canopy	entire fl	100m seg/15m gap			100m seg/no gap		
	mean	mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)	19.854	12.185	12.192	0.838	1787	12.185	0.859	2024
basal area,seen (sq.m/ha)	34.620	17.808	17.983	1.204	5	18.022	1.216	6
basal area,unseen(sq.m/ha)	7.480	4.049	4.083	0.254	5	4.091	0.257	6
basal area,total(sq.m/ha)	42.100	21.857	22.066	1.458	5	22.113	1.474	6
mer.volume,seen(cu.m/ha)	342.000	140.402	142.820	12.710	5	143.396	12.783	6
mer.volume,unseen(cu.m/ha)	50.730	27.007	27.248	1.818	5	27.300	1.840	6
mer.volume,total(cu.m/ha)	392.720	167.410	170.068	14.527	5	170.696	14.621	6

2. 75m segments

		Airborne Laser						
		75m equation						
		entire fl	75m seg/15m gap			75m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		12.185	12.291	1.272	1728	12.186	0.972	2024
basal area,seen (sq.m/ha)		19.528	20.042	1.771	6	19.782	1.285	7
basal area,unseen(sq.m/ha)		3.204	3.381	0.479	6	30298.000	0.340	7
basal area,total(sq.m/ha)		22.733	23.423	2.250	6	23.080	1.624	7
mer.volume,seen(cu.m/ha)		159.316	166.809	20.498	6	163.265	14.559	7
mer.volume,unseen(cu.m/ha)		26.091	27.156	3.090	6	26.645	2.208	7
mer.volume.total(cu.m/ha)		185.407	193.965	23.588	6	189.910	16.767	7

3. 50m segments

		Airborne Laser						
		50m equation						
		entire fl	50m seg/15m gap			50m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		12.185	12.818	1.542	1577	12.186	1.175	2024
basal area,seen (sq.m/ha)		20.275	21.700	1.990	8	20.794	1.497	11
basal area,unseen(sq.m/ha)		3.164	3.732	0.654	8	3.421	0.488	11
basal area,total(sq.m/ha)		23.439	25.432	2.643	8	24.215	1.985	11
mer.volume,seen(cu.m/ha)		162.853	182.619	23.734	8	171.453	17.753	11
mer.volume,unseen(cu.m/ha)		21.983	25.924	4.294	8	23.859	3.198	11
mer.volume,total(cu.m/ha)		184.836	208.543	28.027	8	195.312	20.949	11

4. 25m segments

		Airborne Laser						
		25m equation						
		entire fl	25m seg/15m gap			25m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		12.185	11.726	1.128	1283	12.186	1.027	2024
basal area,seen (sq.m/ha)		16.004	16.193	1.520	13	17.024	1.364	21
basal area,unseen(sq.m/ha)		2.228	2.396	0.446	13	2.665	0.398	21
basal area,total(sq.m/ha)		18.231	18.590	1.965	13	19.689	1.761	21
mer.volume,seen(cu.m/ha)		115.687	121.172	16.693	13	131.017	14.927	21
mer.volume,unseen(cu.m/ha)		13.875	15.649	3.270	13	17.709	2.919	21
mer.volume,total(cu.m/ha)		129.562	136.818	19.958	13	148.726	17.841	21

A. Tileran:

1. Parametric:

b. no transformation (thru origin): model: V=a(h)

1. 100m segments

		Airborne Laser						
	Simulated	100m equation						
	Canopy	entire fl	100m seg/15m gap			100m seg/no gap		
	mean	mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)	19.854	12.185	12.192	0.838	1787	12.185	0.859	2024
basal area,seen (sq.m/ha)	34.620	21.520	21.534	1.480	5	21.521	1.516	6
basal area,unseen(sq.m/ha)	7.480	4.445	4.448	0.306	5	4.445	0.313	6
basal area,total(sq.m/ha)	42.100	25.965	25.981	1.785	5	25.966	1.830	6
mer.volume,seen(cu.m/ha)	342.000	214.751	214.888	14.766	5	214.765	15.133	6
mer.volume,unseen(cu.m/ha)	50.730	30.303	30.323	2.084	5	30.305	2.135	6
mer.volume,total(cu.m/ha)	392.720	245.054	245.211	16.850	5	245.071	17.268	6

2. 75m segments

		Airborne Laser						
		75m equation						
		entire fl	75m seg/15m gap			75m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		12.185	12.291	1.272	1728	12.186	0.972	2024
basal area,seen (sq.m/ha)		21.653	21.841	2.260	6	21.655	1.728	7
basal area,unseen(sq.m/ha)		4.332	4.369	0.452	6	4.332	0.346	7
basal area,total(sq.m/ha)		25.984	26.211	2.713	6	25.987	2.073	7
mer.volume,seen(cu.m/ha)		219.094	221.001	22.872	6	219.112	17.480	7
mer.volume,unseen(cu.m/ha)		29.959	30.220	3.128	6	29.961	2.390	7
mer.volume,total(cu.m/ha)		249.053	251.221	25.999	6	249.074	19.870	7

3. 50m segments

		Airborne Laser						
		50m equation						
		entire fl	50m seg/15m gap			50m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		12.185	12.818	1.542	1577	12.186	1.175	2024
basal area,seen (sq.m/ha)		21.106	22.204	2.671	8	21.108	2.036	11
basal area,unseen(sq.m/ha)		4.586	4.824	0.580	8	4.586	0.442	11
basal area,total(sq.m/ha)		25.692	27.028	3.252	8	25.694	2.479	11
mer.volume,seen(cu.m/ha)		212.098	223.128	26.843	8	212.117	20.462	11
mer.volume,unseen(cu.m/ha)		30.999	32.611	3.923	8	31.002	2.991	11
mer.volume,total(cu.m/ha)		243.097	255.739	30.766	8	243.118	23.452	11

4. 25m segments

		Airborne Laser						
		25m equation						
		entire fl	25m seg/15m gap			25m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		12.185	11.726	1.128	1283	12.186	1.027	2024
basal area,seen (sq.m/ha)		20.382	19.615	1.887	13	20.383	1.717	21
basal area,unseen(sq.m/ha)		4.244	4.084	0.393	13	4.244	0.358	21
basal area,total(sq.m/ha)		24.626	23.700	2.279	13	24.628	2.075	21
mer.volume,seen(cu.m/ha)		204.700	197.001	18.948	13	204.717	17.248	21
mer.volume,unseen(cu.m/ha)		29.536	28.426	2.734	13	29.539	2.489	21
mer.volume,total(cu.m/ha)		234.236	225.427	21.682	13	234.256	19.737	21

A. Tileran:

2. Nonparametric

a. transformation (with intercept): model: ln(V)=a(h)+b

1. 100m segments

		Airborne Laser							
	Simulated	100m equation							
	Canopy	entire fl	100m seg/15m gap			100m seg/no gap			
	mean	mean	mean	stan. err.	n	mean	stan. err.	n	
height,all pulses(m)	19.854	12.185	12.192	0.838	1787	12.185	0.859	2024
basal area,seen (sq.m/ha)	34.620	15.932	16.144	1.262	5	16.193	1.272	6
basal area,unseen(sq.m/ha)	7.480	4.224	4.261	0.275	5	4.269	0.278	6
basal area,total(sq.m/ha)	42.100	20.155	20.405	1.537	5	20.462	1.550	6
mer.volume,seen(cu.m/ha)	342.000	146.038	148.358	12.679	5	148.907	12.761	6
mer.volume,unseen(cu.m/ha)	50.730	6.834	7.515	2.147	5	7.689	2.129	6
mer.volume,total(cu.m/ha)	392.720	152.871	155.873	14.819	5	156.596	14.885	6

2. 75m segments

		Airborne Laser						
		75m equation						
		entire fl	75m seg/15m gap			75m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		12.185	12.291	1.272	1728	12.186	0.972	2024
basal area,seen (sq.m/ha)		23.849	24.142	1.353	6	23.976	1.000	7
basal area,unseen(sq.m/ha)		2.898	3.099	0.603	6	3.002	0.432	7
basal area,total(sq.m/ha)		26.747	27.240	1.955	6	26.977	1.432	7
mer.volume,seen(cu.m/ha)		174.355	180.764	19.354	6	177.651	13.881	7
mer.volume,unseen(cu.m/ha)		25.670	27.270	4.018	6	26.530	2.828	7
mer.volume,total(cu.m/ha)		200.024	208.034	23.372	6	204.181	16.708	7

3. 50m segments

		Airborne Laser						
		50m equation						
		entire fl	50m seg/15m gap			50m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		12.185	12.818	1.542	1577	12.186	1.175	2024
basal area,seen (sq.m/ha)		24.195	24.814	1.102	8	24.337	0.835	11
basal area,unseen(sq.m/ha)		6.409	6.618	0.356	8	6.463	0.269	11
basal area,total(sq.m/ha)		30.604	31.433	1.458	8	30.800	1.104	11
mer.volume,seen(cu.m/ha)		174.293	188.538	20.272	8	179.489	15.200	11
mer.volume,unseen(cu.m/ha)		11.656	20.779	7.662	8	16.728	5.690	11
mer.volume,total(cu.m/ha)		185.949	209.317	27.785	8	196.216	20.774	11

4. 25m segments

		Airborne Laser						
		25m equation						
		entire fl	25m seg/15m gap			25m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		12.185	11.726	1.128	1283	12.186	1.027	2024
basal area,seen (sq.m/ha)		15.562	15.697	1.386	13	16.445	1.244	21
basal area,unseen(sq.m/ha)		4.202	4.209	0.316	13	4.374	0.284	21
basal area,total(sq.m/ha)		19.764	19.905	1.702	13	20.819	1.528	21
mer.volume,seen(cu.m/ha)		141.077	143.531	14.709	13	151.707	13.185	21
mer.volume,unseen(cu.m/ha)		28.342	28.414	2.272	13	29.608	2.042	21
mer.volume,total(cu.m/ha)		169.419	171.945	16.979	13	181.315	15.225	21

A. Tileran:

2. Nonparametric

b. no transformation (thru origin): model: V=a(h)

1. 100m segments

	Simulated Canopy	Airborne Laser						
		100m equation						
		entire fl	100m seg/15m gap			100m seg/no gap		
	mean	mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)	19.854	12.185	12.192	0.838	1787	12.185	0.859	2024
basal area,seen (sq.m/ha)	34.620	18.596	18.608	1.279	5	18.597	1.310	6
basal area,unseen(sq.m/ha)	7.480	6.223	6.227	0.428	5	6.224	0.439	6
basal area,total(sq.m/ha)	42.100	24.819	24.835	1.707	5	24.821	1.749	6
mer.volume,seen(cu.m/ha)	342.000	194.099	194.223	13.346	5	194.112	13.678	6
mer.volume,unseen(cu.m/ha)	50.730	32.334	32.355	2.223	5	32.336	2.279	6
mer.volume,total(cu.m/ha)	392.720	226.433	226.578	15.570	5	226.449	15.956	6

2. 75m segments

		Airborne Laser						
		75m equation						
		entire fl	75m seg/15m gap			75m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		12.185	12.291	1.272	1728	12.186	0.972	2024
basal area,seen (sq.m/ha)		18.266	18.425	1.907	6	18.268	1.457	7
basal area,unseen(sq.m/ha)		4.239	4.276	0.443	6	4.240	0.338	7
basal area,total(sq.m/ha)		22.506	22.702	2.349	6	22.508	1.796	7
mer.volume,seen(cu.m/ha)		176.386	177.922	18.414	6	176.401	14.073	7
mer.volume,unseen(cu.m/ha)		159.984	161.377	16.701	6	159.997	12.764	7
mer.volume,total(cu.m/ha)		336.370	339.298	35.115	6	336.399	26.837	7

3. 50m segments

		Airborne Laser						
		50m equation						
		entire fl	50m seg/15m gap			50m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		12.185	12.818	1.542	1577	12.186	1.175	2024
basal area,seen (sq.m/ha)		17.873	18.803	2.262	8	17.875	1.724	11
basal area,unseen(sq.m/ha)		5.384	5.664	0.681	8	5.385	0.519	11
basal area,total(sq.m/ha)		23.258	24.467	2.943	8	23.260	2.244	11
mer.volume,seen(cu.m/ha)		184.285	193.869	23.323	8	184.301	17.779	11
mer.volume,unseen(cu.m/ha)		34.306	36.090	4.342	8	34.309	3.310	11
mer.volume,total(cu.m/ha)		218.591	229.959	27.665	8	218.610	21.088	11

4. 25m segments

		Airborne Laser						
		25m equation						
		entire fl	25m seg/15m gap			25m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		12.185	11.726	1.128	1283	12.186	1.027	2024
basal area,seen (sq.m/ha)		17.784	17.116	1.646	13	17.786	1.499	21
basal area,unseen(sq.m/ha)		4.605	4.432	0.426	13	4.605	0.388	21
basal area,total(sq.m/ha)		22.389	21.547	2.072	13	22.391	1.887	21
mer.volume,seen(cu.m/ha)		177.411	170.739	16.422	13	177.426	14.949	21
mer.volume,unseen(cu.m/ha)		20.240	19.479	1.873	13	20.242	1.705	21
mer.volume,total(cu.m/ha)		197.651	190.218	18.295	13	197.668	16.654	21

B. LaSelva-5m:

1. Parametric:

a. transformation (with intercept): model: ln(V)=a(h)+b

1. 100m segments

		Airborne Laser						
	Simulated	100m equation						
	Canopy	entire fl	100m seg/15m gap			100m seg/no gap		
	mean	mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)	9.940	22.901	22.958	0.628	13687	22.919	0.614	15703
basal area,seen (sq.m/ha)	26.870	53.971	55.912	2.507	35	55.900	2.423	40
basal area,unseen(sq.m/ha)	1.400	4.641	4.937	0.315	35	4.940	0.306	40
basal area,total(sq.m/ha)	28.270	58.612	60.849	2.822	35	60.840	2.728	40
dry biomass,seen(kg/ha)	156,680	493,708	540,217	42,405	35	541,203	41,729	40
dry biomass,unseen(kg/ha)	5,318	30,027	34,152	3,395	35	34,272	3,382	40
dry biomass,total(kg/ha)	161,997	523,735	574,369	45,795	35	575,476	45,104	40

2. 75m segments

		Airborne Laser						
		75m equation						
		entire fl	75m seg/15m gap			75m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		22.901	22.983	0.646	13077	22.919	0.614	15703
basal area,seen (sq.m/ha)		51.258	53.495	2.160	44	53.571	2.239	53
basal area,unseen(sq.m/ha)		7.074	7.786	0.513	44	7.847	0.556	53
basal area,total(sq.m/ha)		58.332	61.281	2.672	44	61.418	2.793	53
dry biomass,seen(kg/ha)		448,099	493,950	31,046	44	498,899	34,298	53
dry biomass,unseen(kg/ha)		44,858	53,530	4,796	44	54,869	5,618	53
dry biomass,total(kg/ha)		492,957	547,480	35,826	44	553,768	39,893	53

3. 50m segments

		Airborne Laser						
		50m equation						
		entire fl	50m seg/15m gap			50m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		22.901	23.027	0.667	12085	22.920	0.583	15703
basal area,seen (sq.m/ha)		62.172	69.704	4.417	61	68.823	3.542	80
basal area,unseen(sq.m/ha)		5.438	6.557	0.571	61	6.431	0.447	80
basal area,total(sq.m/ha)		67.610	76.261	4.986	61	75.254	3.987	80
dry biomass,seen(kg/ha)		530,081	672,922	67,837	61	655,948	51,764	80
dry biomass,unseen(kg/ha)		34,487	48,302	6,139	61	46,585	4,569	80
dry biomass,total(kg/ha)		564,568	721,224	73,963	61	702,535	56,321	80

4. 25m segments

		Airborne Laser						
		25m equation						
		entire fl	25m seg/15m gap			25m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		22.901	22.581	0.654	9812	22.920	0.502	15703
basal area,seen (sq.m/ha)		70.976	92.890	8.658	99	93.956	6.465	159
basal area,unseen(sq.m/ha)		7.668	11.780	1.449	99	11.839	1.085	159
basal area,total(sq.m/ha)		78.644	104.670	10.102	99	105.796	7.545	159
dry biomass,seen(kg/ha)		1,324,632	4,893,504	1,239,791	99	4,697,695	982,517	159
dry biomass,unseen(kg/ha)		104,136	519,952	150,022	99	494,227	121,006	159
dry biomass,total(kg/ha)		1,428,768	5,413,464	1,389,574	99	5,191,863	1,103,292	159

B. LaSelva-5m:

1. Parametric:

b. no transformation (thru origin): model: V=a(h)

1. 100m segments

		Airborne Laser						
	Simulated Canopy	100m equation						
		entire fl	100m seg/15m gap			100m seg/no gap		
	mean	mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)	9.940	22.901	22.958	0.628	13687	22.919	0.614	15703
basal area,seen (sq.m/ha)	26.870	56.438	56.579	1.549	35	56.484	1.514	40
basal area,unseen(sq.m/ha)	1.400	2.884	2.891	0.079	35	2.886	0.077	40
basal area,total(sq.m/ha)	28.270	59.321	59.470	1.628	35	59.369	1.591	40
dry biomass,seen(kg/ha)	156,680	350,266	351,142	9,613	35	350,548	9,395	40
dry biomass,unseen(kg/ha)	5,318	11,794	11,823	324	35	11,803	316	40
dry biomass,total(kg/ha)	161,997	362,060	362,966	9,937	35	362,351	9,711	40

2. 75m segments

		Airborne Laser						
		75m equation						
		entire fl	75m seg/15m gap			75m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		22.901	22.983	0.646	13077	22.919	0.614	15703
basal area,seen (sq.m/ha)		55.138	55.337	1.555	44	55.183	1.479	53
basal area,unseen(sq.m/ha)		3.451	3.463	0.097	44	3.454	0.093	53
basal area,total(sq.m/ha)		58.589	58.800	1.652	44	58.637	1.571	53
dry biomass,seen(kg/ha)		339,479	340,702	9,573	44	339,756	9,104	53
dry biomass,unseen(kg/ha)		13,832	13,882	390	44	13,843	371	53
dry biomass,total(kg/ha)		353,311	354,584	9,963	44	353,600	9,475	53

3. 50m segments

		Airborne Laser						
		50m equation						
		entire fl	50m seg/15m gap			50m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		22.901	23.027	0.667	12085	22.920	0.583	15703
basal area,seen (sq.m/ha)		55.764	56.071	1.625	61	55.811	1.420	80
basal area,unseen(sq.m/ha)		2.945	2.961	0.086	61	2.947	0.075	80
basal area,total(sq.m/ha)		58.709	59.032	1.711	61	58.758	1.495	80
dry biomass,seen(kg/ha)		354,250	356,202	10,323	61	354,544	9,019	80
dry biomass,unseen(kg/ha)		12,275	12,342	358	61	12,285	312	80
dry biomass,total(kg/ha)		366,525	368,543	10,681	61	366,828	9,331	80

4. 25m segments

		Airborne Laser						
		25m equation						
		entire fl	25m seg/15m gap			25m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		22.901	22.581	0.654	9812	22.920	0.502	15703
basal area,seen (sq.m/ha)		57.751	56.946	1.648	99	57.799	1.265	159
basal area,unseen(sq.m/ha)		3.203	3.159	0.091	99	3.206	0.070	159
basal area,total(sq.m/ha)		60.954	60.105	1.740	99	61.005	1.335	159
dry biomass,seen(kg/ha)		384,617	379,258	10,977	99	384,934	8,425	159
dry biomass,unseen(kg/ha)		13,511	13,323	386	99	13,523	296	159
dry biomass,total(kg/ha)		398,128	392,581	11,362	99	398,457	8,721	159

B. LaSelva-5m:

2. Nonparametric

a. transformation (with intercept): model: ln(V)=a(h)+b

1. 100m segments

		Airborne Laser						
	Simulated	100m equation						
	Canopy	entire fl	100m seg/15m gap			100m seg/no gap		
	mean	mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)	9.940	22.901	22.958	0.628	13687	22.919	0.614	15703
basal area,seen (sq.m/ha)	26.870	47.349	48.544	1.783	35	48.521	1.720	40
basal area,unseen(sq.m/ha)	1.400	3.895	4.110	0.248	35	4.110	0.240	40
basal area,total(sq.m/ha)	28.270	51.243	52.654	2.030	35	52.631	1.959	40
dry biomass,seen(kg/ha)	156,680	453,021	490,574	35,676	35	491,250	34,966	40
dry biomass,unseen(kg/ha)	5,318	54,332	63,741	7,416	35	64,041	7,423	40
dry biomass,total(kg/ha)	161,997	507,353	554,315	43,056	35	555,290	42,344	40

2. 75m segments

		Airborne Laser						
		75m equation						
		entire fl	75m seg/15m gap			75m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		22.901	22.983	0.646	13077	22.919	0.614	15703
basal area,seen (sq.m/ha)		51.796	54.016	2.160	44	54.088	2.237	53
basal area,unseen(sq.m/ha)		7.877	8.714	0.595	44	8.787	0.647	53
basal area,total(sq.m/ha)		59.673	62.730	2.753	44	62.874	2.881	53
dry biomass,seen(kg/ha)		500,791	560,731	38,225	44	568,104	42,926	53
dry biomass,unseen(kg/ha)		23,471	28,024	2,490	44	28,759	2,938	53
dry biomass,total(kg/ha)		524,262	588,756	40,710	44	596,863	45,855	53

3. 50m segments

		Airborne Laser						
		50m equation						
		entire fl	50m seg/15m gap			50m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		22.901	23.027	0.667	12085	22.920	0.583	15703
basal area,seen (sq.m/ha)		58.737	64.918	3.794	61	64.183	3.065	80
basal area,unseen(sq.m/ha)		5.298	6.320	0.533	61	6.205	0.419	80
basal area,total(sq.m/ha)		64.034	71.237	4.326	61	70.388	3.483	80
dry biomass,seen(kg/ha)		542,216	694,078	71,568	61	675,935	54,461	80
dry biomass,unseen(kg/ha)		41,891	61,835	8,672	61	59,301	6,389	80
dry biomass,total(kg/ha)		584,107	755,914	80,209	61	735,238	60,823	80

4. 25m segments

		Airborne Laser						
		25m equation						
		entire fl	25m seg/15m gap			25m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		22.901	22.581	0.654	9812	22.920	0.502	15703
basal area,seen (sq.m/ha)		76.626	102.496	9.993	99	103.576	7.464	159
basal area,unseen(sq.m/ha)		8.060	12.540	1.569	99	12.592	1.175	159
basal area,total(sq.m/ha)		84.686	115.036	11.557	99	116.167	8.635	159
dry biomass,seen(kg/ha)		678,667	1,224,511	183,506	99	1,217,923	138,639	159
dry biomass,unseen(kg/ha)		44,528	94,880	16,598	99	93,543	12,634	159
dry biomass,total(kg/ha)		723,195	1,319,392	200,076	99	1,311,467	151,247	159

B. LaSelva-5m:

2. Nonparametric

b. no transformation (thru origin): model V=a(h)

1. 100m segments

		Airborne Laser						
	Simulated Canopy	100m equation						
		entire fl	100m seg/15m gap			100m seg/no gap		
	mean	mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)	9.940	22.901	22.958	0.628	13687	22.919	0.614	15703
basal area,seen (sq.m/ha)	26.870	58.748	58.869	1.612	35	58.796	1.576	40
basal area,unseen(sq.m/ha)	1.400	2.955	2.962	0.081	35	2.957	0.079	40
basal area,total(sq.m/ha)	28.270	61.703	61.858	1.693	35	61.753	1.655	40
dry biomass,seen(kg/ha)	156,680	295,911	396,652	8,121	35	296,150	7,937	40
dry biomass,unseen(kg/ha)	5,318	5,453	5,466	150	35	5,457	146	40
dry biomass,total(kg/ha)	161,997	301,364	302,119	8,271	35	301,607	8,083	40

2. 75m segments

		Airborne Laser						
		75m equation						
		entire fl	75m seg/15m gap			75m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		22.901	22.983	0.646	13077	22.919	0.614	15703
basal area,seen (sq.m/ha)		60.671	60.889	1.711	44	60.720	1.627	53
basal area,unseen(sq.m/ha)		1.990	1.997	0.056	44	1.991	0.053	53
basal area,total(sq.m/ha)		62.660	62.886	1.767	44	62.712	1.680	53
dry biomass,seen(kg/ha)		298,135	299,209	8,407	44	298,378	7,995	53
dry biomass,unseen(kg/ha)		3,976	3,990	112	44	3,979	107	53
dry biomass,total(kg/ha)		302,110	303,198	8,519	44	302,357	8,102	53

3. 50m segments

		Airborne Laser						
		50m equation						
		entire fl	50m seg/15m gap			50m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		22.901	23.027	0.667	12085	22.920	0.583	15703
basal area,seen (sq.m/ha)		58.781	59.105	1.713	61	58.830	1.496	80
basal area,unseen(sq.m/ha)		0.236	0.237	0.007	61	0.236	0.006	80
basal area,total(sq.m/ha)		59.017	59.342	1.720	61	59.067	1.502	80
dry biomass,seen(kg/ha)		295,847	297,476	8,621	61	296,093	7,532	80
dry biomass,unseen(kg/ha)		6,781	6,818	198	61	6,787	173	80
dry biomass,total(kg/ha)		302,628	304,295	8,819	61	302,879	7,704	80

4. 25m segments

		Airborne Laser						
		25m equation						
		entire fl	25m seg/15m gap			25m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		22.901	22.581	0.654	9812	22.920	0.502	15703
basal area,seen (sq.m/ha)		51.041	50.330	1.457	99	51.084	1.118	159
basal area,unseen(sq.m/ha)		3.247	3.202	0.093	99	3.250	0.071	159
basal area,total(sq.m/ha)		54.288	53.532	1.549	99	54.334	1.189	159
dry biomass,seen(kg/ha)		244,245	240,841	6,970	99	244,450	5,350	159
dry biomass,unseen(kg/ha)		2,899	2,859	83	99	2,902	64	159
dry biomass,total(kg/ha)		247,144	243,700	7,053	99	247,352	5,414	159

C. LaSelva-20m:

1. Parametric:

a. transformation (with intercept): model: ln(V)=a(h)+b

1. 100m segments

		Airborne Laser						
		100m equation						
		entire fl	100m seg/15m gap			100m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)								
basal area,seen (sq.m/ha)								
basal area,unseen(sq.m/ha)								
basal area,total(sq.m/ha)								
dry biomass,seen(kg/ha)								
dry biomass,unseen(kg/ha)								
dry biomass,total(kg/ha)								

2. 75m segments

		Airborne Laser						
		75m equation						
		entire fl	75m seg/15m gap			75m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)								
basal area,seen (sq.m/ha)								
basal area,unseen(sq.m/ha)								
basal area,total(sq.m/ha)								
dry biomass,seen(kg/ha)								
dry biomass,unseen(kg/ha)								
dry biomass,total(kg/ha)								

3. 50m segments

		Airborne Laser						
	Simulated Canopy	50m equation						
		entire fl	50m seg/15m gap			50m seg/no gap		
	mean	mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)	16.830	20.548	19.869	1.795	2598	20.550	1.477	3374
basal area,seen (sq.m/ha)	30.780	33.236	33.128	1.746	13	33.771	1.467	17
basal area,unseen(sq.m/ha)	3.140	4.295	4.404	0.603	13	4.618	0.516	17
basal area,total(sq.m/ha)	33.910	37.531	37.532	2.348	13	38.389	1.981	17
dry biomass,seen(kg/ha)	177,471	199,937	200,907	14,635	13	206,200	12,412	17
dry biomass,unseen(kg/ha)	12,499	16,810	17,503	2,364	13	18,328	2,042	17
dry biomass,total(kg/ha)	189,970	216,746	218,410	16,991	13	224,527	14,446	17

4. 25m segments

		Airborne Laser						
		25m equation						
		entire fl	25m seg/15m gap			25m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		20.548	20.473	1.368	2100	20.551	1.132	3374
basal area,seen (sq.m/ha)		35.635	39.533	4.381	21	40.117	3.298	34
basal area,unseen(sq.m/ha)		6.201	8.328	1.783	21	8.542	1.237	34
basal area,total(sq.m/ha)		41.836	47.861	6.147	21	48.659	4.526	34
dry biomass,seen(kg/ha)		213,217	250,680	35,903	21	255,162	25,942	34
dry biomass,unseen(kg/ha)		26,775	39,275	9.811	21	40,198	6,478	34
dry biomass,total(kg/ha)		239,993	289,955	45,626	21	295,360	32,375	34

C. LaSelva-20m:

1. Parametric

b. no transformation (thru origin): model: V=a(h)

1. 100m segments

		Airborne Laser						
		100m equation						
		entire fl	100m seg/15m gap			100m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m) basal area,seen (sq.m/ha) basal area,unseen(sq.m/ha) basal area,total(sq.m/ha) dry biomass,seen(kg/ha) dry biomass,unseen(kg/ha) dry biomass,total(kg/ha)								

2. 75m segments

		Airborne Laser						
		75m equation						
		entire fl	75m seg/15m gap			75m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m) basal area,seen (sq.m/ha) basal area,unseen(sq.m/ha) basal area,total(sq.m/ha) dry biomass,seen(kg/ha) dry biomass,unseen(kg/ha) dry biomass,total(kg/ha)								

3. 50m segments

		Airborne Laser						
	Simulated Canopy	50m equation						
		entire fl	50m seg/15m gap			50m seg/no gap		
	mean	mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)	16.830	20.548	19.869	1.795	2598	20.550	1.477	3374
basal area,seen (sq.m/ha)	30.780	37.101	35.875	3.241	13	37.105	2.667	17
basal area,unseen(sq.m/ha)	3.140	3.984	3.852	0.348	13	3.985	0.286	17
basal area,total(sq.m/ha)	33.910	41.085	39.728	3.589	13	41.089	2.954	17
dry biomass,seen(kg/ha)	177,471	215,491	208,369	18,826	13	215,512	15,492	17
dry biomass,unseen(kg/ha)	12,499	15,617	15,101	1,364	13	15,618	1,123	17
dry biomass,total(kg/ha)	189,970	231,108	223,469	20,190	13	231,131	16,615	17

4. 25m segments

		Airborne Laser						
		25m equation						
		entire fl	25m seg/15m gap			25m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		20.548	20.473	1.368	2100	20.551	1.132	3374
basal area,seen (sq.m/ha)		37.903	37.764	2.523	21	37.907	2.087	34
basal area,unseen(sq.m/ha)		4.388	4.372	0.292	21	4.389	0.242	34
basal area,total(sq.m/ha)		42.291	42.136	2.815	21	42.296	2.329	34
dry biomass,seen(kg/ha)		221,779	220,961	14,762	21	221,802	12,212	34
dry biomass,unseen(kg/ha)		18,103	18,036	1,205	21	18,105	997	34
dry biomass,total(kg/ha)		239,882	238,998	15,966	21	239,907	13,209	34

C. LaSelva-20m:

2. Nonparametric

a. transformation (with intercept): model: ln(V)=a(h)+b

1. 100m segments

		Airborne Laser						
		100m equation						
		entire fl	100m seg/15m gap			100m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)								
basal area,seen (sq.m/ha)								
basal area,unseen(sq.m/ha)								
basal area,total(sq.m/ha)								
dry biomass,seen(kg/ha)								
dry biomass,unseen(kg/ha)								
dry biomass,total(kg/ha)								

2. 75m segments

		Airborne Laser						
		75m equation						
		entire fl	75m seg/15m gap			75m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)								
basal area,seen (sq.m/ha)								
basal area,unseen(sq.m/ha)								
basal area,total(sq.m/ha)								
dry biomass,seen(kg/ha)								
dry biomass,unseen(kg/ha)								
dry biomass,total(kg/ha)								

3. 50m segments

		Airborne Laser						
	Simulated Canopy	50m equation						
		entire fl	50m seg/15m gap			50m seg/no gap		
	mean	mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)	16.830	20.548	19.869	1.795	2598	20.550	1.477	3374
basal area,seen (sq.m/ha)	30.780	34.932	34.889	2.296	13	35.723	1.918	17
basal area,unseen(sq.m/ha)	3.140	2.265	2.178	0.383	13	2.050	0.314	17
basal area,total(sq.m/ha)	33.910	37.197	37.067	1.919	13	37.773	1.612	17
dry biomass,seen(kg/ha)	177,471	218,545	216	19,747	13	223,363	16,725	17
dry biomass,unseen(kg/ha)	12,499	0	4,395	2,162	13	3,870	1,738	17
dry biomass,total(kg/ha)	189,970	218,545	220,762	18,039	13	227,233	15,368	17

4. 25m segments

		Airborne Laser						
		25m equation						
		entire fl	25m seg/15m gap			25m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		20.548	20.473	1.368	2100	20.551	1.132	3374
basal area,seen (sq.m/ha)		36.111	40.118	4.515	21	40.713	3.394	34
basal area,unseen(sq.m/ha)		7.923	11.549	2.869	21	11.873	1.941	34
basal area,total(sq.m/ha)		44.035	51.666	7.344	21	52.586	5.314	34
dry biomass,seen(kg/ha)		219,198	261,760	39,622	21	266,568	28,332	34
dry biomass,unseen(kg/ha)		11,138	13,251	1,980	21	13,493	1,420	34
dry biomass,total(kg/ha)		230,336	275,012	41,602	21	280,061	29,751	34

C. LaSelva-20m:

2. Nonparametric

b. no transformation (thru origin): model: V=a(h)

1. 100m segments

		Airborne Laser						
		100m equation						
		entire fl	100m seg/15m gap			100m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)								
basal area,seen (sq.m/ha)								
basal area,unseen(sq.m/ha)								
basal area,total(sq.m/ha)								
dry biomass,seen(kg/ha)								
dry biomass,unseen(kg/ha)								
dry biomass,total(kg/ha)								

2. 75m segments

		Airborne Laser						
		75m equation						
		entire fl	75m seg/15m gap			75m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)								
basal area,seen (sq.m/ha)								
basal area,unseen(sq.m/ha)								
basal area,total(sq.m/ha)								
dry biomass,seen(kg/ha)								
dry biomass,unseen(kg/ha)								
dry biomass,total(kg/ha)								

3. 50m segments

		Airborne Laser						
	Simulated	50m equation						
	Canopy	entire fl	50m seg/15m gap			50m seg/no gap		
	mean	mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)	16.830	20.548	19.869	1.795	2598	20.550	1.477	3374
basal area,seen (sq.m/ha)	30.780	31.565	30.522	2.758	13	31.568	2.269	17
basal area,unseen(sq.m/ha)	3.140	5.948	5.752	0.520	13	5.949	0.428	17
basal area,total(sq.m/ha)	33.910	37.513	36.273	3.277	13	37.517	2.697	17
dry biomass,seen(kg/ha)	177,471	190,048	183,767	16,603	13	190,067	13,663	17
dry biomass,unseen(kg/ha)	12,499	17,481	16,903	1,527	13	17,482	1,257	17
dry biomass,total(kg/ha)	189,970	207,528	200,669	18,130	13	207,549	14,920	17

4. 25m segments

		Airborne Laser						
		25m equation						
		entire fl	25m seg/15m gap			25m seg/no gap		
		mean	mean	stan. err.	n	mean	stan. err.	n
height,all pulses(m)		20.548	20.473	1.368	2100	20.551	1.132	3374
basal area,seen (sq.m/ha)		33.435	33.312	2.225	21	33.438	1.841	34
basal area,unseen(sq.m/ha)		2.498	2.489	0.166	21	2.498	0.138	34
basal area,total(sq.m/ha)		35.933	35.800	2.392	21	35.936	1.979	34
dry biomass,seen(kg/ha)		179,410	178,748	11,941	21	179,429	9,879	34
dry biomass,unseen(kg/ha)		26,008	25,912	1,731	21	26,011	1,432	34
dry biomass,total(kg/ha)		205,418	204,661	13,673	21	205,439	11,311	34

The most notable features of Table 14 are the large differences between the simulated canopy measurements and the airborne laser estimates of same. Ideally, the simulated and airborne values for basal area, volume, and biomass should be close since, purportedly, both are characterizing the same tropical forest. For instance, previous research found laser and ground estimates were within 7-8% of one another on a 6000 ha study area in southwestern Georgia (Nelson et al., 1988b). Percent error varies appreciably within study area due to the effects of the different treatment combinations, but for Tileran, airborne estimates are roughly half of the simulated values. The LaSelva-5m data exhibit just the opposite trend; airborne estimates are typically at least twice as large as the simulated values. The LaSelva-20m data also depict a situation where the airborne laser overestimates results simulated using ground reference transect data. Here, however, the airborne estimates are typically only 10-30% larger than the simulation values. No consistent downward transformation bias is noted in the Table 14 results. The under and overestimates associated with the three data sets may be directly tied to differences between simulated and airborne laser estimates of average canopy height, the independent variable used to predict basal area, volume, and biomass (as reported in Tables 10, 12, and 14). The airborne laser results presented in Table 14, then, must be evaluated on the basis of stability of the estimates as well as the accuracy of the estimates.

The following observations, based on the results presented in Table 14, are constrained to qualitative or empirical notations or to paired t-tests across factor levels.

1. Concerning the processing of the entire airborne laser flightline versus segmenting the flightline with 15m gaps between segments versus segmenting the flightline without gaps, the following points are made:

a. When segmenting an airborne laser flightline, the presence of a 15m gap reduces spatial autocorrelation between adjacent segments. The mitigation of autocorrelation is noted in the smaller standard errors of those estimates developed by segmenting without a 15m gap. A one-sided paired t-test of the standard errors for total volume, Tileran, across all treatment levels, showed that standard errors were significantly larger when a 15m gap was incorporated ($p<0.001$, $n=16$). Though not significant at a 95% level of confidence, the presence of a 15m gap increased standard errors of the LaSelva total biomass estimates ($0.05<p<0.10$, $n=24$). Excessive variability of the biomass estimates and their associated standard errors developed using transformed data in conjunction with short segment lengths increases the p-value. Considering total basal area and total volume and biomass, the standard errors of estimate calculated without a gap are smaller than those standard errors with gap in 34 of 40 cases. Standard errors increase up to ~ 40% when a 15m gap is left between segments.

b. Airborne laser flightline segmentation permits the computation of variances of the estimates. These variances are in all likelihood underestimated due to the fact that the presence of a gap only mitigates, but does not remove, autocorrelation. The gap size selected, 15m, is, approximately, the size of the average tree crown diameter at Tileran and LaSelva. A gap is placed between airborne laser segments to reduce the likelihood that adjacent segments are measuring portions of the same tree crown.

c. Segmentation results are not consistently any more or less accurate than those estimates developed when the entire flightline is processed, unless a natural log model is involved. Consider only those models developed based on untranformed data. Except for round-off, estimates based on the entire flightline and those based on segmented laser data without gap are the same since all pulses within a flightline are considered in both cases. A segmented airborne flightline will produce results as accurate as other treatment combinations as long as a natural log model is not considered. A gap should be left between segments to mitigate autocorrelation.

2. Concerning an optimal ground segment length: In order to develop files of simulated laser and ground reference metrics, the ground transects were divided into segments. Four lengths were tested, 25, 50, 75, and 100m. The results of three tables, 7, 8, and 14, are considered in an attempt to identify an optimal segment length. Consider:

a. Selection of a particular segment length involves consideration of two offsetting trends. First, shorter lengths are better from the standpoint of mitigating fieldwork. The shorter the segment length, the less the effort needed to collect the ground data needed to produce a given number of observations for regression analysis. Offsetting this pragmatic trend is a more theoretical one; each segment should, ideally, be a small replicate of the general canopy structure of the forest. The longer the segment, the larger the number of tree crowns involved in canopy simulation, and the lower the likelihood of outliers that might be caused by the influences of single, large trees, or conversely by a lack of trees.

b. The results presented in Table 14 suggest that no convincing argument can be made for a particular segment length. The effect of segment length on the accuracy of the basal area, volume, and biomass estimates depends greately on the levels of the other factors involved. Segment length is an important and significant factor, but its effects do not appear to be linear or consistent. Viewing the LaSelva-20m results, the 25m segment estimates are much more variable and less accurate than the 50m estimates given a model developed using transformed data. This observation carries some weight since the simulated and airborne laser height estimates are closest for this data set, hence airborne estimates of basal area and biomass are comparable to the ground reference figures. However, as seen in Table 7, the strength of the predictive relationship is only 0.1-0.2 for basal area and biomass. The low R^2, however, is due in part to the low sample size and limited ranges of biomasses and canopy height. Note also that the LaSelva-5m results at 25m are extremely variable, differing by over 2000% across treatment combinations. The 25m segment length induces variation and instability due in part due to the effects of highly influential observations. Lengths longer than 25m should be considered as the basic sampling unit in these tropical forests.

c. As noted previously, Figure 22 illustrates a large jump in mean square error across all three study areas as segment length is reduced to 25m from 50m. This notable, consistent increase in MSEs, in conjunction with the variability of airborne laser estimates of basal area and biomass at LaSelva, suggest that a 50m length rather than 25m should be used to segment the ground transect data in order to develop observations for regression analysis. Likewise this same segment length should be used to segment the airborne laser data (with a 15m gap) to produce estimates of basal area, volume, and biomass.

3. Concerning the use of nonparametric versus parametric regression techniques: The results do not suggest any apparent, consistent advantage to either analysis technique in terms of the accuracy of prediction. As noted in Figure 22 and by definition (since parametric analysis is a least squares fitting procedure), the parametric prediction equations always have a lower mean square error. However, nonparametric techniques are especially useful in situations where highly influential observations (HIOs) may be expected. Though there is no consistent, quantitative evidence in these results to suggest the sole use of nonparametric regression approaches, the nonparametric simple linear regression procedure will mitigate the effects of statistical outliers on the prediction equations. The

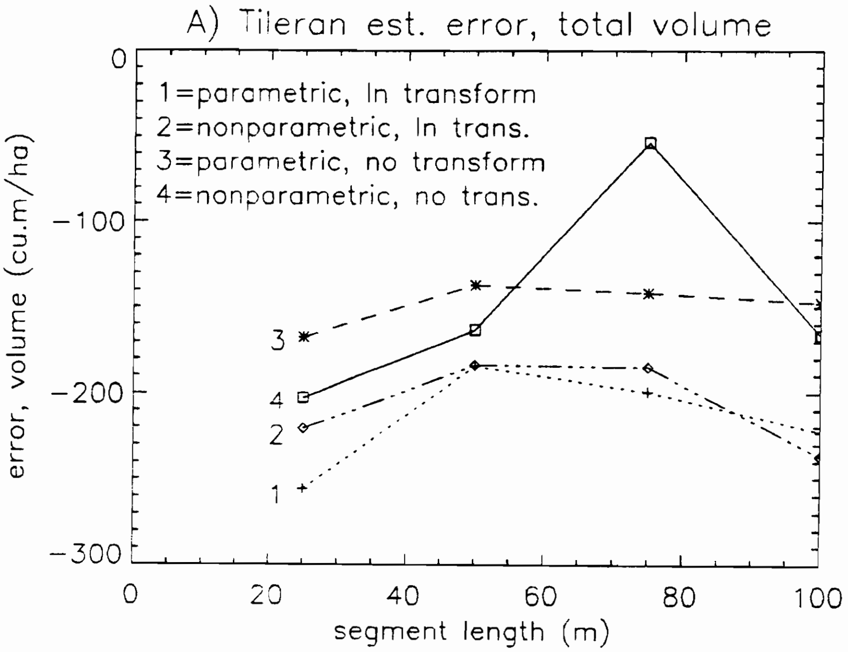
use of nonparametric regression procedures should be considered 1) to corroborate parametric findings; and 2) to provide a level of statistical safety should HIOs be present.

4. Concerning model use, using a natural log transform with an intercept term versus untransformed data with regression forced through the origin: Of the four factors considered, this decision (to transform or not to transform) is by far the most important. Two-sided paired t-tests were performed to compare Tileran total volume estimates developed with and without the log transformation, and to compare LaSelva-5m and 20m total biomass estimates developed with and without the natural log transformation. Considering the LaSelva-5m and 20m results jointly, t-test results showed that the natural log transformation significantly exacerbates ($p < 0.01$, $n=36$) an inherent overestimation problem. At Tileran, the response is reversed. The natural log model consistently, significantly exacerbates ($p < 0.002$, $n=24$) an inherent underestimation problem. Figure 23 illustrates the fact that the absolute estimation error, with few exceptions, increases if a natural log model is employed. No consistent, downward transformation bias is noted in the results involving the log transform, hence no bias adjustment (as proposed by Miller, 1984) is applied. In summary, the use of the natural log model improves regression fit, but prediction errors are amplified when regression calculations are inverted (via exponentiation) to estimate basal area, volume, or biomass

The results of the simple linear regression analyses are used to identify factors and levels of those factors which may be used to estimate ground metrics of interest. Though these estimates are inaccurate due to large discrepancies between the attributes of the forest canopy sensed by the airborne laser and those developed via simulation, factors have been identified which produce excessive error and obvious instabilities. Based on the results in Tables 8 and 14 and keeping in mind that shorter segment lengths reduce fieldwork, laser-based estimates of basal area and volume or biomass should be produced 1) using a ground segment length greater than 25m; 2) by dividing airborne laser data into 50m segments with a 15m gap between segments to reduce autocorrelation; and 3) by using untransformed data with the regression forced through the origin. The use of parametric versus nonparametric regression techniques depends on the analyst's expectations of the presence and importance of HIOs, and on the availability of nonparametric regression software.

IV.C. Multiple Linear Regression Analyses

The simple linear regression findings are used to limit the scope of the multiple linear regression analyses. In this section, only 50m segments are considered, and the airborne laser data are processed with a 15m gap between each 50m segment. Only parametric analyses are considered due to the fact that nonparametric multiple linear regression software is not readily available. Both models - transformed data with intercept and untransformed data forced through the origin - are tested.



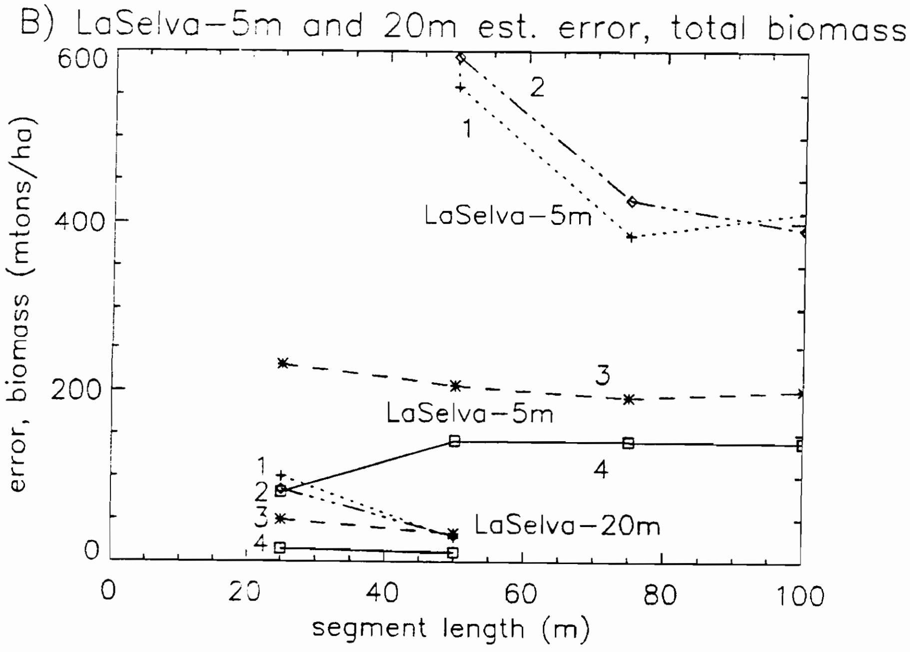


Figure 23. A plot of estimation error (airborne laser estimate minus ground reference estimate) at different segment lengths (with 15m gap). Parametric and nonparametric results are considered as are transformed and untransformed models for A) Tileran, total merchantable volume; and B) LaSelva-5m and -20m total above-ground biomass.

The results of the all possible subsets regressions indicate that the following generic models could be applied to the three study area data sets. As stated in the Procedures section, these models are not optimal for a particular study site. Rather, simulated laser variables common to all three study areas were selected in the hopes of identifying models which might be more generally applied to tropical forests.

$$b_s = f(\bar{h}_a, c_c)$$

$$b_t = f(\bar{h}_a, c_c)$$

$$b_u = b_t - b_s$$

$$v_s = f(\bar{h}_c, c_a, c_c)$$

$$v_t = f(\bar{h}_c, c_a, c_c)$$

$$v_u = v_t - v_s.$$

In some cases, specifically on the LaSelva-20m ground transects, the models involving untransformed data for basal area and biomass were overfit, as evidenced by high variance inflation factors. Normally one or more variables responsible for the multicollinearity would be dropped from the offending models in order to reduce the inflated variances of the coefficients. However, in order to maintain the generic nature of the models across study areas, the VIFs were noted, accepted, and no changes were made to the model.

Table 15 presents the results of the multiple linear regression analysis. Parametric multiple linear regression techniques were used to calculate coefficients for models involving both transformed and untransformed data. Using the models described above, airborne laser data were processed in 50m segments with 15m gaps between each segment to estimate basal area, volume, and biomass on each of the three study sites. As noted with the simple linear regression results, Tileran airborne estimates under-represent ground reference basal area and volume by approximately a factor of two. The LaSelva-5m airborne results, conversely, overestimate ground reference data by approximately a factor of two. The

Table 15. Multiple linear regression estimates of seen, unseen, and total basal area, merchantable volume, and total above-ground dry biomass for the A) Tileran, B) LaSelva-5m, and C) LaSelva-20m study areas. Two models are considered: untransformed data forced through the origin and transformed data with an intercept. Only parametric regression techniques are used to estimate MLR coefficients.

A. Tileran

	Simulated Canopy		Tileran		
			50 m seg/15m gap		
	mean	stan. err.	mean	stan. err.	n
Parametric, no transformation (thru origin)					
Height, all pulses(m)	19.854	0.510	12.818	1.542	1577
Height, canopy hits(m)	22.910	0.220	14.447	1.547	1383
coef. var.,all pulses	0.520	0.040	0.481	0.049	1577
coef. var., canopy hits	0.320	0.010	0.281	0.033	1383
canopy closure(%)	86.670	2.150	87.698	2.825	1577
BA,seen(sq.m/ha)	34.620		21.228	3.020	8
BA,unseen(sq.m/ha)	7.480		4.402	0.736	8
BA,total(sq.m/ha)	42.100		25.630	3.754	8
Vol,seen(cu.m/ha)	342.000		171.958	44.999	8
Vol,unseen(cu.m/ha)	50.730		26.617	5.785	8
Vol,total(cu.m/ha)	392.720		198.575	50.744	8
Parametric, transformed data (with intercept)					
Height, all pulses(m)			12.818	1.542	1577
Height, canopy hits(m)			14.447	1.547	1383
coef. var.,all pulses			0.481	0.049	1577
coef. var., canopy hits			0.281	0.033	1383
canopy closure(%)			87.698	2.825	1577
BA,seen(sq.m/ha)			23.734	2.209	8
BA,unseen(sq.m/ha)			4.080	0.723	8
BA,total(sq.m/ha)			27.813	2.927	8
Vol,seen(cu.m/ha)			204.787	26.420	8
Vol,unseen(cu.m/ha)			33.343	5.778	8
Vol,total(cu.m/ha)			238.130	32.100	8

B. LaSelva-5m

	Simulated Canopy		LaSelva-5m		
			50 m seg/15m gap		
	mean	stan. err.	mean	stan. err.	n
Parametric, no transformation (thru origin)					
Height, all pulses(m)	9.940		23.027	0.667	12085
Height, canopy hits(m)	17.830		24.189	0.640	11461
coef. var.,all pulses	1.102		0.360	0.027	12085
coef. var., canopy hits	0.443		0.276	0.014	11461
canopy closure(%)	55.120		94.836	1.242	12085
BA,seen(sq.m/ha)	26.870		46.207	0.879	61
BA,unseen(sq.m/ha)	1.400		3.515	0.135	61
BA,total(sq.m/ha)	28.270		49.721	0.990	61
Biomass,seen(kg/ha)	156,680		295,123	7,571	61
Biomass,unseen(kg/ha)	5,318		11,189	540	61
Biomass,total(kg/ha)	161,997		306,311	8,071	61
Parametric, transformed data (with intercept)					
Height, all pulses(m)			23.027	0.667	12085
Height, canopy hits(m)			24.189	0.640	11461
coef. var.,all pulses			0.360	0.027	12085
coef. var., canopy hits			0.276	0.014	11461
canopy closure(%)			94.836	1.242	12085
BA,seen(sq.m/ha)			56.344	2.618	61
BA,unseen(sq.m/ha)			5.806	0.422	61
BA,total(sq.m/ha)			62.150	2.967	61
Biomass,seen(kg/ha)			322,626	27,557	61
Biomass,unseen(kg/ha)			22,561	1,525	61
Biomass,total(kg/ha)			345,186	28,634	61

C. LaSelva-20m

	Simulated Canopy		LaSelva-20m		
			50 m seg/15m gap		
	mean	stan. err.	mean	stan. err.	n
Parametric, no transformation (thru origin)					
Height, all pulses(m)	16.830		19.869	1.795	2598
Height, canopy hits(m)	19.850		20.919	1.469	2410
coef. var.,all pulses	0.541		0.394	0.073	2598
coef. var., canopy hits	0.309		0.293	0.032	2410
canopy closure(%)	84.740		92.763	3.461	2598
BA,seen(sq.m/ha)	30.780		33.498	1.465	13
BA,unseen(sq.m/ha)	3.140		4.059	0.548	13
BA,total(sq.m/ha)	33.910		37.557	1.450	13
Biomass,seen(kg/ha)	177,471		215,214	13,340	13
Biomass,unseen(kg/ha)	12,499		16,794	2,403	13
Biomass,total(kg/ha)	189,970		232,008	13,187	13
Parametric, transformed data (with intercept)					
Height, all pulses(m)			19.869	1.795	2598
Height, canopy hits(m)			20.919	1.469	2410
coef. var.,all pulses			0.394	0.073	2598
coef. var., canopy hits			0.293	0.032	2410
canopy closure(%)			92.763	3.461	2598
BA,seen(sq.m/ha)			37.597	3.644	13
BA,unseen(sq.m/ha)			3.804	0.745	13
BA,total(sq.m/ha)			41.401	3.506	13
Biomass,seen(kg/ha)			219,756	23,074	13
Biomass,unseen(kg/ha)			14,744	2,963	13
Biomass,total(kg/ha)			234,499	22,877	13

LaSelva-20m airborne estimates overestimate ground reference data by approximately 23%, reflecting the fact that simulated and airborne laser metrics most closely agree along this 20m wide transect.

Table 16 compares SLR and MLR results for total basal area and total volume or total biomass. In order to develop error terms across all three study areas, the merchantable volume estimates of Tileran were "scaled" to the biomass metric tonnage figures of LaSelva-5m and LaSelva-20m study areas. Tileran's volume estimates were converted to metric tons/ha by dividing by a generic wood density of 2.0 m^3/metric ton.[7] Two error terms are calculated, average absolute error and root mean square error. Formulas follow:

Average Absolute Error:

$$\text{average}\ |\text{error}| = \frac{\sum_{i=1}^{3} \left| \begin{matrix}\text{laser}\\ \text{estimate}\end{matrix} - \begin{matrix}\text{ground}\\ \text{reference}\end{matrix} \right|}{3}$$

Mean Square Error:

$$\text{mse} = \frac{\sum_{i=1}^{3} \left(\begin{matrix}\text{laser}\\ \text{estimate}\end{matrix} - \begin{matrix}\text{ground}\\ \text{reference}\end{matrix} \right)^2}{3}$$

The predictive linear models are developed using both parametric and nonparametric regression techniques. The linear models involving transformed data (i.e., the natural log of basal area, volume, or biomass) include an intercept term; models involving untransformed data are forced through the origin. The smallest error terms identify the approach which produces estimates of basal area, volume, and biomass closest to the ground reference data. A number of observations can be made based on the results presented in Table 16.

- That procedure which produces estimates of total basal area with the lowest errors employs multiple linear regression techniques and untransformed data with the regression forced through the origin. The best procedure to estimate total volume and biomass utilizes simple linear regression ($\bar{h}_a$ as the independent variable) to manipulate untransformed data with the regressions forced through the origin.

- The three most accurate procedures for estimating total basal area involve untransformed data with the model forced through the origin. The two most accurate procedures for estimating total volume and biomass likewise involve untransformed data with the model forced through the origin.

- Coefficients of determination were calculated for the parametric analyses where the model contains an intercept (i.e., transformed data). In four of six cases considered, R^2 values increased between 0.1 and 0.3 with the addition of one or two variables.

---

[7] This conversion factor is equivalent to a current-volume specific gravity of 0.5 (see Husch et al., 1972). This generic density figure for Costa Rican tropical woods is equivalent to the density of wood found in Black Cherry, Sycamore, Red Maple, and Sweetgum (Harford County Forestry Board, no date).

Table 16. Summary of airborne laser estimates of total basal area and total merchantable volume or total dry biomass for the Tileran, LaSelva-5m, and LaSelva-20m study sites. Simple linear and multiple linear regression models are considered. Two error terms are considered, average absolute error and root mean square error. The Tileran merchantable volume estimates are scaled to the LaSelva biomass estimates by multiplying volumes by a density factor of 0.5 mtons/m^3 of wood. Simple linear regression models use $\bar{h}_a$ as the independent variable. The multiple linear regression model variables are reported in Section IV.C.

Total Basal Area (m^2/ha):

		Simple Linear Regression					Multiple Linear Regression		
		Nonparametric		Parametric			Parametric		
study site	ground reference	no trans., thru 0	trans., w/intcpt	no trans., thru 0	trans., w/intcpt estimate	R^2	no trans., thru 0	trans., w/intcpt estimate	R^2
Tileran	42.100	24.467	31.433	27.028	25.432	0.510	25.630	27.813	0. 553
LaSelva-5m	28.270	59.342	71.237	59.032	76.261	0.422	49.721	62.150	0.511
LaSelva-20m	33.910	36.273	37.067	39.728	37.532	0.160	37.557	41.401	0.422
avg. |error|		17.023	18.930	17.217	22.760		13.856	18.553	
sqrt(mse)		20.672	25.625	20.061	29.406		15.755	21.665	

Total Volume (m^3/ha) - Tileran:
Total Biomass (mtons/ha) - LaSelva-5m, LaSelva-20m:

		Simple Linear Regression					Multiple Linear Regression		
		Nonparametric		Parametric			Parametric		
study site	ground reference	no trans., thru 0	trans., w/intcpt	no trans., thru 0	trans., w/intcpt estimate	R^2	no trans., thru 0	trans., w/intcpt estimate	R^2
Tileran	392.720	229.959	209.317	255.739	208.543	0.647	198.575	238 .130	0.660
LaSelva-5m	161.997	304.295	755.914	368.543	721.224	0.548	306.311	345.186	0.739
LaSelva-20m	189.970	200.669	220.762	223.469	218.410	0.199	232.008	234.499	0.480
avg. |error|		78.126	238.804	102.845	226.585		94.475	101.671	
sqrt(mse)		94.844	347.417	127.115	327.630		103.307	117.637	

The results in Table 16 make a strong case for recommending models which use data without any transformations and without a linear intercept term (i.e., forced through the origin). It is only common sense that multiple independent variables should be considered to develop predictive models. In future studies, models should be tailored via variable selection procedures and considerations of multicollinearity to the specific study area.

There are significant costs associated with developing models which predict basal area, volume, and biomass using untransformed data where the regression is forced through the origin. The costs are basically limitations imposed on statistical inference; these limitations, in turn, are related to the assumptions which need to be made in order to calculate various regression metrics. Least squares procedures are used to estimate the linear regression coefficients, and use of least squares requires that 1) the x_i are nonrandom and measured "without error"; and 2) $E(\varepsilon_i) = 0$. Under these assumptions, the least squares regression coefficients are unbiased. In order to calculate variances of the coefficients, the following additional assumptions are needed: 3) the variances are homogeneous at all levels of x_i encountered; and 4) the ε_i are uncorrelated. Finally, to test hypotheses and to calculate confidence intervals, a final assumption is needed: 5) the ε_i are independently and identically distributed (iid) $N(0,\sigma^2)$ (Myers, 1990).

In most real-world situations, one expects the variability of basal area, volume, and biomass to vary predictably at different levels of the regressors. Such variation is noted in Figure 21, where the ranges of basal area, volume, and biomass increase as average canopy height increases. Data are transformed using the natural log function to control variances in an attempt to meet assumptions 3, 4, and 5. The use of untransformed data, in most cases, will render assumptions 3, 4, and 5 invalid. Statistical inference is not precluded, but special variance - covariance estimators must be used to overcome the effects of heteroskedasticity (nonconstant variances at different levels of the independent variable(s)). White (1980) proposes a biased, yet consistent variance - covariance estimator which may be used to estimate the variances of the coefficients, for calculation of confidence limits, and for hypothesis testing.

A second, more minor, concession caused by the use of untransformed data where the linear model is forced through the origin is an inability to calculate a coefficient of determination, a handy metric for determining model fit. The R^2 metric in the zero intercept case is not directly comparable to the R^2 metric where the model has an intercept term due to redefined total and regression sums of squares. Myers

(1990) explains the R^2 - zero intercept limitations and proposes an alternate, more comparable measure, though this alternative has its own limitations.

The net effect of these results and this discussion is that the model employing untransformed data where the regresson is forced through the origin consistently provides the most accurate predictions of basal area, volume, and biomass. However, the use of this model imposes limits on statistical inference, and those limits should be understood by the analyst. The data will not prevent one from calculating variances, performing various F and t tests, and establishing confidence intervals. These calculations are, however, suspect, and conclusions which arise based on these inferences would likewise be suspect.

The results of the SLR and MLR intercomparisons indicate that: 1) basal area is most accurately predicted using parametric MLR techniques in conjunction with untransformed data and a linear model with no intercept term; 2) volume and biomass are most accurately predicted using nonparametric SLR techniques with $\bar{h}_a$ as the sole independent variable, no data transformation, no intercept term. The specific regression equations, given below, were developed for 50m segments assuming an elliptical canopy shape. $\bar{h}_a$ is measured in meters; c_c is unitless.

Basal Area (m^2/ha):

Tileran: $b_s = 1.85849(\bar{h}_a) - 9.24416(c_c)$

$b_t = 2.28949(\bar{h}_a) - 13.24367(c_c)$

LaSelva-5m: $b_s = 1.57933(\bar{h}_a) + 35.63618(c_c)$

$b_t = 1.75597(\bar{h}_a) + 33.63514(c_c)$

LaSelva-20m: $b_s = 0.74540(\bar{h}_a) + 63.85038(c_c)$

$b_t = 1.01798(\bar{h}_a) + 59.11086(c_c)$

Volume (m^3/ha) or Biomass (kg/ha):

Tileran: $v_s = 15.12449(\bar{h}_a)$

$v_t = 17.94005(\bar{h}_a)$

LaSelva-5m: $v_s = 12918.74(\bar{h}_a)$

$v_t = 13214.83(\bar{h}_a)$

LaSelva-20m: $v_s = 9248.82(\bar{h}_a)$

$v_t = 10099.45(\bar{h}_a)$

IV.D. Line Intercept Sampling

Line Intercept Sampling is a statistical framework which permits one to effectively increase the dimensionality of the measurement. Measurements along a line, i.e., length of intercept, are readily converted to an estimate of the areal coverage of the object(s). An along-track distance measure with a corresponding depth (or height) measurement, i.e., canopy profile area, is easily converted to an estimate of the volume of particular objects. These metrics, line intercept length and canopy area/ha (equivalent to canopy density) and canopy profile area and canopy volume/ha (equivalent to average height, $\bar{h}_a$), are measured directly by the airborne laser. Converting these measurements to ones useful to foresters requires additional assumptions and computations.

In order to utilize LIS techniques to estimate basal area, biomass, and volume, four ratios must be predicted over a range of canopy conditions. Either 1) r_b, s_b, r_v, and s_v must be predictable (or, better yet, stable); or 2) r_b, t_b, r_v, and t_v must be predictable or stable (see Table 5 for term definitions). The first set of ratios calculates seen and unseen metrics (eg., basal area) and sums these two estimates to develop a total. The second set of ratios caculates seen and total metrics and differences these estimates to get the unseen component. The first task, then, is to 1) go through a variable selection procedure; and 2) determine the predictability/stability of each of the six ratios, r_b, s_b, t_b, r_v, s_v, and t_v. That subset of the two listed above which exhibits the best model fits in terms of coefficients of determination is selected for subsequent processing to generate airborne laser estimates of basal area, volume, and biomass.

Scatterplots were generated to determine the stability of the ratios and to determine if any of the eight simulated laser metrics could be used to predict any of the six ratios on each of the three study sites. In all cases, an elliptical canopy shape was assumed, and a 50 m ground segment length was utilized. One hundred forty four scatterplots were generated (3 study areas x 8 independent variables x 6 dependent variables, eg., see Figure 18) and checked for linearities or for curvilinear responses which would suggest the use of a data transformation. No single simulated laser variable exhibited any consistent linear trend with any of the nine LIS ratios. None of the ratios were constant across the range of canopy conditions simulated. No consistent curvilinear trends were noted.

As with the multiple linear regression analyses, SAS procedure REG (SAS, 1989) was used to parametrically assess all possible laser variable combinations to see if certain laser variable subsets could be used to predict any of the nine LIS ratios. The following generic models were developed based on 1) the results of the all possible subsets regressions; and 2) multicollinearity statistics. All linear models consider untransformed data; none of the ratios were transformed logarithmically. All models contain an intercept.

Basal Area:

$$r_b = f(g, s_w, c_c)$$

$$s_b = f(g, s_w)$$

$$t_b = f(g, s_w, c_c)$$

Volume or Biomass:

$$r_v = f(g, s_w, c_c)$$

$$s_v = f(g, s_w)$$

$$t_v = f(g, s_w, c_c)$$

These models are not necessarily optimal. Independent variables commonly significant on all three study areas are included in the models in an attempt to make the models more general.

Table 17 reports the coefficients of determination calculated for the models listed above. The coefficients of determination presented in Table 17 are small, the eight simulated laser metrics explain little variation in any of the LIS ratios. R^2 range from 0.09 - 0.44 and average 0.25. Logarithmic transformation of the LIS ratios was considered; the results are not presented because the transformation did nothing to consistently improve model fit. The results indicate that the LIS ratios are not linearly, consistently predictable using the eight simulated laser variables considered in this analysis.

Table 17. Coefficients of determination (R^2) for generic models used to predict line intercept sampling ratios.

LIS (dependent) variable	independent variables	Tileran	LaSelva-5m	LaSelva-20m
<u>Basal Area:</u>				
r_b	g, s_w, c_c	0.160	0.283	0.328
s_b	g, s_w	0.109	0.109	0.444
t_b	g, s_w, c_c	0.287	0.267	0.253
<u>Volume or Biomass:</u>				
r_v	g, s_w, c_c	0.091	0.322	0.432
s_v	g, s_w	0.116	0.131	0.410
t_v	g, s_w, c_c	0.140	0.309	0.360

These results do not suggest that LIS techniques are useless in terms of their applicability to airborne laser surveys. LIS procedures as outlined by Kaiser (1983) can be used to provide directly any of the following airborne laser metrics <u>without the need to identify individual tree crowns in the airborne laser data:</u> average canopy height, all pulses ($\bar{h}_a$), canopy profile area (p), canopy density (g), average canopy heights, canopy hits ($\bar{h}_c$), canopy area, and canopy volume, and the variances of each. The results presented in Table 17 make it clear that, at least in the case of the two Costa Rican tropical forests considered, estimates of basal area, volume, and biomass cannot be derived reliably using LIS techniques.

IV.E. Effects of Canopy Shape

Unless canopy shapes are noted in the field, the analyst must assume a generic shape or shapes for the canopy simulation. The shape of the simulated tree canopies directly affects the simulated laser metrics, thereby affecting subsequent predictive regressions, and estimates of basal area, volume, and biomass. For a particular ground transect, an assumption of a particular canopy shape can affect simulated laser measurements - specifically number of overstory versus understory trees, hence number of trees intercepted, average canopy height, and canopy height variability (Nelson and Yaya, 1990). Throughout this investigation, an elliptical canopy shape has been assumed. The purpose of this section of research is to document and quantify the effects of the canopy shape assumption on 1) the predictive regressions, and 2) on estimates of basal area, volume, and biomass on each of the three study areas.

Building on the results of the SLR and MLR analyses, regression equations were developed based on the following models:

Parametric Multiple Linear Regression, no transformation, no intercept, 50m segment/15m gap:

$$b_s = f(\bar{h}_a, c_c)$$

$$b_t = f(\bar{h}_a, c_c)$$

$$v_s = f(\bar{h}_c, c_c, c_a)$$

$$v_t = f(\bar{h}_c, c_c, c_a)$$

Nonparametric Simple Linear Regression, no transformation, no intercept, 50m segment/15m gap:

$$b_s, b_t, v_s, v_t = f(\bar{h}_a)$$

Regression coefficients were calculated for each of these equations for the four different canopy shapes and a fifth, randomized assortment of these four shapes on the three study areas. The regression equations are used to predict basal area, volume, and biomass using airborne laser data.

Table 18 reports the slope values for the nonparametric regressions forced through the origin. As noted in these nonparametric results, the effects of canopy shape are predictable; four shapes, cone, parabola, ellipse, and sphere, plus a fifth random assortment of these shapes, are considered in this study. Regressions developed using these different shapes indicate that as the angularity of the conic shape decreases, i.e., as the crown becomes more rounded, the slope of the linear regressions decreases. So as the canopy shape becomes more rounded, a given height as measured by the airborne laser accounts for less and less basal area, volume, and biomass. As the assumed canopy shape is changed from cone to parabola to ellipse to sphere, airborne estimates decrease for a given airborne laser transect. This section of research measures and reports the extent of this decrease.

Table 18. Nonparametric regression coefficients of average canopy height, all pulses ($\bar{h}_a$) for different canopy shapes using 50m segments. $b_t = a_1 (\bar{h}_a)$; $v_t = a_2 (\bar{h}_a)$; n = 10 for Tileran; n = 60 for LaSelva-5m; n = 7 for LaSelva-20m.

A. Total Basal Area:

	slope		
canopy shape	Tileran	LaSelva-5m	LaSelva-20m
cone	2.254	3.000	1.960
parabola	2.047	2.701	1.786
ellipse	1.914	2.525	1.710
sphere	1.824	2.350	1.679
random	2.059	2.722	1.780

B. Total Volume (Tileran) or Total Biomass (LaSelva-5m, LaSelva-20m):

	slope		
canopy shape	Tileran	LaSelva-5m	LaSelva-20m
cone	20.809	15264.0	11787.9
parabola	18.915	13955.4	10729.7
ellipse	17.778	13006.7	9991.1
sphere	17.036	12359.0	9514.0
random	18.420	13914.7	10033.8

The effect of the canopy shape assumption on nonparametric simple linear regression slopes is marked. On the average, basal area, volume, and biomass slopes assuming a conical shape are 22.9% larger (range: 16.7 - 27.7%) than the comparable slope calculated assuming a spherical canopy shape. Slope differences tranlate directly to differences in airborne laser estimates of basal area, volume, and biomass in SLR models without an intercept.

Table 19 reports differences in airborne estimates of basal area, volume, and biomass due to canopy shape assumptions.[8] For a particular study site, the same segments of airborne laser data are processed using predictive equations developed under five different assumptions of canopy shape. The two regression technique/model combinations identified as most useful in terms of producing accurate, stable results (see Table 16) are considered. These analyses are conducted only to compare canopy shapes. If

---

[8] Some airborne laser estimates in Table 14 and Table 19 which should be exactly the same are, in fact, not (for instance, note Tileran, nonparametric, no transformation, 50m/15m gap). A programming change was made to handle the mechanics of random number generation differently. Files of simulated metrics made before this change, the results of which are reported in Table 14, and after (Table 19), differ due to the fact that different patterns of uniform random numbers were used to develop the simulated laser metrics.

Table 19. Airborne laser estimates of seen, unseen, and total basal area and merchantable volume or total above-ground dry biomass developed using five different canopy shape assumptions. Two models are considered: A) a simple linear model using average height, all pulses as the independent variable, no intercept; and B) a parametric multiple linear regression model, no transformation, through origin.

A. Nonparametric simple linear regression, no transformation, through origin.

1. Tileran

	Simulated	Airborne Laser Estimate									
		cone		parabola		ellipse		sphere		random	
	Canopy	mean	stan. err.	mean	stan. err.	mean	stan. err.	mean	stan. err.	mean	stan. err.
Height, all pulses(m)	19.854	12.818	1.542	12.818	1.542	12.818	1.542	12.818	1.542	12.818	1.542
BA,seen(sq.m/ha)	34.620	23.802	2.864	20.705	2.491	18.796	2.261	17.622	2.120	19.787	2.380
BA,unseen(sq.m/ha)	7.480	5.093	0.613	5.535	0.666	5.737	0.690	5.755	0.692	6.610	0.795
BA,total(sq.m/ha)	42.100	28.896	3.476	26.240	3.157	24.534	2.951	23.377	2.812	26.397	3.176
mer.volume,seen(cu.m/ha)	342.000	236.800	28.488	213.473	25.682	192.120	23.113	175.512	21.115	185.872	22.361
mer.volume,unseen(cu.m/ha)	50.730	29.931	3.601	28.988	3.487	35.759	4.302	42.863	5.157	50.234	6.043
mer.volume,total(cu.m/ha)	392.720	266.731	32.089	242.461	29.169	227.879	27.415	218.376	26.271	236.106	28.404

2. LaSelva-5m

	Simulated	Airborne Laser Estimate									
		cone		parabola		ellipse		sphere		random	
	Canopy	mean	stan. err.	mean	stan. err.	mean	stan. err.	mean	stan. err.	mean	stan. err.
Height, all pulses(m)	9.940	23.027	0.667	23.027	0.667	23.027	0.667	23.027	0.667	23.027	0.667
BA,seen(sq.m/ha)	26.870	69.074	2.002	62.203	1.803	58.142	1.685	53.292	1.545	61.900	1.794
BA,unseen(sq.m/ha)	1.400	0.000	0.000	0.000	0.000	0.000	0.000	0.823	0.024	0.789	0.023
BA,total(sq.m/ha)	28.270	69.074	2.002	62.203	1.803	58.142	1.685	54.115	1.568	62.689	1.817
Biomass,seen(kg/ha)	156,680	344,625	9,988	313,815	9,095	292,249	8,470	276,475	8,013	307,830	8,922
Biomass,unseen(kg/ha)	5,318	6,855	199	7,534	218	7,253	210	8,113	235	12,580	365
Biomass,total(kg/ha)	161,997	351,481	10,187	321,349	9,313	299,503	8,680	284,588	8,248	320,410	9,286

3. LaSelva-20m

	Simulated	Airborne Laser Estimate									
		cone		parabola		ellipse		sphere		random	
	Canopy	mean	stan. err.	mean	stan. err.	mean	stan. err.	mean	stan. err.	mean	stan. err.
Height, all pulses(m)	16.830	19.869	1.795	19.869	1.795	19.869	1.795	19.869	1.795	19.869	1.795
BA,seen(sq.m/ha)	30.780	34.841	3.148	31.588	2.854	29.357	2.652	28.344	2.561	31.844	2.877
BA,unseen(sq.m/ha)	3.140	4.106	0.371	3.891	0.352	4.626	0.418	5.020	0.454	3.520	0.318
BA,total(sq.m/ha)	33.910	38.947	3.519	35.479	3.205	33.983	3.070	33.364	3.014	35.364	3.195
Biomass,seen(kg/ha)	177,471	216,714	19,580	196,746	17,776	182,573	16,495	172,770	15,610	183,459	16,575
Biomass,unseen(kg/ha)	12,499	17,503	1,581	16,445	1,486	15,943	1,440	16,267	1,470	15,905	1,437
Biomass,total(kg/ha)	189,970	234,217	21,161	213,191	19,262	198,516	17,936	189,037	17,079	199,364	18,012

B. Parametric multiple linear regression, no transformation, through origin

1. Tileran

	Simulated	Airborne Laser Estimate									
	Simulated	cone		parabola		ellipse		sphere		random	
	Canopy	mean	stan. err.	mean	stan. err.	mean	stan. err.	mean	stan. err.	mean	stan. err.
Height, all pulses(m)	19.854	12.818	1.542	12.818	1.542	12.818	1.542	12.818	1.542	12.818	1.542
Height, canopy hits(m)	22.910	14.447	1.547	14.447	1.547	14.447	1.547	14.447	1.547	14.447	1.547
coef. var.,all pulses	0.520	0.481	0.049	0.481	0.049	0.481	0.049	0.481	0.049	0.481	0.049
coef. var., canopy hits	0.320	0.281	0.033	0.281	0.033	0.281	0.033	0.281	0.033	0.281	0.033
canopy closure(%)	86.670	87.698	2.825	87.698	2.825	87.698	2.825	87.698	2.825	87.698	2.825
BA,seen(sq.m/ha)	34.620	27.128	3.740	23.739	3.132	21.679	2.839	20.229	2.559	23.639	2.631
BA,unseen(sq.m/ha)	7.480	4.038	0.543	4.525	0.682	4.414	0.756	4.335	0.869	4.118	0.969
BA,total(sq.m/ha)	42.100	31.167	4.283	28.264	3.813	26.094	3.591	24.563	3.415	27.757	3.564
mer.volume,seen(cu.m/ha)	342.000	242.767	49.333	200.347	46.298	175.015	44.751	152.226	44.477	204.095	40.183
mer.volume,unseen(cu.m/ha)	50.730	20.420	4.108	26.365	5.479	26.484	5.970	26.413	6.595	22.520	7.702
mer.volume,total(cu.m/ha)	392.720	263.187	53.357	226.713	51.751	201.499	50.642	178.639	50.857	226.615	46.996

2. LaSelva-5m

		Airborne Laser Estimate									
	Simulated	cone		parabola		ellipse		sphere		random	
	Canopy	mean	stan. err.	mean	stan. err.	mean	stan. err.	mean	stan. err.	mean	stan. err.
Height,all pulses(m)	9.940	23.027	0.667	23.027	0.667	23.027	0.667	23.027	0.667	23.027	0.667
Height, canopy hits(m)	17.830	24.189	0.640	24.189	0.640	24.189	0.640	24.189	0.640	24.189	0.640
coef. var.,all pulses	1.102	0.360	0.027	0.360	0.027	0.360	0.027	0.360	0.027	0.360	0.027
coef. var., canopy hits	0.443	0.276	0.014	0.276	0.014	0.276	0.014	0.276	0.014	0.276	0.014
canopy closure(%)	55.120	94.836	1.242	94.836	1.242	94.836	1.242	94.836	1.242	94.836	1.242
BA,seen(sq.m/ha)	26.870	51.484	1.030	47.093	0.907	46.091	0.874	45.440	0.891	45.416	0.824
BA,unseen(sq.m/ha)	1.400	3.219	0.117	3.619	0.136	3.449	0.133	3.088	0.118	3.272	0.122
BA,total(sq.m/ha)	28.270	54.704	1.133	50.712	1.022	49.540	0.983	48.528	0.992	48.688	0.919
Biomass,seen(kg/ha)	156,680	342,635	8,662	310,717	7,742	295,225	7,425	279,231	7,204	295,768	6,615
Biomass,unseen(kg/ha)	5,318	10,313	426	12,175	561	11,091	547	9,628	446	10,776	481
Biomass,total(kg/ha)	161,997	352,948	9,066	322,892	8,261	306,316	7,925	288,858	7,620	306,544	7,038

3. LaSelva-20m

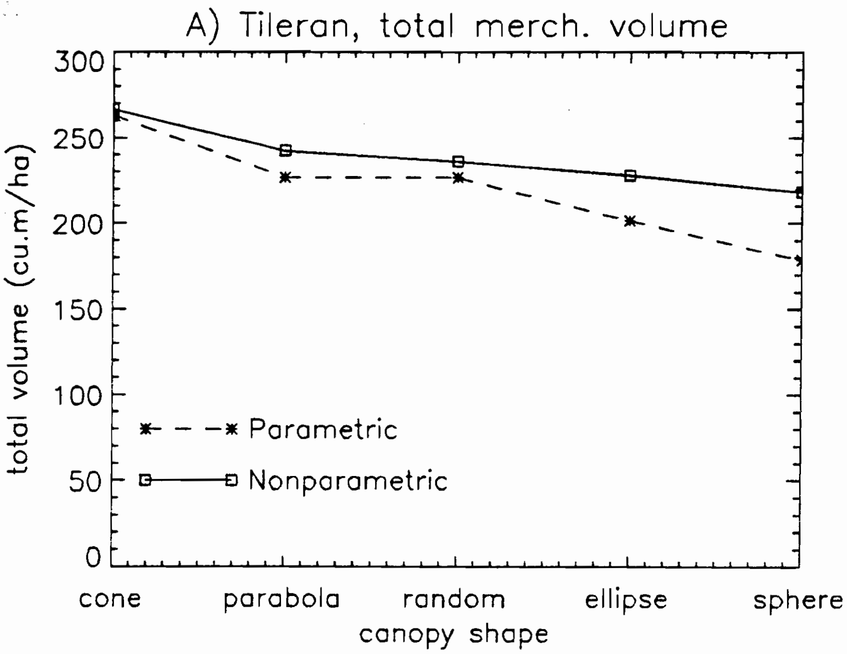
		Airborne Laser Estimate									
	Simulated	cone		parabola		ellipse		sphere		random	
	Canopy	mean	1.795	mean	stan. err.	mean	stan. err.	mean	stan. err.	mean	stan. err.
Height, all pulses(m)	16.830	19.869	1.795	19.869	1.795	19.869	1.795	19.869	1.795	19.869	1.795
Height, canopy hits(m)	19.850	20.919	1.469	20.919	1.469	20.919	1.469	20.919	1.469	20.919	1.469
coef. var.,all pulses	0.541	0.394	0.073	0.394	0.073	0.394	0.073	0.394	0.073	0.394	0.073
coef. var., canopy hits	0.309	0.293	0.032	0.293	0.032	0.293	0.032	0.293	0.032	0.293	0.032
canopy closure(%)	84.740	92.763	3.461	92.763	3.461	92.763	3.461	92.763	3.461	92.763	3.461
BA,seen(sq.m/ha)	30.780	34.844	1.371	33.486	1.351	33.655	1.378	33.714	1.522	33.837	1.346
BA,unseen(sq.m/ha)	3.140	4.064	0.621	3.997	0.662	3.812	0.649	3.515	0.554	3.647	0.603
BA,total(sq.m/ha)	33.910	38.908	1.621	37.483	1.513	37.466	1.484	37.229	1.447	37.484	1.581
Biomass,seen(kg/ha)	177,471	200,890	8,561	194,450	8,733	192,236	8,906	197,981	9,195	196,978	9,433
Biomass,unseen(kg/ha)	12,499	17,042	2,466	16,585	2,617	16,032	2,647	18,566	2,293	15,738	2,384
Biomass,total(kg/ha)	189,970	217,932	10,202	211,035	10,266	208,268	10,467	216,548	10,568	212,716	10,889

simulated and airborne laser data metrics were not so disparate, these analyses could be used to infer the predominant tropical forest canopy shape. However, simulated and airborne data sets, especially Tileran and LaSelva-5m, are very different. These differences limit the scope of the inference to a comparative study of the effects of canopy shape. The results presented in Table 19 for total basal area and total volume or biomass are illustrated in Figure 24. The following observations are noted:

- In general, the more conical the canopy shape assumed, the higher the mean estimate of basal area, volume, and biomass, and the larger the standard error of that estimate. On average, the airborne laser estimates of total basal area assuming an elliptical crown are 3.9% larger than spherical estimates, parabolic total basal area estimates are 9.0% larger than spherical estimates, and conical estimates are 19.3% larger than estimates developed assuming a spherical crown, regardless of regression/model approach. Considering total volume or biomass, elliptical estimates are 4.9% larger, parabolic estimates are 12.1% larger, and conical estimates are 23.3% larger than estimates developed assuming a spherical canopy shape.
- In general, a uniformly random distribution of canopy shapes produces estimates similar to parabolic or elliptical assumptions. Standard errors are likewise, in general, equivalent.

The assumption of a particular canopy shape, then, is not trivial. An inappropriate assumption concerning canopy shape can induce a bias in estimates of basal area, volume, or biomass, and this bias may be on the order of 25%. Ideally, the canopy shape of each tree sampled along a ground transect would be noted during an inventory. Realistically, the canopy shape of an individual tree may be difficult to ascertain in the field, especially in closed canopy situations where the general canopy level is 25 or more meters over the observer's head (eg., a tropical moist forest). A more practical approach may be to note a general canopy shape for the forest in which the ground team is working, and then note obvious exceptions based on direct visual observation and/or based on known growing habits of particular species.

Templates for the canopy shapes considered in this study may be found in any textbook which covers analytic geometry, including most beginning level calculus textbooks. Illustrations of different geometric shapes as they apply to tree canopies are also given in an article by Mawson et al. (1976). Mawson actually characterizes a tree canopy as a function of two geometric shapes, a profile shape (crosssectional shape from the side) - circle, triangle, neiloid, parabola, ellipse - and a plan shape (cross-sectional shape from the top) - circle, ellipse, triangle. Their tree canopy constructs, then, are more sophisticated than ones used in this study, in part because they're interested in characterizing leafy area volume rather than, like the laser, characterizing that volume between the top of the canopy and the ground.



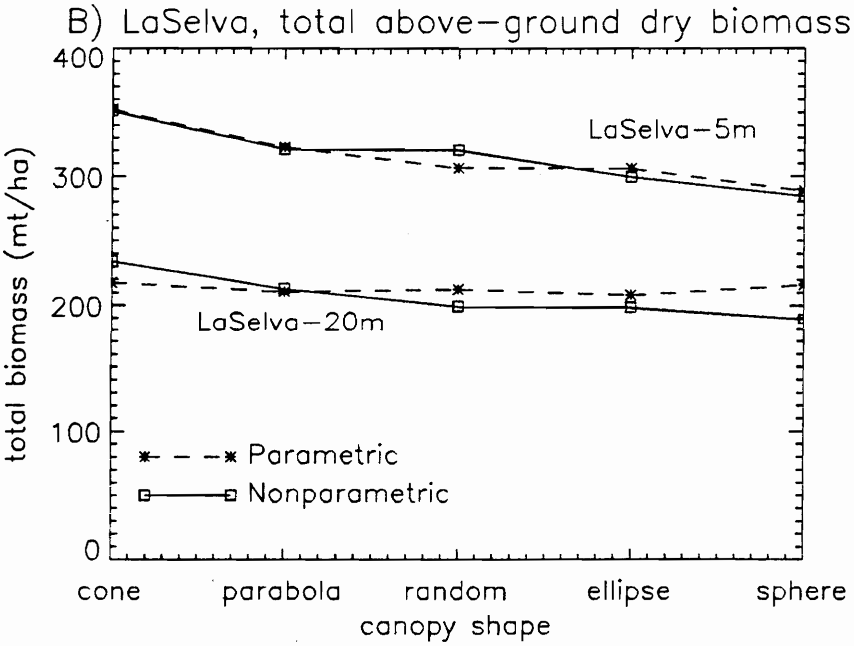


Figure 24. The effects of canopy shape on estimates of A) total merchantable volume, Tileran; and B) total above-ground dry biomass, LaSelva-5m and LaSelva-20m. Both nonparametric SLR - no intercept, results, and parametric MLR - no intercept, results are illustrated.

V. SUMMARY AND CONCLUSIONS

An airborne laser profiling system is a reconnaissance level forest inventory tool which should be employed in situations where the study area is large (100s of kilometers on a side), inhospitable, inaccessible, and/or where little or no mensurational information exists (Nelson et al., 1988b). The airborne laser system acquires forest canopy height information along the flight transects, and various metrics related to canopy height are used to predict basal area, volume, and biomass. Tree height is a relatively poor predictor of these ground metrics; tree diameter squared or basal area is the independent variable of choice. As such, an airborne laser system is not a good site-specific mapper (Nelson et al., 1988b); rather, lasers should be used to assess large tracts or regions where prediction errors can be averaged with the expectation that errors sum to zero, and variation in the estimates can be mitigated by sample size.

Lasers for forest mensuration are best employed in situations where little inventory information exists and where site access may be limited due to challenges imposed by terrain, climate, and/or a lack of infrastructure. Areas such as uncolonized sections of the Amazon, the Congo, the circumpolar boreal forests, and the tropical forests of Southeast Asia are examples of candidate areas for airborne laser surveys. Previous studies have depended on accurate registration of airborne laser and ground survey data in order to develop regression relationships between airborne laser and ground metrics. Colocation of the laser and ground transects within 2-3 meters has, in the past, required the presence of ground control features throughout the study area, features that could be definitively located in the laser data and on the ground. Transects are established between control points; the more closely spaced the control points, the more accurate the location of the laser track on the ground. Areas where lasers might be flown to best effect typically have few ground control features. The Amazon, the Congo, the boreal forests are characterized as broad expanses of unbroken, often closed-canopy forest where registration based on visual clues would be impossible across broad reaches of the region.

Colocation of the laser line and ground transect as a prerequisite to data analysis is limiting even in the Global Positioning System (GPS) era. Currently, the full constellation of 24 operational GPS satellites are in orbit. Despite access to a fully operational system, GPS is not a panacea for ground transect location because of the following: 1) GPS rover units (handheld receivers) may not function well under a forest canopy. GPS satellite signals do not consistently penetrate vegetation, and a GPS rover needs line-of-sight to the satellites to lock and locate. The denser the canopy, i.e., the less sky seen in the direction of the satellites, the worse the locational performance.[9] 2. GPS location accuracies on the order of a meter or less are available only if a) the operator has P-code access (in general only available to the US military or with US Department of Defense approval) and can spend approximately 4 hours per point to replicate and average a particular ground location; or b) an accessible base station has been set up for differential positioning. 3. The GPS position of the laser aircraft (GPS receiver aboard the aircraft) along with aircraft inertial navigation system data must be deconvolved to calculate GPS positions of the laser trace on the ground. The estimation of forest metrics of interest becomes much easier if the ground sampling phase can be divorced from the airborne laser sampling phase, both in terms of distance and time.

The overall purpose of this research is to define this inventory procedure which utilizes airborne laser data but divorces the air and ground sampling phase. The single most important constraint on this research is that the airborne laser inventory method defined should <u>not</u> require coregistration of the airborne laser data and the ground inventory data. This constraint is imposed to make airborne laser survey techniques amenable to those areas where they could be applied to best effect.

To this end, a simulator was written which uses mapped stand data and digitally recreates the height structure of the top of the forest canopy. The two-dimensional height array is randomly transected to develop simulated laser metrics which are matched directly with the corresponding ground metrics of interest. Subsequent statistical analyses result in regression equations which predict ground metrics such as basal area, volume, and biomass as a function of simulated laser measurements. These equations, in turn, are used in tandem with airborne laser data acquired over the study site to estimate the ground metrics.

In the course of the study, a number of factors are addressed in an attempt to identify a useful, though not necessarily optimal, method for developing airborne laser estimates of basal area, volume, and

---

[9] Bolstad (1994, personal communication) reports that points may be located with 2-5m accuracy with a 6-channel handheld unit and a local base station (within 10 km) in forests with leaf area indices up to 8 or 9. Hardwood forests are, in general, more tranparent to the GPS signals than conifers.

biomass. Modeling concerns, including variable selection and the utility of data transforms, are addressed. Segment length (both ground transect and airborne laser), autocorrelation concerns, regression techniques, utility of LIS techniques, ground transect width, and the effects of different canopy shape assumptions are all considered. The net result of the study is the definition of a specific sampling system that has been tested in the tropical forests of Costa Rica. Estimation errors, i.e., the difference between the airborne estimates and the comparable ground estimates, which are appreciable in this study, are quantified. Based on the findings in this study, an inventory approach is defined which integrates airborne laser data and ground surveys without the need for colocation. The specific suggestions which follow - segment length, gap length, specific regression models, regression coefficient estimation techniques, canopy shape assumptions - are provided as a starting point, a foundation. Future studies may confirm, refute, alter, refine, and expand upon these suggestions; these results merely define one useful starting point for subsequent studies.

This final chapter is divided into three sections. The first section reports and summarizes notable, specific results and conclusions developed over the course of the study. The second section describes, based on the findings presented in the first section, this analyst's view of a workable airborne laser survey procedure for tropical forest inventory. The third section reports issues that need further study.

V.A. Conclusions

There are any number of very practical aspects of a forest inventory plan - number of plots, size of plots, spacing, location, pattern, diameter limits, volume or biomass determination, measurement accuracies, instrumentation, method of measurement, etc. - that must be addressed in order to collect coherent, statistically sound data. In this study, the issues involve methods used to process field and airborne laser data that were already collected.

One mechanical sampling question that had to be addressed deals with the definition of an appropriate length to segment ground transect data.[10] Ground transect data are segmented in order to increase

---

[10] Ideally, this question should be considered and answered prior to the ground data collection. DeVries (1986) states that line plot data from a single transect may be treated as a single observation or, if segmented, as a single stage or two-stage cluster (a single stage cluster if the transect is tesselated, a two-stage cluster if gaps are left between transects). Cluster sampling assumes that data from more than one transect are collected. Transects are considered

the number of simulated laser-ground observations available for regression analyses. From a field efficiency standpoint, the shorter the segment length, the higher the number of regression observations per unit field effort - not a small consideration. From a variability standpoint, the longer the segment length, the greater the area averaged, the less the influence of single large trees, the less the likelihood of producing a highly influential observation that is purely a function of sample size, and the better the fit of the predictive regression. Four segment lengths are considered - 25, 50 75, and 100m.

- Of the four segment lengths considered, the 50m segment length is optimal for these Costa Rican forests. The 50m segment length is the shortest of the four considered which did not consistently exhibit high mean square errors when laser-ground regressions were calculated.

The ground transect sample width is important from the standpoint of capturing the spatial characteristics of the overstory tropical forest. Three different ground data sets were utilized in this study, the 5m wide LaSelva data, the 10m wide Tileran data, and the 20m wide LaSelva data. The forests sensed by the airborne laser were very different from those simulated using the Tileran and LaSelva-5m ground transect data. Airborne laser and simulated laser metrics were comparable on the LaSelva-20m sites. The Tileran airborne-ground discrepancies, a situation where the ground transect data were used to simulate a canopy approximately twice as tall as that canopy measured by the airborne laser, are a puzzle and are not addressed. No field visit was paid to this study area. A field visit was paid to LaSelva where simulated canopy heights developed based on a 5m wide ground transect are approximately half the height and canopy density of the forest measured by the airborne laser. During a visit to LaSelva, it was discovered that approximately 61.7% of the trees that make up the overstory canopy over a particular line transect through the forest are missed when the width of the transect is 5m. 31.3% are missed at LaSelva when the transect width is increased to 10m; 10.0% are missed at 15m; none are missed when the transect width is increased to 20m. This phenomenon of missed trees is mitigated in the computer simulation by assuming that adjacent plots are identical to the actual transect being simulated. However, simulated laser metrics developed using replicated 5m wide transect data grossly underestimate conditions actually sensed by the airborne laser. This problem is not noted using 20m wide LaSelva transect data. DeVries (1986,

---

clusters because segments within a cluster are spatially correlated. In this study, this spatial correlation can only be acknowledged and mitigated by leaving gaps between segments, due to the limited number of of ground transects collected (2 transects for the LaSelva-5m data, one transect each at Tileran and in the vicinity of the LaSelva-20m ground line).

pg 27) notes that 10m wide transects are the norm for tropical forest inventory. The following statements are speculative, but preliminary work indicates some merit:

- It is hypothesized that a 5m wide ground transect is too thin to capture the overstory height characteristics of a Costa Rican primary tropical forest. Computer height arrays generated with the 5m ground transect data simulate canopies which are much more open (50% vs 90%) and, on the average, shorter than the actual forest canopy. The trees which comprise the primary tropical forest overstory are widely enough dispersed that a 5m wide transect will include less than half of those individual tree canopies responsible for the height structures sensed by an airborne laser. Twenty meter wide transects are wide enough to include all trees necessary for a simulation. In this study, a 10 meter width is questionable, but historically defensible. It is recommended that tropical forest ground transects be <u>at least</u> 10m wide; primary forests may require widths on the order of 20m in order to capture the spatial variability of the overstory trees.

- Computer simulations which include the effects of adjacency (i.e., which account for the effects of the tree canopies at the transect boundaries) lessen the impact of, but do not solve this problem of a sampling width inadequate for overstory height simulation. It is hypothesized that a duplicated 5m wide transect does not carry the same <u>spatial</u> information as a 10m wide transect. Modeling adjacency, i.e., duplicating, triplicating transect information, is not a substitute for increasing the actual ground transect width.

The airborne laser profiling data can be segmented, or estimates can be developed based on entire flightlines. The same segment length used to partition the ground transect data is used to segment airborne laser data since the former is used to simulate the latter.

- Given that the ground transect data is segmented at 50m intervals, the airborne laser data should be processed in segments of the same length. No benefit is noted by considering an entire airborne laser flightline as a single observation <u>unless</u> the dependent variable is transformed logarithmically. In this situation, consideration of the entire flightline markedly reduces estimation error by mitigating the effects of HIOs, which are most prevalent at the shorter segment lengths.

- Airborne laser data should be segmented in order to develop estimates of variability. As noted previously, from a statistical standpoint, each airborne laser transect should be treated as an individual observation, since segments along a transect will be spatially autocorrelated. If aircraft time is available, perhaps an inventory plan could be devised where 10's of transects could be flown over the region. In lieu of multiple transects being flown, a few flightlines may be acquired and divided (segmented) and treated as individual observations. Spatial autocorrelation can be mitigated by considering discontinuous segments.

- A 15m gap, a distance slightly larger than the average canopy diameter of a tree, was left between segments in order to reduce height autocorrelation at the level of the individual tree crown. Nelson et al. (1988b) pointed out that airborne laser autocorrelation effects could be noted at the tree and stand level. An argument could be made that stand level autocorrelation is, in fact, a structural characteristic of the forest that should be captured in terms of variance computations. Gaps greater than the average canopy diameter of a tree, then, may be unnecessary as long as the airborne laser flightlines have captured a representative sample of the forest height structure.

By far the most important factor to be considered in the development of estimates of basal area, volume, and biomass is the use or nonuse of a logarithmic transformation. In short, do <u>not</u> logarithmically transform. The dependent variable is transformed to control variance at different levels of the independent variables considered. The logarithmic transformation, then, helps satisfy two of the assumptions needed to use a maximum likelihood regression estimator (that of equal variances at all levels of x_i, and that of uncorrelated errors, e_i), and also results in a better model fit (lower MSE, higher R^2). Previous studies (Maclean and Krabill, 1986; Nelson et al., 1988a,b) have relied on this logarithmic model exclusively. The results of this study are conclusive in showing that errors in airborne laser estimates of basal area, volume, and biomass increase dramatically when the logarithmic model is utilized in these Costa Rican tropical forests. Maclean, Krabill, and Nelson worked in second growth forests of the eastern U.S., relatively small, relatively homogeneous (in fact Nelson's studies were carried out in forests which were predominantly plantations). Outliers were not a problem, and inverting a logarithmic transformation to develop estimates of basal area, volume, and biomass did not lead to excessive errors in prediction. The logarithmic models performed well. In Costa Rica, especially in multilevel, all-aged primary tropical forests, one single tree can hold a tremendous amount of wood and account for a substantial amount of basal area. The occasional inclusion of such trees in 50m sampling segments 5, 10, or 20m wide result in highly influential observations which can affect the predictive equations, increasing MSEs, lowering R^2, and making a logarithmically transformed data set that much more attractive. Likewise, these towering superemergents can dominate sections of airborne laser data; exponentiating estimates of ln(basal area or biomass) lead to gross overestimates. The shorter the segment length, the greater the overestimation problem. In short, in forests which exhibit large spatial variability, where alot of wood/height/girth is tied up in a relatively few large trees, the log transform can exacerbate prediction error. Somewhat paradoxically, the log transform will do this while at the same time improving model fit. In this case, the best predictive model is not necessarily the one with the highest R^2 and the lowest MSE.

- In primary tropical forests, at the least, be leary of using a logarithmic transformation to control variability and to improve model fit. The log model is prone to gross overestimation which is exacerbated at the shorter segment lengths. It is suggested that no transformations should be used in order to improve prediction accuracy. The price paid for not using the log transform is 1) larger MSEs; 2) lower R^2 (i.e., a poorer predictive model fit); and 3) an inability to make valid statistical inferences concerning the significance of the overall regression and individual regression coefficients unless a robust variance-covariance estimator is used.[11]

---

[11] If heteroskedasticity is present, the variance-covariance matrix of the coefficient estimates (i.e., $\hat{\sigma}^2(X'X)^{-1}$) is biased and inconsistent. Robust estimators are available; White (1980) reports on a biased yet consistent estimator of the variance-covariance matrix under conditions where the variability of Y increases with increasing values of X.

In this study, it made little difference whether coefficients were developed using parametric or nonparametric regression techniques. The coefficients, for the most part, differed little. Nonparametric regression techniques have two distinct advantages relative to parametric techniques: 1) Nonparametric regression mitigates the effects of highly influential observations, and should be considered in situations (eg., tropical forests) where HIOs may be prevalent. 2) Nonparametric inferences concerning the statistical significance of the linear regression and individual linear coefficients can be made even if the e_i are not normally distributed. As with parametric inference, however, the e_i must be iid, hence nonparametric inference is as much at a disadvantage as parametric techniques in situations where the variability of Y is a function of X. Nonparametric regression techniques are often not considered, in part due to a decreased statistical efficiency if parametric assumptions are met, a lack of familiarity and understanding on the part of many scientists, a lack of access to nonparametric regression software, and ready access to commercial parametric software. Parametric regression is the standard, the norm. Intuitively, however, nonparametric regression approaches provide more robust answers in terms of being less sensitive to outliers, in terms of a lack of concern over data normality, and in terms of a lack of concern over equality of variances. But the fact is that few commercial statistical software company (eg., SAS, BMDP, SPSS) have nonparametric multiple linear regression capability, and until this situation changes, parametric analyses will continue to be the primary analytical tool.

Generic models were developed on each of the three study areas. Useful variables included average canopy height, all pulses ($\bar{h}_a$), average canopy height, canopy hits only ($\bar{h}_c$), and the coefficients of variation of these two height metrics (c_a, c_c). No other variables were found to be consistently useful (all eight laser variables entered regressions at some point on some study areas) for prediction of basal area, volume, and biomass. In keeping with a finding previously noted that the natural log transform exacerbates prediction error, the following multiple linear models were developed; all forced through the origin:

$$b_s,\ b_t = a_1\ (\bar{h}_a) + a_2\ (c_c)$$

$$v_s,\ v_t = a_3\ (\bar{h}_c) + a_4\ (c_a) + a_5\ (c_c)$$

- Subsequent studies would do well to conduct variable selection procedures on specific study areas(s) considering $\bar{h}_a$, $\bar{h}_c$, c_a, c_c, g, and s_w as independent, linear variables. Special attention must be paid to collinearity problems during regression formulation, since many of these metrics may be highly correlated. Realize that:

$$\bar{h}_c = (\bar{h}_a)\ (g/100.0)$$

$$s_w = (\bar{h}_a)\ (s_c)$$

$$\text{and}\quad c_a = \frac{s_a}{\bar{h}_a}\ ,\quad c_c = \frac{s_c}{\bar{h}_c}$$

These and other mathematical interrelationships cause significant multicollinearity problems. Also, $\bar{h}_a$ and $\bar{h}_c$ are highly correlated and often interchangable as a rule in tropical forest settings, with correlations approaching 1.0 as the forest canopy closes.

- Regressions involving untransformed data logically should be forced through the origin. Scatterplots in this study support this contention.

Based on these results, an inventory procedure can be defined which can be used to assess tropical forest basal area, volume, or biomass using airborne laser data without depending on ground registration of the airborne laser track. The procedure produced estimates within 10-20% of ground-estimated values in that situation where ground-measured canopy conditions were accurately simulated. In situations where canopy conditions are not accurately reproduced in the computer, estimates suffer, and the basal area, volume, and biomass inaccuracies are directly tied to the degree to which the simulated canopy architecture represents reality. The procedure outlined below should be viewed as a strawman, a starting point for an inventory process specific to a given study area.

V.B. Airborne Laser Inventory Procedure

Large area, reconnaissance level airborne laser surveys of tropical forests will have certain characteristics in common. Based on the results of this study and work done by other researchers, a general sampling approach may be outlined in order to produce areal estimates of volume or biomass. Basal area may also be estimated using these techniques, but the predictive equations for basal area have been found to be, in general, weaker, hence the resultant estimate will be less reliable.

Stratification should be considered in order to 1) to pare down a large study area into managable units, and 2) to divide the tract into sectors of relatively homogeneous forest. Regressions to predict volume or biomass will be strongest and most accurate if the species composition and canopy architecture, more specifically forest height structure, are uniform throughout the stratum and the ground sample transects are located in the same forest type. In the Amazon, for instance, preliminary stratification could be done based on 1) vegetation maps which separate nonforest (eg., savannah, wooded grassland) from forest (UNESCO, 1980); 2) soil type maps to further differentiate major forest categories; and 3) land use maps based on current satellite data analysis to identify areas which have been cleared for ranching or which have been colonized.

The orientation of the airborne laser flightlines would vary with stratum shape, but in general, laser transects should be set up as a grid with 1) straight flightlines intersecting at right angles; and 2) with flightlines not aligned with any linear hydrologic, topographic, cultural, vegetative, or edaphic feature. Simulation studies have also shown that a star pattern (three lines 60 degrees apart) will reduce bias in the estimates (Hazard and Pickford, 1986). Within a stratum, if possible, enough transects should be flown such that each can be treated as a single observation or (more likely) so that large enough gaps can be left between consecutive segments to minimize spatial autocorrelation at the crown and stand level. In a study in southwestern Georgia, the effects of what may have been forest stand autocorrelation were noted at approximately 250-300m (Nelson et al., 1988b, Fig. 5). In this Georgia study, autocorrelation effects due to individual tree canopies was mitigated with an 11m gap, the equivalent metric implemented in this study is 15m.

Ground data will be collected and regressions developed in each stratum. On the order of 20 to 30 20m x 50m ground transects (or 15m x 50m transects, slightly larger than the average canopy diameter of a tree) should be collected in each stratum; each will serve as a single observation to develop regressions between simulated laser and ground metrics. The locations of these transects will no doubt be a function of accessibility; the individual transects should be discrete rather than clustered along a centerline. As much as possible the 20 to 30 transects should encompass all of the height-volume or height-biomass variability of a given stratum. Multivariate regression techniques, either parametric or nonparametric, are used to relate average height-all pulses, average height-canopy pulses, and their coefficients of variation to volume or biomass. The linear models are forced through the origin, and the dependent variable should not be transformed. Multicollinearity of the independent laser variables will be a potentially significant problem, since the two height metrics considered are usually highly correlated and may be interchangeable in closed canopy forests.

The ground data collection effort will be appreciable, even in relatively small studies. Considering only one stratum, on the order of 1.5 km of 15-20m wide mapped stand data (up to 3 ha) are necessary to develop the predictive equations. Though not tested in this study, it is suggested that the field effort can be reduced appreciably by using a line intersect ground sampling procedure such as one employed by Meeuwig and Budy (1981) to estimate pinyon-juniper biomass in the southwestern U.S. All trees whose crowns are intersected by a vertical plane through the 50m ground transect are located and sampled. Items noted on each tree intersected include all of the metrics needed to simulate the tree's canopy in the computer, i.e., (X,Y) location, total height, height to first branch, canopy diameter, canopy shape (profile). In addition, these line intersect data can also be used to develop estimates of the ground metrics of interest (biomass, volume, basal area) along the same line. A single transect is run across the simulated data to produce the simulated laser metrics. Results not presented have shown that the simulated height mean and variance estimates from multiple segments traversed by a single simulated laser transect are very similar to estimates developed on those same multiple segments using multiple simulated laser transects. This approach, or one similar to it, should significantly cut back on the amount of field work needed by greatly reducing the number of stems that have to be mapped. Crown measurements are taken and tree bole locations are noted on all trees whose crowns are intercepted by the vertical plane.

The regressions developed for each stratum are used to analyze the airborne laser data collected on that stratum. Airborne laser transects or portions of transects separated by gaps are treated as single observations. These observations are handled as a stratified random sample to develop stratum level and study area level estimates of biomass, volume, or basal area. In areas with little comprehensive information on regional renewable resources, the airborne laser system can provide reconnaissance level non-site specific estimates of the amount of woody and/or leafy material above-ground.

V.C. Future Directions

In two of the three ground data sets, the airborne laser either grossly overestimated or grossly underestimated ground-measured basal area and biomass. This simulation work is by no means proven technology. It is this analyst's opinion that there is no reason that the approach (canopy simulation, regression development, application to airborne laser data) should not work, yet the fact remains that two out of three ground data sets were markedly at odds with similar metrics measured using the airborne laser (eg., average height, canopy density). Prior to dedicating significant resources to large area inventory, additional work should be done, perhaps in temperate forests closer to home, so that ground-laser discrepancies can be definitively addressed. Pacific Coast old growth Doug fir/true fir forests would be a useful test case from the standpoint of incorporating an all-aged height distribution and large trees which, like a tropical forest, could present problems with highly influential observations.

Prior to a field test though, additional computer work must be done to refine ground sample unit sizes. This study identified an appropriate ground transect length to be greater than 25m; 50m, the next largest length considered in this study, became the "correct" answer by default, though the number lies between 25 and 50m. A sensitivity analysis incorporating transect lengths in 5m increments between 25 and 50m is in order. MSE rates for the different regressions should be noted and a length selected as best for Costa Rican primary tropical forests such as those found at LaSelva and Tileran.

A second study involving Monte Carlo simulations will be done in an attempt to determine the appropriate ground-sample width which should be used in order to maintain the spatial information needed to accurately simulate the forest canopy height structure. In the course of this study, it was hypothesized that a 5m wide ground transect was an inadequate stand mapping tool. Many of the dominant and codominant trees which would be sensed by an airborne laser lay beyond the boundaries of this fixed area plot. There is a difference between sampling to characterize the spatial distribution of trees and sampling to develop areal estimates. The simulation study will address the appropriate ground sample width needed to adequately characterize canopy height in the primary tropical forests of Costa Rica.

The first two items mentioned above are aimed at refining the dimensions of an appropriate fixed area ground plot which can be used to map primary tropical forest. In an attempt to reduce the amount of field work needed to develop simulated laser measures of canopy height, field tests comparing a line intersect sampling procedure to the "standard", a fixed area ground transect, are in order. Simulated laser measurements developed using both approaches may be compared to each other and to laser measurements developed using more comprehensive mapped stand data. If the standard laser measurements (eg., heights and coefficients of height variation) prove to be statistically indistinquishable, and if the amount of field effort is much reduced using line intersect sampling, then this more efficient procedure should be substituted. Work savings might be channeled into increasing sample size to develop more robust predictive equations.

Once the appropriate field procedures are worked out, a designed experiment, both in terms of aircraft overflights and in terms of ground sampling procedures, should be conducted on a large study area with the following characteristics: 1) on the order of 100,000 ha or greater; 2) encompasses significant areas of primary forests, secondary forests, and cleared or open areas; 3) requires stratification; 4) independent estimates of biomass, volume, and/or basal area available; 5) allometric biomass or volume equations available; 6) a stable government. This study would be, in essence, a dress rehearsal for large-area airborne laser surveys of the tropics. Candidate study areas include US national parks and/or national forests, especially those of the Pacific Northwest and inland empire; Vancouver Island or national parks in southwestern Canada; forests in southern Mexico, Belize, or Costa Rica. Estimates developed using the airborne laser may be compared with independent estimates from the literature or from a separate field sampling exercise, eg., horizontal point samples or fixed area plots.

Assuming comparable results in the designed field experiment, the technique should be considered for reconnaissance level surveys of remote forested areas worldwide. The technique, if successful, may be applied to the tropical forests of South America, Africa, southeast Asia, and the circumpolar boreal forests to estimate above-ground carbon stores. These results may refine components of the global carbon budget. More importantly, these estimates may serve as carbon store benchmarks for change detection purposes. Airborne aircraft will, or should, be outfitted with GPS receivers whenever any major study is undertaken. Flightlines can be coarsely repeated (within 100m) at future dates to detect regional biomass accumulations or losses <u>without</u> the need to repeat the ground phase. Such technology may prove most useful to detect changes in the above-ground carbon stores of the northern boreal forests, where the most rapid and significant climate and vegetation changes are expected over the next 200 years.

V. LITERATURE CITED

1. Aldred, A.H., and Bonner, G.M., (1985), Application of airborne lasers to forest surveys, Info. Rep. PI-X-51, Tech. Info. and Dist. Center, Petawawa National Forestry Inst., Chalk River, Ontario. 62 p.

2. Arp, H., Griesbach, J.C., and Burns, J.P. (1982), Mapping in Tropical Forests: A New Approach Using the Laser APR, Photo. Eng. and Remote Sens. 48(1): 91-100.

3. Avery, T.E., and Burkhart, H.E., (1983), Forest Measurements, 3rd Ed., McGraw-Hill, New York. 331 p.

4. Berisford, C.W., Mizel, R.F., Kudon, L.H., and Ware, K.D., (1985), Line Intersect Sampling Technique for Estimating Ips Populations in Logging Residue, Integrated Pest Management Research Symposium: The Proceedings, Asheville, N.C., General Tech. Report SO-56, U.S.F.S. Southern Forest Exp. Stat., New Orleans, LA. p. 13-25.

5. Canfield, R.H., (1941), Application of the Line Interception Method in Sampling Range Vegetation, Jour. of Forestry 39(4): 388-394.

6. Currie, D.H., Shaw, V.L., and Berch, F.G., (1989), Integration of Laser Rangefinder and Multispectral Video Data for Forest Measurements, Proc., IGARSS '89, 12th Canadian Symposium on Remote Sensing, IEEE, Vancouver, B.C., Canada, Vol.4, 2382-2384.

7. De Vries, P.G., (1973), A general theory on line intersect sampling, with application to logging residue inventory, Mededelingen Landbouw Hogeschool No. 73-11, Wageningen, Netherlands. 23 p.

8. De Vries, P.G., (1974), Multi-Stage Line Intersect Sampling, Forest Science 20(2): 129-133.

9. De Vries, P.G., (1986), Sampling Theory for Forest Inventory, Springer-Verlag, New York. 399 p.

10. De Vries, P.G., and Van Eijnsbergen, A.C., (1973), Line intersect sampling over populations of arbitrarily shaped elements, Mededelingen Landbouw Hogeschool No. 73-19, Wageningen, Netherlands. 21 p.

11. Hallikainen, M.J., Hyyppa, J., and Tares, T., (1990), Measurements of Forest Backscattering Profiles Using an 8-Channel Ranging Scatterometer, Proc. 10th International Geoscience and Remote Sensing Symposium, IEEE, Washingon, D.C., Vol. II, 1217-1220.

12. Hansen, M.H., (1985), Line Intersect Sampling of Wooded Strips, Forest Science 31(2): 282-288.

13. Hazard, J.W., and Pickford, S.G., (1986), Simulation Studies on Line Intersect Sampling of Forest Residue, Part II, Forest Science 32(2): 447-470.

14. Hickman, G.D., and Hogg, J.E., (1969), Application of an Airborne Pulsed Laser for Near Shore Bathymetric Measurements, Remote Sens. of Envir. 1: 47-58.

15. Hoge, F.E., Swift, R.N., and Frederick, E.B. (1980), Water Depth Measurement Using an Airborne Pulsed Neon Laser System, Applied Optics 19(6): 871-883.

16. Holdridge, L.R., Grenke, W.C., Hastheway, W.H., Liang, T., and Tosi, J.A. Jr., (1971), Forest Environments in Tropical Life Zones, A Pilot Study, Pergamon Press, New York. 732 p.

17. Hollander, M., and Wolfe, D.A., (1973), Nonparametric Statistical Methods, John Wiley and Sons, New York, NY, 503 pgs.

18. Husch, B., Miller, C.I., and Beers, T.W., (1972), Forest Mensuration, 2nd Ed., Ronald Press Co., New York, NY, 410 pgs.

19. Hyyppa, J., Hallikainen, M., and Hyyppa, H., (1992), Forest Inventory Using Forest Canopy Profiles, Proc., IGARSS '92, International Space Year: Space Remote Sensing, IEEE, Houston, Texas, Vol. II: 1216-1219.

20. Hyyppa, J., (1993), Development and Feasibility of Airborne Ranging Radar for Forest Assessment, ISBN 951-22-1888-7, Helsinki University of Technology, Laboratory of Space Technology, Espoo, Finland, 112 pgs.

21. Jackson, T.J., and Ritchie, J.C., (1988), Airborne Laser Profile Data for Measuring Ephemeral Gully Erosion, Photo. Eng. and Remote Sens. 54(8): 1181-1185.

22. Jensen, J.R., Hodgson, M.E., Mackey, H.E. Jr., and Krabill, W., (1987), Correlation Between Aircraft MSS and LIDAR Remotely Sensed Data on A Forested Wetland, Geocarto International 4: 39-54.

23. Kaiser, L., (1983), Unbiased Estimation in Line-Intercept Sampling, Biometrics 39: 965-976.

24. Krabill, W.B., Collins, J.G., Link, L.E., Swift, R.N., and Butler, M.L. (1984), Airborne Laser Topographic Mapping Results, Photo. Eng. and Remote Sens. 50(6): 685-694.

25. Lieberman, D., Hartshorn, G.S., Lieberman, M., and Peralta, R., (1990), Forest Dynamics at LaSelva Biological Station, 1969-1985, In Four Neotropical Rainforests, Gentry, A.H. (ed.), Yale University Press, New Haven, Conn., pp. 509-521.

26. Lieberman, M., and Lieberman, D., (1990), Patterns of density and dispersion of forest trees, In LaSelva: Ecology and Natural History of a Neotropical Rainforest, McDade, L.A., Bawa, K.S., Hespenheide, H.A., and Hartshorn, G.S. (eds.), University of Chicago Press, Chicago, Illinois.

27. Link, L.E., and Collins, J.G., (1981), Airborne Laser Systems Use in Terrain Mapping, 15th Int. Symp. on Remote Sens. of Envir., Ann Arbor, MI. pp. 95-110.

28. Lojan, L., (1966), Una formula para estimar volumenes en un bosque tropical humedo, Turrialba 16(1): 67-72.

29. Maclean, G.A. (1982), Timber Volume Estimation Using Cross-Sectional Photogrammetric and Densitometric Methods, Masters Thesis, University of Wisconsin, Madison. 227 p.

30. Maclean, G., and Martin, G.L., (1984), Merchantable timber volume estimation using cross-sectional photogrammetric and densitometric methods, Canadian Jour. of Forest Research 14(6): 803-810.

31. Maclean, G.A., and Krabill, W.B. 1986. Gross-Merchantable Timber Volume Estimation Using an Airborne LIDAR System, Canadian Jour. of Remote Sens. 12(1): 7-18.

32. 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, NEFES, Upper Darby, PA, 9 pgs.

33. Meeuwig, R.O., and Budy, J.D., (1981), Point and Line-Intersect Sampling in Pinyon-Juniper Woodlands, Gen. Tech. Report INT-104, U.S.F.S., Intermountain Forest and Range Experiment Station, Ogden, Utah. 38 p.

34. Miller, D.M., (1984), Reducing Transformation Bias in Curve Fitting, The American Statistician 38(2): 124-126.

35. Myers, R.H., (1990), Classical and Modern Regression with Applications, 2nd ed., PWS-Kent Publishing Co., Boston, MA., 488 pgs.

36. Nelson, R.F., Krabill, W.B., and Maclean, G.A., (1984), Determining Forest Canopy Characteristics Using Airborne Laser Data, Remote Sens. of Envir. 15: 201-212.

37. Nelson, R., Krabill, W., and Tonelli, J., (1988a), Estimating Forest Biomass and Volume Using Airborne Laser Data, Remote Sens. of Envir. 24: 247-267.

38. Nelson, R., Swift, R., and Krabill, W., (1988b), Using Airborne Lasers to Estimate Forest Canopy and Stand Characteristics, Jour. of Forestry 86(10): 31-38.

39. Nelson, R., and Yaya, J., (1990), Simulating Forest Stands to Derive Laser Profiling Metrics, Proc., 10th IGARSS, IEEE, Washington, D.C., Vol. II: 1221-1226.

40. Paine, D.P., (1977), Introduction to Aerial Photography for Natural Resource Management, O.S.U. Book Stores, Inc., Corvallis, Oregon. 314 p.

41. Pickford, S.G., and Hazard, J.W., (1978), Simulation Studies on Line Intersect Sampling of Forest Residue, Forest Science 24(4): 469-483.

42. Pielou, E.C., (1983), Line Intercept Sampling; Line Intersect Sampling, In Encyclopedia of Statistical Sciences, Vol. 5, Kotz, S., and Johnson, N.L. (eds.), John Wiley and Sons, New York, NY, pp. 70-72.

43. Richie, J.C., and Jackson, T.J., (1989), Airborne Laser Measurements of the Surface Topography of Simulated Concentrated Flow Gullies, Transactions of the ASAE 32(2): 645-648.

44. Ritchie, J.C., Jackson, T.J., Everitt, J.H., Escobar, D.F., and Davis, M.R., (1990), Using Airborne Laser Profiler Data to Determine Shrub Cover in a South Texas Rangeland, Proc., 10th IGARSS, IEEE, Washington, D.C., Vol. I: 577-580.

45. Sader, S.A., (1987), Airborne Remote Sensing of Forest Biomes, Geocarto International 1: 9-18.

46. SAS Institute Inc., (1989), SAS/STAT User's Guide, Ver. 6, 4th Ed., Vol. I, II, SAS Institute, Inc., Cary, NC, 943 pgs, 846 pgs, resp.

47. Schreier, H., Lougheed, J.,Gibson, J.R., and Russell, J., (1984), Calibrating an Airborne Laser Profiling System, Photo. Eng. and Remote Sens. 50(11): 1591-1598.

48. Schreier, H., Lougheed, J., Tucker, C., and Leckie, D., (1985), Automated Measurements of Terrain Reflection and Height Variations Using an Airborne Infrared Laser System, Int. Jour. of Remote Sens. 6(1): 101-113.

49. Smith, D.V., (1969), Timber Volume Estimation using Kelsh Plotter and Densitometry, Ph.D. Thesis, State University of New York at Syracuse, New York. 182 p.

50. Solodukhin, V.I., Kulyasov, A.G., Utenkov, B.I., Zhukov, A.Ya., Mazhugin, I.N., Emel'yanov, V.P., and Korolev, I.A., (1977a), S"emka profilya krony dereva s pomoshch'yu lazernogo dal'nomera (Drawing the crown profile of a tree with the aid of a laser), Lesnoe Khozyaistvo 2: 71-73.

51. Solodukhin, V.I., Zhukov, A.Ya., Mazhugin, I.N., Bokova, T.K., and Polezhai, V.M., (1977b), Vozmozhnosti lazernoi aeros"emki profilei lesa (Possibilities of laser aerial photography of forest profiles), Lesnoe Khozyaistvo 10: 53-58.

52. Solodukhin, V.I., Mazhugin, I.N., Zhukov, A.Ya., Narkevich, V.I., Popov, Yu.V., Kulyasov, A.G., Marasin, L.E., and Sokolov, S.A., (1979), Lazernaya aeros"emka profilei lesa (Laser aerial profiling of forest), Lesnoe Khozyaistvo 10: 43-45.

53. Solodukhin, V.I., Zheludov, A.V., Mazhugin, I.N., Bokova, T.K., Shevchenko, K.V., (1985), The processing of laser profilograms for forest mensuration, Lesnoe Khozyaistvo 12: 35-37.

54. Spurr, S.H., (1960), Photogrammetry and Photo-interpretation, 2nd Ed., Ronald Press, New York.

55. Stolyarov, D.P., and Solodukhin, V.I., (1987), Laser Forest Survey, Izvestiya Vysshikh Uchebnykh Zavedenii, Lesnoi Zhurnal 5: 8-15.

56. UNESCO, (1980), Vegetation Map of South America, 1:5,000,000, FAO/UNESCO, Rome Italy.

57. Van Wagner, C.E., (1968), The Line Intersect Method In Forest Fuel Sampling, Forest Science 14(1): 20-26.

58. Warren, W.G., and Olsen, P.F., (1964), A Line Intersect Technique for Assessing Logging Waste, Forest Science 10(3): 267-276.

59. Warren, W.G., (1990), Line Intersect Sampling: An Historical Perspective, In State-of-the-Art Methodology of Forest Inventory: a Symposium Proceedings, Labau, V.J., and Cunia, T., (eds.), USDA-Forest Service Gen. Tech. Report PNW-GTR-263, USFS-Pacific Northwest Region, Portland, OR, pp. 33-38.

60. White, H., (1980), A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, Econometrics 48: 817-838.

VITA

Name:	Ross F. Nelson
Present Position:	Physical Scientist Code 923, Biospheric Sciences Branch, NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771
Research Experience:	The use of AVHRR and Landsat digital data for monitoring tropical and subtropical forest conversion, development of statistical and image processing procedures for subcontinental and continental resource assessment, use of airborne laser data to characterize the forest canopy and to estimate forest biomass and timber volume, use of aircraft and satellite scanner data for delineating insect defoliation.
Education:	1974 - B.S. Forestry University of Maine, Orono, Maine 1979 - M.S. Forestry/Remote Sensing Purdue Univerity, West Lafayette, Indiana
Previous Positions:	1977 - 1979 Research Assistant Purdue University

Ross F. Nelson

Ross F. Nelson