AIRBORNE LIDAR-BASED ESTIMATES OF TROPICAL FOREST STRUCTURE AND GROUND TOPOGRAPHY IN A MOUNTAINOUS AREA OF THE BRAZILIAN ATLANTIC FOREST

Veronika Leitold

Master's Dissertation of the Graduate Course in Remote Sensing, guided by Dr. Yosio Edemir Shimabukuro, approved on April 16, 2014.

URL of the original document:
<http://urlib.net/8JMKD3MGP5W34M/3FU3382>

sid.inpe.br/mtc-m21b/2014/03.17.04.12-TDI

AIRBORNE LIDAR-BASED ESTIMATES OF TROPICAL FOREST STRUCTURE AND GROUND TOPOGRAPHY IN A MOUNTAINOUS AREA OF THE BRAZILIAN ATLANTIC FOREST

Veronika Leitold

Master's Dissertation of the Graduate Course in Remote Sensing, guided by Dr. Yosio Edemir Shimabukuro, approved on April 16, 2014.

URL of the original document:
<http://urlib.net/8JMKD3MGP5W34M/3FU3382>

INPE

São José dos Campos

2014
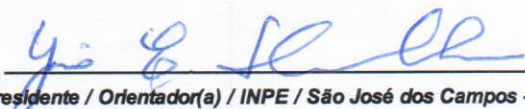
Aprovado (a) pela Banca Examinadora em cumprimento ao requisito exigido para obtenção do Título de **Mestre** em

Sensoriamento Remoto
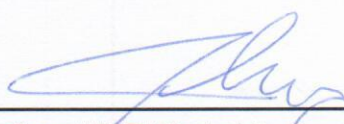
Dr.   Yosio Edemir Shimabukuro

Presidente / Orientador(a) / INPE / São José dos Campos - SP

Dr.   Luiz Eduardo Oliveira e Cruz Aragão

Membro da Banca / INPE / São José dos Campos - SP

Dr.   Thiago Sanna Freire Silva

Membro da Banca / INPE / São José dos Campos - SP

Dr.   Michael Maier Keller

Convidado(a) / USDA Forest Service / Estados Unidos - USA

Este trabalho foi aprovado por:

( ) maioria simples

(x) unanimidade

Aluno (a):   *Veronika Leitold*

São José dos Campos, 16 de Abril de 2014

Mata Atlântica – Serra do Mar

ACKNOWLEDGEMENTS

ABSTRACT

Quantification of tropical forest biomass and characterization of forest structure at fine scales is critical for a better understanding of the role of tropical ecosystems in the global carbon cycle. LiDAR remote sensing is a powerful tool for assessing 3D vegetation structure and estimating aboveground forest biomass, provided that LiDAR measurements penetrate dense forest vegetation to generate accurate estimates of surface topography and canopy heights. Dense tropical forest canopies present various challenges for LiDAR remote sensing, especially in areas of steep topography where much of the remaining Atlantic Forest is concentrated. Airborne LiDAR data were acquired from a commercial provider for a region of the Serra do Mar State Park in the state of São Paulo, Brazil, a mountainous area with steep slopes covered by mature tropical dense forest. Digital terrain models (DTMs) derived from all LiDAR data were compared to 35 ground control points measured with survey grade GNSS receivers, post-processed with differential correction. The two LiDAR-based terrain models were extremely accurate, with mean signed errors of 0.19 m (± 0.97 m) and 0.18 m (± 0.95 m) compared to ground points. Random thinning of the original LiDAR point density (20 points/m^2) decreased the accuracy of the terrain models, with signed errors rising to 0.38 m (± 1.32 m), 1.12 m (± 2.04 m), 1.59 m (± 3.13 m) and 3.21 m (± 3.12 m) as point density was reduced to 8, 4, 2 and 1 points/m^2, respectively. Offsets between LiDAR DTMs and ground data in submontane areas were consistently higher than those in montane areas, possibly reflecting the varying complexity of the terrain and the effects of variable ranging distance. Canopy heights calculated from the thinned LiDAR data also differed significantly from canopy heights estimated with the full LiDAR density. Mean canopy surface height decreased by 3%, 8%, 16% and 25% as return density was reduced to 8, 4, 2 and 1 points/m^2, respectivel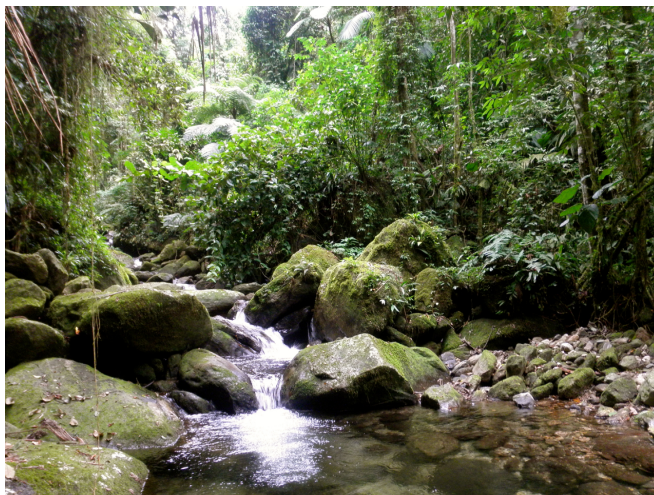y. The magnitude of change in canopy height was greater in submontane plots (range of 0.79-6.08 m) than in montane plots (range of 0.60-4.24 m) within all data density classes. Low variation in vegetation surface model elevations with reduced data density indicated that the decrease in canopy heights was due to the difficulty to characterize topography precisely and not the inability to capture the top heights of trees. Metrics of canopy structure and terrain characteristics derived from the full-density LiDAR data were significantly related to aboveground biomass components in 1-ha permanent plots at the study site. Canopy gap fraction showed close correspondence with aboveground biomass loss ($R^2 = 0.87$) and net change ($R^2 = 0.91$), while mean canopy surface height explained 43% of the variation in total aboveground biomass and 68% of the variation in total basal area. Both canopy gap fraction and rates of mortality were lower on steep terrain than on gentler slopes, suggesting that steeper slopes might provide more favorable conditions (nutrient, water and light availability) for tree growth and biomass accumulation than flat terrain. Given the

growing emphasis on the use of airborne LiDAR for forest management and conservation efforts (REDD+), the results of this study highlight the importance of careful survey planning and consistent sampling frames for accurate quantification of aboveground biomass stocks and dynamics. In mountainous terrain under closed-canopy tropical forest, such as the Atlantic Forest of the Serra do Mar, low-density LiDAR coverage will underestimate biomass and might not characterize truthfully the spatial heterogeneity of forest structure.

ESTIMATIVAS DA ESTRUTURA DA FLORESTA TROPICAL E DA TOPOGRAFIA DO TERRENO EM UMA ÁREA MONTANHOSA DA MATA ATLÂNTICA BRASILEIRA UTILIZANDO LIDAR AEROTRANSPORTADO

RESUMO

Quantificar a biomassa florestal tropical e caracterizar a estrutura da floresta em escalas finas é fundamental para compreender melhor o papel dos ecossistemas tropicais no ciclo global de carbono. O sensoriamento remoto por LiDAR é uma ferramenta poderosa para avaliar a estrutura da vegetação em 3D e estimar a biomassa acima do solo, considerando que as medições do LiDAR penetram na vegetação densa da floresta gerando estimativas precisas da topografia e das alturas das árvores. Florestas tropicais com dossel denso apresentam vários desafios para o sensoriamento remoto por LiDAR, especialmente em áreas de topografia acidentada, onde grande parte das remanescentes da Mata Atlântica se encontra. Dados de LiDAR aerotransportado foram adquiridos sobre o Parque Estadual da Serra do Mar - SP, Brasil, em uma área montanhosa com declives íngremes cobertas por floresta tropical densa. Modelos digitais de terreno derivados do LiDAR foram comparados com 35 pontos de controle medidos com receptores GNSS de alta precisão, pós-processados com correção diferencial. Os dois modelos de terreno obtidos do LiDAR mostraram-se extremamente precisos, com erros médios de 0,19 m (± 0,97 m) e 0,18 m (± 0,95 m) em relação aos pontos de controle. Redução aleatória da densidade original de pontos de LiDAR (20 pontos/m^2) resultou em diminuição da precisão dos modelos de terreno: os erros médios subiram para 0,38 m (± 1,32 m), 1,12 m (± 2,04 m), 1,59 m (± 3,13 m) e 3,21 m (± 3,12 m) com as densidades de 8, 4, 2 e 1 pontos/m^2, respectivamente. As diferenças de elevação entre o modelo de terreno e os pontos de controle em áreas submontanas foram consistentemente mais altas do que as diferenças em áreas montanas, possivelmente refletindo a complexidade variável do terreno e os efeitos da altura do vôo. A altura do dossel calculada a partir dos dados de LiDAR com densidade reduzida diferiu também significativamente da altura do dossel estimada a partir dos dados com densidade original. A altura média da superfície do dossel diminuiu de 3%, 8%, 16% e 25% com densidades de 8, 4, 2 e 1 pontos/m^2, respectivamente. A magnitude da mudança na altura do dossel foi maior nas parcelas submontanas (entre 0,79-6,08 m) do que nas montanas (entre 0,60-4,24 m) dentro de todas as classes de densidade de dados. A baixa variação das elevações do modelo da superfície da vegetação com dados reduzidos indicou que a diminuição da altura do dossel foi devido a dificuldade de caracterizar a topografia e não a incapacidade de capturar as alturas máximas das árvores. Métricas de estrutura do dossel e características do terreno derivadas dos dados de LiDAR (densidade original) foram bons

preditores de componentes da biomassa acima do solo em parcelas permanentes de 1 hectare dentro da área de estudo. A fração de aberturas no dossel foi fortemente relacionada a perda de biomassa acima do solo ($R^2 = 0,87$) e a variação líquida ($R^2 = 0,91$), enquanto que a altura média da superfície do dossel explicou 43% da variação na biomassa total acima do solo e 68% da variação na área basal total. A fração de aberturas no dossel e as taxas de mortalidade foram menores em terrenos íngremes do que em encostas suaves, sugerindo que encostas mais íngremes podem proporcionar condições mais favoráveis (de nutrientes, de água e disponibilidade de luz) para o crescimento das árvores e acúmulo de biomassa. Dada a crescente ênfase sobre o uso de LiDAR aerotransportado para manejo florestal e esforços de conservação (REDD+), os resultados deste estudo destacam a importância do planejamento cuidadoso de levantamentos futuros com amostragem consistente para a quantificação precisa dos estoques e dinâmica da biomassa acima do solo. Em terreno montanhoso coberto com densa vegetacão tropical, como a Mata Atlântica da Serra do Mar, dados de LiDAR de baixa densidade vão subestimar a biomassa e podem não caracterizar verdadeiramente a heterogeneidade espacial da estrutura da floresta.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ALTM   Airborne Laser Terrain Mapping
ASPRS   American Society for Photogrammetry and Remote Sensing
ATIN   Adaptive Triangulated Irregular Network
CHM   Canopy Height Model
DART   Discrete Anisotropic Radiative Transfer
DSM   Digital Surface Model
DTM   Digital Terrain Model
EMBRAPA   Brazilian Enterprise for Agricultural Research
FAPESP   Foundation for Research Support of the State of São Paulo
FUNARBE   Arthur Bernardes Foundation
GIS   Geographic Information System
G-LiHT   Goddard's LiDAR, Hyperspectral and Thermal Imager
GLONASS   Global Navigation Satellite System
GPS   Global Positioning System
GSFC   Goddard Space Flight Center
INCRA   National Institute for Colonization and Agrarian Reform
INPE   National Institute for Space Research
InSAR   Interferometric Synthetic Aperture Radar
LAS   LASer file format
LiDAR   Light Detection and Ranging
NASA   National Aeronautics and Space Administration
NAVSTAR   Navigation Signal Timing and Ranging
PESM   Serra do Mar State Park
REDD   Reduced Emissions from Deforestation and Forest Degradation
SAR   Synthetic Aperture Radar
SBET   Smoothed Best Estimate of Trajectory
TIN   Triangular Irregular Network
TM   Thematic Mapper
USFS   United States Forest Service
UTM   Universal Transverse Mercator
WGS84   World Geodetic System 1984

LIST OF SYMBOLS

a.s.l.	above sea level
AGB	Aboveground Biomass (Mg ha^{-1})
AGH	Aboveground Height (m)
CBH	Circumference at Breast Height (cm)
DBH	Diameter at Breast Height (cm)
D	Diameter (cm)
df	Degrees of freedom
H	Height (m)
H_0	Null hypothesis
H_1	Alternative hypothesis
MCH	Mean Canopy Height (m)
MSE	Mean Signed Error
OLS	Ordinary Least Squares
R^2	Coefficient of determination
Res.SE	Residual Standard Error
RMSE	Root Mean Square Error
stdev	Standard deviation
TCH	Top-of-Canopy Height (m)
z	Elevation (m)
λ	Scaling exponent of the power-law distribution
$\zeta(\lambda)$	Riemann Zeta function
μ	Mean
ρ	Wood density (g cm^{-3})
σ	Standard deviation

CONTENTS

1. INTRODUCTION

The Atlantic Forest is the second most widespread tropical rainforest in Brazil. It is considered a hotspot of biodiversity and endemism, and it provides significant ecosystem services, such as maintenance of natural carbon stocks and water flow regulation (MYERS et al., 2000). As a result of extensive fragmentation and degradation under anthropogenic pressure over the course of the past centuries, almost 90% of the original Atlantic Forest vegetation has been lost (RIBEIRO et al., 2009). The remaining forest cover is currently distributed over approximately 245,000 fragments, the largest of which is located in the *Serra do Mar* along the coastal mountains of the state of São Paulo, Brazil, where only the steepest slope areas were left intact (VIEIRA et al., 2008). This single fragment contains about 1,110,000 ha of continuous forest, accounting for 7% of the total remaining extent of the Atlantic Forest today, and storing a substantial amount of carbon aboveground (RIBEIRO et al., 2009).

Mapping the extent and geographic distribution of the Atlantic Forest remnants has been the focus of various studies over the past decades, but detailed information on the three-dimensional structure and related function of the vegetation in these areas still remains limited (VIEIRA et al., 2008). The effects of short, steep elevational gradients on tropical forest structure and growth patterns, for example, are not well known. Significant differences in forest size structure and biomass stocks have been found among sites over short distances along an elevational gradient in the *Serra do Mar* by Alves and collaborators (2010). Their results suggest that local topographic variation and associated differences in light environment and nutrient availability are likely to have a large effect on growth patterns and biomass accumulation in this coastal Atlantic Forest. For a deeper ecological understanding of the structure and function of this highly diverse and increasingly threatened ecosystem, further research using improved techniques is needed.

One of the most difficult components to quantify in complex tropical forest ecosystems is overall forest structure. The vertical structure of the forest, defined as "the bottom to top configuration of aboveground vegetation within a forest stand" is particularly difficult to assess and quantify over large and remote areas or on complex terrain, although it is certainly important for a better understanding of forest functioning (ZIMBLE et al., 2003). Changes in vertical forest structure affect both microclimatic patterns and processes directly, and have been shown to impact the behavior and distribution of various species in forest ecosystems (BROKAW; LENT, 1999). Understanding how environmental factors constrain species distributions and affect growth and mortality rates is fundamental to predicting the likely impacts of changes in land-use and global climatic conditions on tropical rainforests.

Evidence suggests that light limitation driven by canopy closure and gap distribution plays a strong role in ecosystem dynamics in tropical forests, controlling tree regeneration, forest dynamics and forest diversity (UNGER et al., 2013). Of all environmental factors affecting tropical forest communities (e.g. temperature, light, water, nutrients), sunlight is considered one of the most limiting resources for plant growth (PEARCY, 2007). Within species, higher light interception tends to lead to higher growth rates and enhanced survival (KRUGER; VOLIN, 2006), as well as greater canopy openness and larger branch extension (STERCK; BONGERS, 2001). The productivity and ecosystem role of individuals of different sizes may also depend on the varying levels of light penetration and shading within the forest canopy (STARK, 2012). Developing a detailed understanding of the role of light in forest dynamics requires precise information on the distribution of vegetation elements (leaves, branches, stems) in the forest canopy, as well as on the spatial variation of light itself.

Traditional methods for assessing forest canopy structure are based either on field measurements or photographic and/or photogrammetric interpretations (WULDER, 1998). The drawback of field-based inventories is that they cannot provide spatially continuous information over a large area, while the usefulness of photo interpretation in mountainous terrain is hampered by different illumination and shading effects. The use of Light Detection and Ranging (LiDAR) instrumentation in its various forms for forestry applications is one of the most developed non-traditional capabilities of the technology. The ability of small-footprint airborne LiDAR to penetrate gaps in forest canopies and reach the ground below allows deriving detailed digital terrain models (DTMs), as well as the characteristics of the overlaying vegetation (MAIER et al., 2008). This active optical remote sensing method, therefore, offers a novel way of describing forest structure in three dimensions, and allows for measuring the variation in vegetation distribution at fine scales.

Research shows that various forest structural attributes can be directly retrieved from 3D LiDAR point cloud data, such as canopy height and thickness, subcanopy topography, and the vertical distribution of canopies (LEFSKY et al., 2002a; LOVELL et al., 2003; PARKER et al., 2004). LiDAR metrics derived from the raw point cloud can be subsequently used in empirical models to predict biophysically important forest attributes such as basal area, mean stem diameter and aboveground biomass (HOLMGREN et al., 2003; NAESSET, 2002; LEFSKY et al., 1999a; NILSSON, 1996). Moreover, by examining the three-dimensional distribution of canopy elements and gaps, indices of spatial structure can be developed with direct relevance to important ecological parameters such as light transmittance, while the canopy gap distribution itself is useful in predicting understory conditions (GAULTON; MALTHUS, 2010). It is important to emphasize that the extraction of vegetation characteristics from LiDAR data can be largely influenced by variable measurement conditions, such as laser point density or footprint size, therefore, models have to be calibrated and validated with field data to be able to produce accurate estimates

of forest parameters. This has been a challenging task in tropical forests given the complexity of the forest structure and the high spatial and temporal variability in vegetation density.

The primary goal of this study is to assess the potential of small-footprint airborne LiDAR to provide the necessary high-resolution information for an in-depth understanding of the structure of the forests on Atlantic coastal hillslopes and for accurate estimation of aboveground forest biomass dynamics in mountainous terrain. With its complex topography along a steep elevational gradient and covered by dense multilayered forest canopy, the *Serra do Mar* is unlike most of the areas considered in previous LiDAR forestry studies. Additionally, this site is of particular interest from an ecological point of view, given the existence of a permanent plot network of 1-ha forest inventory plots within the *Serra do Mar* State Park that was installed in 2006 and monitored since then by the BIOTA Project. This valuable biometry survey database from BIOTA in combination with the high-density airborne LiDAR data collected by the Sustainable Landscapes Project in a 1000-ha area overlapping the field plots allows for an in-depth study of canopy structure and biomass dynamics in the Atlantic Forest of the *Serra do Mar*. Based on the available data, this work specifically addresses three main objectives:

[1] The first objective is to evaluate the accuracy of two different DTMs derived from a set of airborne LiDAR data collected in the mountainous area of the *Serra do Mar* in the coastal Atlantic Rainforest in SE Brazil. Samples of the LiDAR DTM raster grids are compared with manually measured control points using survey-grade differential GPS. Statistical analysis of the measured and modeled elevations is performed, and the error sources are investigated.

[2] The second objective of the study is to assess the changes in accuracy of the DTMs generated after synthetic thinning of the LiDAR data to four predefined return density levels, and how the changes in terrain accuracy affect the corresponding LiDAR-derived canopy height metrics. The goal is to answer the question: what is the minimum LiDAR return density required to generate a DTM at an acceptable accuracy level for derivation of forest structure metrics and biomass estimates at a mountainous site like the *Serra do Mar*?

[3] The third objective is to investigate the relationships between LiDAR-derived canopy structure (canopy height and gap fraction) and field-based estimates of aboveground biomass and biomass change (gain, loss and net change) in submontane and montane forest areas of the *Serra do Mar* and to examine how these relationships are influenced by terrain characteristics, such as elevation, slope and aspect, and related environmental factors (e.g. cloud cover, illumination).

The dissertation document is divided into seven sections and organized as follows: after the introduction in Chapter 1, the fundamentals of LiDAR remote sensing are presented in Chapter 2, followed by the description of the study area in Chapter 3. Chapters 4, 5 and 6 each address one of the specific objectives listed above, with each section including a brief introduction, materials and methods, results and discussion, and conclusions specific to the given topic. Finally, Chapter 7 summarizes the main findings of the previous three chapters and concludes the work as a whole, also giving recommendations for future research.

2. LiDAR FUNDAMENTALS

2.1 Sensor characteristics and mode of operation

Light Detection and Ranging (LiDAR) is a commercially available active remote sensing technology that, like aerial photography in the 1970s and 1980s, is rapidly expanding throughout the world for use across a range of environmental sectors (ASNER et al., 2010; HUDAK et al., 2009). The principal advantages of the technology over conventional surveying methods are the provision of precise x, y, z measurements, the high level of automation, rapidity of coverage, and fast delivery time for often extremely large datasets, whereas the main limitation for research utilizing this technology remains the high cost of the instrument, operation, and associated software (HERITAGE; LARGE, 2009).

LiDAR sensors utilize the properties of scattered light to determine the range of distant objects. The working mechanism of the system consist in firing a narrow laser beam at a distant surface and recording the time required for the emitted pulse to travel from the laser source to the target object and back to the detector (so-called *time-of-flight*). A highly accurate clock is used to measure this time, and the distance is calculated by the following relationship: Distance = (Speed of Light × Time of Flight) ÷ 2. The laser system can be operated either from the ground, pointing up, or from airborne or space-borne platforms, pointing down. Typical flying height for airborne LiDAR is 1 - 2 km, and the system is used in combination with GPS instruments and inertial navigation systems for locating the source of the return signal in three dimensions and correcting for the effects of pitch, roll and yaw introduced by the platform motion.

The type of information collected from the LiDAR return signals distinguishes two broad categories of sensors, (i) discrete-return devices that typically measure one, two or a small

number of heights, and have a footprint size of 0.2 m - 0.9 m (*small-footprint LiDAR*), and (ii) waveform-recording sensors that measure the amount of energy returned to the sensor for a series of equal time intervals, and have an average footprint size of 8 m - 70 m (*large-footprint LiDAR*).

For terrestrial applications and vegetation studies, the pulse wavelength is usually in the near infrared part of the electromagnetic spectrum (with the most common being 1064 nm), where vegetation reflectance is high, as this guarantees a relatively strong return signal. Also, atmospheric transmittance is high at these wavelengths, ensuring minimal loss of signal from atmospheric scattering and absorption. One drawback of working in this range of wavelength is absorption by clouds, which impedes the use of these devices during overcast conditions (LEFSKY, 2002a). Also, because the power of the laser diminishes as the square of the distance traveled, laser power determines the effective operating range of a LiDAR system, as well as the extent to which it will penetrate vegetation canopies.

The major operational specifications of a LiDAR system include:

(1) *scanning frequency*, the number of pulses emitted by the laser instrument in 1 second;

(2) *scanning pattern*, the spatial arrangement of pulse returns determined by the mechanism used to direct pulses across the flight line;

(3) *beam divergence* that occurs as the distance between the laser instrument and a plane that intersects the beam axis increases;

(4) *scanning angle*, the angle by which the beam axis is directed away from the focal plane of the laser sensor;

(5) *scanning swath* determined by the aboveground flight height and the scanning angle;

(6) *footprint diameter* which is a function of both beam divergence and the above-target flight height;

(7) *pulse length,* which is the duration of the pulse;

(8) *number of returns per pulse*;

(9) *footprint spacing*, which, along with the beam divergence, determines the spatial resolution of LiDAR data;

(10) various *discretization settings* that are specifications integral to the processing of the backscattered energy of a pulse to identify individual returns, control the minimum energy amplitude necessary to produce a return and determine the minimal distance between consecutive returns from the same pulse (GATZIOLIS; ANDERSEN, 2008).
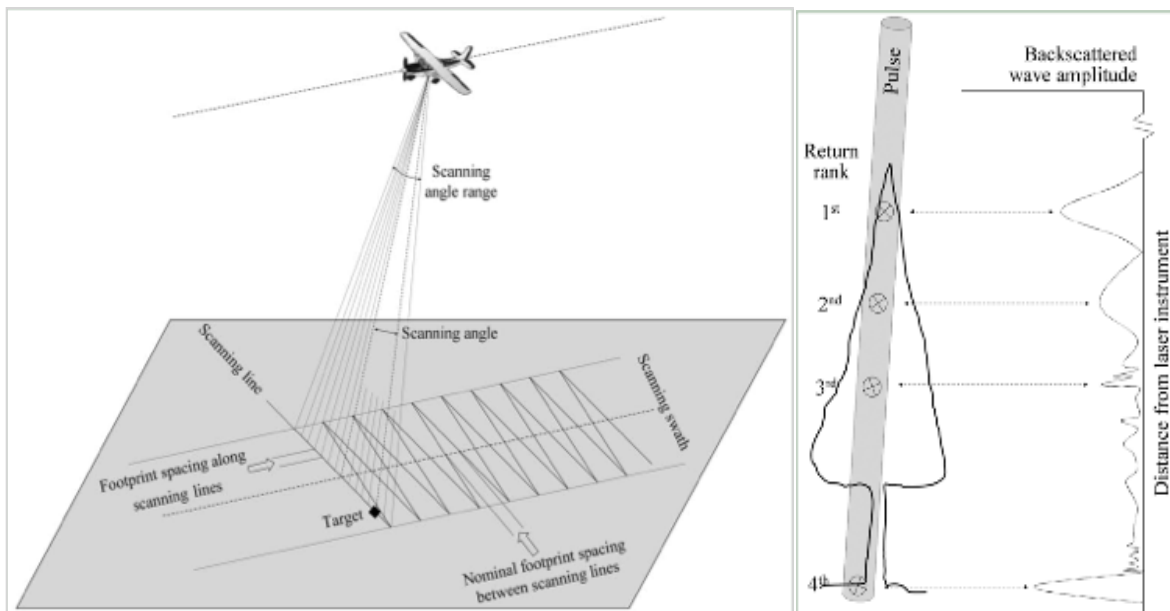
Figure 2.1 Illustration of the scanning attributes of LiDAR data acquisition (on left) assuming seesaw scanning pattern and flightline parallel to the ground; and the discretization process used to identify individual returns by processing the backscattered energy of a laser pulse (on right).
Source: Adapted from Gatziolis and Andersen (2008).

2.2 Data attributes, processing and products

Small-footprint LiDAR data comprise a set of return coordinates in three dimensions, commonly referred to as *point cloud*, with each return usually carrying attribute values that relate either to that return or to the pulse from which the return was generated:

(1) *pulse density* is a direct function of the footprint spacing over a hypothetical flat plane, and is the most consistent measure of the spatial resolution of a LiDAR data set;

(2) *return density* is the mean number of returns in the data set present in a unit square area in 2D, typically 1 m^2;

(3) *return intensity* describes the strength of the beam backscattering pertaining to the return in question and depends on the reflectance properties of the target, hence, it can potentially be used in target discrimination;

(4) *return number* is the rank of a return among those generated from one beam.

Other attributes, which a return inherits from its parent beam, include the *scan angle* (º) and those sometimes assigned at the data post-processing phase such as *indices to flight lines*, or *GPS time*, which is an indication of the precise time that a pulse was emitted and can be used as a unique identifier for a pulse (GATZIOLIS; ANDERSEN, 2008).

LiDAR point cloud data are generally stored in LAS binary file format following the American Society for Photogrammetry and Remote Sensing (ASPRS) format standards (RENSLOW, 2012), including x-coordinate, y-coordinate, elevation, return number, intensity and scan angle for each return, at a minimum. Additional information is usually organized in the form of metadata, and often contains spatial geographic information system (GIS) layers with the spatial extent of the data acquisition, flight lines, the date and time range, the model and characteristics of the LiDAR instrument, etc.

The processing of raw LiDAR data involves the following basic steps and products:

(1) *Classification* of non-ground and ground returns through the application of a classification model or filtering algorithm to the point cloud data;

(2) *Digital Terrain Model* (DTM) generation through interpolation of the classified ground returns, representing the ground elevation surface;

(3) *Above Ground Height* (AGH) assigned to each classified non-ground point relative to the DTM surface by subtracting the interpolated ground surface elevation from the point's elevation;

(4) *Canopy Height Model* (CHM) generation through interpolation of elevations using all or first LiDAR returns, representing the surface of the outer forest canopy in the landscape;

(5) *Analysis* through which various metrics are derived from the LiDAR point cloud representing the spatial distribution of points. These parameters fall into two groups: (i) metrics calculated on numeric values (i.e., intensity, elevation, canopy height), and (ii) metrics representing density of points (i.e., canopy density, stratified canopy density, etc.) (EVANS et al., 2009).

3. STUDY AREA

The study area is located within the São Paulo State Park of *Serra do Mar* (PESM) (23º34'S and 45º02'W; 23º17'S and 45º11'W) in southeast Brazil (Figure 3.1). It is characterized by complex terrain along an altitudinal gradient (0-1200 m a.s.l.) and is covered by the dense vegetation of the Atlantic Forest. The state park contains 47,500 ha of diverse tropical ecosystems, including mangroves, *restinga* (sandy coastal plain forests), and humid tropical forests subdivided into vegetation types by altitude – lowland, submontane and montane forest – from sea level up to 1200 m elevation (SMA, 1998). The forest canopy is irregular, with mean heights ranging from 17 to 22 m and rarely emergent trees reaching 35 m. Due to this irregularity of the canopy, the amount of light that gets through sets conditions for the development of hundreds of epiphytic species (JOLY et al., 2012).

Mean annual rainfall in the area is approximately 2500 mm with no dry season, and from 400 m above sea level up to the top of slopes the mountains are covered by a dense fog almost daily (JOLY et al., 2012). Monthly average temperature ranges from 17.6 °C to 24.7 °C varying according to altitude (SCARANELLO et al., 2012). Soils in lowland, submontane and montane forest sites are classified as sandy-loam *Inceptisols* resting on a crystalline basement with predominance of gneiss, granite and migmatite (ALVES et al., 2010).

Within the study area, the proportion of trees and palms shows an inverse relationship, wherein the proportion of trees decreases and that of palms increases with altitude. Across the site, trees prevail (71 - 90%), followed by palms (10 - 25%) and tree ferns (0 - 4%), and results presented by Joly and colleagues (2012) show an extraordinary richness and diversity of tree species of the Atlantic Rainforest in this region (Shannon-Weiner index

ranging from 3.96 to 4.48 along the hillslope). The most rich families in these forests sites are: Myrtaceae, Rubiaceae, Fabaceae, and Lauraceae, while common canopy tree species include: *Hieronyma alchorneoides* Allemão, *Virola bicuhyba* (Schott ex Spreng.) Warb, *Eriotheca pentaphylla* (Vell.) A. Robyns, *Sloanea guianensis* (Aubl.) Benth, *Cryptocaria mandiocanna* Meisn. and *Ecclinusa ramiflora* Mart. (SCARANELLO et al., 2012).
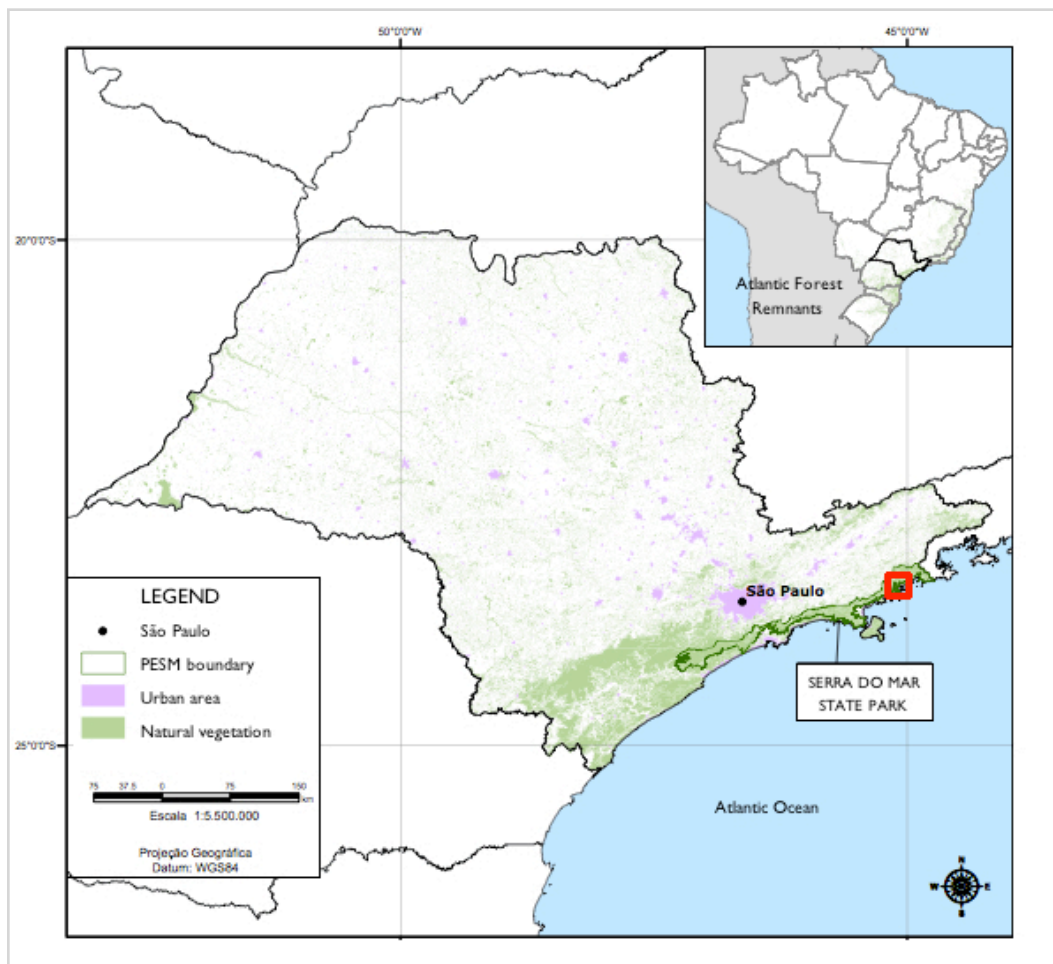
Figure 3.1 Geographical position of the *Serra do Mar* State Park (PESM) – solid green line and shading – and the location of the study area within its bounds – indicated in red.
Source: Serra do Mar State Park Management Plan, Secretary of Environment of the State of São Paulo, Forestry Institute, 2006.

The approximately 1000-hectare area within the PESM where LiDAR data were collected for this study (Figure 3.2) contains nine permanent forest inventory plots that belong to a 14-plot network established along the altitudinal transect in the *Serra do Mar* (JOLY et al., 2012) to evaluate forest diversity and ecosystem function variation. Out of these nine plots within the LiDAR coverage, one plot is located in the lowland forest at an elevation of 100 m (Plot F), four plots in the submontane forest at an average elevation of 250 m (Plots G, H, I and J), and four plots in the montane forest at about 1000 m a.s.l. (Plots K, L, M and N). The permanent plots were designed such that they each have a projected area of one hectare, being rectangular in shape and with 100-meter-long edges.
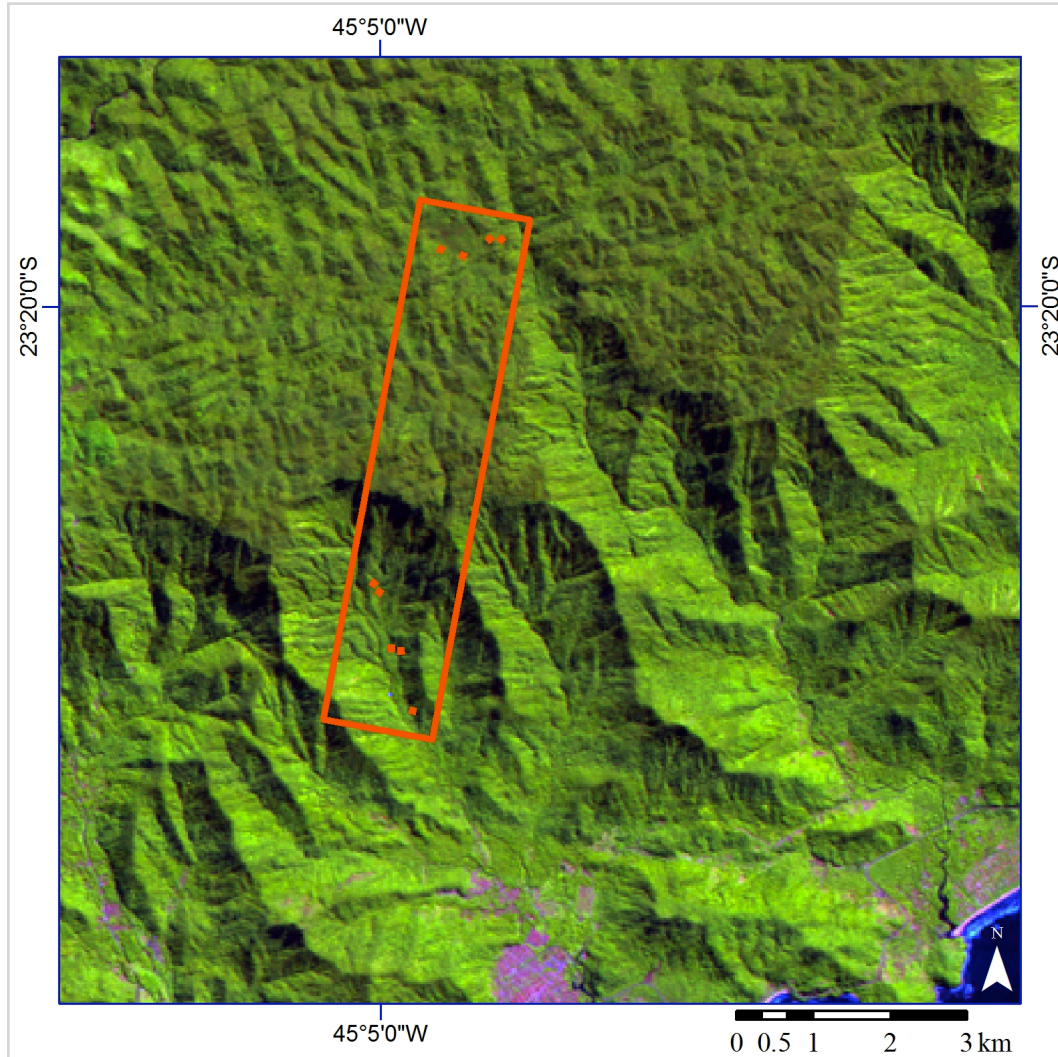
Figure 3.2 Location and extent of the LiDAR coverage – red rectangle – and the nine permanent field plots – red points – within this area of the Atlantic Forest over a section of Landsat-TM scene 218/76 (path/row) to the north of the municipality of Ubatuba.
Source: INPE *Catálogo de Imagens*.

4. DIGITAL TERRAIN MODEL ACCURACY ASSESSMENT

4.1 Background

Digital Terrain Models (DTMs) are a computational representation of the ground surface topography, commonly obtained by remote sensing techniques such as stereo photogrammetry, interferometric synthetic aperture radar (SAR), and more recently, by airborne laser technology or LiDAR. The most common forms of DTMs are raster grids, triangular irregular networks (TINs), and contour line models (LIU, 2008), and they are used in various applications, ranging from research through education to resource management. For forestry studies in particular, LiDAR is capable of characterizing both terrain and vegetation structure effectively. However, as all LiDAR-derived vegetation metrics (e.g. tree height) are calculated relative to a DTM surface, any error in the DTM will propagate to affect the accuracy of the derived vegetation metrics (TINKHAM et al., 2012). It is necessary, therefore, to evaluate the errors associated with the LiDAR-derived DTM to then be able to characterize the overlying vegetation accurately.

The three basic types of error that should be taken into account for accurate DTM generation from LiDAR data are: (1) sensor-specific errors associated with the navigation, positioning and LiDAR systems that occur during data acquisition; (2) geometrical errors related to the flight altitude and scan angle or the local topography; and (3) errors arising during the post-processing steps, such as point classification or surface interpolation (SU; BORK, 2006). The quality of the DTM is a measure of how accurate the elevation is at each pixel of a grid (absolute accuracy) and how accurately the morphology is represented (relative accuracy). DTM accuracy estimates based on airborne laser scanning data are typically provided by comparing LiDAR heights against a finite sample of check point

coordinates from an independent source of higher accuracy, supposing a normal distribution of the derived height differences or errors (AGUILAR; MILLS, 2008).

The accuracy of LiDAR-derived DTMs can differ significantly across the topographic spectrum as well as with different land cover types. Over open areas with relatively flat terrain, it is common to achieve elevation accuracies below 0.15 m Root Mean Square Error (RMSE) (COBBY et al., 2001; HODGSON et al., 2005; SPAETE et al., 2010). In a study evaluating DTM accuracy for six different land-cover types, Hodgson and Bresnahan (2004) observed RMSE values ranging from a low of 0.17 to 0.19 m in pavement and low grass classes to a high of 0.26 m in a deciduous forest. In areas covered by dense vegetation, DTM elevation errors tend to increase because fewer LiDAR beams reach the ground resulting in fewer ground points for DTM surface interpolation (CLARK et al., 2004). Several studies have assessed LiDAR-derived DTM accuracy in temperate coniferous, deciduous and mixed forests, reporting RMSE values that range between 0.32 m and 1.22 m (REUTEBUCH et al., 2003; KRAUS; PFEIFER, 1998; HODGSON et al., 2003). However, there have been relatively few studies providing rigorous assessment of elevation accuracy under the complex multilayered canopy of tropical rain forests. One study by Clark et al. (2004) reported a DTM accuracy of 0.58 m RMSE in open-canopy flat areas of an old-growth Costa Rican rain forest, and overall RMSE of 2.29 m when steep slopes and multilayered dense vegetation areas were also considered.

4.2 Material and methods

4.2.1 LiDAR dataset

LiDAR data were collected by the GEOID Laser Mapping company (Belo Horizonte, MG) in April 2012 as part of the Sustainable Landscapes joint project of EMBRAPA, the Arthur

Bernardes Foundation (FUNARBE) and the United States Forest Service (USFS). The study area was overflown with an Optech ALTM 3100 laser scanner instrument at an average flying altitude of 1600 m a.s.l., covering a rectangular strip of the surface (about 1.5 km x 7 km) with a total area of approximately 1000 ha. The high-resolution discrete-return LiDAR data were collected in the conditions detailed in Table 4.1 below. Average pulse density was 12 pulses/m^2, resulting in an average return density of 20 points/m^2 and average return spacing of 0.22 m.

Table 4.1 Laser system parameters.

Parameter	Specification
Positioning system	POS AV™ 510 (OEM) - GPS/GNSS/L-Band receiver
Horizontal accuracy	≤ 50 cm (1:1000 scale; PEC "A"); 1σ
Vertical accuracy	≤ 15 cm; 1σ
System frequency (PRF)	50 kHz
Scan frequency	25 Hz
Scan angle (FOV)	≤ 20°
Data recording	up to 4 returns/pulse
Average flight altitude	1600 m a.s.l.
Beam divergence	0.25 mrad (1/e)
Overlap between flight lines	30%

Raw LiDAR data were delivered by GEOID in standard LAS format with the following parameters for each registered return: point number, x and y location, elevation, intensity, pulse number, and return number. Because LiDAR datasets are large, the data points were resampled into smaller files (tiles) for more efficient handling and processing. The company also provided the navigation files of the aerial survey trajectory in SBET (Smoothed Best Estimate of Trajectory) file format, from which the individual flight lines could be extracted.

4.2.2 LiDAR processing

The LiDAR data for the study area were processed following two different methodologies (Figure 4.1): one by the data provider, GEOID, and one by the G-LiHT research group at NASA Goddard Space Flight Center (COOK et al., 2013). Flight line calibration was performed first (PosPac software) to adjust parameters such as heading, roll, pitch and height. This is an important step, given that flight line misalignment is considered one of the major sources of error in LiDAR data (SHRESTHA et al., 2010). Next, height filtering was performed on the point cloud to select ground points from the data set – a critical step for DTM generation from LiDAR data (LIU, 2008). To accomplish this, GEOID used the Adaptive Triangulated Irregular Network (ATIN) filter of the TerraScan software, while G-LiHT applied a progressive morphological (PM) filter. The Adaptive TIN filter is based on the relationship between surface angles and elevation differences of a TIN surface (a so-called slope-based filter) that is initiated with a set of seed points and applied iteratively to classify ground points (see AXELSSON, 2000). The PM filter is used to identify objects in grayscale images based on spatial structure, and works with dilation and erosion in combination with opening and closing operators to separate ground points from non-ground ones (see ZHANG et al., 2003). Point classification was followed in both cases by Delaunay triangulation to create a TIN of the filtered ground returns. Finally, the TIN was used to interpolate the ground elevations onto a 1-meter raster grid, thus obtaining the DTM (denoted as "GEOID DTM" for the first and "G-LiHT DTM" for the second processing approach).

```
        original LiDAR point cloud
                    |
           calibration of flight lines
                    |
         _____|_____
        |                       |
GEOID   Adaptive TIN filter     Progressive Morphological filter   G-LiHT
        (Axelsson)              (Zhang et al.)
        |                       |
        ground returns (1)      ground returns (2)
        |                       |
        Delaunay triangulation  Delaunay triangulation
        |                       |
        TIN of ground points (1) TIN of ground points (2)
        |                       |
        Interpolation (1m)      Interpolation (1m)
        DTM (1)                 DTM (2)
```
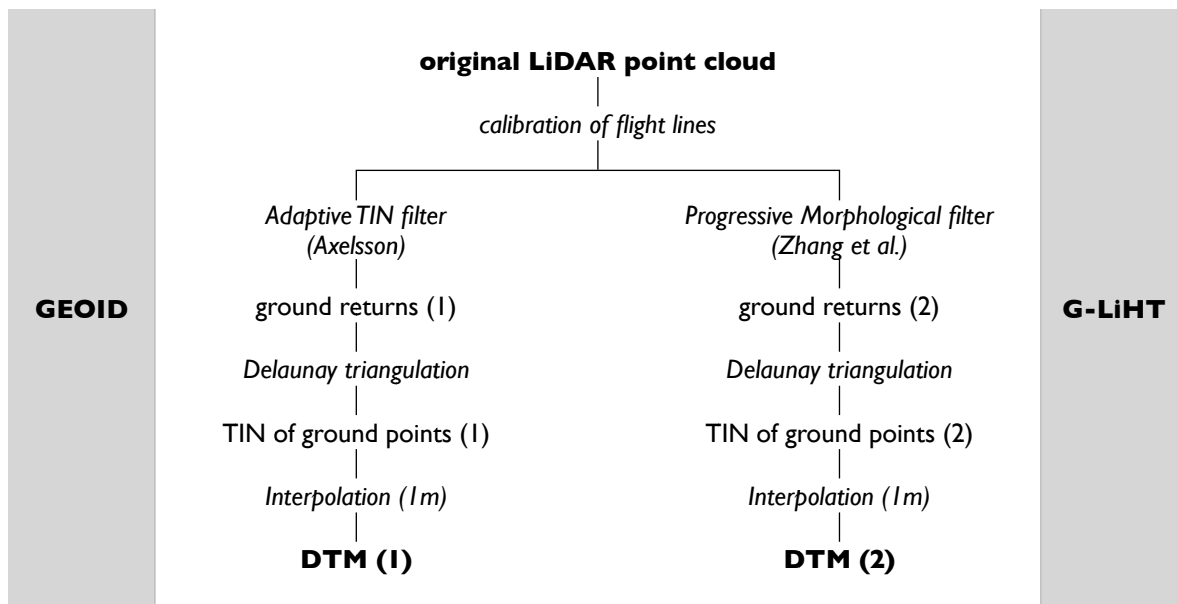
Figure 4.1 LiDAR data processing steps for DTM generation as performed in the two methodologies, GEOID and G-LiHT.

4.2.3 Ground data acquisition

Ground survey data collected in June 2013 within the study area were treated as a reference dataset for LiDAR DTM validation. A total of 36 points were measured under closed forest canopy in the predominantly hilly terrain along the altitudinal transect, marking the corner points of the nine permanent forest inventory plots located within the LiDAR coverage. The differential GPS equipment used for the measurements consisted of two Topcon HiPer (L1/L2) receivers, one used as a base and the other as rover. These are surveying-grade dual-frequency units capable of receiving both NAVSTAR and GLONASS signals. Raw data at the unknown points were collected for 20 - 35 minutes on average (60 minutes at locations where terrain or dense overstory vegetation shaded satellite signal). Simultaneous measurements were made at a registered point of known position for subsequent differential correction – the survey marker (INCRA "ABE M0693") is located at the Santa Virginia

base station in the PESM, in an open area within less than 10 km of the forest plots. Post-processing of the GPS data was performed to produce the estimated position of the unknown points. Out of the 36 control points, 35 were measured with success, and 30 of these met the required submeter accuracy ($\sigma < 1$ m) in all three coordinates x, y, z (UTM easting, northing and elevation). The remaining 5 points were less accurate ($\sigma < 2.2$ m). The GPS system parameters and measurement conditions during the survey are summarized in Table 4.2.

Table 4.2 GPS system parameters and survey conditions.

Parameter	Specification
GPS system	Topcon HiPer L1 / L2 receiver
Horizontal accuracy	3 mm + 0.5 PPM
Vertical accuracy	5 mm + 0.5 PPM
System frequency	20 Hz
Linear units	meters
Angular units	degrees
Datum	WGS84
Projection	UTM Zone 23 South
Geoid	MAPGEO 2010
Base Reference Point	INCRA "ABE M0693"
Number of points measured with success	35 (out of 36 total)
Points with $\sigma < 1$ m (x,y,z)	30
Points with $1 \leq \sigma < 2.2$ m (x/y/z)	5
Precision (RMSE) Easting	range: 0.006 - 2.130 m; mean = 0.473 m
Precision (RMSE) Northing	range: 0.006 - 1.876 m; mean = 0.225 m
Precision (RMSE) Elevation	range: 0.019 - 2.195 m; mean = 0.469 m

4.2.4 Statistical analysis of the datasets

In the following analysis, the differences between the LiDAR-derived elevation data and the GPS survey data are described and evaluated. For each one of the 35 valid control points i, the corresponding DTM cell is found by spatial intersection, and the cell value extracted. Next, the difference in elevation (Δz) between the survey point z_{GPS} and the respective DTM cell z_{DTM} is calculated by:

$$\Delta z_i = z_{DTMi} - z_{GPSi} \tag{4.1}$$

As described by Su and Bork (2006), Mean Signed Error (MSE) and Root Mean Square Error (RMSE) have been commonly used to assess the accuracy of LiDAR-derived DTMs. Mean signed error can be useful to identify the tendency for under- or over-estimation of elevations, and is calculated by the equation:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} \Delta z_i \tag{4.2}$$

where n is the number of reference points and Δz_i is the elevation difference between DTM and survey data at each point. RMSE is generally calculated to determine the overall mean elevation accuracy of a DTM, and is expressed by the equation:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - x_i)^2} \tag{4.3}$$

where n is the number of reference points, y_i is the predicted value calculated from the regression model (DTM elevation) and x_i is the observed value (measured GPS elevation) at each checkpoint i.

To determine if the difference between the two sets of height points (DTM vs. GPS elevations) is statistically significant, a two-sided t-test is performed assuming a normal distribution of the error. The t-test standardizes the values in the dataset and compares the

transformed data with a theoretical student t-distribution (STAL et al., 2011). A null hypothesis (H_0) was set up stating that the height difference (Δz) between the LiDAR DTM and the control points is equal to zero

$$H_0 : \mu_{\Delta z} = 0 \qquad (4.4)$$

while the alternative hypothesis

$$H_1 : \mu_{\Delta z} \neq 0 \qquad (4.5)$$

stated that the height difference is not equal to zero. The null hypothesis was evaluated with a confidence level of 95%.

Given the significant variation in terrain elevation across the study area (from about 100 m a.s.l. up to 1100 m a.s.l.) and the relatively constant flying altitude during the LiDAR survey (~1600 m a.s.l.), the sensor height above the ground varied significantly between different regions of the 1000 ha study area (Figure 4.2). The mean ranging distance between the sensor and the ground surface was ~660 m for the montane region on top of the plateau, while it was about twice as large (~1320 m) for the submontane region. Since a larger LiDAR ranging distance results in a larger footprint on the ground (due to beam divergence), variation of sensor height above the ground can influence the measurement results, such as laser point density, penetration, ground detection, and calculated metrics (MORSDORF et al., 2008). In the present case, the LiDAR footprint diameter increases to about twice its size between the montane and submontane regions, going from ~0.16 m to ~0.33 m (0.25 mrad beam divergence).
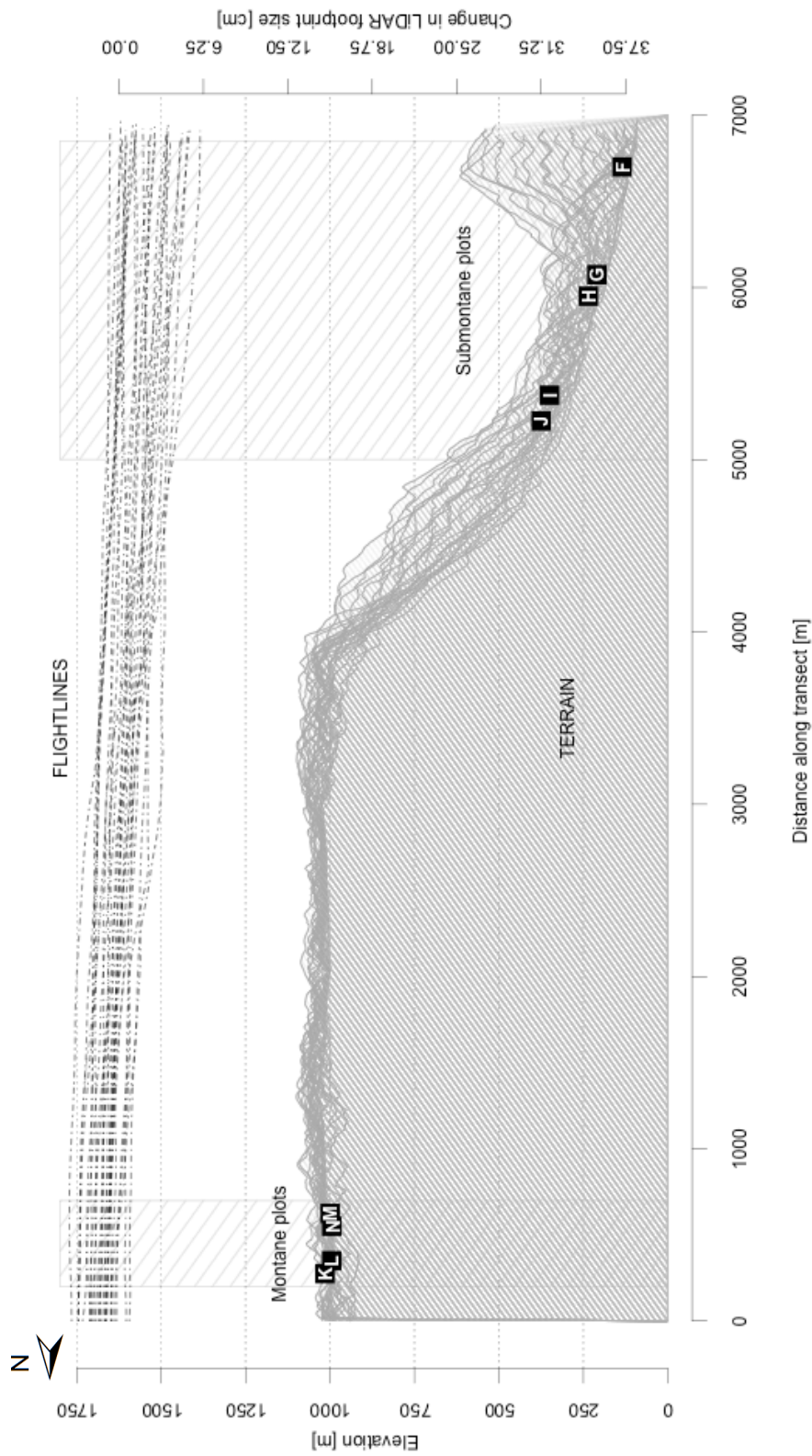
Figure 4.2 Graphical overlay of vertical transects along the length of the study area including the 26 individual flight-lines and the underlying terrain elevation. The control points in the montane forest plots (K–N) lie at an average elevation of 1000 m, while those in the submontane plots (F–J) are located between 100-350 m elevation. North is to the left of the figure, and the y-axis on the right-hand side shows the increase in LiDAR footprint size with growing distance from the airborne sensor.

To assess the effect of different ranging distances (i.e. variable footprint size) on DTM error across the study area, the control points were grouped into two elevational classes, and the error distributions between the groups were compared. This allowed also for investigating the influence of DTM accuracy on the LiDAR-derived canopy heights of different forest types (submontane and montane forest), as the type of vegetation is strongly associated with elevation along this altitudinal gradient.

To test if the means of the errors associated with the specified elevation classes are statistically different (with the assumption that canopy structure does not affect the detection of the ground surface), a two-sided t-test was performed using the following hypotheses:

$$H_0 : \mu_1 - \mu_2 = 0 \tag{4.6}$$

$$H_1 : \mu_1 - \mu_2 > 0 \tag{4.7}$$

where μ_1 is the mean in elevation error for the submontane class (lower elevation) and μ_2 is the mean error for the montane class (higher elevation).

4.3 Results

The comparison between elevations measured in the field (GPS elevation) and those extracted from the LiDAR-derived DTMs (G-LiHT and GEOID DTM) indicated a good fit between these two sources of terrain data (Figure 4.3). The regression analysis resulted in very high R^2 values (0.99 for both the G-LiHT and GEOID DTMs) when compared to the GPS survey data. Calculated RMSE values were 0.97 m for the G-LiHT DTM and 0.95 m for GEOID.
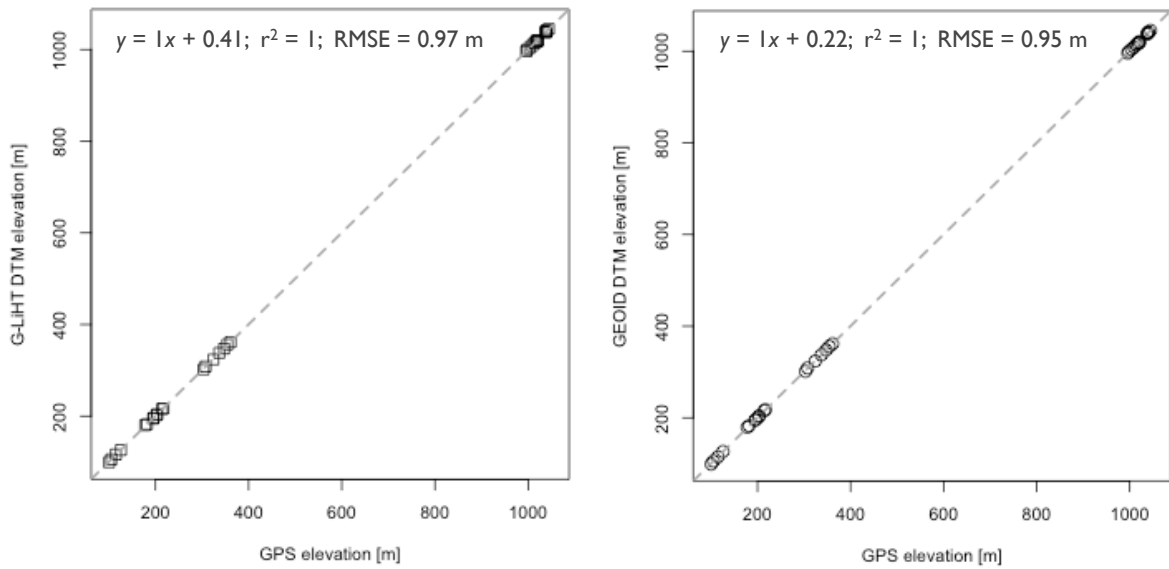
Figure 4.3 Comparison of GPS elevations with DTM elevations for the two models, G-LiHT (on the left) and GEOID (on the right).

The error analysis of elevations using all 35 valid control points resulted in a mean signed error of 0.19 ± 0.97 m (μ ± σ) for the G-LiHT DTM and 0.18 ± 0.95 m for the GEOID DTM (Figure 4.4). The DTM elevations were higher on average than the GPS elevations in both cases, with the difference being quite consistent between the two DTMs. The slightly smaller standard deviation and the narrower error range associated with the GEOID DTM could be related to a more rigorous LiDAR point filtering when the terrain model was generated. Considering that there is uncertainty in calculating the LiDAR DTM (vertical 1σ = 0.15 m on flat terrain) as well as error in the GPS measurements, this 0.18 m elevation difference indicates a very good agreement between field data and terrain model.
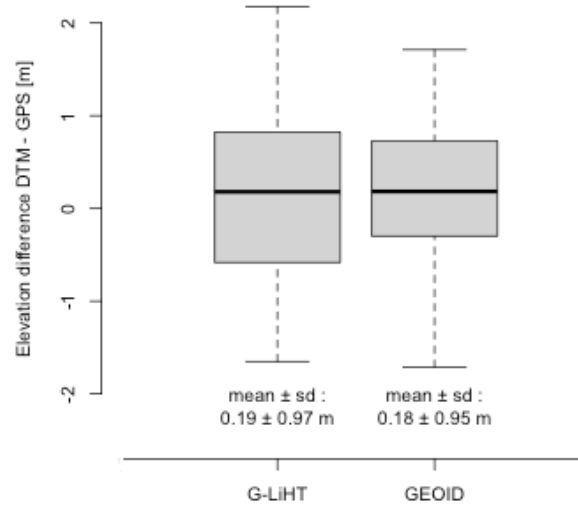
Figure 4.4 Distribution of the errors between GPS and DTM elevations for G-LiHT and GEOID. Mean error and standard deviation values are consistent between the two models, with a slightly smaller range of errors in the GEOID case.

Using only the 30 most accurate control points ($\sigma < 1$ m) for comparison, the mean signed error dropped to less than half, resulting in 0.072 ± 0.895 m and 0.074 ± 0.895 m difference of terrain elevations for G-LiHT and GEOID respectively. The error statistics for the complete point set ($n = 35$) and the set of best points ($n = 30$) are summarized in Table 4.3. Based on a one-sided t-test performed with the 30 most accurate control points, the null hypothesis that the mean of errors is not significantly different from zero could not be rejected with a confidence level of 95% for both the GEOID (p-value = 0.6532) and the G-LiHT DTM (p-value = 0.6624).

Table 4.3 Summary statistics from the error analysis comparing field-measured GPS terrain elevations with LiDAR-derived DTM elevations for two different models (G-LiHT and GEOID).

Error statistics (Δz) in meters							
Data type		range	min	max	mean	stdev	rmse
G-LiHT	best 30	3.829	-1.651	2.178	0.072	0.895	0.883
	ALL 35	3.829	-1.651	2.178	0.189	0.965	0.970
GEOID	best 30	3.516	-1.951	1.565	0.074	0.895	0.884
	ALL 35	3.663	-1.951	1.712	0.182	0.945	0.949

To examine whether there was significant difference between the elevation errors associated with points located at lower altitudes (submontane plots) as opposed to those associated with points located at a higher altitude (montane plots), first the normality of errors was assessed using normal Q-Q (quantile) plots. Visual evaluation of the graphs (Figure 4.5) indicated that the elevation errors are normally distributed, since the majority of the data points fell on the 1:1 line with few points falling further from the line at the low and high ends of the distributions. The errors associated with the G-LiHT DTM showed better linearity than the error points from the GEOID DTM. In both models, the Kolmogorov-Smirnov test confirmed that error distributions were not significantly different from the normal distribution (G-LiHT p-value = 0.923 and GEOID p-value = 0.261).
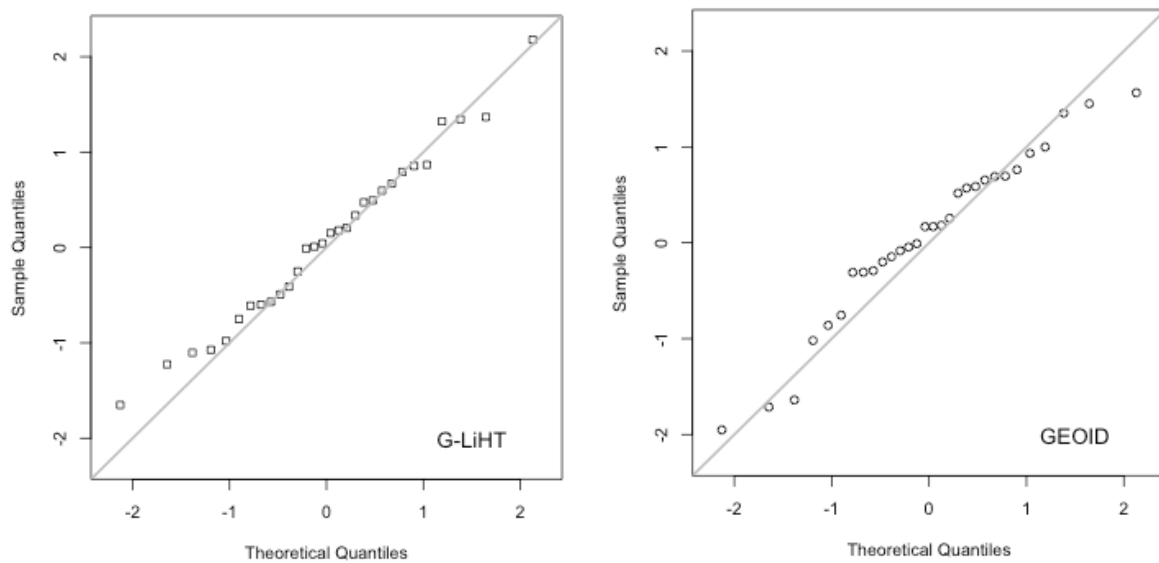
Figure 4.5 Normal Q-Q plots evaluating the linearity of errors suggest that the data are normally distributed for both the G-LiHT and the GEOID DTM.

Calculated mean signed errors for submontane vs. montane areas revealed a positive discrepancy between DTM and GPS elevation at a lower altitude (0.23 ± 0.88 m for G-LiHT and 0.13 ± 0.94 m for GEOID) indicating that the LiDAR-derived data could slightly overestimate terrain elevation in this area; while there was a negative, though smaller, discrepancy observed at a higher altitude (-0.14 ± 0.90 m for G-LiHT and -0.004 ± 0.868 m for GEOID) between DTM and GPS elevations indicating a potential underestimation of terrain elevation by the LiDAR. The relevant statistics are summarized in Table 4.4.

Table 4.4 Summary statistics from the error analysis of terrain elevations (G-LiHT and GEOID DTMs vs. GPS) comparing error sign and magnitude between submontane and montane areas.

Data type		range	min	max	mean	stdev	rmse
		Error statistics (Δz) in meters					
G-LiHT	submontane	3.403	-1.225	2.178	0.230	0.883	0.887
	montane	2.995	-1.651	1.344	-0.135	0.904	0.879
GEOID	submontane	3.516	-1.951	1.565	0.134	0.938	0.920
	montane	3.164	-1.713	1.451	-0.004	0.868	0.834

Based on a two-sided t-test performed with the two sets of errors, submontane and montane, the null hypothesis that the true difference in means is not significantly greater than zero could not be rejected with a confidence level of 95% (p-value = 0.139). This indicates that there was no statistically significant effect of flying height (i.e. variable footprint size) on DTM errors across our study area, while the different vegetation classes (submontane vs. montane) also didn't seem to affect significantly the accuracy of the LiDAR-derived DTM. Additionally, these results reinforce the fact that canopy structure at this study site does not affect the acquisition of ground points.

4.4 Discussion

The overall DTM accuracy of 0.18 m mean vertical error and 0.95 m RMSE observed in this study is in agreement with previous findings of LiDAR forest studies (KRAUS; PFEIFER, 1998; HODGSON et al., 2003; REUTEBUCH et al., 2003; CLARK et al., 2004), and demonstrates the ability of small-footprint systems to accurately characterize the ground surface in complex mountainous terrain covered by dense tropical vegetation. This ability to generate a highly accurate terrain model despite such a challenging environment can be attributed to the high point density with which the LiDAR data was sampled (20

points/m^2 on average). Typical LiDAR data densities used for forest research and management purposes have been within the range of 0.5 - 4 points/m^2 (ANDERSEN et al., 2006; GONZALEZ et al., 2010; GATZIOLIS; ANDERSEN, 2008), occasionally reaching a higher value of 10 to 12 points/m^2 (SÄYNÄJOKI et al., 2008; HUDAK et al., 2012). The effect of LiDAR point density on the accuracy of the generated DTM has been evaluated by several studies, and most of them agree that the error increases exponentially as the point spacing increases (JAKUBOWSKI et al., 2013). This observation was tested in a synthetic data thinning experiment using the LiDAR data from this study, and the results are presented in the next chapter (Chapter 5).

In addition to the effect of point density on the accuracy of LiDAR-derived DTMs, the impact of other factors such as flying altitude (i.e. sensor elevation) and terrain variability have also been investigated in a number of studies. It has been demonstrated that as flying altitude increases, the percentage of pulses that penetrate the forest canopy decreases, reducing the chances of getting a return from the ground surface (TINKHAM et al., 2012). The lower penetration power has been explained by the broadening of the LiDAR footprint (as a result of beam divergence) which causes a weakening in the laser energy (HYYPPÄ et al., 2005). The present analysis comparing DTM accuracy in the submontane vs. montane regions of the study area showed better agreement of the model with the ground reference elevations at the higher altitudes, where the LiDAR system was closer to the ground surface than at the lower altitudes, where ranging distance was twice as large on average.

One possible explanation of the above observation can be the effect of the change in the pulse's footprint size and laser energy with flying height, leading to loss of penetration power further from the LiDAR system. This hypothesis is consistent with the evidence that shows different proportions of ground returns in our data in the submontane and montane areas. The average percentage of points classified as ground returns in the original data set

was 3% in the montane area, while only 1% in the submontane area. With a larger number of ground points available for surface interpolation, it is possible to derive a DTM that represents the true ground surface with higher accuracy. Additionally, terrain complexity has also been mentioned as a potential cause for the variation in DTM accuracy across landscapes (LIU, 2008). The steeper slopes and more variable topography in the submontane region might be harder to capture by the LiDAR system than the generally more homogeneous montane terrain on top of the plateau. By optimizing flight line configuration at the time of data collection (e.g. constant flying height above ground, even point distribution), the observed difference between DTM accuracy in submontane vs. montane areas could potentially be minimized, and the variability of the terrain surface represented more faithfully. Future airborne LiDAR surveys in similarly complex mountainous terrain would benefit from such recommendations for mission planning to ensure high data quality for forestry applications.

4.5 Conclusions

Based on the major findings of this analysis, the following conclusions can be drawn:

[1] The strong agreement between differential GPS elevations and digital terrain model elevations suggests that the LiDAR-derived DTM can be considered as a highly accurate representation of the true terrain surface.

[2] The similarity of the sign and magnitude of errors associated with the two DTMs (GEOID and G-LiHT) suggests that the two different processing methodologies resulted in equally valid digital terrain models.

[3] Error magnitudes associated with DTM elevations vary between areas of submontane and montane forest possibly caused by the different ranging distances above these areas, as well as the varying degree of terrain complexity across the study area.

[4] The difference between DTM accuracy in the submontane vs. montane regions was not statistically significant, suggesting that the variation in flying altitude or terrain complexity did not strongly affect the ability of the LiDAR to accurately characterize the ground surface.

[5] The high accuracy of the LiDAR-derived DTM shown in the present study provides reasonable confidence that potential error propagation from DTM to the calculated vegetation metrics is minimal.

5. LiDAR DATA THINNING EXPERIMENT

5.1 Motivation

Given that the LiDAR data used in this study had a higher average return density (20 points/m^2) than is common for forest research purposes and biomass evaluations (0.5 - 4 points/m^2), random thinning of the original LiDAR point cloud was performed to assess the changes in DTM accuracy with increased point spacing of the data and the corresponding effects on LiDAR-derived forest structure metrics. The main goal of this experiment was to answer the questions:

[1] Is exponential increase in LiDAR-derived DTM error observed with increased point spacing, as has been described before in the literature? (JAKUBOWSKI et al., 2013)

[2] How does the density of ground points change with increased thinning of the LiDAR point cloud?

[3] What is the effect of increased point spacing on the variability of LiDAR-derived vegetation metrics, specifically that of mean canopy height?

[4] What is the minimum LiDAR return density required for generating a DTM with acceptable accuracy for canopy height models that contribute to aboveground biomass estimations?

5.2 Material and methods

5.2.1 LiDAR thinning

The original LiDAR point cloud data were reduced to the four pre-defined return densities of 8, 4, 2 and 1 points/m^2 following a thinning algorithm applied by the G-LiHT group at NASA GSFC. These specific data densities were chosen to represent a range of values most

frequently quoted in LiDAR forestry studies. The basic processing steps for data thinning included:

(1) assessing the point density in the original point cloud;

(2) assessing the return numbers (1st, 2nd up to 4th) associated with each point in the data set;

(3) thinning of the original point cloud by random selection of the desired number of returns while preserving the original return number ratios (i.e. relative number of 1st, 2nd, 3rd and 4th returns).

The resulting data sets were then processed to produce three different data products for each thinning density: DTM (Digital Terrain Model), CHM (Canopy Height Model) and DSM (Digital Surface Model) raster layers with 1-meter grid resolution, representing the terrain surface, the canopy heights above ground, and the outer surface of the forest vegetation in absolute elevation (a.s.l.), respectively. The new DTM raster grids were created from the reduced data using the same G-LiHT methodology described in the previous chapter. The CHM rasters were also generated using the G-LiHT algorithm by selecting the highest LiDAR return in every 1-meter grid cell, building a TIN based on these points, and interpolating the canopy heights on a 1-meter raster grid (COOK et al., 2013). The DSMs of the outer canopy were produced from only the first-return points in the LiDAR point cloud that were processed using the BCAL LIDAR Tools open-source software package. The processing steps described here are summarized in form of a flowchart in Figure 5.1.

original LiDAR point cloud

data thinning

density-8 (D8) point cloud
density-4 (D4) point cloud
density-2 (D2) point cloud
density-1 (D1) point cloud

G-LiHT Progressive Morphological filter

BCAL LiDAR Tools raster products tool

ground returns

non-ground returns

first returns only

assign above ground height (AGH) to points

find highest point in each grid cell

find highest point in each grid cell

Delaunay triangulation

Delaunay triangulation

Delaunay triangulation

TIN of ground points

TIN of canopy heights

TIN of surface elevation

Interpolation (1m)

Interpolation (1m)

Interpolation (1m)

DTM - D20
DTM - D8
DTM - D4
DTM - D2
DTM - D1

CHM - D20
CHM - D8
CHM - D4
CHM - D2
CHM - D1

DSM - D20
DSM - D8
DSM - D4
DSM - D2
DSM - D1

Figure 5.1 LiDAR data processing steps for data thinning.

5.2.2 Analysis of the thinned datasets

The accuracy of the DTMs generated after data thinning was evaluated using the same approach as with the full-density DTM. The elevation of the 35 ground control points measured with differential GPS was compared to the DTM elevations of thinned data at the same locations. The calculated differences were expressed in the form of Mean Signed Error and Root Mean Square Error (as of Equations 4.2 and 4.3 in the previous chapter). The error statistics corresponding to the four different thinning levels were compared both for the study area as a whole and for the montane and submontane forest regions separately. The variation in DTM elevation accuracy across variable terrain elements (i.e. steep slopes and flat plateaus) was examined and described.

The total number of LiDAR returns as well as the number of ground returns was assessed from the reduced-density point clouds for each permanent field plot location separately. The ground point density (points/m^2) and the fraction of ground returns out of all returns was calculated for each thinning level within the nine 1-ha plots to quantify the change in commission errors resulting from the ground classification algorithm.

To examine the effect of data thinning on the canopy height estimates and to illustrate the potential impact on estimates of aboveground biomass (AGB), plot-level LiDAR metrics and field-based biomass data were analyzed. Samples (1 ha) were extracted from the whole-area full-density and thinned CHM raster grids corresponding to the permanent plot locations. The plot-level means of the CHM raster samples were calculated, and these were compared for the different data density levels to assess any changes in mean canopy height. Additionally, estimates of the outer canopy surface (DSM) and underlying DTM on the 1-ha plot scale were examined to evaluate the source of elevation biases in the thinned LiDAR data.

Forest inventory data were used from the Serra do Mar permanent plot network (BIOTA Project, see Joly et al., 2012) to calculate field-based AGB estimates in the nine plots following the methodology applied by Alves and colleagues (2010) (see detailed description of the field data and processing in Chapter 6). A linear model was developed to predict AGB based on plot-level mean canopy surface heights derived from the full-density LiDAR data. This regression equation was then used to generate biomass estimates based on the thinned LiDAR datasets with mean canopy surface height as the predictor, and the resulting values were compared across the different data densities.

5.3 Results

The results observed in the data thinning analysis were consistent with the expected trend of decreasing DTM elevation accuracy with increasing levels of data reduction. When compared with the set of GPS control point elevations, mean signed errors of the thinned DTM elevations increased from a low of 0.19 ($\pm$ 0.97) m to a high of 3.21 ($\pm$ 3.12) m between the original data density of 20 returns/m^2 and the highest level of thinning of 1 return/m^2. Figure 5.2 shows a comparison of the error distributions associated with each of the five different data density levels analyzed: the original density of 20 returns/m^2 (D20) and the thinned densities of 8, 4, 2 and 1 returns/m^2 (denoted D8, D4, D2 and D1, respectively). DTM elevations were higher than the GPS elevations in all cases, with increasing error magnitudes as point spacing increased. This consistent overestimation of the true ground elevation results from the incorrect classification of vegetation features as ground surface by the point filtering algorithm. Calculated RMSE values showed a similar increasing trend with decreasing data density, ranging from a low of 0.97 m for the highest density DTM to a high of 4.45 m for the lowest data density of 1 return/m^2.
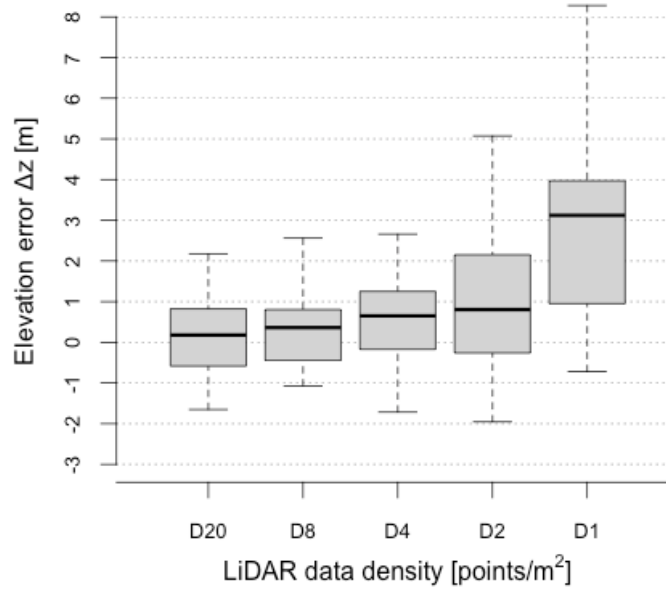
Figure 5.2 Distribution of the errors between GPS and DTM elevations with data density levels of 20, 8, 4, 2 and 1 returns/m² (D20, D8, D4, D2 and D1, respectively).

A comparison between the elevation errors associated with montane and submontane sites in the thinned DTMs showed generally larger error magnitudes in the submontane region than in the montane area for all the data densities tested. This observed difference between elevation classes became larger with increased levels of data thinning: with 20 returns/m² the mean signed error difference was 0.31 m between submontane and montane areas, while it increased to 2.64 m when data density dropped to 1 return/m². Similarly to the trend observed in mean signed errors, RMSE values also followed an increasing pattern, with growing difference between submontane and montane DTM accuracy as data density was reduced. Figure 5.3 shows the RMSE values of elevation difference between the thinned DTMs and the reference GPS data calculated separately for the submontane vs. montane regions and also for the study area as a whole. With the highest data density, submontane and montane RMSE values are practically equal (< 0.1 m difference), while with lower data densities, montane RMSE stays relatively more stable and submontane

RMSE rises more quickly (0.76 to 3.08 m difference). The elevation error statistics based on thinned data are summarized in Table 5.1.
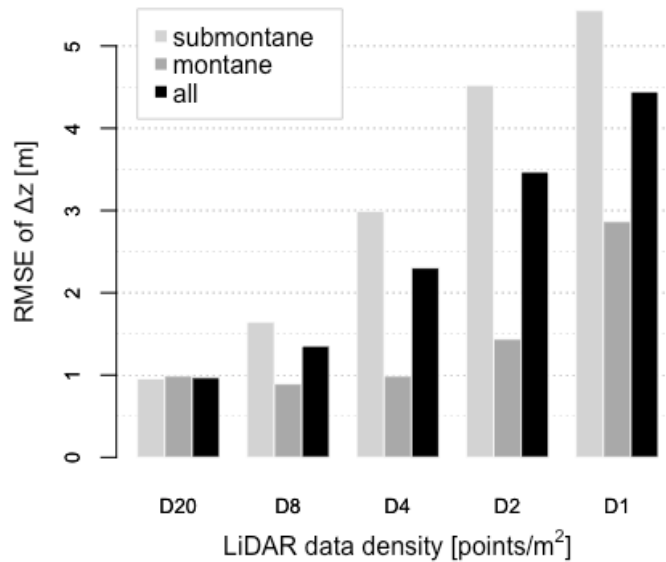
Figure 5.3 Comparison of RMSE values in the DTMs based on the five data density levels of 20, 8, 4, 2 and 1 returns/m^2 (D20, D8, D4, D2 and D1, respectively).

Table 5.1 Summary statistics from the DTM error analysis after data thinning. Signed error and RMSE values are shown for density levels of 20, 8, 4, 2 and 1 returns/m^2. Error statistics were calculated for the study area as a whole, and separately for two elevation classes.

Data type		Error statistics (Δz) in meters				
		min	max	mean	stdev	rmse
D20	submontane	-1.225	2.178	0.330	0.919	0.950
	montane	-1.651	1.862	0.020	1.020	0.988
	ALL	**-1.651**	**2.178**	**0.188**	**0.965**	**0.970**
D8	submontane	-2.876	4.510	0.538	1.597	1.645
	montane	-1.073	1.852	0.189	0.899	0.892
	ALL	**-2.880**	**4.510**	**0.380**	**1.320**	**1.354**

Error statistics (Δz) in meters						
Data type		min	max	mean	stdev	rmse
D4	submontane	-1.717	6.977	1.810	2.450	2.992
	montane	-0.940	2.253	0.300	0.969	0.987
	ALL	**-1.720**	**6.980**	**1.120**	**2.040**	**2.303**
D2	submontane	-1.955	14.620	2.359	3.963	4.523
	montane	-1.390	3.334	0.665	1.315	1.437
	ALL	**-1.960**	**14.620**	**1.590**	**3.130**	**3.470**
D1	submontane	0.459	14.046	4.417	3.244	5.431
	montane	-0.721	7.492	1.775	2.322	2.866
	ALL	**-0.720**	**14.050**	**3.210**	**3.120**	**4.445**

To better understand the spatial variability of the DTM elevation errors across the landscape, a transect line was drawn along the center of the study area and DTM elevations were sampled from the 1-meter raster grids for all data densities. The difference between the cell values of the full-density DTM extracted along the transect line and the corresponding cell values of each thinned DTM was calculated and the elevation errors plotted (Figure 5.4). In general, the elevation difference between full-density and thinned DTMs was larger at lower altitudes, along the hillslope and in the valley, and smaller on top of the plateau. The magnitude of the difference increased with increased data thinning throughout the whole area, and the spatial distribution of the errors was associated with the level of complexity of the terrain in all DTMs examined. Where the terrain surface was more accentuated (i.e. greater rate of change of elevation), the corresponding difference in full-density vs. thinned DTM values was also larger, while with a smoother terrain surface, the associated DTM differences were smaller in magnitude.
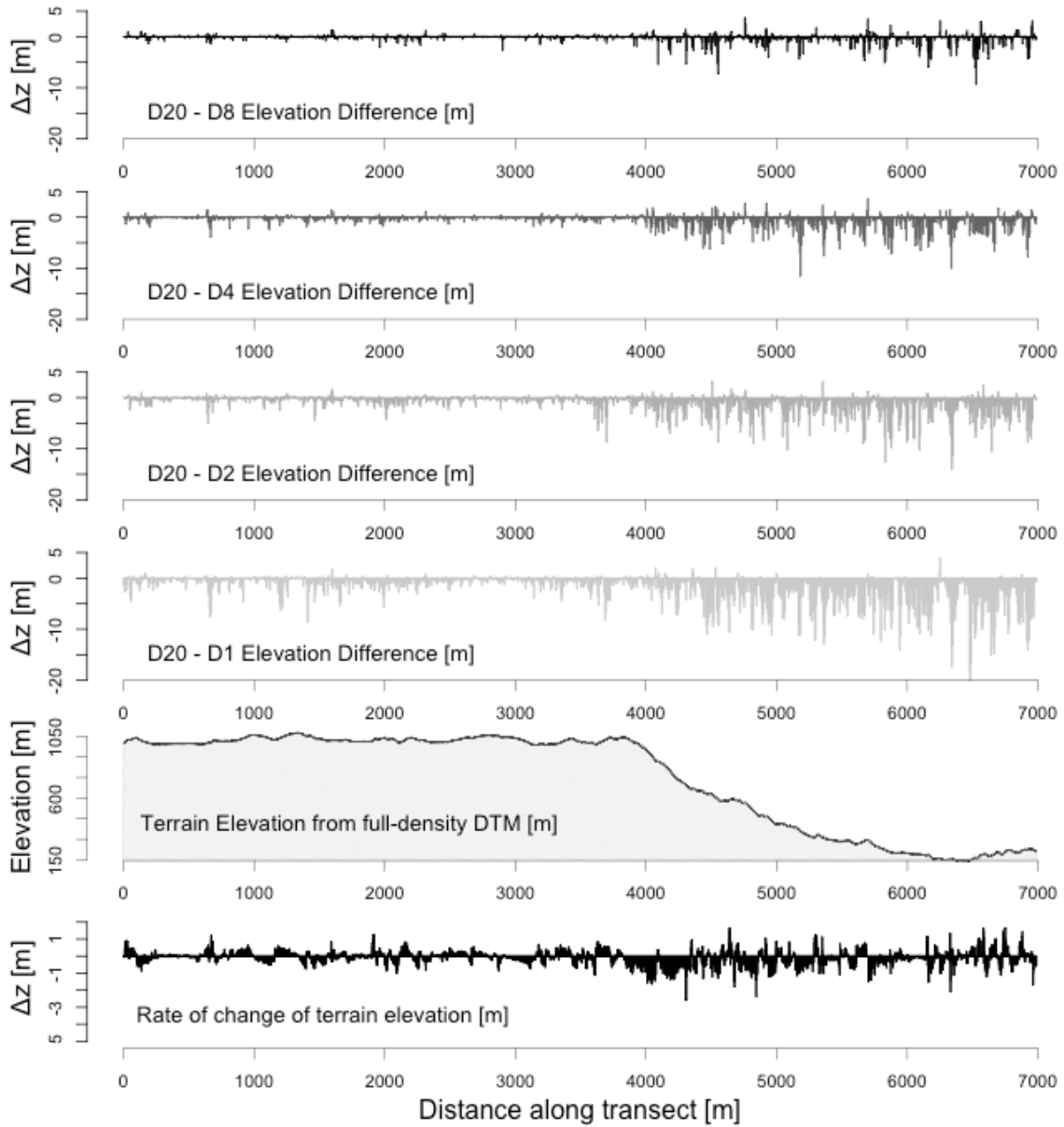
Figure 5.4 Elevation differences between the original DTM generated from the full-density data (D20) and the thinned DTMs (D8, D4, D2, D1) extracted from a 1-m grid along the central line of the study area. A vertical transect of the corresponding terrain elevations extracted from the original DTM along the same central line is shown for reference, as well as the calculated rate of change of the terrain elevation along the transect.

As it was expected, the density of ground points (expressed in points/m^2) followed a decreasing trend with increased data thinning (Table 5.2). From a high of 0.43 ($\pm$ 0.26) points/m^2 in the full-density data, the density of ground returns dropped by an order of magnitude to 0.05 ($\pm$ 0.01) points/m^2 in the lowest-density data. On average, montane plots had a higher ground point density than submontane plots, but this difference became less apparent with increased levels of thinning. When looking at the proportion of ground returns in relation to all returns in the thinned data sets within the whole study area, a fractional increase was observed from a low of 1.91 ($\pm$ 1.04) % ground returns in the original data set up to a value of 4.79 ($\pm$ 0.87) % in the highest level of data thinning. Comparing the percentage of ground points in montane vs. submontane plots showed an increase from 3% to 5.5% in montane plots and from 1% to 3% in submontane plots between the full data density and the thinnest data.

Table 5.2 Ground point density (points/m^2) and fraction of ground returns out of all returns in the submontane and montane forest areas, and in the study area as a whole after various levels of data thinning.

Ground return density after data thinning		
Data type	**points / m^2** mean ± stdev	**% of all returns** mean ± stdev
D20 submontane	0.228 ± 0.031	1.047 ± 0.079
D20 montane	0.684 ± 0.177	2.985 ± 0.293
D20 **ALL**	**0.431 ± 0.264**	**1.908 ± 1.039**
D8 submontane	0.108 ± 0.006	1.354 ± 0.068
D8 montane	0.262 ± 0.026	3.290 ± 0.333
D8 **ALL**	**0.177 ± 0.083**	**2.215 ± 1.042**
D4 submontane	0.067 ± 0.007	1.677 ± 0.173
D4 montane	0.150 ± 0.015	3.752 ± 0.387
D4 **ALL**	**0.104 ± 0.045**	**2.599 ± 1.126**
D2 submontane	0.050 ± 0.005	2.489 ± 0.234
D2 montane	0.091 ± 0.010	4.550 ± 0.517
D2 **ALL**	**0.068 ± 0.023**	**3.405 ± 1.143**

Ground return density after data thinning		
Data type	**points / m^2** mean ± stdev	**% of all returns** mean ± stdev
submontane	0.043 ± 0.006	4.247 ± 0.639
DI montane	0.054 ± 0.006	5.459 ± 0.630
ALL	**0.048 ± 0.009**	**4.786 ± 0.872**

The comparison of mean canopy surface height values derived from the original and thinned CHM rasters showed that lower point density data increasingly underestimate canopy height within the 1-ha plot samples (Figure 5.5). With the original data density of 20 returns/m^2, mean canopy height ranged between 19.52 and 22.91 m across the nine locations (Plots F-N). With increasing levels of data thinning, the mean canopy height decreased on average by 0.70 m, 1.75 m, 3.40 m and 5.26 m for data densities of 8, 4, 2 and 1 returns/m^2, equivalent to a reduction of 3%, 8%, 16% and 25% in mean canopy height, respectively, for the increasing levels of thinning. The observed magnitude of canopy height change was generally larger for the submontane plots (F, G, H, I and J), resulting in mean decrease of 0.79 m, 1.99 m, 3.93 m and 6.08 m with the increasing thinning levels. In comparison, the mean canopy height change in the montane plots (K, L, M and N) was relatively smaller in magnitude with values of 0.60 m, 1.45 m, 2.73 m and 4.24 m for the data densities of 8, 4, 2 and 1 returns/m^2, respectively.
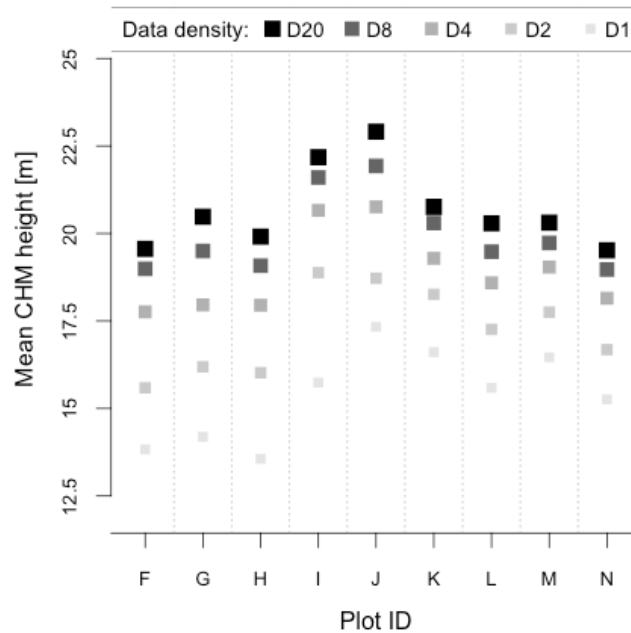
Figure 5.5 Mean canopy surface heights associated with the field plot locations (F - N) based on CHMs generated from original and thinned LiDAR data (D20, D8, D4, D2 and D1 indicate the different data density levels).

Examination of the Digital Surface Models (DSMs) at the permanent field plot locations showed little variation in the outer vegetation surface elevation with the different levels of data thinning. The associated change in the Digital Terrain Model elevations was much more apparent, indicating the increasing difficulty in detecting the ground surface as data density was reduced. Figure 5.6 shows the DSM raster for the montane forest plot "N" generated from the five different LiDAR data densities. It can be observed that the canopy surface became slightly more rugged with increased data thinning, but the overall canopy surface elevation and shape did not change significantly. In comparison, the terrain surface showed larger changes with increased levels of thinning. Specifically, because vegetation features are incorrectly classified as ground by the filtering algorithm in the thinned data sets, the overall effect was a positive bias in the ground elevation, which translated into

lower canopy heights with decreasing data density (see third column of tiles in Figure 5.6). The same effects were observed for the submontane Plot "H" in Figure 5.7. Here, the positive bias in ground elevations was even more apparent (as was generally observed in the submontane area), resulting in more significant shrinkage in mean canopy height values as data density was reduced.
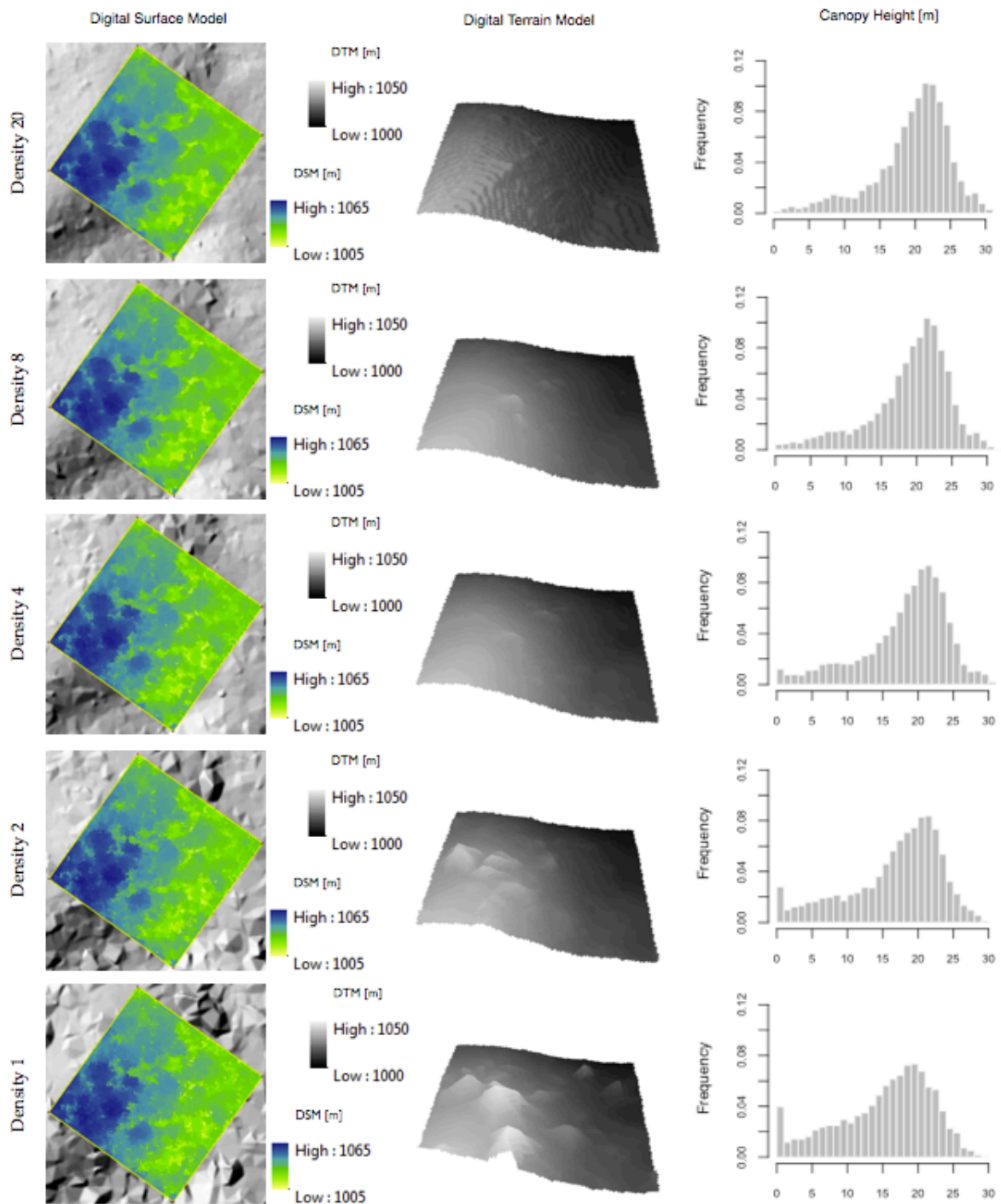
Figure 5.6 Comparison of changes in Digital Surface Model heights, Digital Terrain Model heights, and associated canopy height distributions with data thinning in the montane forest Plot N.
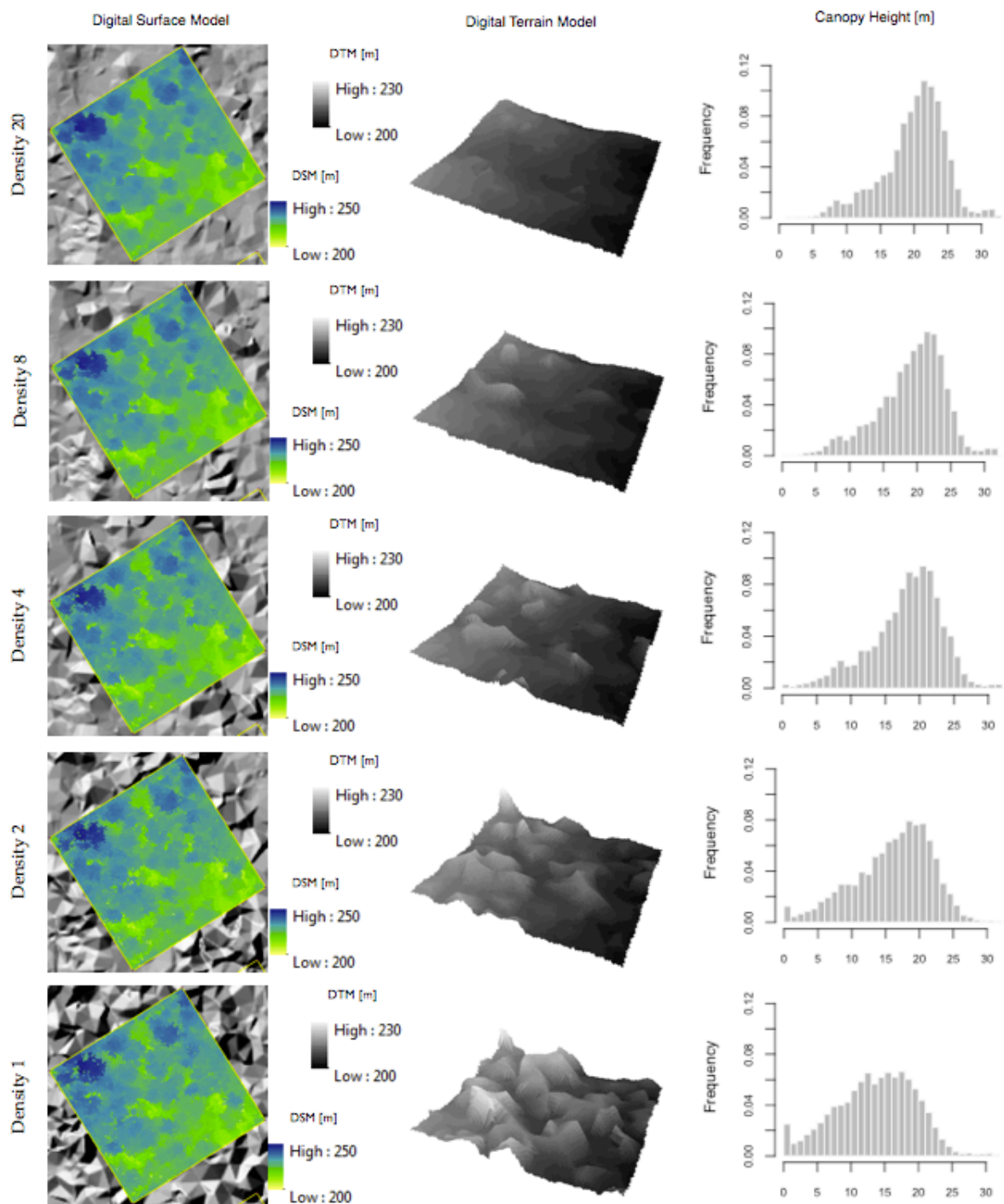
Figure 5.7 Comparison of changes in Digital Surface Model heights, Digital Terrain Model heights, and associated canopy height distributions with data thinning in the submontane forest Plot H.

Using the simple regression model developed based on mean canopy surface height (mCH) to predict aboveground biomass (AGB = 24.13 × mCH - 204.76; R^2= 0.43; RMSE = 30.0 Mg ha^{-1}), significant decreases were observed in modeled AGB values with data thinning (Figure 5.8). Aboveground biomass estimates (mean ± standard deviation across nine permanent plots) for the different thinning levels ranged from 295.3 (± 27.9) Mg ha^{-1} with full-density LiDAR data to 168.2 (± 31.5) Mg ha^{-1} with the lowest data density of 1 return/ m^2. This means that a 1 - 5 m bias in mean canopy height from incorrect ground detection may lead to errors in AGB estimates on the order of 15 - 125 Mg ha^{-1}. For LiDAR data densities below 4 returns/m^2, especially, the bias in height estimates translates into aboveground biomass errors significantly higher than the model error of ~30 Mg ha^{-1}. These results indicate that approaches relying on mean canopy height to estimate aboveground biomass may be particularly sensitive to DTM errors that arise from variability in LiDAR sampling density.
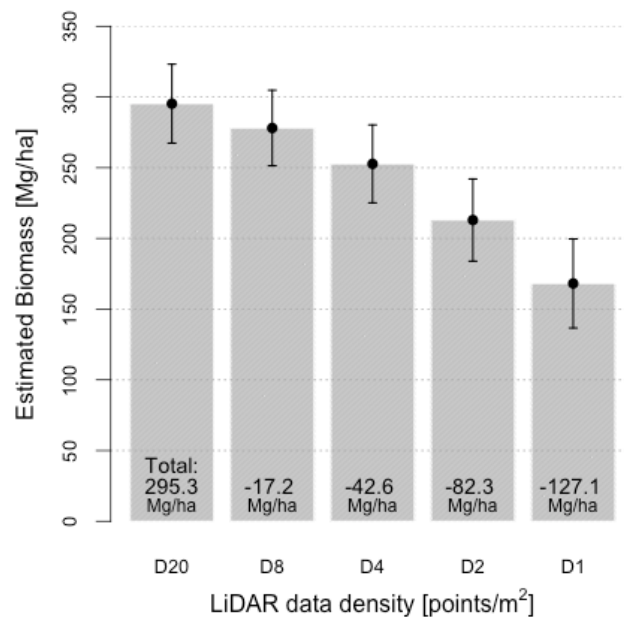
Figure 5.8 Aboveground biomass estimates (mean ± standard deviation across nine permanent plots) for the different data densities predicted with a linear model using mean canopy surface height as input.

5.4 Discussion

Accurate characterization of the ground surface is critical for LiDAR vegetation studies, since the estimation of vegetation heights is calculated relative to the associated bare earth surface. Consequently, any error present in the DTM also introduces error in the canopy height calculations, ultimately leading to erroneous estimation of related forest metrics, such as aboveground biomass. Previous work has demonstrated that over vegetated surfaces, there tends to be an upwards bias in LiDAR elevation data, which was also observed here. This consistent upwards error of the DTM elevations detected within the study area suggests that, as opposed to having an unbiased noise in the canopy height distribution, there is a directional bias that leads to underestimation of canopy heights.

The analysis of the LiDAR-derived CHMs in nine permanent plots showed that the exponential increase in mean DTM error observed with increasing levels of data thinning was coupled with significant decrease in canopy heights. The DSMs were examined to investigate whether the main cause of this change in canopy heights is the inaccuracy of the ground elevations or the inability of the LiDAR to capture the highest point of tree crowns, or a combination of both these factors. No significant change was observed in the DSM heights, which suggests that even with low point density, it is possible to capture the highest points of tree crowns and generate a canopy surface model representative of the true outer vegetation surface. The quick deterioration of ground elevation accuracy, on the other hand, points to the inability to capture the complexity of the terrain when data density is low.

Additionally, it was observed that the magnitude of the directional bias associated with the DTM (upward shift) as well as the CHM (downward shift) varied between the submontane and montane areas, being consistently larger in the submontane forest than in the montane

forest. Such difference would lead to varying degrees of underestimation of aboveground biomass across the study area, where the level of uncertainty is essentially linked to the accuracy of the DTM in any given location, as well as the capacity to accurately calculate canopy heights relative to the DTM. The results of this work suggest that canopy height estimation is most sensitive to changes in DTM accuracy, therefore it is recommended that high density data be used for terrain characterization, especially in mountainous areas with complex topography. Based on the analyses presented here, it becomes evident that improved DTMs and CHMs can increase the accuracy and precision in the derived biomass estimates, benefiting forest research and management applications.

Knowing the amount and detailed spatial distribution of aboveground forest biomass is of primary importance for calculating carbon sources and sinks over time. Precise estimates of aboveground biomass pools are also required to support efforts to mitigate climate change by Reduced Emissions from Deforestation and Forest Degradation (REDD+) (ZOLKOS et al, 2013). Airborne LiDAR has been successfully used to derive aboveground biomass estimates in a range of forest ecosystems (e.g. LEFSKY et al., 1999b; LEFSKY et al., 2002b; DRAKE et al., 2003; NAESSET, 2002; NAESSET; GOBAKKEN, 2008; ASNER et al., 2009). Typical approaches to predict AGB with LiDAR data are based on regression models linking LiDAR metrics to field-based plot-level estimates of biomass and using the derived equations to scale up biomass to larger areas. The LiDAR-derived metrics most frequently used as proxy for biomass include mean or maximum canopy height (e.g. CLARK et al., 2011; MASCARO et al., 2011a; ASNER et al., 2011); vertical canopy profile metrics, such as height percentiles and variance of heights (NI-MEISTER et al., 2010; d'OLIVEIRA et al., 2012) and top-of-canopy height (ASNER; MASCARO, 2014).

Building on recent research efforts to develop generalized biomass allometries for tropical trees (CHAVE et al., 2005; FELDPAUSCH et al., 2012), current work is aimed at

developing a universal model to relate LiDAR canopy height to field-measured aboveground carbon density, accounting for regional variations in climate regime and forest type. Asner and collaborators (2011b; 2014) use three variables to quantify the amount of carbon stored in aboveground biomass in all forests (height from LiDAR and basal area and wood density based on forest inventory). Their work seeks to use simple LiDAR-derived metrics – mean canopy height (MCH) or top-of-canopy height (TCH) – combined with generalized basal area and wood density from limited field studies to produce economical estimates of forest carbon stocks with a minimum of site-specific calibration. A general LiDAR approach like this could greatly improve mapping of aboveground carbon stocks and monitoring carbon emissions over large areas. However, certain limitations also exist. The results of the data thinning experiment presented in this study demonstrate how difficult it could be to apply the same biomass model, relying on a single LiDAR metric (MCH or TCH), across a heterogeneous landscape with both flat and sloped terrain, like the *Serra do Mar*. The study also demonstrates how point density could enhance these issues if flying elevation or coverage differs between study sites (e.g. submontane and montane areas). Multivariate regression models using several LiDAR metrics as input variables might be more suitable in this case to explain variations of biomass across the landscape and among different forest types. Future research will aim at exploring and understanding these LiDAR-biomass relationships at the *Serra do Mar* site in greater detail.

5.5 Conclusions

Based on the major findings of this thinning experiment, the following conclusions can be drawn:

[1] The observed accuracy of the DTMs generated from thinned data decreased with increasing point spacing, but it remained relatively high until point densities as low as 4 returns/m^2.

[2] Point densities below 4 returns/m^2 quickly led to very high vertical error in the resulting DTM, rendering it inadequate for subsequent study of the vegetation.

[3] Canopy heights calculated from the LiDAR-derived CHM decreased significantly with decreasing return density. The magnitude of this change varied between submontane and montane forests, indicating a greater underestimation of canopy height in the submontane area.

[4] DSM heights did not change significantly with data thinning, demonstrating the ability of small-footprint airborne LiDAR to accurately capture the top heights of trees.

[5] Consistent overestimation of ground elevation in LiDAR-derived DTMs can introduce directional bias in canopy height calculations, which can result in significant underestimation of biomass if LiDAR return density is too low.

[6] Single- and multivariate regression models need to be explored to investigate which method and what metrics are most suitable for establishing LiDAR-biomass predictive relationships for tropical forests in complex terrain, such as the *Serra do Mar* site.

6. ABOVEGROUND BIOMASS AND DYNAMICS

6.1 Background

One important question in tropical forest ecology is how vegetation structure, composition and productivity are affected by the variation in local environmental conditions (LOSOS; LEIGH, 2004). Small-scale natural disturbances, for example tree falls, greatly influence the light regime in the forest canopy, and other abiotic factors, such as soil structure and moisture regime can also contribute to habitat or microhabitat differences. At least some plant species respond strongly to these environmental variations, such as gap-loving pioneer trees, while other plant species may respond more subtly by creating a shift in species composition between gap and shade or across soil boundaries (CONDIT, 2004). In the last few decades, extensive research has been conducted on gap phase dynamics, examining the mechanisms of forest regeneration and tree growth. As a result, the importance of tree fall gaps in creating habitat heterogeneity and affecting overall forest dynamics is generally accepted, especially for tropical ecosystems (DENSLOW, 1980; WHITMORE, 1984; MABBERLEY, 1992).

Less attention, however, has been paid to other factors such as topographic variation, soil type, and the efficiency of plants utilizing available resources (POULSEN, 1996; CLARK et al., 1998; PALMIOTTO, 1998). Tropical forest composition, structure and diversity characteristically undergo changes along elevational gradients (VÁZQUEZ; GIVNISH, 1998; HOMEIER et al., 2008). General trends observed with increasing altitude include overall decline in forest height, tree species diversity and live aboveground biomass, while stem density tends to increase with elevation (e.g. AIBA; KITAYAMA, 1999; LEUSCHNER et al., 2007; LIEBERMAN et al., 1996). In addition to altitude, topography of the substrate underlying a forest is an important factor that can influence species

composition and vegetation structure by providing microhabitat heterogeneity (e.g. CLARK et al., 1999; HARMS et al., 2001; PHILLIPS et al., 2003). The slope of the forest floor, for example, affects water drainage and the leaching of nutrients (ASHTON; HALL, 1992), and a sloping plane can also provide more light resources and greater access to space for trees to grow.

Many studies have used forest plots to investigate the structure of tropical forests, showing differences in tree density and basal area across slopes, plateaus, ridges and valleys. Manokaran and LaFrankie (1990) proposed that the impact of topographic variation may differ across size classes, though the form of this relationship is unclear. Additionally, the effects of horizontal gradients on vegetation in tropical mountains have not been studied extensively, and the mechanisms underlying elevational and topographical changes of forest structure and diversity are also not well known (HOMEIER et al., 2008). To gain a better understanding of the effects of topography on vegetation pattern and dynamics it is necessary to formulate and test hypotheses and create predictive models for the topographic drivers of landscape dynamics (DORNER et al., 2002). Multivariate regression techniques can be employed to build predictive models of topographic influence and estimate the degree to which patterns are determined by topography, while spatial correlograms and Mantel tests offer alternatives to classical statistical methods when spatial scaling of associations between vegetation mosaic and physical landscape is of concern (LEGENDRE; FORTIN, 1989; LEDUC et al., 1992).

This analysis aims to evaluate whether LiDAR-derived metrics of canopy structure (gap fraction, mean canopy height) and topographic variation (elevation, slope, aspect) can serve as good predictors of different components of aboveground biomass dynamics (gain, loss, net change) as well as the overall biomass stock and basal area in the Atlantic Forest of the *Serra do Mar*. Specifically, the following research questions will be investigated:

[1] What is the distribution of canopy heights in the permanent plots and how does it compare to canopy structure in the surrounding landscape?

[2] Is canopy height a good predictor of field-estimated aboveground biomass stocks and basal area?

[3] What is the distribution of gap size and gap area in the permanent plots and is it representative of the gap distribution in the surrounding landscape?

[4] How does gap fraction compare to field estimates of biomass gain, loss and net change in 1-ha permanent sample plots?

[5] Does terrain slope and/or aspect have a significant influence on biomass dynamics and/or canopy structure?

6.2 Material and methods

6.2.1 Field surveys and biomass dynamics

Live aboveground biomass variation in the study area was assessed from tree diameter survey data collected in the *Serra do Mar* permanent plot network (Table 6.1) over a 6-year period starting in 2006/07. The survey protocol followed standard tropical forest methodologies for tree growth measurements (ALVES et al., 2010): all trees ≥ 4.8 cm DBH (Diameter at Breast Height, 1.30 m) were tagged and mapped to a horizontal distance of ± 10 cm in each 1-ha permanent plot, and tree circumference was measured within ± 1 mm. Following the initial census in 2006/07, repeated forest inventories were carried out in these permanent plots in 2008/09 and 2011/12, the last of which coincided in time with the LiDAR data collection.

Table 6.1 Summary information about the permanent plots along the Atlantic Rainforest altitudinal gradient, Serra do Mar State Park – SP. (Adapted from JOLY et al., 2012)

Project/ Funding	BIOTA Functional Gradient Project/ FAPESP Funding (proc. No. 03/12595-7)
Location (State, Conservation Unit, Municipality and Geographical Coordinates)	Northeastern region of the State of São Paulo, Serra do Mar State Park – Picinguaba and Santa Virginia Centers, municipality of Ubatuba (Restinga, Lowland and Submontane) and São Luiz do Paraitinga (Montane) – 23°19'31" - 23°22'52" S and 44°49'55" - 45°05'02" O
Year of installation	2006/2007
Measurement periodicity	Biennial
Biome/ Forest formation	Restinga and Dense Atlantic Rainforest in the Lowland, Submontane and Montane areas
Form	Square plots
Plot size (m²) **– Dimensions** (m x m)	1 ha (100 x 100 m), divided into 10 x 10 m subplots
Quantitative variables	CBH (Circumference at Breast Height) of all trees, palms and treeferns with DBH (Diameter at Breast Height) ≥ 4.8 cm; total height; stem height; x-y position inside the plot
Qualitative variables	Crown Illumination Index; Trunk decomposition grade (when dead); Crown quality; Presence of bamboo and liana in the crown; Trunk quality; Floristic identification
Number of plots installed	14
Sampled area (ha)	4 ha in each of the Lowland, Submontane and Montane areas + 1 ha Lowland disturbed forest + 1 ha Restinga
Additional Details	All plots were georeferenced and individuals permanently marked with numbered aluminium plates located 30 cm above the point of measurement (Breast Height = 1.30 m)

Forest carbon stocks can be estimated using biomass allometric models that are developed by harvesting and weighing trees to determine their dry mass and allow for the calculation of aboveground biomass based on tree size. The simplest allometric models are of the form:

$$\text{AGB} = a \times D^{\,b} \tag{6.1}$$

where AGB is aboveground biomass (kg), D is stem diameter (cm), and a and b are parameters estimated empirically. In equations developed from field measurements at a

single site, diameter can usually explain the majority of variation in the aboveground biomass of individual trees. Yet, to obtain regionally comparable biomass estimates in highly heterogeneous tropical forests, it is necessary for the allometric model to incorporate additional parameters that account for evolutionary and environmental variations of forest characteristics at regional scales. Recent research focusing on biomass allometric models in tropical forests has shown that the inclusion of tree height as an additional size covariate, as well as the inclusion of one environmental factor and one functional trait (annual precipitation and wood specific gravity) produced a highly robust generic model leading to accurate estimates of tree biomass (CHAVE et al., 2005).

Given the significant variability of the Atlantic Forest biome along the coast of southeast Brazil (VIEIRA et al., 2008) and in lack of an allometric model developed specifically for the old-growth forest of the *Serra do Mar* site, this study applied Chave's et al. (2005) pan-tropical biomass allometry to estimate live aboveground biomass of all trees in the 1-ha permanent plots. This allometric model is based on the following relationship:

$$\text{AGB}_{\text{tree}} = (\ 0.0509 \times \rho \times DBH^{2} \times H\ )$$ (6.2)

where AGB_{tree} is aboveground tree biomass (kg), ρ is species level wood density (g cm^{-3}), *DBH* is tree diameter at breast height measured in the field (cm), and *H* is total tree height (m) estimated from the DBH values based on stand-specific allometric equations that were developed in the study area by Scaranello and colleagues (2012). Live aboveground biomass estimates for palms were calculated using the allometric equation developed by Hughes (1997, cited in ALVES et al., 2010):

$$\text{AGB}_{\text{palm}} = \{\exp[\ 0.9285 \ln(DBH^{2}) + 5.7236\ ] \times 1.050001\} \div 10^{3}$$ (6.3)

59

where AGB_{palm} is aboveground palm biomass (kg), and the only input variable is *DBH*, the diameter at breast height of the palm. The biomass of tree ferns was considered to have only negligible contribution to total aboveground biomass estimates (0.5% or less) in the area of interest, therefore tree ferns were not included in the present study (L. F. ALVES, 2013 personal communication).

Plot-level aboveground biomass estimates were calculated by summing up the individual tree and palm biomass values within each plot for the separate forest inventory dates (2008/09 and 2011/12) and converting these values to mass per unit area, expressed in Mg ha^{-1}. Total basal area was also calculated on the plot level (m^2 ha^{-1}) from the DBH data of the third survey (2011/12). Estimated components of aboveground biomass dynamics over the 3.5-year time period between the second and third surveys included:

(i) AGB gain from growth and recruitment;
(ii) AGB loss to mortality; and
(iii) AGB net change (growth + recruitment – mortality)

all expressed in the units Mg ha^{-1} $year^{-1}$.

It is important to note that several different approaches exist for calculating changes in aboveground biomass, distinguished primarily by the methods applied to quantify biomass gains due to recruitment and losses due to mortality (e.g. CLARK et al., 2001; CHAVE et al., 2003; MALHI et al., 2004). In this study, recruit biomass was estimated using the measured DBH of all recruits at the time of the forest survey and dividing the sum by the time elapsed between surveys. For quantifying mortality, an individual found dead at the time of the forest survey was considered to have lost its total biomass during the time period since the last survey. Given that the time period considered in this study was longer than one year and that the exact point in time when the individual died was unknown, it was

inevitable to introduce some uncertainties in the estimates of yearly mortality rates (L.F. ALVES, 2013 personal communication).

6.2.2 LiDAR-derived terrain and canopy metrics

The Canopy Height Model (CHM) generated from all vegetation returns in the full-density LiDAR point cloud following the G-LiHT methodology (COOK et al., 2013) described in section 5.2.1 was used in this part of the study (Figure 6.1). The LiDAR data were associated with forest plots based on the differential GPS measurements previously described in Section 4.2.3. Samples corresponding to the 1-hectare permanent field plots were extracted from the whole-area CHM raster dataset, and the distributions of canopy heights (1-m^2 raster cells) within these plots were described and compared. To examine the broader landscape surrounding the 1-ha field plots, canopy height distributions were also generated for the greater submontane and montane forest areas. The two forest regions were distinguished based on elevation, with submontane forests occupying the lower-lying hillslopes between 100 - 500 m elevation and montane forests lying at 900 m or higher on top of the plateau (see Figure 6.1). In addition, descriptive statistics were generated from the LiDAR-derived DTM for the 1-ha permanent plot locations, including terrain elevation, slope and aspect – all with 1-m raster resolution.
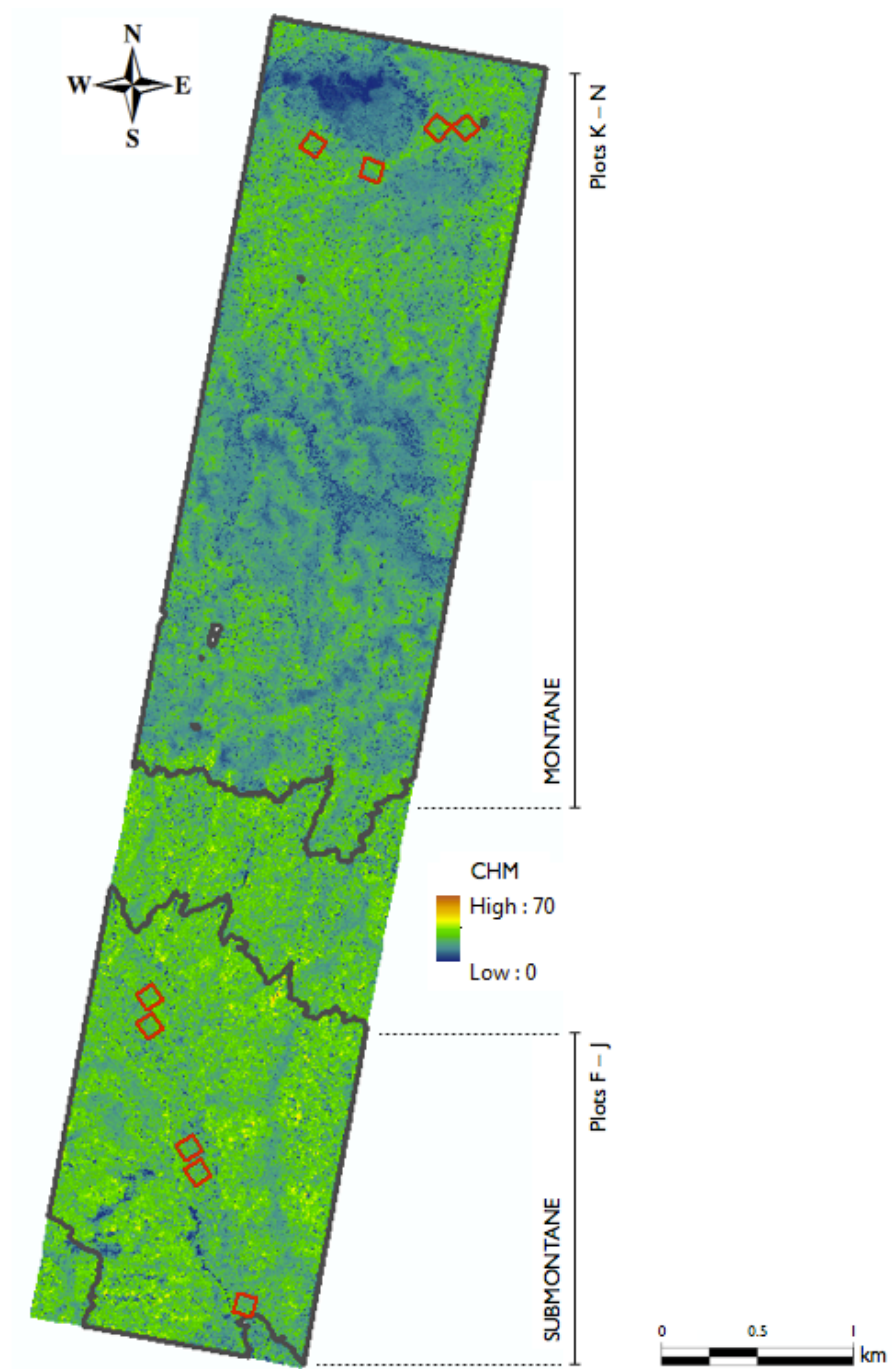
Figure 6.1 Canopy Height Model (CHM) from LiDAR in the 1000-ha study area. The nine permanent field plots are highlighted in red; thick grey boundary lines indicate the submontane and montane regions used for landscape-scale comparisons.

In order to investigate the disturbance regime of forest canopies in the permanent plots, gap size-frequency distributions were generated from the LiDAR CHM. Canopy gaps were identified using a definition similar to that of Brokaw (1982), where a gap is considered to be an opening in the forest canopy that extends down to a defined height above the ground surface. Here, a range of height thresholds were tested (from 1m up to 20m in 1-m increments) for gap identification in the CHM raster image. Two height definitions that optimized the sensitivity of gap fraction to components of biomass dynamics were selected for further analysis. The CHM image was processed in a GIS environment following the 5-step approach described in Asner et al. (2013) to distinguish canopy gaps:

(1) *Identify a pixel that is < the minimum-height threshold. Assign this pixel a unique label.*

(2) *Evaluate the 8 neighbors of the pixel identified in step 1. If they are < minimum-height threshold, and therefore part of the same canopy gap, assign them the same label.*

(3) *Repeat Step 2. until there are no neighbors < the minimum-height threshold. This completes the identification of a single canopy gap.*

(4) *Repeat Steps 1 - 3 until there are no remaining unlabeled pixels in the image.*

(5) *Count the number of pixels within each uniquely-labeled canopy gap.*

The gap statistics derived from the CHM (i.e. number and size of gaps) were further analyzed and visualized in the R statistical environment (R Core Team, 2011) using the source code and detailed instructions provided by Asner et al. (2013). Gap size-frequency distributions were generated based on the Zeta distribution, which is a discrete power-law probability density used to describe the size-frequency of canopy gaps (ASNER et al., 2013). As stated by the authors, for the Zeta distribution with parameter λ, the probability that gap size takes the integer value k is:

$$f(k) = k^{-\lambda} / \zeta(\lambda) \qquad\qquad (6.4)$$

where ζ(λ) is the Riemann Zeta function, and is undefined for λ = 1. Maximum likelihood estimates (MLE) of the scaling exponent (λ) were calculated, and the resulting λ values were used to describe the disturbance regime of forest vegetation in the permanent plots. Values of λ typically range from 1.0 - 3.0 in forests (ASNER et al., 2013), where λ lower than 2.0 is indicative of large-scale forest disturbance (predominance of large-sized gaps), while λ above the 2.0 threshold suggests that disturbance occurs on much smaller scales (FISHER et al., 2008).

Plot-level LiDAR metrics were merged with the field-based biomass data for regression modeling in order to analyze how LiDAR-derived terrain and canopy indices (slope, aspect, gap fraction and canopy height) are related to aboveground biomass and dynamics.

6.3 Results

6.3.1 Terrain characteristics

Figure 6.2 shows the 3D perspective view of the permanent field plots as well as histograms of the ground slope distributions at the 1-meter raster resolution extracted from the DTM. As it can be observed, the nine field plots are characterized by varying levels of steepness. The lower-lying submontane Plots F, G and H (elevation range: 99 - 217 m) as well as the montane Plot N have gentler slopes of 13.7º, 13.8º, 11.1º and 16.6º, respectively, while the rest of the plots (I and J, located between 300 - 400 m elevation, and K, L and M between 994 - 1064 m) are steeper in comparison with slopes ranging from 25.3º to 27.8º. In terms of the compass orientation of the downward slope (i.e. terrain aspect) there is also large variation both among plots and within plots (Figure 6.3). Overall, the submontane plots are oriented towards the southern directions, and montane plots are northward-facing.
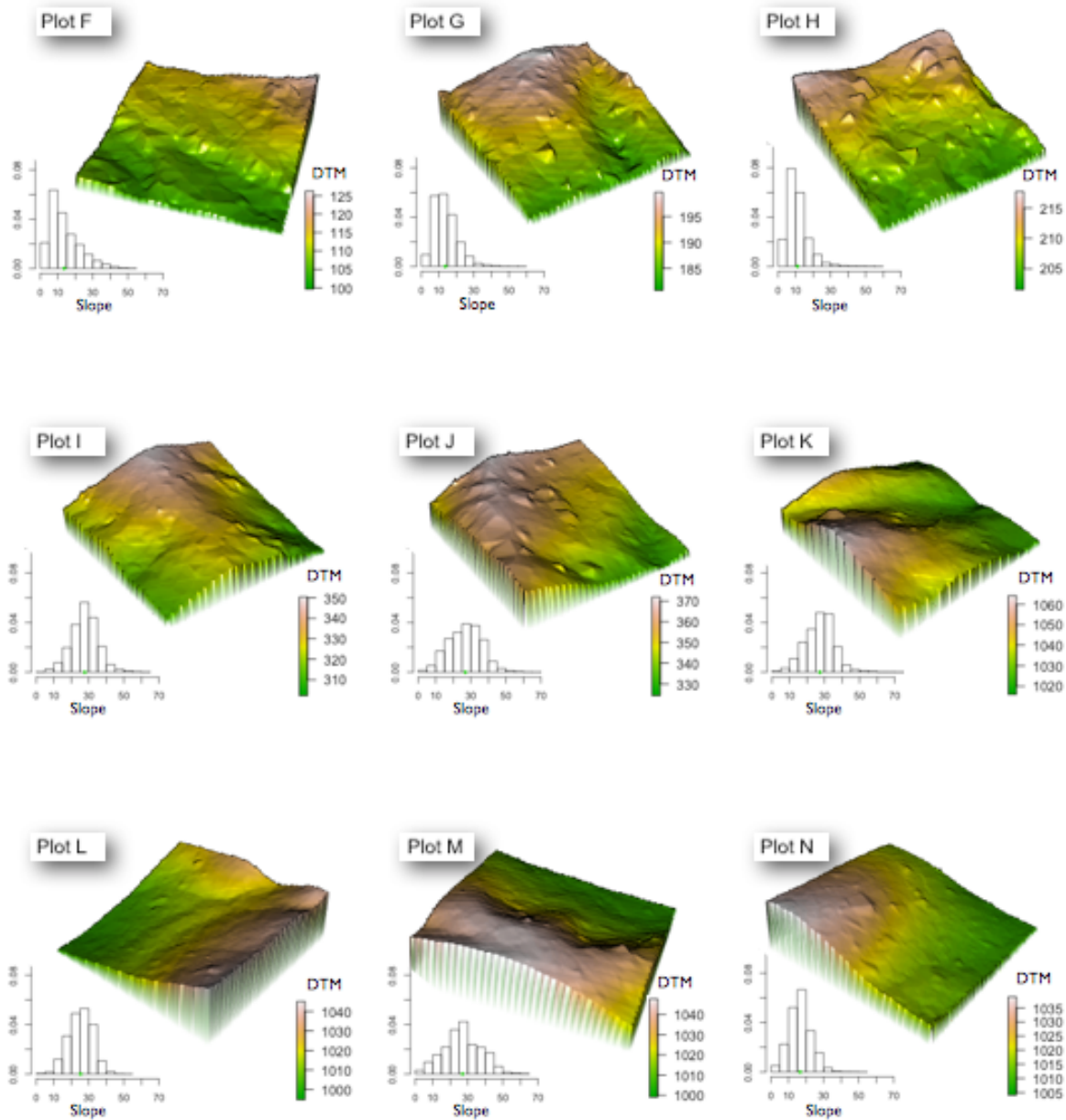
Figure 6.2 Terrain elevation from DTM (3D perspectives) and frequency distribution of ground slopes (histograms) in the nine permanent field plots: F–J submontane and K–N montane. Note that each plot is represented with its own color bar (elevation in meters); color schemes are not inter-comparable among plots.
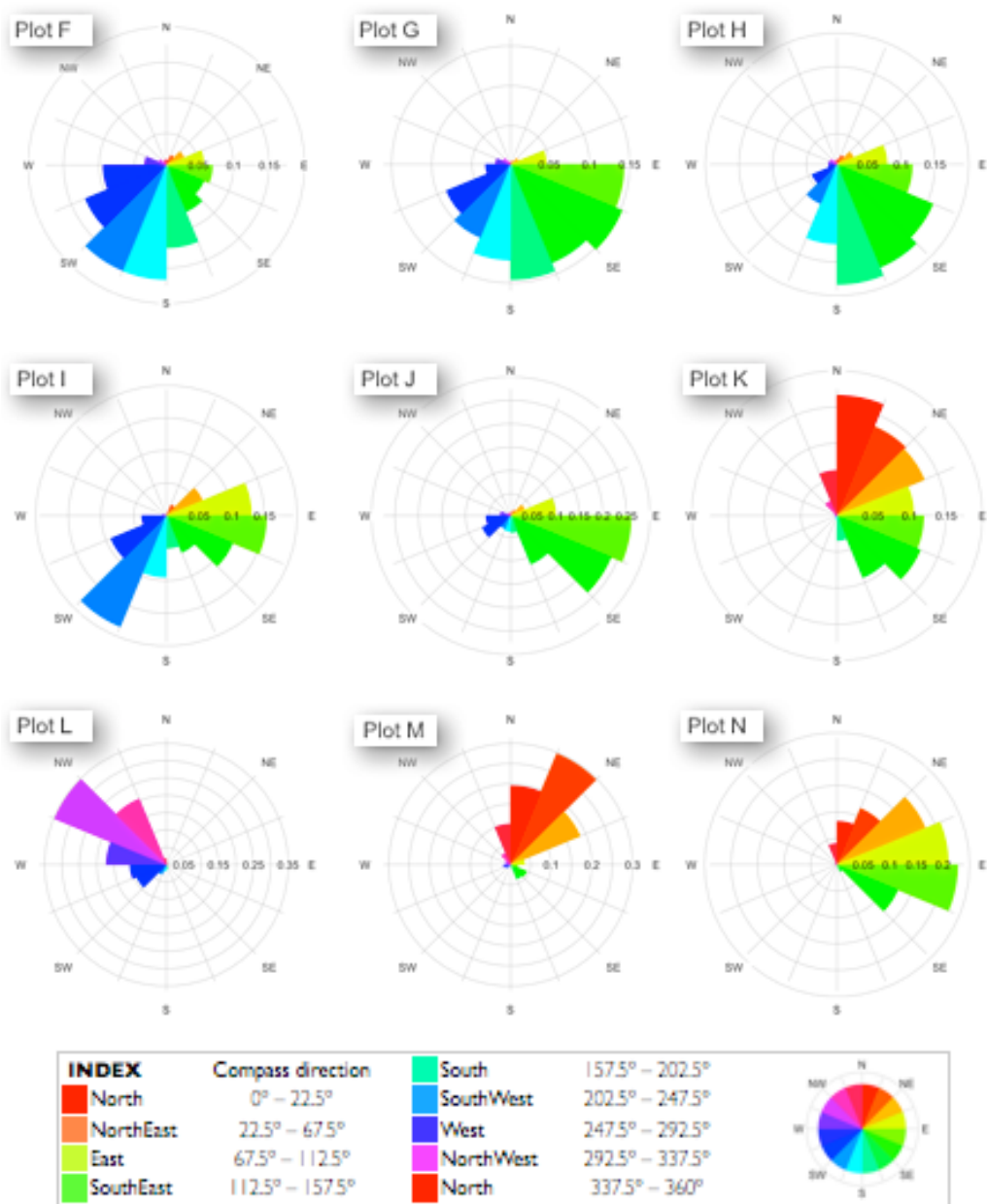
Figure 6.3 Distribution of the compass directions of the terrain aspect (at 1-m raster resolution) in the permanent plots: F–J submontane and K–N montane.

6.3.2 Canopy heights and gap distributions

The 1-ha samples extracted from the 1-m resolution CHM raster for the nine permanent plot locations are shown in Figure 6.4. The corresponding distribution of canopy heights (i.e. distribution of CHM raster cell values) in each plot is displayed in Figure 6.5, and associated descriptive statistics are summarized in Table 6.2 along with the summary of terrain characteristics. Canopy heights reach their maxima at 30.9 to 37.9 meters, and most plots have their mean around 20 meters, except for Plots I and J, where the means are higher (22.2 and 22.9 meters, respectively). Standard deviations of the canopy heights range from 4.1 meters in Plot I to 5.4 meters in Plot N. Pairwise comparisons of the canopy height distributions in plots (using the *TukeyHSD* function in R with 95% confidence level) showed statistically significant differences (p-value < 0.001) among most plots. The only pairs where the difference was not statistically significant were: Plots F and N (p-value = 0.998), Plots L and M (p-value = 0.999), Plots M and G (p-value = 0.202) and L and G (p-value = 0.124), indicating that the distributions of canopy heights in these plots follow similar patterns.
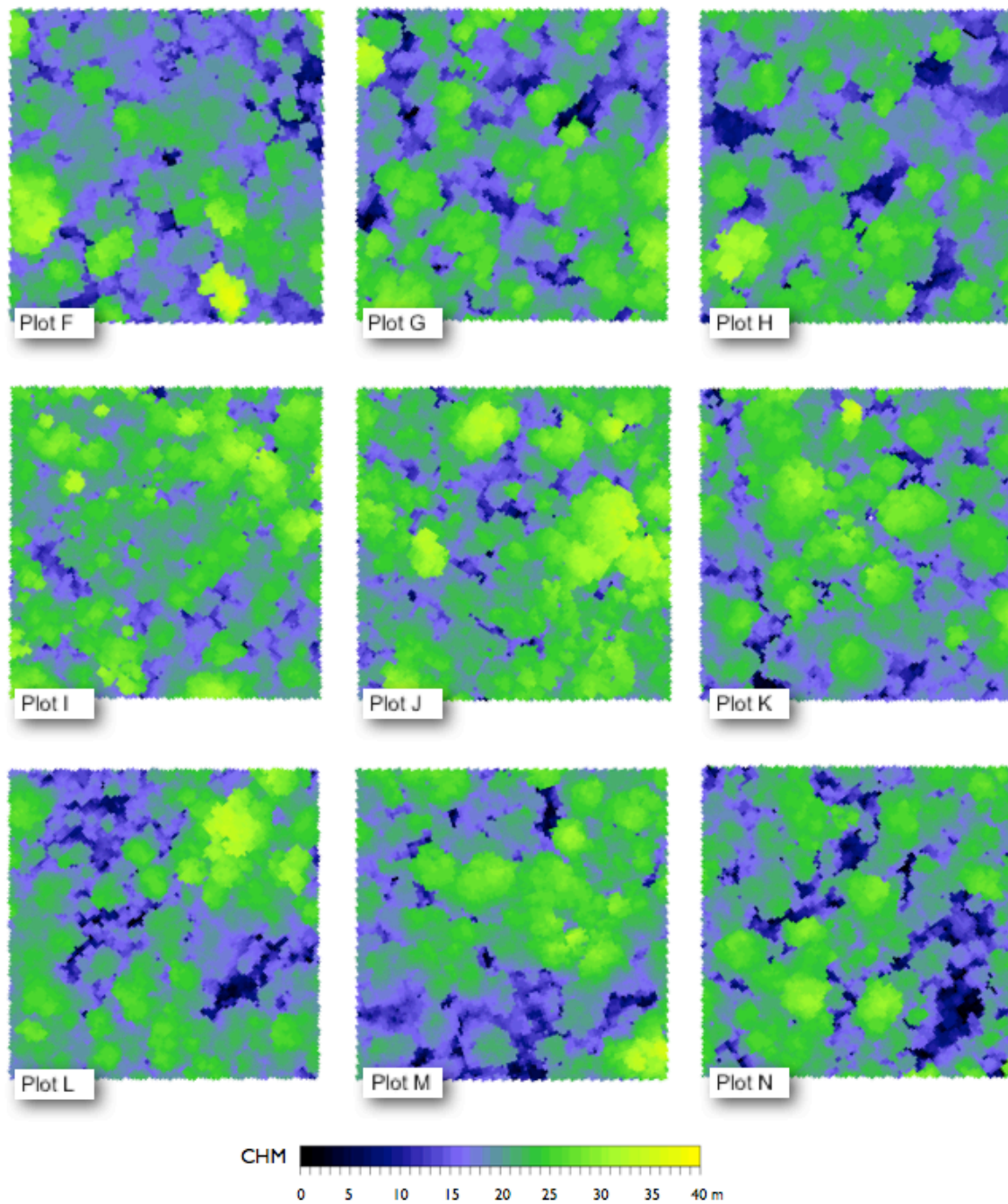
Figure 6.4 Canopy height from CHM at 1-m spatial resolution within the nine permanent Plots F–N. Light shades of green and yellow mark the tallest tree crowns in the canopy, while dark purple shades indicate the locations within the plots where canopy openings were identified.
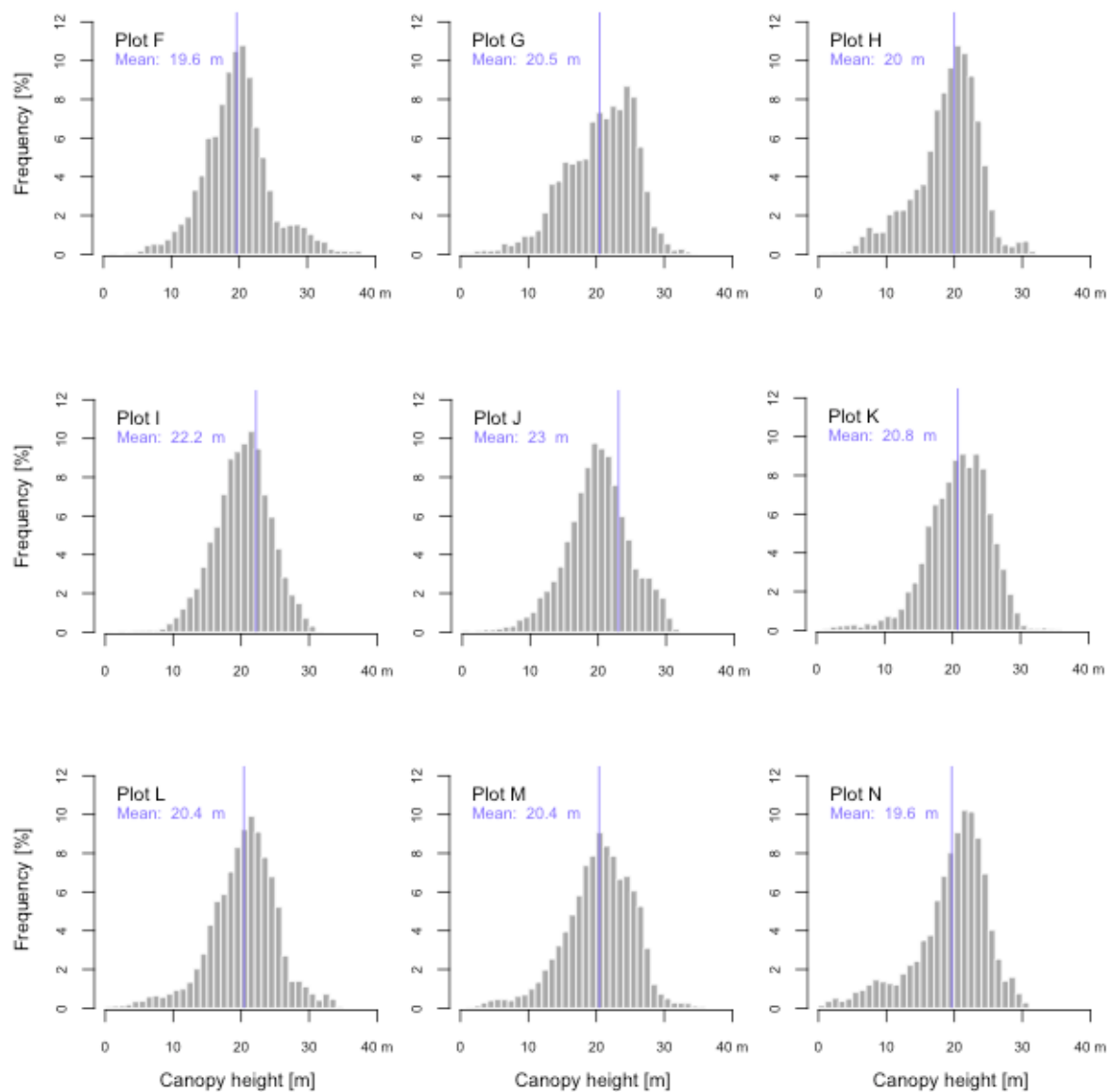
Figure 6.5 Height frequency distributions from the 1-m resolution LiDAR Canopy Height Model (CHM) for the nine permanent plots: F–J submontane and K–N montane.

Table 6.2 LiDAR-derived canopy height metrics and terrain characteristics in permanent Plots F–N.

Plot ID		min – max [m]	mean ± stdev [m]	mode [m]	Elevation [m] min – max	Slope [°] mean ± stdev	Aspect [°] mode (dir)
		Canopy Height			**Terrain Characteristics**		
SUB MON TANE	F	0.8 – 37.7	19.6 ± 4.9	20.1	99 – 126	13.7 ± 8.7	200 (S)
	G	0.2 – 34.3	20.6 ± 5.2	24.7	180 – 199	13.8 ± 6.8	128 (SE)
	H	1.4 – 32.5	19.9 ± 4.7	21.8	201 – 217	11.1 ± 6.2	170 (S)
	I	2.6 – 33.1	22.2 ± 4.1	23.7	302 – 350	27.8 ± 7.9	214 (SW)
	J	3.7 – 34.8	22.9 ± 4.7	23.7	324 – 372	26.9 ± 9.7	105 (E)
MON TANE	K	0.0 – 35.2	20.8 ± 4.6	21.3	1015 – 1064	27.4 ± 8.6	18 (N)
	L	0.2 – 34.4	20.4 ± 4.9	21.7	994 – 1044	25.3 ± 6.7	308 (NW)
	M	0.0 – 35.6	20.4 ± 5.1	20.8	999 – 1047	27.0 ± 10.7	39 (NE)
	N	0.2 – 30.9	19.6 ± 5.4	21.9	1004 – 1038	16.6 ± 6.2	108 (E)

Height frequency distributions from the 1-m resolution CHM for the broader submontane and montane forest areas surrounding the permanent plots are shown in Figure 6.6. Canopy heights in the submontane landscape show a symmetrical distribution with values centered around a mean of 22.5 m (± 5.9 m standard deviation). In comparison, the distribution of canopy heights in the montane region was significantly different (Kolmogorov-Smirnov test, p-value < 2.2e-16), with a mean value of 16.9 m (± 6.7 m standard deviation) and characterized by a heavy lower tail. Two-sample Kolmogorov-Smirnov tests were used to compare plot-level canopy height distributions to their surrounding landscapes. In almost all cases, the distribution of canopy heights in the 1-ha sample plot was significantly different from the canopy height distribution in the broader landscape (p-value < 2.2e-16). There was only one plot overall where canopy heights showed some correspondence to the distribution in the landscape: Plot I (p-value = 0.3094), located in the submontane area.
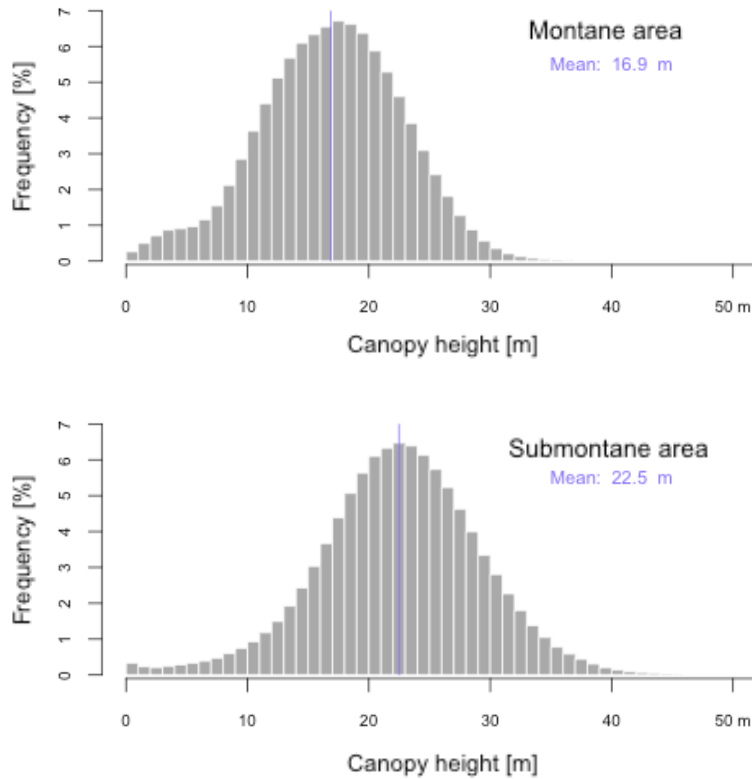
Figure 6.6 Height frequency distributions for the montane and submontane forest areas from the 1-m resolution LiDAR Canopy Height Model (CHM).

Canopy gap sizes and frequencies (Table 6.3) were quantified using two alternate gap definitions: (i) gaps extending down to 14 m above the forest floor and having a contiguous area of at least 10 m^2 – hereby denoted as "*mid-canopy gaps*" – that showed greatest sensitivity to aboveground biomass loss and net change, and (ii) gaps extending down to 5 m above ground and having a minimum area of 1 m^2 – denoted as "*low-canopy gaps*" – that were most sensitive to biomass gain. The maximum gap height thresholds tested for this optimization ranged from 1 m to 20 m in 1-meter increments, while the minimum gap area cutoff values tested were 1, 5, 10, 25, 50 and 100 m^2. The selection criteria for finding the best fit height and area thresholds to biomass components were largest R^2 and smallest RMSE values.

Table 6.3 Canopy gap fraction and scaling exponent values in the 1-ha sample plots and their surrounding landscapes calculated from the CHM raster using two different gap definitions (*low-canopy gap* and *mid-canopy gap*).

PLOT ID	Low-canopy Gaps (H < 5m; A > 1m^2)					Mid-canopy Gaps (H < 14m; A > 10m^2)				
	Gap count	Mean size [m^2]	Max. size [m^2]	% Gap	Scaling exp. (λ)	Gap count	Mean size [m^2]	Max. size [m^2]	% Gap	Scaling exp. (λ)
F	2	5.0	8	0.10	2.53	19	48.8	260	9.27	1.49
G	6	4.3	13	0.62	1.66	16	66.1	254	10.57	1.44
H	7	3.0	5	0.21	1.92	10	109.2	367	10.92	1.50
I	2	2.0	2	0.04	2.14	9	24.1	83	2.17	1.61
J	3	2.7	3	0.08	1.86	9	26.8	69	2.41	1.66
K	8	7.9	43	0.63	1.99	20	28.0	73	5.60	1.55
L	12	4.8	25	0.69	1.93	17	49.1	213	8.34	1.53
M	11	6.9	24	0.76	1.72	11	87.6	522	9.64	1.53
N	25	5.3	52	1.71	1.71	19	67.7	380	12.86	1.54
SUBMONTANE	7.9 /ha	15.8	1,492	1.25	1.72	12.2 /ha	68.5	6,786	8.37	1.55
MONTANE	15.7 /ha	19.9	42,865	3.13	1.69	10.6 /ha	281.3	521,308	29.82	1.57

Using the low-canopy gap definition, very few canopy openings were found in the submontane plots (4 gaps on average) and slightly more in the montane plots (14 gaps per plot on average). Mean gap size with this definition was 4.65 m^2 overall, and the largest gap reached an area of 52 m^2 in the montane Plot N. Gap fraction was highest in Plot N (1.71%), lowest in Plots I, J, F and H (0.04 - 0.21%), and intermediate in Plots G, K, L and M (0.62 - 0.76%). The calculated scaling exponent (λ) had an average value of 1.94 across all plots, with a low of 1.66 in Plot G and high of 2.53 in Plot F. With the mid-canopy gap definition 15 gaps were found on average per plot, distributed evenly across all plots. Mean gap size ranged from 24.1 to 109.2 m^2 in Plots I an H, respectively, and the largest gap found was 522 m^2 in area (Plot M). Plot N still had the highest gap fraction (12.86%) with

this definition among all plots, but Plot G, H, M and F had almost as high values (9.27 - 10.92 %). The lowest gap fractions were observed in Plots I (2.17%) and J (2.41%). The scaling exponent with the 14-m gap definition was more uniform overall (mean of 1.54), ranging between 1.44 (Plot G) and 1.66 (Plot J).

When compared with the plot-scale gap size, count and gap fraction values, the landscape-scale gap statistics showed variable correspondence. Mean and maximum gap sizes were significantly larger on the landscape scale than in the 1-ha plots, and gap fraction values also followed this trend (see Table 6.3 above). Additionally, low-canopy gap statistics showed larger discrepancies between the plot and landscape scales than mid-canopy gap statistics. Scaling exponent values with the low-canopy gap definition were 1.72 and 1.69 in submontane and montane regions, respectively, significantly lower than what was observed on the plot level. Using the mid-canopy gap definition, λ values were 1.55 and 1.57 in the submontane and montane regions, respectively, which correspond strongly to the mean value observed on the plot level. When comparing the submontane with the montane landscape (Figure 6.7), more gaps were found in the latter region, and those gaps were also larger in size. Both low-canopy and mid-canopy gap fractions were significantly higher in the montane landscape than in the submontane one.
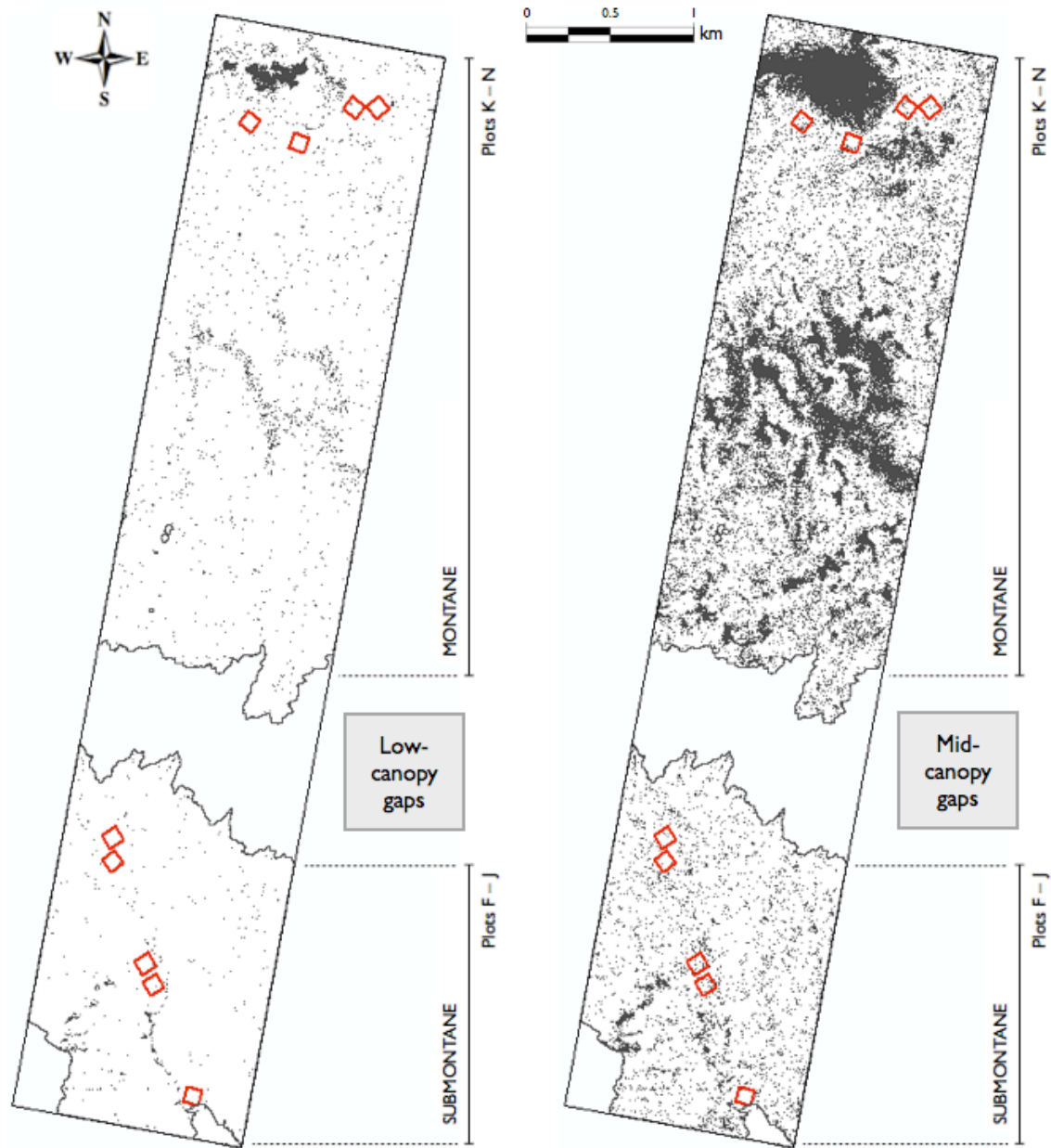
Figure 6.7 Distribution of low-canopy gaps (left) and mid-canopy gaps (right) within the 1000-ha LiDAR coverage in the *Serra do Mar*. Permanent field plots are highlighted in red; submontane and montane landscape boundaries are in grey.

6.3.3 Aboveground biomass dynamics

Total live aboveground biomass (AGB) in 1-ha plots at the time of the third survey (2011/12) ranged from 214.2 Mg ha^{-1} to 348.6 Mg ha^{-1} across all nine locations considered. Submontane plots had an overall mean AGB value (286.4 ± 46.5 Mg ha^{-1}) lower than montane plots (306.5 ± 39.9 Mg ha^{-1}). Overall net biomass change calculated from a 3.5-year period (2008/09 - 2011/12) was small and negative in Plots G, H and N, close to zero in Plot M, and positive in Plots F, I, J, K and L. Among the plots that had positive net change, I and J showed the largest increase in total AGB (3.1 Mg ha^{-1} y^{-1} and 2.7 Mg ha^{-1} y^{-1}). Mean annual AGB gain ranged from a low of 3.9 Mg ha^{-1} y^{-1} (Plot M) to a high of 5.7 Mg ha^{-1} y^{-1} (Plot F); on average, submontane plots showed higher rates of biomass gain (5.1 ± 0.6 Mg ha^{-1} y^{-1}) than montane plots (4.3 ± 0.6 Mg ha^{-1} y^{-1}). Calculated rates of mortality in the same time period were more variable across the nine plots, ranging form a minimum loss of 1.8 Mg ha^{-1} y^{-1} (Plot J) to a maximum of 5.7 Mg ha^{-1} y^{-1} (Plot N), however, when averaged within submontane and montane forests, mortality rates were not significantly different between these two areas (3.8 ± 1.6 Mg ha^{-1} y^{-1} and 3.8 ± 1.4 Mg ha^{-1} y^{-1} in submontane and montane forests, respectively). Mean basal area in the submontane and montane areas showed little difference with overall values of 39.3 (± 6.3) m^2 ha^{-1} in the submontane and 39.4 (± 4.0) m^2 ha^{-1} in the montane forest. A comparison of individual plots showed more variation: Plot F had the lowest basal area with 31.7 m^2 ha^{-1}, Plots G, H, L and N had intermediate values around 36 m^2 ha^{-1}, while the highest basal areas were observed in Plots M, I and J, reaching 46.5 m^2 ha^{-1}. Live AGB statistics for the nine plots are summarized in Table 6.4 below.

Table 6.4 Components of aboveground biomass and dynamics in the nine permanent plots F–N between the time period 2008/09-2011/12 (2nd and 3rd forest inventories).

	Submontane plots						Montane plots				
BIOMASS COMPONENTS	F	G	H	I	J	mean ± stdev	K	L	M	N	mean ± stdev
Total Biomass2 (Mg ha⁻¹)	214.2	289.5	268.4	308.5	328.1	**281.7** ± 43.8	326.5	268.4	348.5	276.3	**304.9** ± 38.8
Total Biomass3 (Mg ha⁻¹)	219.8	286.5	267.1	319.9	338.7	**286.4** ± 46.5	332.5	270.5	348.6	274.5	**306.5** ± 39.9
Total Basal Area3 (m² ha⁻¹)	31.7	37.4	35.8	46.5	45.0	**39.3** ± 6.3	41.7	36.1	43.8	35.8	**39.4** ± 4.0
GAIN 2–3 (Mg ha⁻¹ year⁻¹)	5.73	4.22	5.20	5.51	4.57	**5.1** ± 0.6	4.39	3.98	3.87	5.08	**4.3** ± 0.5
LOSS 2–3 (Mg ha⁻¹ year⁻¹)	4.41	4.96	5.51	2.45	1.84	**3.8** ± 1.6	2.38	3.27	3.83	5.71	**3.8** ± 1.4
NET Change 2–3 (Mg ha⁻¹ year⁻¹)	1.32	-0.74	-0.31	3.06	2.74	**1.2** ± 1.7	2.01	0.70	0.04	-0.63	**0.5** ± 1.1
TIME elapsed 2–3 (years)	4.2	4.1	4.1	3.7	3.9	**4.00**	3.0	3.0	2.1	2.9	**2.75**

6.3.4 LiDAR metrics vs. biomass components

Figure 6.8 shows the relationship between demographic components of AGB dynamics and canopy gap fraction in the nine 1-ha permanent plots based on ordinary least squares (OLS) regression. The regression coefficients are summarized in Table 6.5 below. Mid-canopy gap fraction was very strongly related to net aboveground biomass change in the 1-ha plots (R^2 = 0.91, p-value < 0.001): observed net positive change in biomass coincided with areas where mid-canopy gap fraction was low (Plots I, J, K, F and L) and net negative change was observed in plots where mid-canopy gap fraction was high (Plots G, H and N). Mortality rates were also significantly related to mid-canopy gap fraction (R^2 = 0.87, p-value < 0.001), predicting large mid-canopy gap fraction values where aboveground biomass loss was high. Variation in biomass gain did not explain well gap fraction (R^2 = 0.01, p-value = 0.85) when using the mid-canopy definition for comparison, and this relationship improved only slightly with the low-canopy definition (R^2 = 0.1, p-value =

0.42). However, by removing one point from the analysis (Plot N with unusually high gap fraction compared to other plots), 76% of the variation in low-canopy gap fraction was explained by biomass gain (p-value < 0.005).
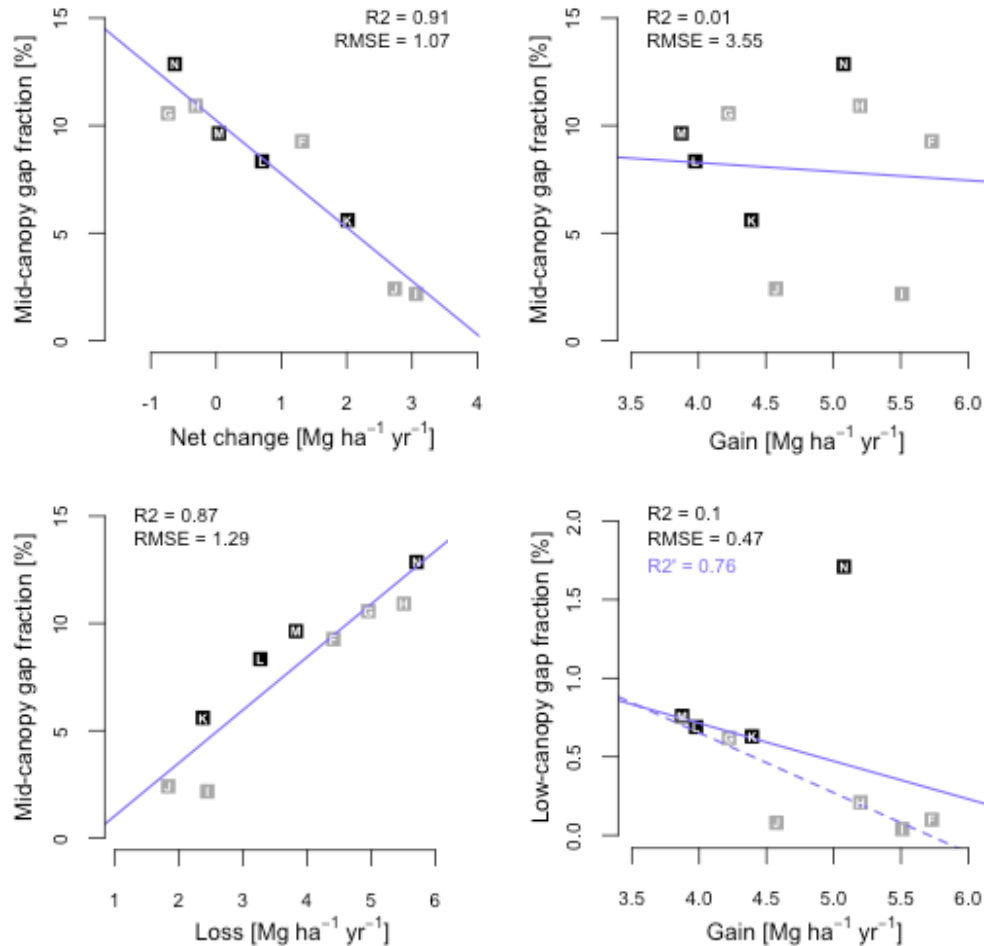
Figure 6.8 Demographic components of aboveground biomass dynamics in relation to canopy gap fraction in the nine 1-ha permanent Plots F–N. Regression lines are plotted in purple, grey squares indicate submontane plots, while montane plots appear in black. Note that biomass gain was compared to gap fraction using both the *mid-canopy* (top right graph) and the *low-canopy* (bottom right graph) gap definitions. (Dotted line in bottom right graph is the regression line after removal of Plot N from the analysis.)

Table 6.5 Regression coefficients between biomass components and canopy gap fraction.

BIOMASS COMPONENTS	Intercept	Slope	R²	p-value	Res. SE	df
Net change	10.24	-2.49	0.91	<0.001	1.22	7
Gain (vs. "mid canopy gap")	9.93	-0.41	0.01	0.85	4.03	7
Gain (vs. "low canopy gap")	1.68	-0.24	0.10	0.42	0.54	7
Loss	-1.45	2.47	0.87	<0.001	1.46	7

Mean canopy height predicted 43% of the variation in total aboveground biomass (p-value = 0.054), with a positive relationship between these two variables (Figure 6.9). Basal area was also positively and significantly related to mean canopy height (R^2 = 0.68, p-value = 0.006).
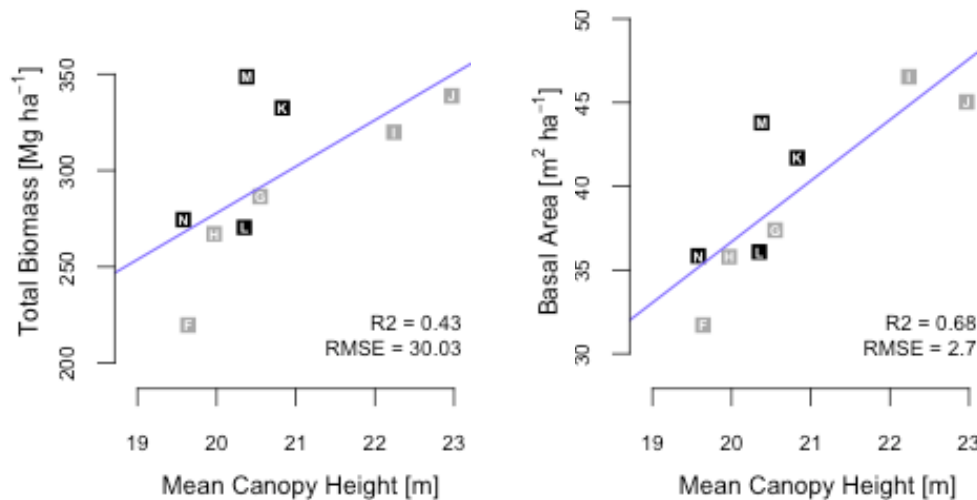
Figure 6.9 Total aboveground biomass and basal area in relation to LiDAR-derived mean canopy height in the nine 1-ha permanent Plots F–N. Regression lines are plotted in purple, grey squares indicate submontane plots, while montane plots appear in black.

To examine the effect of terrain slope on biomass dynamics and related metrics (see Figure 6.10), the nine permanent plots were grouped into two distinct slope classes that emerged from the data. The first group, characterized by gentler slopes (range: 11-16º, mean: 13.8º), includes Plots F, G, H and N, and the second group is composed of Plots I, J, K, L and M, with steeper slopes ranging from 25 to 28º (mean: 26.9º). Mean aboveground biomass as well as mean basal area were both significantly higher in the steeper plots than in the less steep ones (322.0 Mg ha^{-1} and 42.6 m^2 ha^{-1} in comparison with 261.9 Mg ha^{-1} and 35.2 m^2 ha^{-1}). Biomass gain rates did not differ notably between the two slope classes, on the other hand, yearly biomass loss was much lower on average in steeper plots (2.75 Mg ha^{-1} y^{-1}) than in the flatter ones (5.15 Mg ha^{-1} y^{-1}). Net biomass change indicated increase in biomass on steep slopes (+1.71 Mg ha^{-1} y^{-1}), while there was an overall small decrease in biomass observed in the plots with gentler slopes (-0.09 Mg ha^{-1} y^{-1}). Average gap fraction on gentler slopes was twice as large as the mean across the steeper plots (10.9 % vs. 5.6 %, respectively, using the mid-canopy gap definition). Finally, mean canopy height was similar in the two groups, with an average value of 19.9 m for the flatter plots and 21.4 for the steeper slopes, the difference being driven by Plots I and J that had the highest mean canopy heights among all plots. A similar analysis of biomass components and canopy structure using terrain aspect as the predictor variable as opposed to slope did not show any significant relationships at the 1-ha plot scale.
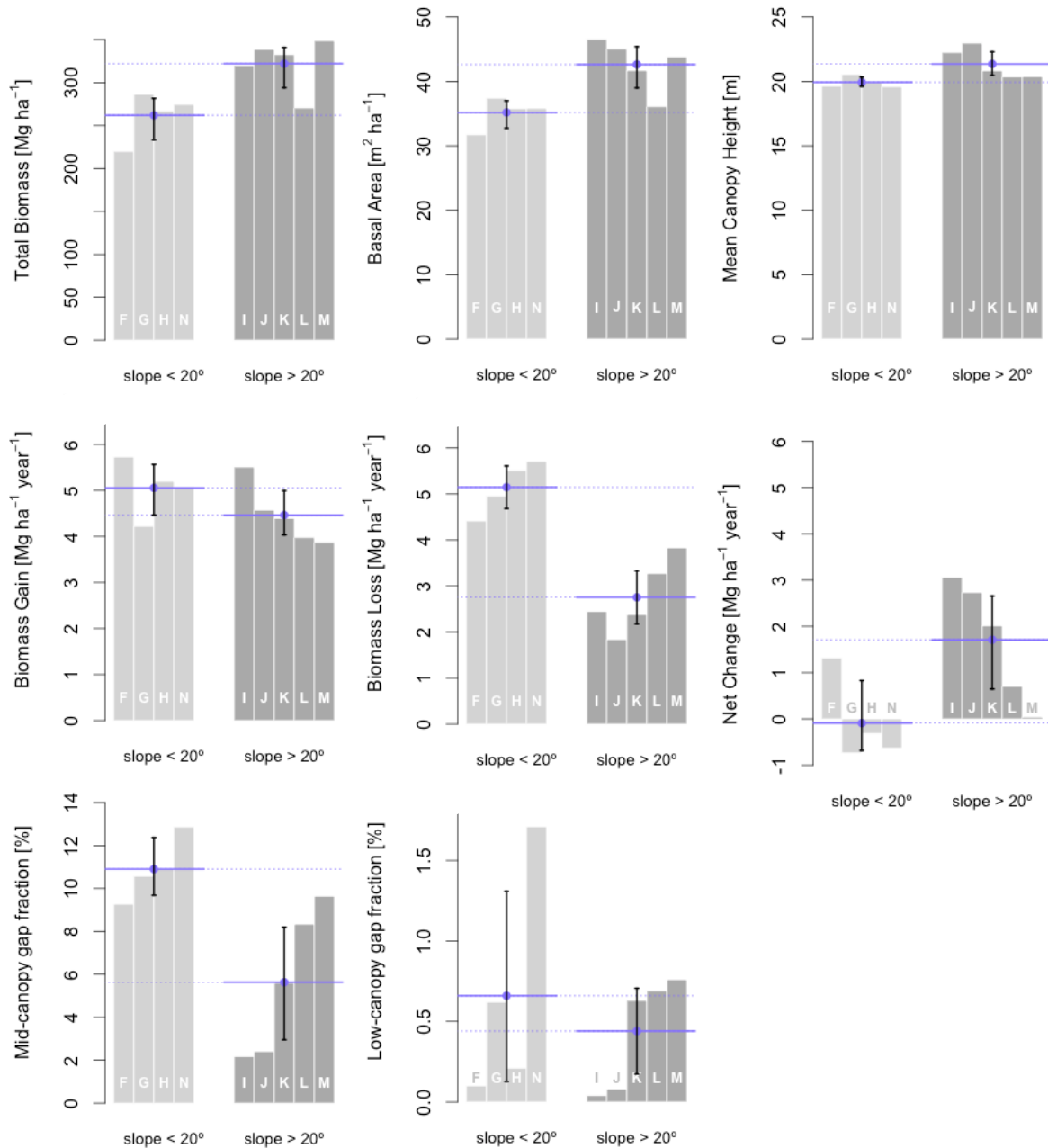
Figure 6.10 Variation of biomass components (total biomass, basal area, gain, loss) and canopy metrics (canopy height and gap fraction) as a function of ground slope in two groups: Plots F, G, H and N with gentler slopes (< 20º) and Plots I, J, K, L and M with steeper slopes (> 20º). Horizontal colored lines indicate the mean values within each slope class, error bars show 95% confidence intervals around the group means and each bar is labeled with the respective plot ID.

6.4 Discussion

LiDAR-derived canopy height distributions shown in this work were not validated directly by field observations. However, a recent study by Scaranello and collaborators (2012) in the *Serra do Mar* permanent plots documented tree heights that were measured using traditional field techniques for a random sample of trees stratified by DBH class. The range of field-measured tree heights was grouped into different altitude classes, and the reported values were 4.3 - 30.0 meters for the 400-m altitude class (i.e. Plots G, H, I and J or "*submontane plots*") and 5.2 - 29.4 meters for the 1000-m altitude class (i.e. Plots K, L, M and N or "*montane plots*"). Though direct comparison between the field-measured and LiDAR-derived tree heights is not appropriate given that the two sets of data were sampled differently, the range of canopy heights captured by the LiDAR (2.0 - 33.6 m and 0.24 - 34.0 m for the lower and higher altitude classes, respectively) corresponds well to those measured in the field.

There was considerable variation in the LiDAR-derived canopy height distributions among the 1-ha sample plots, and these plot-level samples also did not correspond closely with the canopy height distributions in the surrounding broader landscape. Especially in the montane region on top of the plateau, the observed discrepancy between plot-level and landscape-scale canopy structure was more pronounced: the four montane permanent plots capture only a limited area of the forest where the canopy is taller on average than it is across the broader landscape. Similar observations can be made with regard to the distribution of canopy gaps, where the plot-level gap statistics were not strongly representative of the distributions of gap sizes and frequencies in the broader montane and submontane areas. These effects might be explained, in part, by the small size of the permanent plots, indicating that 1-ha samples are not adequate to capture the disturbance regimes present in the landscape. The small number of field plots and their non-random spatial distribution

81

across the landscape are additional limitations that have to be considered before conclusions can be drawn from the analyses presented in this study.

It is also important to emphasize that several different gap definitions exist in the literature and various measurement approaches have been applied for the characterization and quantification of canopy gaps. In this study, gaps were evaluated empirically, and the two different gap definitions adapted for the analysis were based on relationships established between canopy height and aboveground biomass components in 1-ha field plots. Such localized gap definitions might not be appropriate when applied to the wider landscape or different forest sites. Additionally, since canopy gaps are defined in reference to a given height threshold above the ground surface, regional variations in canopy heights might influence gap characterization in individual plots. Therefore, it might be necessary to normalize canopy heights to plot-specific maximum height or upper quantile values before calculating gap fractions to obtain metrics that are inter-comparable across all plot samples.

Results from the regression analysis between aboveground biomass and LiDAR metrics showed that LiDAR-derived mean canopy surface height alone was a good predictor of biomass stocks in the 1-ha permanent plots at the *Serra do Mar* study site. This single LiDAR metric explained 43% of the variation in total aboveground biomass in plots and the associated RMSE value was low (30 Mg ha^{-1}). The predictive relationship might be further improved by using a different LiDAR-derived canopy metric – such as canopy rugosity or the mean height of the vertical vegetation profile – that captures information about the internal structure of the forest canopy. Alternately, multivariate regression techniques could be explored to make use of a variety of LiDAR metrics simultaneously – stepwise variable selection approaches or decision-tree based regression (e.g. *randomForest* in R) might be useful for this purpose.

This study found that field-estimated biomass was higher on steeper slopes and at greater altitudes than in lower-lying or flatter areas, and average canopy height derived from LiDAR also followed the same trend. Although this is contrary to the expectations along altitudinal gradients, where biomass and tree height usually decline with increasing elevation, previous studies have reported similar findings. Recent work along this elevational gradient in the *Serra do Mar* by Alves and colleagues (2010) found that large individuals (≥ 50 cm DBH) were more abundant on steeper slopes, and average AGB was higher in steep plots – consistent with the results of the present study. As another example, Mascaro and colleagues (2011b) also observed a strong positive effect of slope on biomass density on Barro Colorado Island, in Panama. It has been suggested that higher biomass accumulation on steeper slopes could be driven by higher productivity resulting from the interaction between complex topography and light availability in these areas. Still another study, by Castilho and collaborators (2006), found an effect of topography not on total biomass stocks but on tree size distributions, with more small individuals relative to very large individuals on the steeper slopes.

Components of aboveground biomass dynamics estimated from field inventory data over a 3.5-year period proved to be strong predictors of canopy gap fraction in the 1-ha plots. The LiDAR data used in this study for the characterization of canopy structure captured one point in time and therefore served only for describing the *static state* of canopy gaps at the time of the LiDAR overflight. This one-time snapshot of the spatial distribution of static gaps is expected to reflect patterns of tree mortality and related changes in aboveground biomass from the time period preceding the LiDAR survey. This was indeed observed in the data, where the yearly rates of biomass loss and net change were very good indicators of mid-canopy gap fraction ($R^2 = 0.91$ and 0.87, respectively), suggesting a cause and effect relation between mortality patterns and mid-canopy gaps. In comparison, the relationship between static canopy gaps and biomass gain by growth and recruitment was not so

apparent in the data (R^2 = 0.1, though it rose to 0.76 when Plot N was removed from the regression – this one plot might have suffered some disturbance between the field inventory and the LiDAR survey resulting in new gaps not accounted for in the biomass survey). Repeat LiDAR overflights of the study area would provide measurements of the observed canopy openings over time and contribute to a better understanding of the gap formation and closure processes that are closely related to regeneration and tree growth. Information about the spatio-temporal distribution of canopy gaps might also help us understand how environmental factors (e.g. topography, soil, climate) influence forest turnover and biomass dynamics.

An interesting observation of this analysis was the combination of lower rates of mortality and fewer canopy gaps on steeper slopes in comparison with gentler areas. Steeper slopes are generally thought to be characterized by unstable conditions (due to wind exposure, landslides) and increased chances of tree falls and canopy damage. Under such conditions, higher mortality and turnover rates are expected resulting in more canopy openings, which in turn allow for more light penetration into the deeper canopy thus elevating regeneration rates and productivity. The steeper plots examined in this study did not provide evidence for such an unstable dynamic environment. Alternately, the more stable conditions observed on the slopes at the *Serra do Mar* (i.e. low mortality, few gaps) suggest that steeper slopes might provide a more favorable environment than gently sloped areas in terms of nutrient supply and water availability as well as greater access to light resources and space to grow. A simple analysis of biomass components in function of the terrain aspect did not reveal any apparent pattern that would relate the orientation of the 1-ha plots (i.e. solar illumination) to the amount of biomass and/or biomass change.

To examine how fine-scale spatial patterns of topography (slope and aspect) and vegetation structure interactively mediate tree growth by controlling light availability, a radiative

transfer model (e.g. DART), parameterized with LiDAR and field data, could be used to simulate variations of the incident solar illumination as well as the diffuse light environment in the forest canopy. The DART (Discrete Anisotropic Radiative Transfer) model simulates radiative transfer in the Earth-Atmosphere system for any wavelength in the optical domain with a combination of the ray tracing and the discrete ordinate methods (see GASTELLU-ETCHEGORRY et al., 2012). It simulates any landscape, natural or urban, as a 3D matrix of cells that contain turbid material for simulating vegetation and the atmosphere and triangles for simulating translucent and opaque surfaces that make up topography, urban elements and 3D vegetation. DART can use structural and spectral data bases (e.g. atmosphere, vegetation, soil) as input, and direct products of the model include: (i) remote sensing images for any direction and altitude within the atmosphere; (ii) Bidirectional Reflectance Distribution Function; (iii) 3D radiative budget of absorbed, scattered, intercepted and emitted radiation; (iv) 3D scene description (3D LAI, 3D triangle description); and (v) LiDAR ground footprint and backscattered image.

6.5 Conclusions

Based on the main findings of the present analysis, the following conclusions can be drawn:

[1] Plot-based canopy height and canopy gap distributions derived from LiDAR data were variable within submontane and montane forest classes and the small number of these sample plots might not provide a faithful representation of the landscape-scale characteristics of canopy structure.

[2] Canopy height alone was a good predictor of aboveground biomass in the 1-ha permanent plots at the *Serra do Mar* study site. The predictive relationship might be

improved with multivariate regression techniques making use of a variety of LiDAR metrics not solely canopy height.

[3] Processes of aboveground biomass dynamics, principally mortality and net change, were reflected in the LiDAR-derived distribution of static canopy gaps. Multi-temporal LiDAR would allow for the examination of gap dynamics and canopy height changes, possibly more closely related to biomass gain by growth and recruitment.

[4] Terrain slope showed close correspondence with total aboveground biomass, basal area, mortality, gap fraction and net biomass change; while a similar comparison of biomass components with terrain aspect did not reveal any significant relationships at the 1-ha plot scale.

7. SUMMARY AND FINAL REMARKS

Dense tropical forest canopies present various challenges for LiDAR remote sensing, especially in areas of steep topography where much of the remaining Atlantic Forest is concentrated. The main objective of this work was to assess the potential of airborne LiDAR to accurately estimate tropical forest structure and the underlying ground topography in steep mountainous terrain. The study found that the small-footprint LiDAR sensor can be used to characterize the sub-canopy terrain elevation with very high accuracy in the topographically complex *Serra do Mar* region. The accuracy of the LiDAR-derived ground elevations was greatly influenced by the ranging distance and the complexity of the terrain features, and also by the sampling point density of the LiDAR data. The results of the study suggest that in mountainous terrain under closed-canopy tropical forest, dense and uniform LiDAR coverage is essential to faithfully characterize the spatial heterogeneity of forest structure and accurately estimate canopy height. Such a consistent sampling frame allows for improved biomass estimates as well as closer examination of ecological processes related to forest growth, turnover and canopy light availability.

This study integrated fine-scale LiDAR data of canopy structure with field-based measurements of aboveground biomass in order to gain a better understanding of the processes of forest dynamics in the submontane and montane forest areas of the *Serra do Mar* and to examine the influence of terrain characteristics on canopy structure and biomass. LiDAR-derived metrics of canopy surface height and ground topography showed close correspondence with aboveground biomass components in the 1-ha permanent plots. The analyses presented here were essentially based on plot-level samples of three different raster surfaces generated from the LiDAR data: the digital terrain model (DTM), the digital surface model (DSM) and the canopy height model (CHM). Future work will benefit from exploring additional LiDAR metrics of vegetation density and distribution that capture the

internal structural variability of the forest canopy, not just the height of the outer canopy surface. Further research is needed to evaluate whether the plot-scale findings of this study can be extended to the broader landscape (i.e. the entire 1000-ha area of LiDAR coverage) in the *Serra do Mar*. Repeat LiDAR overflights of the study area would also add valuable information about the landscape-scale patterns of gap phase processes and their role in the biomass dynamics of the Atlantic Forest.

Additionally, the fine-scale LiDAR data from this work could be used in combination with other types of remote sensing data for detailed carbon studies or serve as input for ecological models. The fusion of LiDAR structural data with multispectral optical imagery (e.g. Landsat ETM+) or radar (SAR or InSAR), for example, can aid efforts to scale up biomass estimates from the plot level to the broader landscape or regional scales. Also, high-resolution canopy structural information obtained with LiDAR remote sensing can be used for model simulations of the canopy light environment (e.g. DART) to examine how topography affects the light regime by controlling forest structure in a mountainous environment. Continued research based on the present work at the *Serra do Mar* elevational gradient aims to investigate the spatial and temporal patterns of vegetation productivity and turnover rates, and the seasonal and diurnal variation of light availability in the forest canopy by integrating ground-based, airborne and satellite data and spatial models of ecosystem processes and feedbacks.

REFERENCES

AGUILAR, F. J.; MILLS, J. P. Accuracy assessment of LiDAR-derived digital elevation models. **The Photogrammetric Record**, v. 23, n. 122, p. 148-169, 2008.

AIBA, S.; KITAYAMA, K. Structure, composition and species diversity in an altitude-substrate matrix of rain forest tree communities on Mount Kinabalu, Borneo. **Plant Ecology**, v. 140, p. 139-157, 1999.

ALVES, L. F.; VIEIRA, S. A.; SCARANELLO, M. A.; CAMARGO, P. B.; SANTOS, F. A. M.; JOLY, C. A.; MARTINELLI, L. A. Forest structure and live aboveground biomass variation along an elevational gradient of tropical Atlantic moist forest (Brazil). **Forest Ecology and Management**, v. 260, p. 679-691, 2010.

ANDERSEN, H.-E.; REUTEBUCH, S. E.; McGAUGHEY, R. J. A rigorous assessment of tree height measurements obtained using airborne LIDAR and conventional field methods. **Canadian Journal of Remote Sensing**, v. 32, n. 5, p. 355-366, 2006.

ASHTON, P. S.; HALL, P. Comparisons of structure among mixed dipterocarp forests of north-west Borneo. **Journal of Ecology**, v. 80, p. 459-481, 1992.

ASNER, G. P.; HUGHES, R. F.; VARGA, T. A.; KNAPP, D. E.; KENNEDY-BOWDOIN, T. Environmental and biotic controls over aboveground biomass throughout a tropical rain forest. **Ecosystems**, v. 12, p. 261-278, 2009.

ASNER, G. P.; POWELL, G. V.; MASCARO, J.; KNAPP, D. E.; CLARK, J. K.; JACOBSON, J. High-resolution forest carbon stocks and emissions in the Amazon. **Proceedings of the National Academy of Sciences**, v. 107, p. 16738-16742, 2010.

ASNER, G. P.; MASCARO, J.; MULLER-LANDAU, H. C.; VIEILLEDENT, G.; VAUDRY, R.; RASAMOELINA, M.; HALL, J. S.; VAN BREUGEL, M. A universal airborne LiDAR approach for tropical forest carbon mapping. **Oecologia**, v. 168, n. 4, p. 1147-60, 2011.

ASNER, G. P.; KELLNER, J. R.; KENNEDY-BOWDOIN, T.; KNAPP, D. E.; ANDERSON, C.; MARTIN, R. E. Appendix S1 of Forest canopy gap distributions in the Southern Peruvian Amazon. **PLoS ONE**, v. 8, n. 4, e60875, 2013.

ASNER, G. P.; MASCARO, J. Mapping tropical forest carbon: Calibrating plot estimates to a simple LiDAR metric. **Remote Sensing of Environment**, v. 140, p. 614-624, 2014.

AXELSSON, P. DEM generation from laser scanner data using adaptive TIN models. **International Archives of Photogrammetry and Remote Sensing**, v. 35, part B4/1, p. 236-241, 2000.

BCAL LiDAR Tools ver. 2.x.x-dev9. Idaho State University, Department of Geosciences, Boise Center Aerospace Laboratory (BCAL), Boise, Idaho. Available at: http:// bcal.geology.isu.edu/envitools.shtml. Accessed on Feb. 1, 2014.

BROKAW, N. V. The definition of treefall gap and its effect on measures of forest dynamics. **Biotropica**, v. 14, p. 158-160, 1982.

BROKAW, N. V. L.; LENT, R. A. Vertical structure. In: HUNTER, M. L. (Ed.). **Maintaining biodiversity in forest ecosystems**. Cambridge, UK: Cambridge University Press, 1999. p. 373-399.

CASTILHO, C. V.; MAGNUSSON, W. E.; ARAÚJO, R. N. O.; LUIZÃO, R. C. C.; LUIZÃO, F. J.; LIMA, A. P.; HIGUCHI, N. Variation in aboveground tree live biomass in a central Amazonian Forest: Effects of soil and topography. **Forest Ecology and Management**, v. 234, n. 1-3, p. 85-96, 2006.

CHAVE, J.; CONDIT, R.; LAO, S.; CASPERSEN, J.; FOSTER, R. B.; HUBBELL, S. P. Spatial and temporal variation of biomass in a tropical forest: results from a large census plot in Panama. **Journal of Ecology**, v. 91, p. 240-252, 2003.

CHAVE, J.; ANDALO, C.; BROWN, S.; CAIRNS, M. A.; CHAMBERS, J. Q.; EAMUS, D.; FOLSTER, H.; FROMARD, F.; HIGUCHI, N.; KIRA, T.; LESCURE, J.-P.; NELSON, B. W.; OGAWA, H.; PUIG, H.; RIERA, B.; YAMAKURA, T. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. **Oecologia**, v. 145, n. 1, p. 87-99, 2005.

CLARK, D. A.; BROWN, S.; KICKLIGHTER, D. W.; CHAMBERS, J. Q.; THOMLINSON, J. R.; NI, J.; HOLLAND E. A. Net primary production in tropical forests: an evaluation and synthesis of existing field data. **Ecological Applications**, v. 11, n. 2, p. 371-384, 2001.

CLARK, D. B.; CLARK, D. A.; READ, J. M. Edaphic variation and the mesoscale distribution of tree species in a neotropical rain forest. **Journal of Ecology**, v. 86, p. 101-112, 1998.

CLARK, D. B.; PALMER M. W.; CLARK, D. A. Edaphic factors and the landscape-scale distributions of tropical rain forest trees. **Journal of Ecology,** v. 80, p. 2662-2675, 1999.

CLARK, M. L.; CLARK, D. B.; ROBERTS, D. A. Small-footprint lidar estimation of sub-canopy elevation and tree height in a tropical rain forest landscape. **Remote Sensing of Environment**, v. 91, p. 68-89, 2004.

CLARK, M. L.; ROBERTS, D. A.; EWEL J. J.; CLARK, D. B. Estimation of tropical rain forest aboveground biomass with small-footprint lidar and hyperspectral sensors. **Remote Sensing of Environment**, v. 115, p. 2931-2942, 2011.

COBBY, D. M.; MASON, D. C.; DAVENPORT, I. J. Image processing of airborne scanning laser altimetry data for improved river flood modelling. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 56, n. 2, p.121-138, 2001.

CONDIT, R. Local variation in canopy disturbance and soil structure. In: LOSOS, E. C.; LEIGH, E. G. (Eds.). **Tropical forest diversity and dynamism: findings from a large-scale plot network**. Chicago, USA: The University of Chicago Press, 2004. p. 175-176.

COOK, B. D.; CORP, L. A.; NELSON, R. F.; MIDDLETON, E. M.; MORTON, D. C.; McCORKEL, J. T.; MASEK, J. G.; RANSON, K. J.; LY, V.; MONTESANO, P. M. NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager. **Remote Sensing**, v. 5, p. 4045-4066, 2013.

DENSLOW, J. S. Gap partitioning among tropical rainforest trees. **Biotropica** (Suppl.), v. 12, p. 47-95, 1980.

d'OLIVEIRA, M. V. N.; REUTEBUCH, S. E.; McGAUGHEY, R. J.; ANDERSEN, H-E. Estimating forest biomass and identifying low-intensity logging areas using airborne scanning lidar in Antimary State Forest, Acre State, Western Brazilian Amazon. **Remote Sensing of Environment**, v. 124, p. 479-491, 2012.

DORNER, B.; LERTZMAN, K.; FALL, J. Landscape pattern in topographically complex landscapes: issues and techniques for analysis. **Landscape Ecology**, v. 17, n. 8, p. 729-743, 2002.

DRAKE, J. B.; KNOX, R. G.; DUBAYAH, R. O.; CLARK, D. B.; CONDIT, R.; BLAIR, J. B.; HOFTON, M. Above-ground biomass estimation in closed canopy Neotropical forests using lidar remote sensing: factors affecting the generality of relationships. **Global Ecology and Biogeography**, v. 12, p. 147-159, 2003.

EVANS, J. S.; HUDAK, A. T.; FAUX, R.; SMITH, A. M. S. Discrete Return Lidar in Natural Resources: Recommendations for Project Planning, Data Processing, and Deliverables. **Remote Sensing**, v. 1, p. 776-794, 2009.

FELDPAUSCH, T. R.; LLOYD, J.; LEWIS, S. L.; BRIENEN, R. J. W.; GLOOR, M.; MENDOZA, A. M.; LOPEZ-GONZALEZ, G.; BANIN, L.; SALIM, K. A.; AFFUM-BAFFOE, K.; ALEXIADES, M.; ALMEIDA, S.; AMARAL, I.; ANDRADE, A.; ARAGÃO, L. E. O. C.; MURAKAMI, A. A.; ARETS, E. J. M. M.; ARROYO, L.; AYMARD, G. A.; BAKER, T. R.; BÁNKI, O. S.; BERRY, N. J.; CARDOZO, N.; CHAVE, J.; COMISKEY, J. A.; ALVAREZ, E.; DE OLIVEIRA, A.; DI FIORE, A.; DJAGBLETEY, G.; DOMINGUES, T. F. et al. Tree height integrated into pantropical forest biomass estimates. **Biogeosciences**, v. 9, p. 3381-3403, 2012.

FISHER, J. I.; HURTT, G. C.; THOMAS, R.; CHAMBERS, J. Q. Clustered disturbances lead to bias in large-scale estimates based on forest sample plots. **Ecology Letters**, v. 11, p. 1-10, 2008.

GASTELLU-ETCHEGORRY, J. P.; GRAU, E.; LAURET, N. DART: A 3D Model for Remote Sensing Images and Radiative Budget of Earth Surfaces, Modeling and Simulation in Engineering. In: ALEXANDRU, C. (Ed.). **InTech**, DOI: 10.5772/31315. Available at: http://www.intechopen.com/books/modeling-and-simulation-in-engineering/dart-model-remote-sensing-images-radiative-budget-and-energy-budget-of-earth-surfaces. Accessed on Mar. 15, 2014.

GATZIOLIS, D.; ANDERSEN, H.-E. **A guide to LIDAR data acquisition and processing for the forests of the Pacific Northwest**. Portland, USA: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, 2008. Gen. Tech. Rep. PNW-GTR-768.

GAULTON, R.; MALTHUS T. J. LiDAR mapping of canopy gaps in continuous cover forests: A comparison of canopy height model and point cloud based techniques. **International Journal of Remote Sensing**, v. 31, p. 1193-1211, 2010.

GONZALEZ, P.; ASNER, G. P.; BATTLES, J. J.; LEFSKY, M. A.; WARING, K. M.; PALACE, M. Forest carbon densities and uncertainties from Lidar, QuickBird, and field measurements in California. **Remote Sensing of Environment**, v. 114, p. 1561-1575, 2010.

HARMS, K.; CONDIT, R.; HUBBELL, S. P.; FOSTER R. B. Habitat association of trees and shrubs in a 50-ha neotropical forest plot. **Journal of Ecology**, v. 89, p. 947-959, 2001.

HERITAGE, G. L.; LARGE, A. R. G. (Eds.). **Laser scanning for the environmental sciences**. Chichester, UK: Wiley-Blackwell Publishing Ltd., 2009. 278p.

HODGSON, M. E.; JENSEN, J. R.; SCHMIDT, L.; SCHILL, S; DAVIS, B. An evaluation of LIDAR- and IFSAR-derived digital elevation models in leaf-on conditions with USGS Level 1 and Level 2 DEMs. **Remote Sensing of Environment**, v. 84, p. 295-308, 2003.

HODGSON, M. E.; BRESNAHAN, P. Accuracy of airborne LiDAR-derived elevation: empirical assessment and error budget. **Photogrammetric Engineering and Remote Sensing**, v. 70, n. 3, p. 331-339, 2004.

HODGSON, M. E.; JENSEN, J.; RABER, G.; TULLIS, J.; DAVIS, B. A.; THOMPSON, G.; SCHUCKMAN, K. Evaluation of lidar-derived elevation and terrain slope in leaf-off conditions. **Photogrammetric Engineering and Remote Sensing**, v. 71, n. 7, p. 817-823, 2005.

HOLMGREN, J.; NILSSON, M.; OLSSON, H. Estimation of tree height and stem volume on plots using airborne laser scanning. **Forest Science**, v. 49, p. 419-428, 2003.

HOMEIER, J. The influence of topography on forest structure and regeneration dynamics in an Ecuadorian montane forest. In: GRADSTEIN, S. R., HOMEIER, J.; GANSERT, D. (Eds.). **The tropical mountain forest** – Patterns and processes in a biodiversity hotspot. Universitätsverlag Göttingen Biodiversity and Ecology Series, v. 2, p. 97-107, 2008.

HUDAK, A. T.; EVANS, J. S.; STUART SMITH, A. M. LiDAR Utility for Natural Resource Managers. **Remote Sensing**, v. 1, p. 934-951, 2009.

HUDAK, A. T.; STRAND, E. K.; VIERLING, L. A.; BYRNE, J. C.; EITEL, J. U. H.; MARTINUZZI, S.; FALKOWSKI, M. J. Quantifying aboveground forest carbon pools and fluxes from repeat LiDAR surveys. **Remote Sensing of Environment**, v. 123, p. 25-40, 2012.

HYYPPÄ, H.; YU, X.; HYYPPÄ, J.; KAARTINEN, H.; KAASALAINEN, S.; HONKAVAARA, E.; RÖNNHOLM, P. Factors affecting the quality of DTM generation in forested areas. In: ISPRS WG III/3, III/4, V/3 WORKSHOP "LASER SCANNING 2005", 2005, Enschede, Netherlands. **Electronic Proceedings...** Enschede, Netherlands: ISPRS, 2005. Available at: http://www.helsinki.fi/~korpela/MINV12/DTM_accuracy.pdf. Accessed on Nov. 1, 2013.

JAKUBOWSKI, M. K.; GUO, Q.; KELLY M. Tradeoffs between lidar pulse density and forest measurement accuracy. **Remote Sensing of Environment**, v. 130, p. 245-253, 2013.

JOLY, C. A.; ASSIS, M. A.; BERNACCI, L. C.; TAMASHIRO, J. Y.; CAMPOS, M. C. R.; GOMES, J. A. M. A. et al. Floristic and phytosociology in permanent plots of the Atlantic

Rainforest along an altitudinal gradient in southeastern Brazil. **Biota Neotropica**, v. 12, p. 125-145, 2012.

KRAUS, K.; PFEIFER, N. Determination of terrain models in wooded areas with airborne laser scanner data. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 53, p. 193-203, 1998.

KRUGER, E.; VOLIN, J. Reexamining the empirical relation between plant growth and leaf photosynthesis. **Functional Plant Biology**, v. 33, p. 421-429, 2006.

LEDUC, A.; DRAPEAU, P.; BERGERON, Y.; LEGENDRE, P. Study of spatial components of forest cover using partial Mantel tests and path analysis. **Journal of Vegetation Science**, v. 3, p. 69-78, 1992.

LEFSKY, M. A.; COHEN, W. B.; ACKER, S. A.; PARKER, G. G.; SPIES, T. A.; HARDING, D. Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. **Remote Sensing of Environment**, v. 70, p. 339-361, 1999a.

LEFSKY, M. A.; HARDING, D. J.; COHEN, W. B.; PARKER, G. G. Surface lidar remote sensing of basal area and biomass in deciduous forests of eastern Maryland, USA. **Remote Sensing of the Environment**, v. 67, p. 83-98, 1999b.

LEFSKY, M. A.; COHEN, W. B.; PARKER, G. G.; HARDING, D. J. Lidar remote sensing for ecosystem studies. **BioScience**, v. 52, n. 1, p. 19-30, 2002a.

LEFSKY, M. A.; COHEN, W. B.; HARDING, D. J.; PARKER, G. G.; ACKER, S. A.; GOWER, S. T. Lidar remote sensing of aboveground biomass in three biomes. **Global Ecology and Biogeography**, v. 11, n. 5, p. 393-400, 2002b.

LEGENDRE, P.; FORTIN, M.-J. Spatial pattern and ecological analysis. **Vegetation**, v. 80, p. 107-138, 1989.

LEUSCHNER, C. H.; MOSER, G.; BERTSCH, C.; RÖDERSTEIN, M.; HERTEL, D. Large altitudinal increase in tree root/shoot ratio in tropical mountain forests of Ecuador. **Basic and Applied Ecology**, v. 8, p. 219-230, 2007.

LIEBERMAN, D.; LIEBERMAN, M.; PERALTA, R.; HARTSHORN, G. S. Tropical forest structure and composition on a large-scale altitudinal gradient in Costa Rica. **Journal of Ecology**, v. 84, p. 137-152, 1996.

LOSOS, E. C.; LEIGH, E. G. (Eds.). **Tropical forest diversity and dynamism: findings from a large-scale plot network**. Chicago, USA: The University of Chicago Press, 2004. 645p.

LOVELL, J. L.; JUPP, D. L. B.; CULVENOR, D. S.; COOPS, N. C. Using airborne and ground-based ranging lidar to measure canopy structure in Australian forests. **Canadian Journal of Remote Sensing**, v. 29, p. 607-622, 2003.

LIU, X. Airborne LiDAR for DEM generation: some critical issues. **Progress in Physical Geography**, v. 32, n. 1, p. 31-49, 2008.

MABBERLEY, D. J. **Tropical rain forest ecology**. New York, USA: Chapman and Hall, 1992. 300p.

MAIER, B.; TIEDE, D.; DORREN, L. Characterising mountain forest structure using landscape metrics on LiDAR-based canopy surface models. **Object-Based Image Analysis**, Lecture Notes in Geoinformation and Cartography, p. 625-643, 2008.

MALHI, Y.; BAKER, T. R.; PHILLIPS, O. L.; ALMEIDA, S.; ALVAREZ, E.; ARROYO, L.; CHAVE, J.; CZIMCZIK, C. I.; DI FIORE, A.; HIGUCHI, N.; KILLEEN, T. J.; LAURANCE, S. G.; LAURANCE, W. F.; LEWIS, S. L.; MONTOYA, L. M. M.; MONTEAGUDO, A.; NEILL, D. A.; VARGAS, P. N.; PATINO, S.; PITMAN, N. C.; QUESADA, C. A.; SALOMAO, R.; SILVA, J. N. M.; LEZAMA, A. T.; MARTÍNEZ, R. V.; TERBORGH, J.; VINCETI, B.; LLOYD, J. The above-ground coarse wood productivity of 104 Neotropical forest plots. **Global Change Biology**, v. 10, n. 5, p. 563-591, 2004.

MANOKARAN, N.; LAFRANKIE, J. V. Stand structure of Pasoh Forest Reserve, a lowland rain forest in peninsular Malaysia. **Journal of Tropical Forest Science**, v. 3, p. 14-24, 1990.

MASCARO, J.; DETTO, M.; ASNER, G. P.; MULLER-LANDAU, H. Evaluating uncertainty in mapping forest carbon with airborne LiDAR. **Remote Sensing of Environment**, v. 115, p. 3770-3774, 2011a.

MASCARO, J.; ASNER, G. P.; MULLER-LANDAU, H. C.; van BREUGEL, M.; HALL, J.; DAHLIN, K. Controls over aboveground forest carbon density on Barro Colorado Island, Panama. **Biogeosciences**, v. 8, p. 1615-1629, 2011b.

MORSDORF, F.; FREY, O.; MEIER, E.; ITTEN, K. I.; ALLGÖWER, B. Assessment of the influence of flying altitude and scan angle on biophysical vegetation products derived from

airborne laser scanning. **International Journal of Remote Sensing**, v. 29, n. 5, p. 1387-1406, 2008.

MYERS, N.; MITTERMEIER, R. A.; MITTERMEIER, C. G.; FONSECA, G. A. B.; KENT, J. Biodiversity hotspots for conservation priorities. **Nature**, v. 403, p. 852-858, 2000.

NAESSET, E. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. **Remote Sensing of Environment**, v. 80, p. 88-99, 2002.

NAESSET, E.; GOBAKKEN, T. Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. **Remote Sensing of Environment**, v. 112, p. 3079-3090, 2008.

NILSSON, M. Estimation of tree heights and stand volume using an airborne lidar system. **Remote Sensing of Environment**, v. 56, p. 1-7, 1996.

NI-MEISTER, W.; LEE, S.; STRAHLER, A. H.; WOODCOCK, C. E.; SCHAAF, C.; YAO, T.; RANSON, K. J.; SUN, G.; BLAIR, J. B. Assessing general relationships between aboveground biomass and vegetation structure parameters for improved carbon estimate from lidar remote sensing. **Journal of Geophysical Research**, v. 115, G00E11, 2010.

PALMIOTTO, P. A. The role of specialization in nutrient-use efficiency as a mechanism driving species diversity in a tropical rain forest. D. F. dissertation. **Yale University**, New Haven, CT, 1998.

PARKER, G. G.; HARDING, D. J.; BERGER, M. L. A portable LIDAR system for rapid determination of forest canopy structure. **Journal of Applied Ecology**, v. 41, p. 755-767, 2004.

PEARCY, R. W. Responses of plants to heterogenous light environments. In: PUGNAIRE, F. I.; VALLADARES, F. (Eds.). **Functional plant ecology**. 2.ed. Lincoln, USA: CRC Press, 2007. p. 213-257.

PHILLIPS, O. L.; NUNEZ VARGAS, P.; LORENZO MONTEAGUDO, A.; PENA CRUZ, A.; CHUSPE ZANS, M-E.; GALIANO SANCHEZ, W.; YLI-HALLA, M.; ROSE, S. Habitat association among Amazonian tree species: a landscape-scale approach. **Journal of Ecology**, v. 91, p. 757-775, 2003.

POULSEN, A. D. Species richness and density of ground herbs within a plot of lowland rainforest in north-west Borneo. **Journal of Tropical Ecology**, v. 12, p. 177-90, 1996.

R CORE TEAM. R: A language and environment for statistical computing. **Foundation for Statistical Computing.** Vienna, Austria, 2011. http://www.R-project.org/

RENSLOW, M. S. (Ed.). Manual of Airborne Topographic Lidar. **American Society for Photogrammetry Remote Sensing**, Bethesda, MD, 528p., 2012.

REUTEBUCH, S. E.; McGAUGHEY, R. J.; ANDERSON, H. E.; CARSON, W. W. Accuracy of a high-resolution lidar terrain model under a conifer forest canopy. **Canadian Journal of Remote Sensing**, v. 29, n. 5, p. 527-535, 2003.

RIBEIRO, M. C.; METZGER, J. P.; MARTENSEN, A. C.; PONZONI, F. J.; HIROTA, M. M. The Brazilian Atlantic Forest: How much is left, and how is the remaining forest distributed? Implications for conservation. **Biological Conservation**, v. 142, p. 1141-1153, 2009.

SÄYNÄJOKI, R.; MALTAMO, M.; KORHONEN, K. T. Forest inventory with sparse resolution Airborne Laser Scanning data – a literature review. **Working Papers of the Finnish Forest Research Institute**, 103, 2008. Available at: http://www.metla.fi/julkaisut/ workingpapers/2008/mwp103.htm. Accessed on Oct. 6, 2013.

SCARANELLO, M. A. S.; ALVES, L. F.; VIEIRA, S. A.; CAMARGO, P. B.; JOLY, C. A.; MARTINELLI, L. A. Height-diameter relationships of tropical Atlantic moist forest trees in southeastern Brazil. **Scientia Agricola**, v. 69, p. 26-37, 2012.

SHRESTHA, R.; GLENN, N. F.; SPAETE, L.; TEDROW, L.; GESSLER, P; HUDAK, A. Data flow for LiDAR acquisition and processing. **Poster Presentation**, 2010. Available at: http://www.idaholidar.org/f/reports/Poster-EPSCoR-ID-2010-LiDARWorkflow.pdf. Accessed on Aug. 6, 2013.

SMA - Secretaria do Meio Ambiente. Planos de Manejo das Unidades de Conservação: Parque Estadual da Serra do Mar - Núcleo Picinguaba. **Plano de Gestao Ambiental - Fase I**, São Paulo, 1998.

SPAETE, L. P.; GLENN, N. F.; DERRYBERRY, D. R.; SANKEY, T. T.; MITCHELL, J. J.; HARDEGREE, S. P. Vegetation and slope effects on accuracy of a LiDAR-derived DEM in the sagebrush steppe. **Remote Sensing Letters**, v. 2, n. 4, p. 317-326, 2010.

STAL, C.; NUTTENS, T.; BOURGEOIS, J.; CARLIER, L.; De MAEYER, P.; De WULF, A. Accuracy assessment of a LiDAR digital terrain model by using RTK GPS and total station. **EARSeL eProceedings**, v. 10, n. 8, p. 1-8, 2011.

STARK, S. C. **On the mechanistic connection of forest canopy structure with productivity and demography in the Amazon**. 2012. PhD (Doctoral in Ecology & Evolutionary Biology) - The University of Arizona, Tucson, USA, 2012. Available at: http://hdl.handle.net/10150/265347. Accessed on May 6, 2013.

STERCK, F.; BONGERS, F. Crown development in tropical rain forest trees: patterns with tree height and light availability. **Journal of Ecology**, v. 89, p. 1-13, 2001.

SU, J.; BORK, E. Influence of vegetation, slope, and LiDAR sampling angle on DEM accuracy. **Photogrammetric Engineering and Remote Sensing**, v. 72, n. 11, p. 1265-1274, 2006.

TINKHAM, W. T.; SMITH, A. M. S.; HOFFMAN, C.; HUDAK, A. T.; FALKOWSKI, M. J.; SWANSON, M. E.; GESSLER, P. E. Investigating the influence of LiDAR ground surface errors on the utility of derived forest inventories. **Canadian Journal of Forest Research**, v. 42, p. 413-422, 2012.

UNGER, M.; HOMEIER, J.; LEUSCHNER, C. Relationships among leaf area index, below-canopy light availability and tree diversity along a transect from tropical lowland to montane forests in NE Ecuador. **Tropical Ecology**, v. 54, p. 33-45, 2013.

VÁZQUEZ, G. J. A.; GIVNISH, T. J. Altitudinal gradients in tropical forest composition, structure and diversity in the Sierra de Manantlán. **Journal of Ecology**, v. 86, p. 999-1020, 1998.

VIEIRA, S. A.; ALVES, L. F.; AIDAR, M. P. M.; ARAÚJO, L. S.; BAKER, T.; BATISTA, J. L. F.; CAMPOS, M. C. R.; CAMARGO, P. B.; CHAVE, J.; DELITTI, W. B.; HIGUCHI, N.; HONÓRIO, E.; JOLY, C. A.; KELLER, M.; MARTINELLI, L. A.; DE MATTOS, E. A.; METZKER, T.; PHILLIPS, O.; SANTOS, F. A. M.; SHIMABUKURO, M. T.; SILVEIRA, M.; TRUMBORE, S. E. Estimation of biomass and carbon stocks: the case of the Atlantic Forest. **Biota Neotropica**, v. 8, p. 21-29, 2008.

WHITMORE, T. C. **Tropical rain forest of the Far East**. Oxford, UK: Clarendon Press, 1984. 352p.

WULDER, M. Optical remote-sensing techniques for the assessment of forest inventory and biophysical parameters. **Progress in Physical Geography**, v. 22, n. 4, p. 449-476, 1998.

ZHANG, K.; CHEN, S.; WHITMAN, D.; SHYU, M.; YAN, J.; ZHENG, C. A progressive morphological filter for removing nonground measurements from airborne LiDAR data. **IEEE Transactions on Geoscience and Remote Sensing**, n. 41, p. 872-882, 2003.

ZIMBLE, D. A.; EVANS, D. L.; CARLSON, G. C.; PARKER, R. C.; GRADO, S. C.; GERARD, P. D. Characterizing vertical forest structure using small-footprint airborne LiDAR. **Remote Sensing of Environment**, v. 87, p. 171-182, 2003.

ZOLKOS, S. G.; GOETZ, S. J.; DUBAYAH, R. A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. **Remote Sensing of Environment**, v. 128, p. 289-298, 2013.

APPENDIX A. List of the 36 control points used for DTM validation (coordinates in UTM Z23S) measured with dGPS and their respective precision (RMSE).

	UTM coordinate [m]			Coordinate precision [m]		
Point	Easting	Northing	Elevation	Easting	Northing	Elevation
BASE	485083.39	7419288.63	898.19	na	na	na
F_A0	491797.33	7414215.84	105.23	0.19	0.05	0.11
F_A10	491823.42	7414312.49	114.99	0.59	0.18	0.77
F_K0	491895.88	7414190.08	100.43	0.45	0.31	0.35
F_K10	491918.21	7414286.73	125.91	0.02	0.02	0.03
G_A0	491679.62	7414925.73	177.87	0.01	0.01	0.02
G_A10	491596.94	7414873.50	182.00	1.08	0.28	0.55
G_K0	491626.06	7415010.19	195.93	0.72	0.25	0.26
G_K10	491542.10	7414956.27	195.01	0.40	0.23	0.40
H_A0	491631.34	7415056.81	203.66	0.18	0.09	0.25
H_A10	491547.00	7415003.21	202.85	0.31	0.19	0.41
H_K0	491578.18	7415141.35	214.36	0.02	0.01	0.03
H_K10	491495.26	7415087.52	216.98	0.16	0.06	0.17
I_A0	491351.64	7415637.23	307.67	0.01	0.01	0.02
I_A10	491292.19	7415717.79	324.14	0.26	0.10	0.25
I_K0	491432.43	7415696.24	303.10	0.35	0.15	0.36
I_K10	491372.97	7415776.97	337.09	0.01	0.01	0.02
J_A0	491341.80	7415788.44	354.23	1.53	0.48	1.86
J_A10	491289.83	7415871.69	347.37	0.03	0.03	0.05
J_K0*	491428.64	7415842.64	na	na	na	na
J_K10	491373.56	7415926.61	361.47	0.02	0.01	0.03
K_A0	492936.56	7420396.71	1039.60	0.34	0.19	0.62
K_A10	493012.72	7420459.56	1019.80	0.03	0.02	0.21
K_K0	492999.92	7420319.43	1039.80	1.11	0.48	2.21
K_K10	493077.62	7420382.08	1020.66	0.80	0.46	0.79
L_A0	492789.85	7420382.51	995.84	0.49	0.15	0.42
L_A10	492851.76	7420458.33	1019.96	0.61	0.28	0.47
L_K0	492866.54	7420318.45	1044.60	0.52	0.11	0.39
L_K10	492930.47	7420395.92	1039.56	0.35	0.18	0.53
M_A0	492486.56	7420233.11	1009.22	0.37	0.21	0.48
M_A10	492581.36	7420199.80	997.84	0.35	0.26	0.71
M_K0	492454.77	7420138.28	1038.99	2.13	0.84	1.54
M_K10	492549.30	7420105.76	1019.97	0.32	0.25	0.37
N_A0	492198.98	7420371.08	1013.53	0.30	0.05	0.19
N_A10	492281.08	7420313.21	1003.43	0.01	0.01	0.03
N_K0	492141.62	7420289.42	1036.79	0.41	0.08	0.19
N_K10	492223.29	7420231.12	1017.21	2.11	1.88	1.21

*Coordinates of the missing point "J_K0" were calculated by triangulation using the three other measured corners of the rectangular plot J.

APPENDIX B. Schematic map of the nine field plots – montane (K-N) and submontane (F-J) – with the corner point markers (A0, A10, K0, K10) and corresponding subplot codes (1, 10, 91, 100).