The creation of a forestry relational database and assessment of aboveground biomass variability across Costa Rica using forest management data

Sienna Avril Pearl Svob

Department of Geography

McGill University

Montreal, Quebec, Canada

May 2014

A thesis submitted to McGill University

in partial fulfillment of the requirement of the degree of

Master of Science

Abstract

Tropical forests are a globally important carbon store threatened by rapid rates of deforestation and forest degradation. When estimating these vital carbon stocks, the high spatio-temporal variability of aboveground biomass (AGB) in tropical forests is a large source of uncertainty. Pre-felling inventories (i.e. forest management data) may help to reduce this uncertainty due to their spatial distribution and sampling intensity. Throughout the past decade, over 1000 Costa Rican natural forest management plans, spanning 30 years in age and covering the country's Atlantic coast, northern lowlands, and southern Pacific coast, have been collected and digitized. In this study, I first aim to create a powerful database to ensure the standardization, quality, and analyzability of this unique collection of tree inventory data. Second, I use this database to evaluate the variability of estimated AGB (EAGB) across five regions of Costa Rica.

The final geodatabase efficiently stores the large and diverse forest management dataset through the integration of logical relations and quality checks, permitting flexible data access across 32 tables and over 250,000 tree records. Through the use of quality tests, the database improves overall data integrity. An exploratory analysis of the database's standardized taxonomic data reveals the weaknesses, strengths, and potential ecological applications of forest management data.

Using this database, the EAGB analysis found that, of the regions studied, Costa Rica's Osa peninsula had a significantly higher EAGB (DBH≥30cm) (173.47 (mean) ± 60.23 (SD) Mg ha^{-1}). Further, the analysis indicated that the density of large trees (DBH≥70cm) explained approximately 50% of the variability of EAGB across the five ecosystems studied. Comparing the EAGB of this study to published estimates illustrates that a forest management based approach produces a similar range of values.

Overall, the geodatabase represents the most complete record of forest management practices in Costa Rica to date. This tool will permit a better understanding of logging practices, species-environment relationships, and carbon

stocks in selectively logged forests at a scale necessary to address the conservation and resource management challenges of today. Further, this study describes, the most spatially rich analysis of ground level AGB data in Costa Rica as of yet. Using data from pre-felling inventories, this research shows that EAGB within and among the five ecosystems studied is highly variable. As it is the non-protected areas of the tropics that offer the greatest opportunity to reduce rates of deforestation and forest degradation, logging inventories offer a promising source of data to support mechanisms such as the REDD+ (Reducing Emissions from Tropical Deforestation and Degradation) program.

Résumé

Les forêts tropicales forment un réservoir de carbone d'importance mondiale. Ce réservoir est menacée par un taux rapide de déforestation et dégradation. Lors de l'estimation de ces stocks de carbone vitaux, la variabilité spatio-temporelle de la biomasse hors-sol (AGB) dans les forêts tropicales est une grande source d'incertitude. Les inventaires pré-abattage (c.à.d. données de gestion de la forêt) peuvent aider à réduire cette incertitude en raison de leur distribution spatiale et de l'intensité d'échantillonnage. Durant la dernière décennie, plus de 1000 plans de gestion des forêts naturelles du Costa Rica, couvrant la côte de l'Atlantique, les plaines du nord, et la côte sud du Pacifique du pays ont été collectionnés et numérisés. Dans cette étude, je cherche d'abord à créer une base de données efficace qui permet d'assurer la normalisation, la qualité, et analysabilité de cette collection unique de données des arbres. Deuxièmement, j'utilise cette base de données pour évaluer la variabilité de AGB estimée (EAGB) dans cinq régions du Costa Rica.

La base de données finale entrepose efficacement le vaste ensemble de données de gestion de la forêt par l'intégration des relations logiques et des contrôles de qualité. Cette dernière permet l’accès de données multiples (tel que la grandeur d’arbre, diamètre de troncs) parmi 32 tableaux et plus de 250,000 dossiers d'arbres. Grâce à l'utilisation de tests de qualité, la base de données améliore l'intégrité globale des données. Une analyse exploratoire de données taxonomique normalisée de la base de données révèle les faiblesses, les forces, et les applications écologiques potentielles de données de gestion de la forêt.

En utilisant cette base de données, l'analyse EAGB constate que, dans les régions étudiées, la péninsule d'Osa a un EAGB significativement plus élevé (DHP ≥ 30cm) (173,47 (moyenne) ± 60,23 (SD) Mg ha^{-1}). En outre, l'analyse démontre que la densité de grands arbres (DHP ≥70cm) explique environ 50% de la variabilité de EAGB à travers des cinq écosystèmes étudiés . En comparant la EAGB de cette étude aux estimations publiées, démontre qu'une approche fondée sur la gestion des forêts produit une gamme de valeurs similaires.

Dans l'ensemble, cette base de données représente le dossier le plus complet des pratiques de gestion forestière Costaricaines à ce jour. Cet outil permet une meilleure compréhension des pratiques d'exploitation forestière, les relations espèces-environnement, et des réserve de carbone dans les forêts d'abattage sélectif à l'échelle nécessaire pour relever les défis d'aujourd'hui de gestion et conservation des ressources naturelles. En outre, cette est actuellement l'analyse la plus spatialement riche des données AGB au Costa Rica. En utilisant des données de stocks pré-abattage, cette recherche montre que EAGB dans et entre les cinq écosystèmes étudiés est très variable. Comme ce sont les zones non protégées des tropiques qui offrent la plus grande opportunité de réduire le taux de déforestation et de dégradation des forêts, les inventaires pré-abattage offrent une source prometteuse de données à l'appui des mécanismes tels que le programme REDD+ (Réduction des émissions de la déforestation tropicale et de la dégradation).

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AGB	Aboveground Biomass
AGC	Aboveground Carbon
ACAHN	Arenal Huetar Norte Conservation Area
ACAHN	Central Volcanic Conservation Area
ACLA-C	Carribbean La Amistad Conservation Area
ACOSA	Osa Conservation Area
AP	Parent Tree
ACTO	Tortuguero Conservation Area
C	Cut
CA	Conservation Area
CH	Commercial Height
COP	Conference of Parties
CWD	Coarse Woody Debris
DBH	Diameter at Breast Height
DBMS	Database Management System
EAGB	Estimated Aboveground Biomass
ER	Entity Relationship
FGIS	Forestry Geographic Information Systems
GIS	Geographic Information Systems

GHG	Green House Gas
GPP	Gross Primary Productivity
IPCC	Intergovernmental Panel on Climate Change
LiDAR	Light Detection and Ranging
N/ID	Not Identified
NEE	Net Ecosystem Exchange
NFMP	Natural Forest Management Plan
NPP	Net Primary Productivity
OT	Other
R	Remnant
RaDAR	Radar Detection and Ranging
REDD(+)	Reducing Emissions from Deforestation and Forest Degradation
RSE	Residual Standard Error
SD	Standard Deviation
SQL	Standard Query Language
UN	United Nations

Acknowledgements

The completion of this thesis would not have been possible without the encouragement and support of my supervisors, colleagues, friends, and family. First and foremost, I would like to express my deepest gratitude to my co-supervisors, Dr. Margaret Kalacska and Dr. Pablo Arroyo-Mora, for allowing me to pursue an exciting research topic of great interest to myself. I am thankful to Pablo, whose enthusiasm for the pursuit of knowledge about tropical forests and forest ecology was ceaselessly motivational. I am grateful for his guidance, passion, and endless advice during everything from the development stages to the completion of my research. I am also indebted to Margaret for her support and knowledgeable commentary over the past two years. She was a great role model for a young female scientist due to her exhaustive dedication to her field and her capacity to stay up to date with all things technology (I have yet to fully understand how she finds enough hours in the day). I would like to thank both of supervisors jointly for providing me with an immense number of applicable tools and valuable experiences that will allow me to pursue future opportunities in the vast number of fields related to sustainable management, ecology, and conservation science.

I am forever appreciative to SIREFOR (Maria Isabel Chavarria), SINAC, and the directors and employees of the sub-regional offices of ACAHN, ACLA-C, ACCVC, ACOSA, and ACTO, for their support and permission to access the original hard copies of the NFMPs. I would like to thank Prof. Gail Chmura for serving on my supervisory committee and providing suggestions on my research proposal. I am thankful to all of the students who actively participated in the data collection, entry, and cleaning of the large forest management dataset used in this study. I would also like to mention and express my appreciation to Maria Rodriguez Solis, who not only helped me digitize and collect forest management data in Costa Rica during a yearlong period, but who embraced me into the Costa Rican culture and patiently improved my once broken Spanish.

This research and earlier in field data collection was supported by funding from the geography department's Global Environment and Climate Change Center at McGill university to Dr. J. Pablo Arroyo-Mora, and the Natural Sciences and Engineering Research Council of Canada and the Canada Foundation for Innovation to Dr. Margaret Kalacska, and Mission Airborne Carbon 2013 (supported by the Canadian Space Agency) to both Dr. Margaret Kalacska and Dr. J. Pablo Arroyo-Mora. Further, I would like to acknowledge the ESRI Canada GIS Scholarship Award and the Rathlyn GIS Award for their generous support to my research.

I would like to recognize the help of Chris Firger, who, despite having a background in everything but natural sciences, patiently sat with me and edited a large portion of my thesis, improving both my written communication skills and the quality of this work. I would like to give a special thanks to my father. He taught me the splendor and magnificence of forest landscapes at the ripe age of seven as he strapped a 15 lbs backpack packed with the essentials of life onto my 55 lbs frame - preparing me for a week long hike with himself and my eight year old brother in the Canadian Rocky Mountains. Although the thought of a father bringing his two young children backpacking in the wild may strike a note of fear in some, it was this trip that first sprouted the motivation behind my research. I would also like to thank my brother, who hid four large rocks in my backpack and forgot to mention so until day three of the hike. He taught me to keep on going, regardless of the obstacles dragging you down - *literally*. Finally, I would like to thank my mother, who provided me with the tools and wool socks to pursue my interests and was always willing to give me hers in cases of emergency or missteps into deep muddy puddles.

Contribution of Authors

This thesis is composed of two chapters (chapter 3 and 4) that are written in a manuscript style in order to be submitted to peer-reviewed scientific journals. In both articles, I developed the research objectives, carried out the analysis, interpreted the results, and wrote the manuscripts as first author. My M.Sc. supervisors, Dr. J. Pablo Arroyo-Mora and Dr. Margaret Kalacska, aided in the development and clarification of my research objectives, helped in the organization of data collection, field work, and data digitization. Further, my supervisors have read and critiqued these manuscripts, providing comments that have improved the quality of the research presented. Pablo Arroyo-Mora also provided the set of data collected from two areas of Costa Rica (ACCVC and ACOSA) previous to the start of my thesis. The contribution of the co-authors in both manuscript style chapters are as follows:

Manuscript I (Chapter 3): "The development of a forestry geodatabase for natural forest management plans in Costa Rica" by Sienna A.P. Svob, Pablo Arroyo-Mora, and Margaret Kalacksa (Svob et al., 2014a, *in press at* Forest Ecology and Management). PAM and MK have read, edited, and supplied comments, reviewing this manuscript multiple times before it was submitted for publication. This research idea and partial implementation was developed by PAM.

Manuscript II (Chapter 4): "A wood density and aboveground biomass variability assessment using pre-felling inventory data in Costa Rica" (Svob et al., 2014b, *submitted to* Carbon Balance and Management). To date, PAM and MK have aided in the completion and editing of multiple revisions of this manuscript as it was prepared for submission to a peer-reviewed journal.

1. Introduction

1.1. Research Context

In the face of rising atmospheric CO_2 concentrations and climate change, a great amount of scientific effort is being devoted to quantifying and understanding terrestrial carbon fluxes. Tropical forests play a pivotal role in the global carbon cycle as they store a large portion of the earth's terrestrial carbon and biodiversity stocks (Dirzo and Raven, 2003; Pan et al., 2011). In recent decades, these ecosystems have suffered substantially from deforestation and degradation. It is estimated that 15-35% of the anthropogenic greenhouse gas emissions (GHGs) in the 1990's was due to tropical deforestation alone (Houghton, 2005a).

As these ecosystems decline, a renewed interest in quantifying and understanding spatial and temporal variations in biodiversity and carbon distribution across tropical forests has arisen (Petrokofsky et al., 2011). Further, novel mechanisms, such as the REDD+ (Reducing Emissions from Deforestation and Forest Degradation) program, have been developed to provide financial incentives to developing countries that can show a reduction in rates of deforestation and forest degradation. Today, much of the research on tropical forest carbon stocks relies upon the relationship between aboveground biomass (AGB) and forest carbon. Despite recent efforts to estimate AGB in the tropics (and in turn carbon) (refer Saatchi et al., 2011; Baccini et al., 2012; Asner et al., 2013), a large level of uncertainty in the spatial distribution and accuracy of these estimates remains (Clark and Kellner, 2012; Mitchard et al., 2013). Based on a review of published papers, estimates of tropical AGB and AGC (aboveground carbon) (Table 1.1) range from 20 Mg ha^{-1} in a very dry tropical forest (Brown and Gaston, 1995) to 497 Mg ha^{-1} in a moist tropical forest (Slik et al., 2010). The uncertainty of geospatially explicit AGB estimates is largely rooted in the inadequate spatial distribution and sampling intensity of traditional ecological data sets (Clark et al., 1998, 1999). While landscape scale tropical forest inventories may offer a solution to this uncertainty, they are rare among ecological studies due to their high cost and complex logistics (Greig-Smith, 1983). Moreover, large ecological datasets require strong database management approaches to ensure reproducible and valid data analyses (Le Duc et al., 2007; Condit et al. 2013). Existing pre- felling inventories from selective logging (i.e. forest management) may represent a solution due to their spatial distribution and sampling intensity (Couteron et al., 2003; Arroyo-Mora et al., 2009).

Logging inventories are common to tropical countries and represent an abundant source of data on forest structure and composition (Putz et al., 2001; Couteron et al., 2003; Réjou-Méchain et al., 2011a). With their success in measuring diversity on large spatial scales (Réjou-Méchain et al., 2011a), determining ecological factors that influence forest structure (Couteron et al., 2003), and estimating emission factors under REDD+ (Maniatis et al., 2011), logging inventories may provide a suitable source of forest data, complementing ecological datasets, to estimate reliable baseline carbon stocks. As reliable carbon stock estimates and "sound science" are essential to the realization and success of REDD (Petrokofsky et al., 2011), the inclusion of forest management data will help support Costa Rica's application to the REDD+ program.

1.2. Research Objectives

The overarching objectives of my Master's thesis are: firstly, to create a relational database encompassing the complexity of natural forest management data; and secondly, to assess the variability of aboveground biomass in the tropical forests of Costa Rica using this database. Prior to my research, the forest management data required to estimate aboveground biomass had yet to be fully digitized and structured in a manner that would permit its analysis. For this reason, I first aimed to develop and explore methodologies for the standardization, quality control, and organization of a large forest management dataset. Following the completion of this database, I aimed to assess the variability of aboveground biomass within and between areas of Costa Rica using forest management data. The specific objectives of my research are:

(1) To develop a spatial/tabular database that integrates methods for standardization and quality control, ensuring forest management data is structured in a logical and accessible manner.

(2) To assess the variability of aboveground biomass among and within five Costa Rican conservation areas using forest management data.

This research was conducted in the five following conservation areas: ACLA-C (Caribbean La Amistad Conservation Area), ACAHN (Arenal Huetar Norte Conservation Area), ACTO (Tortuguero Conservation Area), ACCVC (Central Volcanic Conservation Area), and ACOSA (Osa Conservation Area). All NFMPs were collected at sub-regional offices within each CA via digital photography. Tabular NFMP data was extracted and digitized in excel spreadsheets. Spatial NFMP data was digitized in ArcGIS by georeferencing maps of each managed forest unit taken directly from the NFMPs. The NFMPs collected were created between 1983 and 2011. These managed forest units represent the tropical moist, wet, and rain forest lifezones and largely fall within the lowland areas of Costa Rica with a few exceptions within the premontane transition zone.

1.3. Thesis structure

In total, this thesis is composed of five chapters. In the first chapter, I provide an overview of the background of my research, placing my thesis into the context of tropical forest management and the variability of these ecosystems across spatial scales. The second chapter presents a thorough review of the scientific literature on numerous aspects relating to large scale assessments of tropical forest biomass relevant to my M.Sc. research. Both chapters three and four present the outcome of my thesis work in the form of two distinct manuscripts for the submission and publication in peer-reviewed academic journals. More specifically, chapter three describes the development of a forest management database for five Costa Rican conservation areas, providing approaches for the standardization and quality control of large diverse forest datasets. Employing this database and the forest management data it stores, chapter four evaluates the variability of aboveground biomass within and among five Costa Rican conservation areas. In the final chapter of this thesis, chapter five, I conclude by summarizing the major findings of my M.Sc. research, emphasizing the potential impacts of my work in both the scientific research community and Costa Rica. Further, I describe future research opportunities brought to light by my results.

1.4. Tables and Figures

Table 1.1. Examples of multiple published AGB and carbon stock values measured in the tropics. AGB and carbon stock values are reported as a range or a mean value (SD) is the reported standard deviation. All values of AGB and carbon are reported in Mg ha^{-1}.

Study Location	Forest Type	AGB(StD)	Carbon(StD)*	Forest Traits	Reference
Peru (Madre de Dios)	Tropical lowland wet, moist	65.4(15)	32.7(7.5)	Secondary forest	Asner et al., 2010
Peru (Madre de Dios)	tropical lowland wet, moist	81.2(30.8)	35.6(15.4)	Forest degradation (selective logging)	Asner et al., 2010
Costa Rica (Caribbean Region)	very humid tropical	82.2(47.9)	41.1(24.0)	Secondary forest 4-20 years of age	Fonseca et al., 2011
Costa Rica (Caribbean Region)	very humid tropical	92.8(20.8)	42.9(6.0)	Secondary forest 20 years of age	Fonseca et al., 2011
Costa Rica (La Selva Biological Sation)	lowland tropical wet	161-186	80.5-93	Primary Forest	Clark and Clark, 2000
Costa Rica (Osa)	lowland tropical wet	335.12(75.13)	167.56(37.57)	Primary forest	Hofhansl et al., 2012
French Guiana (Nourages Field Station)	lowland tropical wet	356-396	178-198	Primary forest	Chave et al., 2008
Panama (Along Panama Canal)	lowland tropical moist	277.9	138.95	Secondary forest	Drake et al., 2003
Panama (Barro Colorado)	lowland tropical moist	286.8	143.4	Primary forest	Drake et al., 2003
Costa Rica (La Selva Biological Station)	tropical wet	160.5	80.3	Primary forest	Drake et al., 2003
Costa Rica (La Selva Biological Station)	tropical wet	129.4	64.7	Secondary forest of 22 years age	Drake et al., 2003
Costa Rica	tropical, subtropical	182-228	96-114	Primary and secondary forest (entire country)	Saatchi et al., 2011
Peru (High Andes)	moist tropical montane	62.5	31.25	Primary forest	Girardin et al., 2010
Neotropics: Costa Rica (La Selva), Panama (Barro Colorado), Peru (Cocha Cashu), Brazil (North Manaus)	tropical moist, wet	183.9	92	Primary forest	DeWalt and Chave, 2004

*If the carbon stock value was not reported it was calculated as 0.5 times the AGB value provided in the literature. Refer to section 2.8 p. 15 for details on conversion of AGB to carbon.

Table 1.1. Continued.

Study Location	Forest Type	AGB(StD)	Carbon(StD)*	Forest Traits	Reference
Peru (Lowland Amazon)	lowland tropical moist	121	60.5	Primary forest	Girardin et al., 2010
Peru (Cocha Cashu)	lowland tropical moist	383.5(28.9)	191.2(14.5)	Primary forest (entisol soil)	DeWalt and Chave, 2004
Brazil (North Manaus)	lowland tropical moist	190.4(73.6)	95.2(36.8)	Primary forest (spodosol soil)	DeWalt and Chave, 2004
Brazil (Amazon)	tropical moist	290	145	Mainly primary forest	Brown and Lugo, 1992
Brazil, Bolivia	Tropical terra firme, subhumid, moist	24.1	12.1	-	Houghton et al., 2001
Brazil, Columbia, Venezuela, Panama	tropical moist, dry, very dry, wet	290	145	-	Houghton et al., 2001
Borneo	tropical	457.1	228.6	Primary forest	Slik et al., 2010
Amazon: Ecuador, Venezuela, Peru, Guyana, French Guiana, Bolivia, Brazil	lowland tropical	288.6	144.3	Primary forest	Malhi et al., 2006
Brazil (North of Manaus, Amazon)	Lowland tropical moist	356(47)	178(23.5)	Primary and secondary forest	Laurance et al., 1999
Tropical Africa (37 countries)	Lowland moist tropical	299(44.9)	149.5(22.5)	Primary and secondary forest	Brown and Gaston, 1995
Tropical Africa (37 countries)	Lowland seasonal tropical	141(100.1)	70.5(50.1)	Primary and secondary forest	Brown and Gaston, 1995
Tropical Africa (37 countries)	Lowland dry tropical	60(67.2)	30(33.6)	Primary and secondary forest	Brown and Gaston, 1995
Tropical Africa (37 countries)	Lowland very dry tropical	20(22.2)	10(11.1)	Primary and secondary forest	Brown and Gaston, 1995
Tropical Africa (37 countries)	montane moist tropical	105(64.1)	52.5(32.1)	Primary and secondary forest	Brown and Gaston, 1995
Tropical Africa (37 countries)	montane seasonal tropical	37(41.1)	18.5(20.6)	Primary and secondary forest	Brown and Gaston, 1995

*If the carbon stock value was not reported it was calculated as 0.5 times the AGB value provided in the literature. Refer to section 2.8 p. 15 for details on conversion of AGB to carbon.

2. Literature Review

2.1. Literature review summary

This literature review provides a general overview of the present state of knowledge on the distribution and variability of aboveground biomass across the tropics. It begins by describing the role and importance of tropical forests in climate change. After briefly reviewing the carbon cycle of a tropical forest, I describe the United Nations REDD+ program. I then review how pre-felling inventories could help to circumvent the weaknesses associated with developing large scale AGB distribution maps, leading into an overview of the characteristics of forest management data in Costa Rica. I will then explore the most commonly used techniques to manage the large ecological datasets typical of today. Further, I will discuss the methodology used to convert tree measurements extracted from forest inventories to estimates of aboveground biomass. After, I will compare and contrast previous studies reporting the key environmental factors that influence the variability of tropical forest biomass across a range of spatial scales. Finally, I will describe the potential advantages associated with an increased understanding of biomass variability and distribution in tropical forests.

2.2. Tropical forests in the face of climate change

Representing only 7% of the Earth's land surface (Bradshaw et al., 2009), tropical forests house a disproportionately large fraction of the globe's biodiversity, ~60% of known species (Dirzo and Raven, 2003), and forest carbon (~55%) (Pan et al., 2011). Tropical forests are among the most carbon dense ecosystems and store roughly 56% of their carbon in the form of biomass (Pan et al., 2011). Through the processes of deforestation and degradation, the carbon once present in tropical forest biomass is released into the atmosphere as greenhouse gases (GHGs). As tropical forests are being cleared at rapid rates, with over 27.2 million hectares of humid tropical forests cleared from 2000 to 2005 alone (Hansen et al. 2008), they are the second largest source of anthropogenic GHG emissions (Van der Werf et al., 2009). In fact, Baccini et al. (2012) estimate that the net rate of carbon emissions from tropical deforestation and land use change between 2000 and 2010 was 1.0 Pg C yr^{-1}. Tropical forests also supply a number of environmental services beyond carbon storage including: the regulation water flow

(Bradshaw et al., 2007), the reduction of soil erosion rates, and the provision of food, shelter, and pollinators (Sodhi et al., 2007).

With the threat of climate change, international initiatives to decrease the rates of deforestation and forest degradation in the tropics, such as the UN REDD+ program, have been conceived. The first element needed to implement REDD+ is the creation of forest carbon stock baselines (Salimon et al., 2011). An established baseline is essential to monitoring changes in carbon storage through time . In recent years, the quantity of studies devoted to estimating and understanding carbon stocks across the tropics, supporting the REDD+ mechanism, have radically increased (Petrokofsky et al., 2011). More specifically, a significant number of the published AGB and carbon distribution maps have been developed for the Amazon Basin (Houghton et al., 2001; Malhi et al., 2006; Saatchi et al., 2007). Notable limitations of the available AGB and carbon stock maps are their concentration on and use of ground data from undisturbed old growth forests (Houghton et al., 2001; Malhi et al., 2006; Slik et al., 2010; Saatchi et al., 2011). In Costa Rica, no national AGB or carbon stock distribution map with detail at the landscape scale has been developed (Arroyo-Mora *pers. comm*). Assistance in the production of such a map may provide a new found potential to evaluate the variability of AGB and carbon stocks on numerous scales.

2.3. An overview of the carbon cycle in a humid tropical forest[1]

In tropical forests, photosynthetic organisms capture incoming solar radiation and atmospheric carbon (CO_2) for the construction of fresh plant biomass, net primary productivity (NPP), and their own metabolic requirements. Gross primary productivity (GPP), which is a measure of total ecosystem photosynthesis, includes the proportion of fixed carbon consumed by a plant's metabolism and released back into the atmosphere as CO_2 (Chambers et al., 2004). It has been shown that approximately 70% off GPP is respired directly back to the atmosphere by tropical vegetation (Chambers et al., 2004). The component of GPP that is not respired as CO_2 is allocated to the different components of a plant's living biomass (wood, root, and leaf). Malhi et al. (2011) found that on average NPP

---

[1] this brief review, the definition of Lewis (2006) will be adopted where a humid tropical forest is characterized by: (1) a fairly dense stand; (2) mainly evergreen broadleaf trees; (3) a closed canopy; (4) an annual precipitation of 1500 mm or greater; (5) no more than 6 months with less than 100 mm of rainfall; and (3) multiple distinct forest strata.

(described as the quantity of carbon up taken per unit time) in tropical forests is partitioned equally between a plant's wood, leaves, and roots - although notable variability exists amongst stands. In contrast to peatland ecosystems, humid tropical carbon stocks are largely stored in living plant biomass – specifically within trees (Laumonier et al., 2010). It is estimated that humid tropical forests have the greatest *total* NPP and NPP *per unit area* (1,422 g C m^{-2} yr^{-1}) of all the vegetation types on earth, accounting for approximately 35% of the net ecosystem exchange of carbon between the atmosphere and terrestrial vegetation (Melilo et al., 1993). It is important to consider the magnitude and rate of carbon uptake by the living plant biomass in these ecosystems as, it is pivotal to maintain their overall productivity. The living carbon store in plant biomass supplies the dead carbon pool via processes such as litter fall and tree mortality (Potter et al., 1993).

As a large portion of tropical soils are low in fertility, the amount and rate of carbon input from the living carbon pool is essential to the maintenance of the entire ecosystem's productivity (Vitousek and Sandford, 1986). Each of the constituents of the dead carbon pool (coarse woody debris (CWD), soil, slow and passive pools) has a unique turnover rate. These turnover rates control the rate of carbon cycling within humid tropical forests. Carbon supplied to the dead pools is released back to the atmosphere by microbial mediated decomposition of plant and soil organic matter (Cleveland and Townsend, 2006). CWD is colonized by decomposer organisms within forest ecosystems, initiating the process of decomposition (Swift et al., 1979). Respiration rates from CWD vary substantially in tropical forests (0.014 to 1.003 μg C g^{-1} C min^{-1}) based on the density and moisture content of the wood (Chambers et al., 2001). As CWD decomposes it also undergoes fragmentation, supplying carbon to the soil organic pool (Swift et al., 1979). Leaf litter fall and precipitation exert control over soil respiration in tropical forests, providing a flush of decomposable carbon to the soils (Cleveland and Townsend, 2006). Water meditated leaching transports soluble carbon molecules into the soil carbon pool, speeds up CWD decomposition, and breaks down CWD (McMinn and Crossley, 1993). Carbon may also be transported by water out of the forest system via leaching from the soil carbon pool.

The net flux of carbon between tropical forests and the atmosphere (net ecosystem exchange; NEE) is the difference between the inputs and outputs of carbon to the forest system described above. If the net output of carbon from the system is positive, it qualifies as a carbon source while, if it is negative, it is labeled as a carbon sink. Studies of NEE in humid tropical forests have resulted in contrasting conclusions, with a disagreement between their status as carbon sinks (Andraea et al., 2002) or carbon sources (Miller et al., 2011).

2.4.The REDD to REDD+ program

In the face of climate change, an international dialogue on rising greenhouse gas emissions (GHGs) has taken shape in the form of the 1997 Kyoto Protocol and the United Nations Framework Convention on Climate Change (UNFCCC). While the Kyoto Protocol did not include policy specifically aimed at reducing deforestation rates (Pielke et al., 2002), the 2005 UNFCCC included the first exploratory negotiations on the incorporation of tropical forests as a means to reduce GHG emissions from deforestation (UNFCCC, 2005). It was not until 2007, however, through the UNFCCC's adoption of the Bali Action Plan that the urgent need to lower GHGs emitted due to deforestation was officially brought to the international stage. This resolution stimulated parties to introduce methods for the Reduction of Emissions from Deforestation and Forest Degradation (REDD) in developing countries (UNFCCC, 2007). The general purpose of the REDD concept was to provide financial incentives to developing countries that could show a decrease of GHG emissions from forest degradation and deforestation. Therefore, REDD aimed to compensate developing countries for losses of income opportunities associated with a decrease in deforestation rates (Ghazoul et al., 2010). At the 16th UNFCCC Conference of the Parties (COP-16), the parties settled on a policy framework for an expanded Reduction of Emissions from Deforestation and Forest Degradation mechanism, REDD+. Beyond providing developing countries an incentive for the reduction of emissions from deforestation and forest degradation, REDD+ offers a number of additional benefits including biodiversity protection, forest restoration, and sustainable forest management and development (Asner, 2011).

For the REDD+ mechanism to actually function, the development of robust national and sub-national methods to monitor, report, and verify forest carbon stocks is required. Further, reference emission levels will need to be established against which the loss of carbon stocks can be estimated. Essentially, the net change in carbon stock values would be measured against the established baseline to calculate carbon credits to be sold to an international carbon market (Griscom et al., 2009). No agreement, however, has been made on the methodology, criteria, and constitution of a reference level. At a conceptual level, references can be set as a business as usual or a crediting baseline. Particular disagreement has occurred on the issue of whether or not baselines should be set at national levels. Advocates of national baselines at the REDD negotiations (the EU, USA, Canada, Japan and the Rainforest Coalition) asserted that national baselines are the only way to control leakage - the displacement of degradation and deforestation activities within a country (Potvin and Bovarnik, 2008). Contrastingly, a group of Latin American countries led by Columbia argue that national baselines are presently beyond the financial and technical capacity of many developing countries. Due to the importance of reference levels to the success of the REDD program, the UNFCCC COP-14 held an expert meeting on the topic (UNFCCC, 2008). This discussion drew particular attention to uncertainty in data and data quality related to carbon and biomass density estimates, allometric equations, and biomass expansion factors (UNFCCC, 2009).

2.5. Techniques and sources of error in large scale AGB and AGC estimates

Remote sensing and ground level data are two techniques proposed for the production of reliable carbon estimates (Hill et al., 2013). An advantage of remote sensing technology is its ability to collect data covering an entire country. In recent years, a great amount of effort has been placed on the use and integration of LiDAR (Light Detection and Ranging) and RaDAR (Radar Detection and Ranging) for the estimation of AGB (e.g. Saatchi et al., 2011; Asner et al., 2013). LiDAR is a particularly valuable form of remotely sensed data as it is capable of capturing the aspects of forest structure (e.g. tree height) required to estimate AGB (Drake et al., 2003; Saatchi et al., 2011; Zolkos et al., 2013). However, this data still needs to be calibrated and validated with ground truth measurements (Drake et al., 2003). Additionally, some remote sensing instruments may not be sensitive enough to detect differences in AGB among the high-density forest stands (Goetz et al., 2009) common to tropical wet and moist ecosystems. Ground level data from traditional ecological forest inventories may introduce error to carbon stock estimates. Within these inventories, the standard plot size can be quite small (0.1 ha) and there can be a bias towards high biomass (ideal) locations within a forest stand (Brown and Lugo, 1992). According to Clark and

Clark (2000), a minimal plot size of 0.3 ha is required to capture the variability present in a tropical wet forest stand. Recently, Wagner et al. (2010) reported that four 0.25 ha plots are required to estimate the AGB of a neotropical rain forest within a 20% error of the mean. When extrapolated to the landscape scale, an inventory that lacks the variation needed to accurately measure biomass will produce an unreliable estimate (Brown and Lugo, 1992). Due to their spatial distribution and sampling intensity, pre-felling inventories from selective logging (i.e. forest management) may represent a solution to these problems (Couteron et al., 2003). Logging inventories are common to tropical countries and represent a large source of data on forest structure and composition (Putz et al., 2001). With their proven success in measuring diversity on large spatial scales (Réjou-Méchain et al., 2011a) and determining ecological factors that influence forest structure (Couteron et al., 2003), logging inventories have the potential to be a suitable source of forest data to estimate baseline carbon stocks.

2.6. Natural forest management in Costa Rica

In Costa Rica, selective logging inventories are developed under a standardized Natural Forest Management Plan (NFMP) framework (Arroyo-Mora et al., 2009) that has been implemented on private lands since 1996 (Forestry Law 7575). Despite the country's small size (approximately 5100 km^2) it contains a rich diversity of tropical ecosystems ranging from dry to wet forests. Further, Costa Rica is distinct among developing countries due to its forward thinking environmental policies, system of protected areas, and dramatic reduction in rates of deforestation during 1990's (Miller, 2011). As Costa Rica's forest resources are managed under a national system of conservation areas (CAs), NFMPs are handled in local sub regional offices. In total, the country is divided into 11 CAs (Boza, 1993). Within each CA, sub regional offices are responsible for the reception, revision, approval or rejection, and follow up of the forest management plans. Due to the limited resources of sub regional offices, this rich information is often inadequately stored, organized, and in some cases even destroyed (e.g. through floods, fire).

A NFMP is required of a private farm owner who desires to selectively log his *primary* forest stand. In order to be legally approved to log, the owner must hire a certified forester to conduct an inventory and census of the forest stand (Arroyo et al., 2009). In an inventory, every tree with a diameter at breast height (DBH) equal or greater than 30 cm is measured and identified in plots of 0.3 ha. In a census, every tree with a DBH equal or greater than 60 cm is measured and identified throughout the entire forest stand. In order to recover these archives and evaluate their capacity to support forest conservation, Arroyo-Mora et al. (2009) developed a historical forestry geographic information system (FGIS) in a case study of a CA in Northern Costa Rica. Following that original study, my current work encompasses the complete historical FGIS of Costa Rica, including the five CAs where selective logging has been most heavily practiced. Specifically, selective logging has focused in the lowland areas of the Atlantic coast, the northern lowlands, and the southern Pacific coast. The agglomeration of NFMPs is important to the conservation of Costa Rica's forests, as, despite the country's recognition for its environmental policies, the sustainability of its NFMP framework has been subject to little testing and validation (McGinley and Finegan, 2003).

2.7. The management of large forest themed datasets

With modern ecological datasets combining spatial and temporal scales, powerful data infrastructures are a necessity for efficient data storage and analysis (McInthosh et al., 2007). The relational database model, developed by International Business Machines corp. (IBM) in the 1970's, has become an international standard (Codd, 1979). A relational database model recognizes logical associations between information of different types and themes. It is primarily composed of tables, fields (columns), records (rows), relationships, and primary keys (a unique value that provides an "*address*" for each record in the database) (Codd, 1979). Ecological databases following this model have been used across the globe to efficiently store, mine, and analyze data - all while minimizing redundancy (McIntosh et al., 2007). The RAINFOR database is one such model, composed of long-term ecological information from plots established across the Amazonian rainforest (Peacock et al., 2007). The database infrastructure consists of a wide variety of forest themed tables and relationships. For example, the "Plot" table is related to the "Local Climate" table by the primary key "PlotID" (Figure 2.1). As an analysis tool, the database can rapidly assess and graph above ground biomass (AGB) across over 100 plots with SQL (standard query language). Recently, the Center for Tropical Forest Science (CTFS) developed a database for a global network of forest plots (Condit et al. 2013), applying the relational data model and its theories of data normalization (Codd, 1971). Both Condit et al. (2013) and Peacock et al. (2007) emphasize the utility of a normalized relational database for the storage and tracking of taxonomic data. Further, they highlight the importance of developing standards for the spelling of species, genera, and families for the reliable analysis of taxonomic data across a large forest themed dataset. It is also essential that databases establish a system to track changes in taxonomy and unidentified trees as (1) taxonomy is in a constant state of change; (2) tropical trees are frequently unidentified at a species level; and (3) hundreds to thousands of species can exist within a few hectares of a tropical forest (Condit et al. 2013).

Another advantage of relational database management systems is its capacity to ensure quality through the use of built-in data constraints and quality check queries (Le Duc et al., 2007). Data constraints limit the data that can be entered into a given field or table of a database. For example, in their relational database, Le Duc et al. (2007) specify that frond dry mass (g) must be greater than or equal to 0 but less than 1000. Quality check queries are run after data entry to check for errors. In the RAINFOR database, for example, a quality check query ensures that the number of stems alive, dead, and recruited is consistent between inventories of the same forest plot through time (Peacock et al., 2007). Beyond correcting data errors, the possibility to include customized quality assurance procedures in relational database managements systems can help to uncover the rate, source, and consequence of data entry errors (Le Duc et al., 2007). A more recent extension of the relational database model is the geodatabase, which contains spatial information that may be displayed in a geographic information system (GIS) (Zeiler, 1999). GIS databases have been employed, for instance, in Mexico to study land cover change (Mas et al., 2004) and by the US Forestry Service to model appropriate locations for forest management (Loh et al., 1994). In Costa Rica, the large NFMP dataset will not be reliably or easily analyzable without first using these versatile relational database tools. By developing a relational-spatial database tailored to Costa Rican NFMPs as well as customized standardization and quality assurance procedures, it will be feasible to efficiently access data, perform repeatable data analyses, and improve overall data quality.

2.8. Estimating AGB and AGC from tree measurements

Following the creation of a powerful data infrastructure populated with tree level information acquired from NFMPs, it will be possible to analyze and estimate AGB. Chave et al. (2004) outline four steps to determine AGB (the first value needed to determine carbon stocks). The first step, measurement of trees, can be mined from NFMPs. Though this means that only tree data will be contained in the AGB estimates, it has been found throughout tropical forest ecosystems that the largest portion of biomass is held within living trees. Further, trees above 20 cm DBH can contribute up to 80% of the total AGB (Laumonier et al., 2010). Despite the fact that NFMPs only include trees of 30 cm DBH or greater, these estimates will still encompass a substantial and important portion of AGB. Next, tree level measurements are entered into allometric models to calculate AGB. In contrast to temperate ecosystems where species specific allometric models are applied, the large diversity of tree species in tropical forests means that mixed species allometric models are required (Chave et al., 2005). Allometric models are developed by destructively sampling trees and relating directly measured biomass to tree measurements, such as DBH (Brown et al., 1989). It is important to note that allometric models *do not directly* measure AGB but, instead, provide an AGB *estimate* (Clark and Kellner, 2012). Choosing the correct allometric model is a critical step to avoid errors in AGB estimates. Important criteria for model selection include the compatibility of your sample DBH range with that of the model and the ecosystem being studied (specifically because mixed species models are used in the tropics). Extrapolating an allometric model beyond its maximum DBH range can introduce up to 30% error in final AGB estimates (Chave et al., 2004). The third step is to sum biomass per tree across all trees in a given plot. It is in this step that an unrepresentative plot size will strongly affect the accuracy of AGB estimates. In small ecological plots (0.1 ha) large diameter trees represent a specific source of error due to their spatial rarity and significant contribution to total AGB (Brown and Lugo, 1992). The final step in determining AGB is averaging biomass across plots. Having a representative sample of the landscape heterogeneity will produce an average that is truly representative of a given area. Additionally, to allow for regional scale comparisons, AGB estimates must be based on a consistent regression model (Baker et al., 2004).

In recent years, the pantropical allometric models developed by Chave et al. (2005)

have been widely applied across the globe to estimate AGB. In Costa Rica, however, numerous studies of biomass have employed Brown's (1997) equation for wet forests (e.g. Clark and Clark, 2000; Letcher and Chazdon, 2009; Clark et al., 2011; Saatchi et al., 2011) as it was calibrated with data collected at Costa Rica's La Selva Biological Station. For this study, the Brown's (1997) equation had many disadvantages when compared to those developed by Chave et al. (2005). These shortcomings included: (1) the representation of a smaller DBH range; (2) the development of the equation from a smaller sample size; (3) the limited application of the equation outside of Costa Rica (making the comparison of AGB estimates with other countries/studies more complex); and (4) the absence of wood density as a parameter (an aspect of forest structure that varies significantly at regional scales) (Baker et al., 2004; Chave et al., 2009).

Using wood specific gravity, available for many neotropical tree species in online databases and scientific literature (Chave et al., 2005), to calculate tree biomass may help account for variations in biomass among different species. Wood density is an important predictive variable when estimating AGB (Brown and Lugo, 1992; Chave et al., 2005; Keeling and Phillips, 2007; Baker et al., 2009). As it is known to vary among different forest communities (Baker et al., 2004; Chave et al., 2009; Muller-Landau, 2004; Zhang et al., 2011), this is also critical to studying the differences in AGB across a landscape. Despite these findings, wood density has yet to be studied or implemented when estimating AGB across Costa Rica.

Feldpausch et al. (2011, 2012) reported that an additional source of error in AGB estimates results from the exclusion of height as a predictor variable in allometric models. Regardless of this finding, Feldpausch et al. (2012) also reported that the decrease in AGB estimation error occurred only in smaller DBH classes (≥40cm) and not larger ones. Therefore, the greater the proportion of AGB in trees with a DBH less than or equal to 40 cm, the greater the amount of error introduced into AGB estimates when excluding height in allometric models. As NFMP data only includes trees with a DBH≥30cm, the exclusion of height in the models applied to estimate AGB will not greatly increase the error of the reported values.

Once an average AGB density (biomass per area) is calculated, it can be converted to carbon by multiplying the value by 0.5. This relatively simple relationship, where carbon represents 50% of the total tree biomass, is the standard conversion factor throughout the literature (Houghton et al., 2009). However, studies have reported that the fraction of carbon content of AGB fluctuates between tree species. Elias and Potvin (2003) found the proportion of carbon in the biomass of 32 neotropical tree species varied between 44.4%-49.4%. Further, Martin and Thomas (2011) reported that wood carbon content ranged from 41.9%-51.6% based on a sample of 190 neotropical trees. Additionally, the IPCC recommends the use of a 47% conversion factor from biomass to carbon (Gibbs et al., 2007). As this conversion is an additional source of uncertainty, studies often report both AGB and AGC values (e.g. Saatchi et al., 2011). Next, an average AGB or AGC density can be extrapolated to the landscape using a land cover map and/or the application of calibrated models derived from ground level and remotely sensed data (Gibbs et al., 2007).

In the Brazilian Amazon, a mere 5% level of agreement was found between published AGB distribution maps that were produced using seven different extrapolation methodologies (Houghton et al., 2001). In particular, the simpler methodologies applied to extrapolate plot level AGB data, such as krigging and spline interpolation, performed poorly (Houghton et al., 2001; Malhi et al., 2006). The poor performance of these techniques in tropical ecosystems is due to local scale variations in AGB outweighing the variability of AGB found at regional scales (Houghton et al., 2001).

2.9. Factors influencing the spatial distribution of AGB

The resulting AGB distribution maps produced by of the aforementioned techniques may be improved by the inclusion of information on environmental factors known to affect forest structure and carbon stocks (Table 2.1). In fact, recent advances in AGB estimation that combine remotely sensed data, forest plots, and environmental factors (e.g. Asner et al., 2013) have performed much better, producing higher quality and finer scale carbon density maps. Multiple studies have investigated the influence of environmental factors on tropical forest structure, reporting a wide array of variable relationships (Table 2.1). In a wet tropical forest, Clark and Clark (2000) observed the proportion of large diameter trees increased with soil fertility. Further, they found that stem density displayed a positive relationship with slope. Although their study found relationships between forest structure, soil type, and slope, Clark and Clark (2000) concluded that these factors did not explain the low AGB variability (12%) across the forest stand. In contrast, Laurence et al. (1999) reported that soil fertility explained approximately 30% of the AGB variation in a tropical moist forest. These findings were supported by Castilho et al. (2006) who found that AGB was positively related to soil quality and elevation but independent of slope. Similarly, Alves et al. (2010) showed that AGB increased along an elevation gradient in a moist tropical forest. A regional scale study of the Amazon Basin indicated that another environmental variable, dry season length, was an important factor controlling AGB variability (Malhi et al., 2006). In the same study, conflicting with the findings of Castilho et al. (2006) and Laurance et al. (1999), AGB was reported to reach maximum values in slow growing forests found on infertile soils (Malhi et al., 2006). The mean wood density of a forest stand (Baker et al., 2004; Asner et al., 2009) and forest disturbance (Urquiza-Haas et al., 2007) have also been recognized as factors related to AGB variation in the Neotropics. Habitat fragmentation perturbs the cycling of biomass at a forest's edge, speeding up biomass' turnover rate, altering species composition, and ultimately decreasing carbon storage in the living vegetation of a forest stand (Nascimento and Laurance, 2004). Recently, Slik et al. (2013) found that large tree density (DBH≥70cm) explained 70% of the variation in pantropical AGB estimates. Although they found large tree density was positively associated with soil fertility, the dominance of wind dispersed species, and other climatic variables, they reported that the relationship of large tree density to associated factors was inconsistent between continents (Slik et al., 2013). Further, Asner et al. (2013) determined that human activity was the greatest driving force in the variability of AGB across Panama, highlighting the need to measure the impact of human activity on AGB across the tropics. As NFMP data offers a sample of the tropical forests that are being greatly impacted by anthropogenic activities, their inclusion with ecological data will provide a more representative ground level dataset for the estimation of AGB at landscape, regional, and national scales.

2.10. Conclusions

With the carbon market valued at over US$100 billion per year, a great emphasis has been placed on the production of high-quality national carbon stock baselines (Petrokofsky et al., 2011). Therefore, assessing the spatial variation of AGB will assist in the implementation of the REDD+ mechanism, substantially improving estimates of emissions caused by deforestation and forest degradation. In the past, these emissions have been calculated using single biome averages of carbon density (Gibbs et al., 2007)[2]. It is currently well known that carbon stocks vary within biomes and on finer scales (Clark and Clark, 2000; Houghton et al., 2001; Malhi et al., 2004; Laumonier et al., 2010). A former lack of consideration for this variation has resulted in relatively unreliable GHG emission estimates due to tropical deforestation (Baccini et al., 2012). The main objective of my Master's thesis is to assess the variability of aboveground biomass in tropical forests within and among regions of Costa Rica using forest management data. The outcome of my thesis will be especially helpful in Costa Rica, where accurate carbon estimates will improve the country's implementation of the REDD+ program - placing an economic value on living forest stands.

[2] Refer to example of exception in Baccini et al. 2012

2.11 Tables and Figures

Table 2.1. Summary of environmental factors observed to effect (positively or negatively) the spatial distribution of AGB in tropical forests.

Environmental Variable	Reference
Forest Fragmentation and Disturbance	Nascimento and Laurance, 2004; Urquiza-Haas et al., 2007
Soil Type and Characteristics	Laurance et al., 1999; Malhi et al., 2006; Castilho et al., 2006; DeWalt and Chave, 2004; Paoli et al., 2008
Geologic Substrate	Asner et al., 2009
Precipitation	Malhi et al., 2006; Slik et al., 2010; Chave et al., 2004; Slik et al., 2013
Slope	Clark and Clark, 2000; Castilho et al., 2006
Elevation	Alves et al., 2010; Castilho et al., 2006
Forest Type	Houghton et al., 2001; Chave et al., 2005
Wood Density	Baker et al., 2004; Asner et al., 2009
Dry Season Length	Malhi et al., 2006

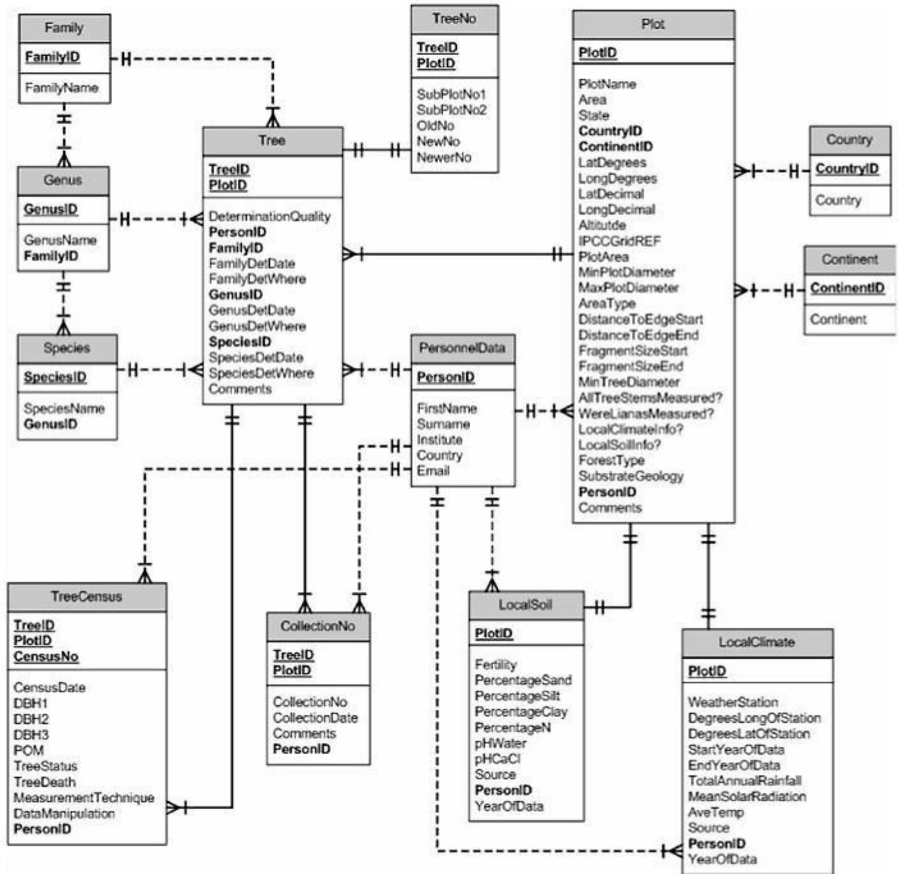

Figure 2.1. Source: Peacock et al. (2007). Schema of the relations between the tables of the RAINFOR database with a list of their respective fields. Primary keys are underlined item in bold are required. Solid lines represent links between primary keys. The relationship between TreeNo and Tree. Plot and LocalSoil, and Plot and LocalClimate are one to one. All other relationships are one to many. (License for use of Figure 2.1 in thesis given in Appendix 1).

3. The development of a forestry geodatabase for natural forest management plans in Costa Rica

Citation: Svob, S., Arroyo-Mora, J.P. and Kalcska, M. (in press) 2014. The development of a forestry geodatabase for natural forest management plans in Costa Rica. *Forest Ecology and Management.*

3.1. Context within thesis

This chapter describes the development of the forestry GIS (FGIS) database for the storage, standardization, quality control, and analysis of a large Costa Rican forest management dataset. The results of this study illustrate the advantages of database development for tropical forest data. Further, an exploratory analysis demonstrates the utility of the database, providing a platform for the efficient and reproducible analysis of taxonomic information for over 250,000 tree records. The main outcome of this chapter is the finalized FGIS database, representing the most complete digital record of natural forest management in Costa Rica to date and a tool for future large scale analyses of tropical forest ecosystems.

3.2. Abstract

Forest management data is available for many tropical countries, representing a large and spatially rich source of tree level data. Over the past decade, we have collected and digitized over 1000 Costa Rican natural forest management plans, spanning 30 years and spread out over approximately 26,700 km^2 along the country's Atlantic coast, northern lowlands, and southern Pacific coast. In order to analyze this unique collection of tree inventory data we developed a system to ensure the standardization, quality control, and reliable management of the dataset. We developed a relational geodatabase, forming logical associations between and within the spatial and tabular components of the forest management data. In this study, we outline the data standardization procedures established to permit the comparison of data across time and space. Further, we describe quality checks built-into the database's functionality to identify and reduce the presence of errors. The final customized forest management geodatabase efficiently stores a large and diverse dataset through the integration of logical relations, quality checks, and flexible data access across 32 tables and

over 250,000 tree records. Through the use of quality tests, the database provides a means to improve overall data integrity and illustrates the magnitude, sources, and types of errors present in the initial dataset. Finally, the value of a comprehensive database for the management of forest data is demonstrated through an exploratory analysis of standardized tree taxonomic information. With this analysis, we begin to explore the potential strengths, weaknesses, and applications of forest management data for future ecological studies (e.g. species diversity assessment). For instance, although most of the forest management data (tree inventories and censuses) is collected using common names in the field, our standardization process has allowed us to depict trends similar to those found in ecological studies (e.g. dominant species for different ecosystems). Overall, our forestry geodatabase represents the most complete record of natural forest management practices in Costa Rica to date.

***3.2.1. Keywords*:** Data management; Logging inventories; Tropical forest; Ecoinformatics; Quality assurance

3.3. Introduction

While tropical forests roughly represent only 7% of global land cover (Bradshaw et al., 2009), they are home to an estimated two-thirds of the globe's terrestrial biodiversity (Gardner et al., 2009) and contain nearly 40% of the earth's carbon biomass (Brown and Lugo, 1992). Due to a rapidly changing climate combined with high rates of deforestation and forest degradation, these globally important carbon and biodiversity stocks are under threat (Malhi and Grace, 2000). As these ecosystems decline, a renewed interest in quantifying and understanding spatial and temporal variations in biodiversity and carbon distribution across tropical forests has arisen (Petrokofsky et al., 2011). However, a high level of uncertainty in the assessment and understanding of environmental factors influencing biomass and biodiversity across spatial scales persists (Houghton et al., 2001; Chave, 2008; Gardner et al., 2009). This uncertainty is primarily rooted in the inadequate spatial distribution and sampling intensity of traditional ecological data sets (Clark et al., 1998, 1999). While landscape scale tropical forest inventories may offer a solution to this uncertainty, they are rare among ecological studies due to their high cost and complex logistics (Greig-Smith, 1983). Moreover, long-term datasets require strong database management approaches to ensure reproducible and valid data analyses (Le Duc et al., 2007; Condit et al., 2013). Existing pre-felling inventories from selective logging (i.e. forest management), on the other hand, may represent a solution due to their spatial distribution and sampling intensity (Couteron et al., 2003; Arroyo-Mora et al., 2009).

Logging inventories are common to tropical countries. They represent a valuable and abundant source of data on forest composition and structure, permitting a variety of ecological questions to be addressed at spatial scales larger than those in traditional ecological inventories (Putz et al., 2001; Couteron et al., 2003; Réjou-Méchain et al., 2011a). Due to their unique scale, logging inventories have allowed researchers to observe the environmental factors controlling species distribution and community composition at regional and landscape scales (Couteron et al., 2003; Réjou-Méchain et al., 2008; Gourlet-Fleury et al., 2011; Fayolle et al., 2012). Research conducted at these large scales provides a link to local scale studies and therefore, insight into the importance of certain environmental factors and processes controlling community composition at different spatial scales (Réjou-Méchain et al., 2011b). The large area sampled by logging inventories also allows for studies to successfully capture the heterogeneity of tropical forest biomass distribution, and in turn forest carbon (Gourlet-Fleury et al., 2011; Maniatis et al., 2011). These studies can help reveal the impact of particular environmental factors (e.g. elevation, soil type, and land-use history) on tropical forest biomass, providing crucial information for future studies attempting to predict the regional and landscape patterns of biomass. Finally, large scale inventories aid in the development of detailed wall-to-wall vegetation and land use maps by offering a rich source of ground level data for the validation and calibration of remotely sensed data (Gond et al., 2013).

In Costa Rica, selective logging inventories are developed under a standardized Natural Forest Management Plan (NFMP) framework (Arroyo-Mora et al., 2009) that has been implemented on private lands since 1996 (Forestry Law 7575). As Costa Rica's forest resources are managed under a national system of conservation areas, NFMPs are handled in local sub regional offices. In total, the country is divided into 11 conservation areas (Boza, 1993). Within each conservation area, sub-regional offices are responsible for the reception, revision, approval or rejection and follow up of the forest management plans. Due to the limited resources of sub-regional offices, this rich information is often inadequately stored, organized, and in some cases even destroyed (e.g. floods, fire).

In order to recover these archives and evaluate their capacity to support forest conservation, Arroyo-Mora et al. (2008) developed a forestry geographic information system (FGIS) in a case study of a conservation area in Northern Costa Rica. Following that original study, this present work encompasses the complete FGIS of Costa Rica, including the five conservation areas where selective logging has been most heavily practiced. Specifically, selective logging has focused in lowland areas in the Atlantic coast, the northern lowlands, and the southern Pacific coast. The main objective of this study is to develop a countrywide tree database from data extracted from natural forest management plans (NFMPs). Our second objective is to describe the broad taxonomic patterns occurring in our study region based on the FGIS database in order to assess the utility of the NFMP dataset.

With modern ecological datasets combining spatial and temporal scales, powerful data infrastructures are a necessity for efficient data storage and analysis (McIntosh et al., 2007). The relational database model, developed by IBM in the 1970's, has become an international standard (Codd, 1979). Relational databases following this model have been used extensively to efficiently store, mine, and analyze ecological data, all while minimizing redundancy (McIntosh et al., 2007). The RAINFOR database is one such model, composed of long-term ecological information from plots established across the Amazonian rainforest (Peacock et al., 2007). More recently, Condit et al. (2013) developed the CTFS database to reduce the presence of integrity errors within a global repeated-measurements forest plot dataset. A newer extension of the relational database model is the geodatabase, which also contains spatial information that may be displayed in a geographic information system (GIS) (Zeiler, 1999). For instance, geographic information databases have been employed in Mexico to study land cover change (Mas et al., 2004) and by the US Forestry Service to model locations for forest management (Loh et al., 1994). This study presents the development and structure of the FGIS geodatabase along with the forest management, taxonomic, and spatial data that it stores. Additionally, it highlights the importance of establishing data standardization and quality assurance procedures when managing large ecological datasets. Finally, it describes the limitations of such a database before providing an example of its potential application in the study of Costa Rica's diverse forested ecosystems.

3.4. Natural forest management plans in Costa Rica

In Costa Rica, a natural forest management plan is a document containing a collection of technical standards combined with data, developed to govern the management activities of a privately owned area of natural forest (management unit or forest unit). The data applied in this study was extracted from NFMPs conducted by certified foresters between 1983 and 2011. Below, we provide a summary of the standard methodology followed and data collected in these NFMPs.

In general, a NFMP document produced by a certified forester contains information on the structure and composition of a forest stand (inventory and census), official land tenure information, the protected and productive areas within the stand, proposed logging roads and timber patios (MINAE, 2008). The document also contains general maps of land holdings including, among other features, the shape of the forest unit, the trees to be logged, and the trees to be preserved as progeny trees. In addition, protected areas in the management unit are defined as areas near water bodies (e.g. streams, springs) and on steep slopes where selective logging is prohibited (MINAE, 2008). The forested area outside of the protected area within the bounds of each management unit is the productive area.

Costa Rican natural forest management plans generally follow a systematic field sampling design. The design consists of transects placed across the forest stand perpendicular to a pre-established baseline. Depending on the terrain and total forest unit area, adjacent transects are separated by 50-100 m and extend the length of the forest stand. Transects are located spatially by at least one point in the field and later mapped. Inventory plots of 0.3 ha (30 by 100 m) are mounted randomly along these transects and all trees with a DBH ≥ 30 cm are recorded within the plots. Despite these standards, however, inventory plots are sometimes biased towards species rich and biomass dense areas of the forest stand (Arroyo-Mora, *pers. obs.*) The number of plots mounted through the inventory is determined so that the sampling error is less than 20% (95% confidence level) for the basal area per hectare of trees with a DBH ≥ 30 cm (MINAE, 2008). Of the data included in this study, the number of inventory plots per NFMP ranged from one to 53. A more complete characterization of the tree population in the forest stand is given by the census. In the census, all trees with a DBH ≥ 60 cm are recorded along the previously mounted transects (50-100 m apart extending the length of the forest stand). Trees recorded in the census are spatially located in the forest stand and classified as trees for harvest or as remnants. The harvest intensity is determined in proportion to the abundance of each tree species. In Costa Rica, the logging intensity must be less than 60% of the number of trees per species with a DBH ≥ 60 cm (MINAE, 2008). Based on the inventory, any tree species represented by less than 0.3 trees per ha of DBH ≥ 30 cm is not subject to harvest, even if it is a commercially viable species. These trees, any non-commercially viable trees, and the remaining 40% of commercially viable trees are classified as either remnant or parent trees (refer to Appendix 2 for detailed definitions). The DBH of all trees sampled by a census or inventory is measured with a diametric tape at a height of 1.3 m from the base of the tree. Commercial height, the height from the base of the tree to the bole of the first branch, is measured for all trees using either a clinometer or hypsometer. When possible, all trees are identified in the field by local experts, referred to as "baquianos", using vernacular names.

3.4.1 NFMP data collection

The NFMP data included in this study is composed of NFMPs that were produced between 1983 and 2011. The NFMP records were collected at sub-regional offices in the following five conservation areas: Caribbean La Amistad (ACLA-C), Arenal Huetar Norte (ACAHN), Tortuguero (ACTO), Central Volcanic Cordillera (ACCVC), and Osa (ACOSA) (Fig. 3.1). At each sub-regional office visited, a digital camera was used to collect photographs of every page of the hard-copy NFMPs. The quality and organization of the hard copy NFMPs varied widely from one conservation area to another; nonetheless, regardless of the completeness of the management plans, we aimed to extract as much information as possible. Following the methodology developed by Arroyo-Mora (2008), the general aspects of the forest management plans (location, forest unit size, productive, protected area extent, etc.), tree inventories, tree censuses, and the lists relating scientific names to common names were digitized into separate Excel spreadsheets. In addition, the map of each NFMP was digitized and georeferenced in ArcGIS (versions 8.0-10.1, ESRI Inc., Redlands, CA, USA).

3.4.2. Data standardization

Data recorded in the excel spreadsheets was formatted uniformly in order to be integrated into a relational database. In addition, the terminology and floristic information reported in the NFMPs was not up to date or consistent (data spanning 30 years over 5 different regions) and required standardization before constructing the database and performing subsequent analysis.

The complexity of the information extracted from the forest management plans (land tenure, inventories, census, location, etc.), required the development of several procedures, each tailored to a particular subset of data, to completely standardize the data. To ensure that values were comparable across all NFMPs, fields were standardized to a single unit of measurement. For example, all DBH measurements were converted into centimeters, commercial heights were converted into meters, and any measurements of area (e.g. productive area) were converted into hectares. A forest management plan identification code (NFMPID), which uniquely identifies each NFMP, was made into a standardized format consisting of six letters from the NFMP's name followed by the year the NFMP was submitted to the forest authority of the state. The statuses of all the individual trees recorded in a census were also standardized, coded as cut, parent tree, remnant, or other. As Costa Rica falls within two Universal Transverse Mercator zones, it was important to standardize the projected coordinate system of the NFMP polygons and any additional spatial datasets. To do so, all projections were transformed into the country's official coordinate grid, the Costa Rican Transverse Mercator.

Taxonomic information also needed to be updated and standardized as it was extracted from NFMPs produced across Costa Rica and over several decades. NFMPs generally only supply floristic information at the genus and species levels. Therefore, family names were extracted after correcting any spelling errors of species and genus names using the Taxonomic Name Resolution Service v3.2 (Boyle et al., 2013). When running the scientific names through the resolution service we employed TROPICOS (Missouri Botanical Garden, 2013) as the taxonomic source for species and genus matching and the Angiosperm Phylogeny Group III (APG III, 2009) as the standard authority for family-genus taxonomy. Any species listed as an out of date synonym was replaced by the currently accepted species name. Specified author(s) and publication information at the family, genus, and species levels was also taken from TROPICOS. Beyond the NFMP data, a national list of scientific names provided by the government of Costa Rica (SIREFOR, 2012 *unpubl.*) was standardized following the above procedure. The national list provides both the common names and corresponding scientific names. Unlike NFMP lists, the national list includes all of the possible common names for a single species and all of the possible species names for a single common name. The national level list is particularly useful when a NFMP list relating common names to scientific names is absent.

3.4.3 Database design

In order to create a relational geodatabase based on NFMPs from five conservation areas in Costa Rica (Fig. 3.1.), we used the following set of rules:

1. Be able to manage spatial information.
2. Be able to form logical associations between NFMP based information of different types and themes.
3. Support the production and storage of metadata.
4. Include methods for quality control and assurance.
5. Be proficient at importing and exporting information in basic file formats, particularly those that can be easily used in statistical analysis software.

After determining the requirements of the geodatabase, an entity relationship model was generated, defining entities and the relationships among entities (Chen, 1976). The entity relationship model was translated into a relational model, which then underwent database normalization, minimizing redundancy and dependency (Codd, 1970). PostgreSQL (version 9.2., PostgreSQL Global Development Group, Berkeley, CA, USA) and the spatial extension PostGIS 2.0 (version 2.0., Refractions Research Inc., Victoria, BC, CA) were selected to construct the geodatabase. PostGIS is supported by a variety of GIS software such as QuantumGIS (version 1.8.0., Open Source Geospatial Foundation, DE, USA) and ArcGIS (version 10.1., ESRI Inc., Redlands, CA, USA), allowing the use of a GIS to display and edit spatial information.

3.4.4. Database structure

3.4.4.1. NFMP level database components

The final database structure is composed of 32 relational tables (Fig.3.2.). The *nfmp_general* table relates to nearly all other tables within the database. This table stores NFMP level information such as the name of the NFMP (farm_name), the sub-regional office where the NFMP's data was collected (sroid), the year the plan was created (start_year), the year it expires (end_year), and whether or not the plan was approved for logging (approved). The table's primary key, the natural forest management plan identification code (nfmpid), uniquely identifies each management plan within the database and forms important relationships with the tree level data. As reported in a NFMP, the *nfmp_general* table stores a minimum, maximum, and average value of slope and elevation for each management unit. The total, managed, productive, and protected area are also recorded in the *nfmp_general* table as an NFMP can have only one value for each of these attributes. The table includes a set of attributes that specify whether an NFMP collected contained a census, inventory, list, or map of the management unit (census, inventory, list, map_farm, map_census, map_inventory). Finally, the table stores timber harvest information such as the number of trees and volume requested for extraction by the forester responsible for the plan (number_tree_requested, volume_requested).

The *forester* table provides information on each forester responsible for the production of an NFMP. The cardinality between the forester and *nfmp_general* tables represents a many-to-many relationship as a forester can develop many NFMPs and a NFMP can be developed by more than one forester. To meet database normalization standards, this many-to-many relationship needed to be broken down into two one-to-many relationships. To do so, an associative table (*nfmp_general_forester*) connecting nfmp_general and forester was incorporated into the database. Each record in *nfmp_general_forester* relates one forester to one NFMP and has a composite primary key consisting of a foreign key referring to *nfmp_general* (nfmpid) and a foreign key referring to *forester* (foresterid). The owner table, which stores an alphanumeric code, was structured in the database following the same logic as the forester dataset.

The *nfmp_general* table links to spatial and tabular data related to the administrative divisions of Costa Rica at the district, canton, and provincial levels. The *nfmp_polygon* table stores a record of the polygon(s) outlining the location of each management unit. All tables storing spatial data contain two fields that refer to the geographic and projected coordinate systems tables. These references ensure that the projected and geographic coordinates of all spatial data is tracked in the database.

3.4.4.2. Tree level database components

The tree level information stored in the *inventory* and *census* tables forms a one-to-many relationship to the *nfmp_general* table. A one-to-many relationship was assigned between these tables because an NFMP can have many trees recorded in either a census or inventory. To form the relationship, the primary key of *nfmp_general* table (nfmpid) was included in the *inventory* and *census* tables as a foreign key. *Census* contains tree level information collected during a census. This includes a tree number, common name, DBH, commercial height, and status. *Inventory* contains tree level information collected during the plot based sampling of an NFMP's productive area. For each tree record, *inventory* stores a plot number, tree number, common name, DBH, and commercial height. Because natural forest management plans do not consist of repeated measurements of the same tree, unlike many ecological inventories (Condit et al., 2013), it was not necessary to separate permanent tree attributes (e.g. common name) and changing traits (e.g. DBH or commercial height) into different tables. In the *census* table, *inventory* table, and throughout the database, auto-increment primary keys were used. Auto-increment primary keys are not reliant on an actual value extracted from an NFMP, permitting the entry of incomplete records (i.e. missing data). This was an important factor to consider when developing the database structure as some of the original hard copy NFMPs were missing data or contained illegible data. To ensure that duplicate entries were not allowed despite the use of auto-increment primary keys, the combination of particular fields was given a unique constraint. For example, in *inventory*, a unique constraint was placed across the plot, tree number, and nfmpid fields. The *census* and *inventory* tables also have a comment field, allowing users to flag potential data issues and enter notes about missing data or data changed during standardization procedures.

The *inventory_details* table was incorporated into the database to manage the problem of incomplete data. The table stores the number of plots included in a NFMP. The table also records the range of DBH values that were sampled in plots of a specific area for each NFMP inventory recorded. This is especially important as some inventories sampled trees of a smaller DBH than legally required (i.e. less than 30 cm).

3.4.4.3. Database management of taxonomy

The *list* table contains common names and corresponding scientific names as reported in the NFMP. Both the *census* and *inventory* tables link to *list* through the combination of their nfmpid and common_name fields. Beyond NFMP level data, *list* also incorporates the national list of common names and corresponding scientific names (SIREFOR, 2012 *unpubl.*). While both the national and NFMP level lists have undergone taxonomic standardization, the original spelling and species names they report are recorded in the attribute original_sciname.

As more than one hundred species can be found in a single hectare of tropical forest, any census, even when conducted by an experienced botanist, will contain unidentified individuals. When dealing with unidentified individuals in the database, we took a conservative approach. This approach did not assume that unidentified individuals or common names without a corresponding scientific name were consistent across or within management plans. Because an unidentified individual or unidentified common name could possibly represent multiple species, we did not categorize unidentified individuals into specific groups. All individuals unidentified at the species, genus, or family levels are solely listed as unidentified at the appropriate taxonomic level. We included the national list in the table to enable users to review all of the potential species a single common name could represent if the common name was not assigned a scientific name by the NFMP.

The standardized species, genera, and family codes stored in the *list* table form links to the taxonomic tables (*family, genus,* and *species*). As the taxonomy of tropical species is constantly changing, the database, applying a similar model to Condit et al. (2013), incorporates specific attributes and tables to track changes in species names through time. In the *species* table, the attribute 'accepted' flags whether a species name is currently in use or not. The *taxonomy_change* table tracks the history of species name changes in the database. One row of the table stores the currently used species, previously used species, and the date the species was considered outdated in the database.

Species, *genus*, and *family* were placed in separate tables to meet the requirements of database normalization. In doing so, spelling errors were prevented and the assignment of a species to an incorrect family or genus was largely avoided. For example, the record *Pentaclethra macroloba* in the *species* table is dependent on the existence and correct spelling of *Pentaclethra* in the *genus* table. The *genus* table is dependent on the *family* table in the same fashion.

The *wood_density* table is linked to the taxonomic tables and reports wood specific gravity at the species level. Wood density values were incorporated into the database to simplify future studies aiming to calculate aboveground biomass with the NFMP database. The *wood_density* table is largely composed of values extracted from the Global Wood Density database (Chave et al., 2009; Zanne et al., 2009). Additional measurements of wood density were extracted from the scientific literature for species sampled by the NFMPs (e.g. Williamson and Wiemann, 2010). We included all wood density values present in the Global Wood Density database for every family recorded in a NFMP or the national level species list to permit future users to tabulate genus and family level wood density values.

3.5. Outcomes

3.5.1. Data summary

The database contains a total 300,181 records. These records represent 595 censuses, 526 inventories, 768 NFMP centroids (geolocated center point for each NFMP management unit) and 768 NFMP polygons all extracted and digitized from roughly 1000 NFMPs. The number of NFMPs recorded per conservation ranges from 150 to 250, with ACTO having the greatest quantity and ACCVC the fewest (Fig. 3.3a.). A simple analysis of the count of NFMPs generated per year across all five conservation areas reveals a surge in NFMP production during the years of 1997 and 1998 (Fig. 3.3b.). This trend follows that found by Arroyo-Mora et al. (2009) in NFMPs collected entirely from ACCVC. At the tree level, the database consists of 253,923 entries: 154,225 from censuses and 99,698 from inventories. Trees identified in either a census or inventory represent 525 species, 300 genera, and 81 families. A large proportion of the species, genera, and families sampled had a frequency of less than 0.01% in the *census* and/or *inventory* tables (Table 3.1). Overall, 395 of the species were sampled in both the inventory and census while 105 species were sampled only in the inventory and 44 species were sampled only in the census. The most species rich families sampled in the census and/or inventory are *Fabaceae*, *Malvaceae*, and *Lauraceae*. Within the taxonomic tables, 1388 species, 417 genera, and 101 families are stored. Of the 645 NFMPs with either a census or inventory present in the database, 99 were missing an NFMP level list (i.e. the list was not present in the original hardcopy NFMP) to relate the common names found within them to scientific names. Of the remaining 546 NFMPs with an NFMP level list, the percentage of trees identified at the species, genus, and family levels within the *census* and *inventory* tables was examined (Table 3.2). A large improvement in the taxonomic identification of trees is found when moving from the species to the genus level. For example, within the *inventory* table there is a mean increase of 11.35% in the fraction of trees identified when moving from the species (69.58%) to genus (80.93%) level. A comparison of the *census* and the *inventory* results reveals that a larger proportion of trees were identified across all taxonomic levels in the census.

3.5.2. Quality assurance

Quality control within the database was first implemented by setting constraints on the data that can be entered into a table's field. These field constraints include, but are not limited to: unique constraints, limiting a field's data type, checks comparing fields, and coding field values. For example, a coded field was used in the census table to limit tree status values to 'cut', 'remnant', 'parent tree', and 'other.' The field constraints incorporated into the database flag and block the entry of illegal data into the database. Overall, constraints on the census and inventory tables detected an average error rate of 0.18% (±0.53%) in the commercial height and DBH measurements digitized for each NFMP census/inventory.

The second type of quality control occurs after data has already been imported into the database. Quality check queries are run to search for spatial and internal tabular inconsistencies as well as rare values. A full list of the quality check queries performed in the database is given in Table 3.3. Any data issues highlighted by the check queries or field constraints can be compared with the original NFMP digital photos to correct for errors that occurred from the digitization level onwards. In-field errors, already present in the NFMPs, were harder to identify and could not be corrected in the database. Instead, any record identified as a potential error incurred during field collection was flagged. For example, if a tree's DBH is measured as 260 cm, a warning message will be included alongside the record. As buttresses, tree irregularities and lianas can impede the correct measurement of DBH, we conservatively assumed that values of such magnitude are potentially the result of infield errors. DBH measurement errors were identified by a series of rare value queries run on the *inventory* and *census* tables. These queries checked for any tree records with a DBH greater than 200 cm and any inventory and census tree records with a DBH less than the appropriate minimum sampling DBH (standardly 30 cm and 60 cm). Overall, the three aforementioned rare value checks returned 7.16% (approx. 18 ,000 records) of all the records stored in the *census* and *inventory* tables. Within the records returned by these rare value queries, 149 data entry errors were uncovered, indicating a 0.78% error rate in the process of digitizing DBH measurements. Most of the records returned by these queries were instances where trees just below the minimum sampling DBH were recorded (e.g. a tree of a 58 cm DBH was recorded in an NFMP's census).

Internal consistency queries were the most helpful form of check queries, finding otherwise undetectable issues with the data. For example, the number of trees with the status of cut in a NFMP's census was compared to the number of trees requested to harvest within the same NFMP. This query returns a false value when the two values are not equal and indicates by how many trees the values differ (Table 3.4). When run, this query tabulated 356 cases where both numbers were equal, 174 cases where values were inconsistent and 65 cases where no results were returned due to missing data. Of the 174 cases where a discrepancy existed, the difference between the counts reported in the tables ranged from one to 327 trees, with the largest frequency of cases (38) differing by only one tree. For each incident with a difference greater than one, the cause of the inconsistency was investigated. This process revealed a range of potential sources of data error which are listed in (Table 3.5). While some types of errors found with this query had little impact on overall data quality, other, more far-reaching errors were identified and corrected. Not digitizing census pages, for example, not only causes inconsistency between the two tables, but also represents a significant omission of tree data. If not corrected, this lack of data quality could impact all analyses that use that NFMP or the tree data within it.

3.5.3. Data exploration

To illustrate the utility of the database a simple analysis of relative species abundance was conducted by querying the database. The results of the analysis revealed the three most abundant species recorded in the *census* or *inventory* tables at the conservation area level (Table 3.6 and Table 3.7). In the *inventory* table, the three most abundant species sampled represent 12.86-42.81% of the trees sampled per conservation area. In the *census* table, the three most abundant species per conservation area represent 20.80-50.73% of the trees sampled per conservation area. In three of the five conservation areas (ACCVC, ACLA-C, ACTO), *P. macroloba* was the most abundant species sampled in both the census and inventory. Of these conservation areas, the largest relative abundance values were found in ACCVC where *P. macroloba* had a relative abundance 0.3317 in the inventory and 0.3595 in the census. Contrastingly, in ACAHN, the most abundant species sampled differed in the inventory and census. Here, *Vochysia ferruginea* was the most abundant species within the census data and *Dialium guianense* was the most abundant species within the inventory data. A comparison of the inventory and census results reveals that differences in species composition are mainly due to the inclusion and reduction in rates of identification of trees in the 30-60 cm DBH range (i.e. smaller trees) in the *inventory* table. Additionally, the magnitude of relative abundance values was greater in the *census*, indicating a greater dominance of particular tree species within larger DBH classes (greater than or equal to 60 cm). In ACOSA, two of the most abundant species (*Brosimum utile* and *Qualea paraensis*) are not present within the top three most abundant species sampled in the other four conservation areas. These results suggest that ACOSA differs the greatest in species composition from the other areas studies. At a conservation area level, the greatest number of species was recorded in census data collected from ACLA-C (294) while the greatest number of genera (201) and families (66) were recorded in census data collected from ACAHN (Table 3.8). Within the inventory table, the greatest number of species (311), genera (164), and families (54) were sampled in ACLA-C (Table 3.8). Despite the variation of total area sampled (sampling effort) between conservation areas, these results indicate how many species, genera, and families were captured in the NFMP data collected from the five different areas.

3.6. Discussion

The design and development of the FGIS geodatabase was tailored to a heterogeneous dataset collected over 30 years across Costa Rica. During the design stages, a number of procedures were developed in order to standardize and efficiently store the dataset. It was particularly important, for example, that scientific names, spellings, and taxonomic standards were made consistent and brought up to date. Additionally, the geographic projections, the statuses of the trees in the census, and the total area sampled per NFMP were all tracked and standardized within the database. The completion of tasks such as these allows the data to be compared and more efficiently analyzed across all NFMPs. The database's analytical functionality is increased further by its capacity to store spatial and tabular datasets. Spatial data is especially valuable in this case, as it allows the data to be utilized in larger scale studies, particularly those that incorporate remote sensing.

A key asset of the database is the inclusion of multiple quality assurance procedures (Le Duc et al., 2007; Peacock et al., 2007; Condit et al., 2013). These procedures reduce the effort expended during data entry/use and improve overall data quality. The quality control system evaluates the entire dataset, finding errors that can be corrected by referencing the original NFMP documents. Further, the quality checks identify errors that are present within the original NFMPs such as missing census pages or inconsistencies between reported numbers. The recognition and flagging of such issues within the database greatly improves the reliability of results from future analyses. Quality checks also help to estimate the rate and variability of human error that occurred during the various stages of data collection (i.e. digitization and data entry) and flag potential in-field measurement errors.

A caveat of using forest management data not addressed by these queries, however, is the reliability of species identification. This is especially pertinent in the tropics where species diversity is high and vernacular names are often used during field data collection (Lam and Klein, 2008; Lacerda and Nimmo, 2010). The translation of these common names to scientific names is error prone as common names can be site, community, or regionally specific and can represent multiple different species (Lacerda and Nimmo, 2010). Additionally, the local experts responsible for species identification in forest management can vary in their level of expertise and can perform differently based on the working conditions (Réjou-Méchain et al., 2011a). As floristic identification is perhaps one of the most significant sources of error in Costa Rican forest management practices, it is subsequently a significant source of error in the FGIS geodatabase. A potential approach to address this issue is the solicitation of expert knowledge for the development of an index on the degree of confidence in floristic identification for each species, genus, and family (Réjou-Méchain et al., 2011a). Future work should also consider the spatial distribution of species recorded in the management plans to locate potential outliers or errors incurred during the translation of the common names to scientific names in different regions. Although there may be issues with the identification of trees in management plans (Lacerda and Nimmo, 2010), our study is a first attempt to organize this massive amount of data.

Over a quarter of a million records can now be queried and analyzed as a result of the development of the FGIS database. Both spatial and temporal trends in forest management can be investigated at a variety of scales. Analyses also reveal potential strengths and weaknesses in the dataset. For example, across all conservation areas, a great improvement in the fraction of trees identified is seen when moving from the species to genus level. Réjou-Méchain et al. (2011a) reported similar findings for commercial inventories conducted in the Central African Republic. In tropical forests, patterns of biodiversity have been studied at higher taxonomic levels to decrease the amount of data errors and noise (Gaston and Williams, 1993; La Torre-Cuadros et al., 2007). As numerous identification errors are the result of confusion between species of the same genus, working at the genus level specifically (e.g. Prinzing et al., 2003; Villaseñor et al., 2005) decreases the overall floristic identification errors. Considering this, analysis at the genus level is a favorable approach to take with the NFMP data. Analyses also expose notable differences between inventory and census data. While inventories capture a more unique set of tree species, censuses are able to identify a greater percentage of the forest stand. Abundance analyses look at how dominant tree species are within a conservation area using inventory and census data. As commercial inventories more accurately identify large trees and common species than small trees and rare species (Réjou-Méchain et al., 2011a), confirmed by our comparison of the inventory and census identification results, these abundance analyses provide insight into the fraction of trees reliably identified. For example, in ACTO, the three most abundant species in censuses represent over 50% of the trees recorded. Keeping in mind that abundant species are more accurately identified, it is conceivable that over 50% of census tree records are reliably identified to the species level in ACTO.

The abundance results driven by the NFMP data are consistent with ter Steege et al. (2013) who showed that, even in hyper diverse forests, there is a hyperdominace of a few species. Further, the NFMP based abundance analysis produced similar results to those reported for Costa Rica in the scientific literature. In Osa Peninsula (ACOSA), studies confirm that *B. utile* is an abundant species (Herwitz, 1981; Cleveland et al., 2004; Huber et al., 2008). In Eastern Costa Rica (ACCVC, ACLA-C, ACTO, ACAHN), multiple studies validate the findings that *P. macroloba, Carapa nicaraguensis* and *V. ferruginea* are dominant canopy species (Herrera and Finegan, 1997; Webb and Peralta, 1998; Webb, 1999). The correspondence of the findings of ecological studies to those found with the commercial logging data emphasizes the potential to exploit the NFMP data in ecological studies. The count of species, genera, and families sampled within and across the conservation areas highlights the biodiversity captured by the logging inventories. A notable proportion of the taxonomic data within the tree records, however, was of low frequency. With this in mind, the absolute counts should be viewed with discretion as less abundant species are more likely to be inaccurately identified in the field (Hanazaki et al., 2007; Kenfack et al., 2007; Réjou-Méchain et al., 2011a).

3.7. Conclusions

In this study we present the most complete digital record of natural forest management in Costa Rica to date, the FGIS geodatabase. Due to the private and detailed nature of the data contained in the FGIS geodatabase it is only available for use by assigned project collaborators and is not publicly accessible. This tool will permit an improved comprehension of carbon stocks and species environment relationships at a scale necessary to address the conservation and forest management problems of today. For example, with the database, the estimation of aboveground carbon stocks within five conservation areas will be made possible, helping to create a national baseline forest carbon stock for the country based on estimated biomass (Svob et al., 2014 *submitted*) from forest management data. This analysis has the potential to improve Costa Rica's REDD+ program (Reducing Emissions for Deforestation and Forest Degradation). The REDD+ program was designed to provide financial incentives to developing countries that can show lower GHG emissions from forest degradation and deforestation (Gibbs et al., 2007), where sustainable forest management is a key component.

The geodatabase will also provide a means to assess species distribution, biodiversity, and species environment interactions at a landscape scale (Arroyo-Mora et al., 2009). With the NFMPs in a standardized digital format, the database may serve also as a tool to evaluate the historical trends and sustainability of natural forest management in Costa Rica. This is a key achievement of the database because, as highlighted by Bradshaw et al. (2009), the sound management of human-modified landscapes is becoming increasingly more essential to the survival of tropical forest biodiversity. This is particularly pertinent in Costa Rica where, despite the country's recognition for its environmental policies, the sustainability of its NFMP framework has been subject to little testing and validation (McGinley and Finegan, 2003). Finally, we provide a framework for geodatabase construction that can be used in other tropical countries where natural forest management data is available

3.8. Tables and Figures

Table 3.1. The families, genera, and species sampled in NFMPs. The values reported are a count of unique families, genera, or species sampled within the *inventory*, *census*, or both tables as: (1) total number and (2) the number of those representing less than 0.01% of the trees sampled.

	Family		Genus		Species	
	Total	< 0.01% tree records	Total	< 0.01% tree records	Total	< 0.01% tree records
Inventory	66	12	229	60	395	158
Census	80	13	292	75	500	174
Total*	81	22	300	105	525	244

*Total includes both the census and inventory data

Table 3.2. The percentage of trees identified per NFMP in the *inventory* and *census* tables at different taxonomic levels.

	Trees identified (%)	
Taxonomic level	Census (mean±SD)	Inventory (mean±SD)
Family	89.95±15.37	81.21±20.43
Genus	89.73±15.38	80.93±20.56
Species	80.94±17.93	69.58±20.53

Table 3.3. List of all quality check queries performed in the database. The tables listed represent the tables over which the query is being run. The name of each query is enumerated and in italics. The column 'Type' refers to whether a query was a spatial (SP), rare value (RV), or internal consistency (IC) check query.

Name Table(s)	Description	Type
1. NFMP polygon within conservation area		
nfmp_polygon; nfmp_general; ca_polygon	Checks whether an NFMP's polygon falls within the bounds of the appropriate conservation area polygon as specified in nfmp_general.	SP
2. NFMP polygon within district/canton/province		
nfmp_polygon; nfmp_general distrito_polygon; canton_ polygon; province_ polygon	Checks whether an NFMP's polygon falls within the bounds of the appropriate province, canton, and district as specified by the district listed in the nfmp_general table.	SP
3. Census dbh greater than 200 cm		
census	Returns census records with a dbh greater than 200 cm.	RV
4. Census dbh less than 60 cm		
census	Returns census records with a dbh less than 60 cm.	RV
5. Inventory greater than 200 cm		
inventory	Returns inventory records with a dbh greater than 200 cm.	RV
6. Inventory dbh less than minimum specified for sampling		
inventory; inventory_details	Returns records with a dbh less than the minimum dbh specified for collection in the inventory of a NFMP.	RV
7. Elevation greater than 300 m		
nfmp_general	Returns NFMP records with an elevation attribute value (max, min, or avg) greater than 300 m.	RV
8. NFMP duration less than 15 years		
nfmp_general	Returns NFMP records where the duration based on the start and end years is less than 15 years.	RV
9. Scientific names in list up to date		
list; taxonomy_change	Returns species from the list table that are outdated and indicates the currently accepted name for the given species.	IC
10. Census trees to be logged and requested trees to log consistent		
census; nfmp_general	For a given NFMP, the number of trees categorized as 'to log' in a census is compared to the number of trees requested to log.	IC
11. Inventory plot count consistent		
inventory, inventory_details	Checks whether the number of unique plots recorded in *inventory* for a single NFMP is consistent with the number of plots specified for the same NFMP in inventory_details.	IC
12. Inventory present		
inventory; nfmp_general	Checks whether an NFMP's inventory recorded as present in nfmp_general is present in *inventory*.	IC
13. Census present		
census; nfmp_general	Checks whether an NFMP's census recorded as present in nfmp_general table is present in the *census*.	IC
14. List present		
list; nfmp_general	Checks whether an NFMP's list recorded as present in nfmp_general is present in the list table.	IC
15. NFMP polygon present		
nfmp_polygon; nfmp_general	Checks whether an NFMP's polygon recorded as present in NFMP in nfmp_general table is present in the nfmp_polygon.	IC

Table 3.4. Example of output from query checking for an internally consistent number of logged trees. The 'trees requested to log' is reported in the *nfmp_general* table. The 'census C tree' is the count of trees reported with a status of 'cut' per NFMP in the *census* table. F (false) in the "values consistent" column indicates an inconsistency between the *census* and *nfmp_general* tables with regards to the number of trees to be logged. The final column indicates by exactly how many trees the two tables differ.

NFMPID	Trees requested to log	Census C trees	Values consistent	Difference of trees to log
orcave1998	84	83	F	1
oresem2003	223	223	T	0
osfear1997	55	56	F	1
osfego1993	190	190	T	0
oshiri1994	25	23	F	2
osroro1998	97	97	T	0
otavar1999	69	69	T	0

Table 3.5. Types of errors detected with quality check query comparing *nfmp_general* and *census* tables. In the "error(s) correctable" column 'T' signifies that the error is correctable and 'F' signifies the error is not correctable. The "impact level" column rates the error type by its level of impact on overall data quality with 'High' having the largest impact and 'Low' having the smallest.

Type of Error	NFMP (count)	Error rate* (%)	Error(s) correctable	Impact level
Duplicate import	8	1.51	T	High
Page(s) not digitized	23	4.34	T	High
Status entry error	31	5.85	T	Low
Error in *nfmp_general* table	31	5.85	T	Low
1-5 trees not entered	4	0.75	T	Low
Difference of 1 tree**	38	7.17	T	Low
Page(s) missing in NFMP	9	1.70	F	Med/High
NFMP 'C' trees unclear	11	2.07	T/F	Low
Within NFMP, numbers Inconsistent	28	5.28	F	Low

*Error rate within the NFMPs cross-checked with original NFMP photos
**Differences of 1 tree could fall into many of the error types listed above but was categorized separately due to its very low impact on data quality.

Table 3.6. The relative abundance of the three most abundant species sampled by inventories categorized per conservation. The relative abundance values are given in brackets below the species name. Relative abundance values range from 0 to 1.

	Conservation Area Species (Relative Abundance)				
Rank	ACAHN	ACCVC	ACLA-C	ACOSA	ACTO
1	*Dialium guianense* (0.0711)	*Pentaclethra macroloba* (0.3317)	*Pentaclethra macroloba* (0.0822)	*Qualea paraensis* (0.0437)	*Pentaclethra macroloba* (0.3138)
2	*Pentaclethra macroloba* (0.0638)	*Carapa nicaraguensis* (0.0341)	*Virola koschnyi* (0.0427)	*Carapa nicaraguensis* (0.0427)	*Carapa nicaraguensis* (0.0918)
3	*Vochysia ferruginea* (0.0436)	*Vochysia ferruginea* (0.0316)	*Prioria copaifera* (0.0389)	*Brosimum utile* (0.0422)	*Pterocarpus hayesii* (0.0225)
Total*	0.1785	0.3974	0.1638	0.1286	0.4281

*The total is the sum of the relative abundance of the three most abundant species.

Table 3.7. The relative abundance of three most abundant species sampled by censuses categorized per conservation area. Relative abundance values range from 0 to 1. The relative abundance values are given in brackets below the species name.

	Conservation Area *Species* (Relative Abundance)				
Rank	ACAHN	ACCVC	ACLA-C	ACOSA	ACTO
1	*Vochysia ferruginea* (0.1249)	*Pentaclethra macroloba* (0.3595)	*Pentaclethra macroloba* (0.0981)	*Brosimum utile* (0.0775)	*Pentaclethra macroloba* (0.3085)
2	*Carapa nicaraguensis* (0.1061)	*Vochysia ferruginea* (0.0758)	*Carapa guianense* (0.0923)	*Qualea paraensis* (0.0713)	*Carapa nicaraguensis* (0.1738)
3	*Dialium guianense* (0.0889)	*Carapa nicaraguensis* (0.0632)	*Prioria copaifera* (0.0574)	*Carapa nicaraguensis* (0.0592)	*Virola koschnyi* (0.0250)
Total*	0.3199	0.4985	0.2478	0.2080	0.5073

*The total is the sum of the relative abundance of the three most abundant species

Table 3.8. Number of families, genera and species sampled by census and inventories conducted within each conservation area. The values represent the count of unique families, genera, and species sampled by censuses and inventories conducted within each conservation area.

	Census				Inventory			
CA	Family Count	Genera Count	Species Count	Area Sampled (km^2)	Family Count	Genera Count	Species Count	Area Sampled (km^2)
ACAHN	66	201	275	29.2	53	138	227	3.4
ACCVC	53	155	211	21.8	47	131	214	1.6
ACLAC	62	194	294	30.3	54	164	311	2.9
ACOSA	52	157	227	23.1	43	140	247	2.3
ACTO	53	133	173	19.0	46	104	181	1.9
Total	80	292	500	123.4	66	229	395	12.2

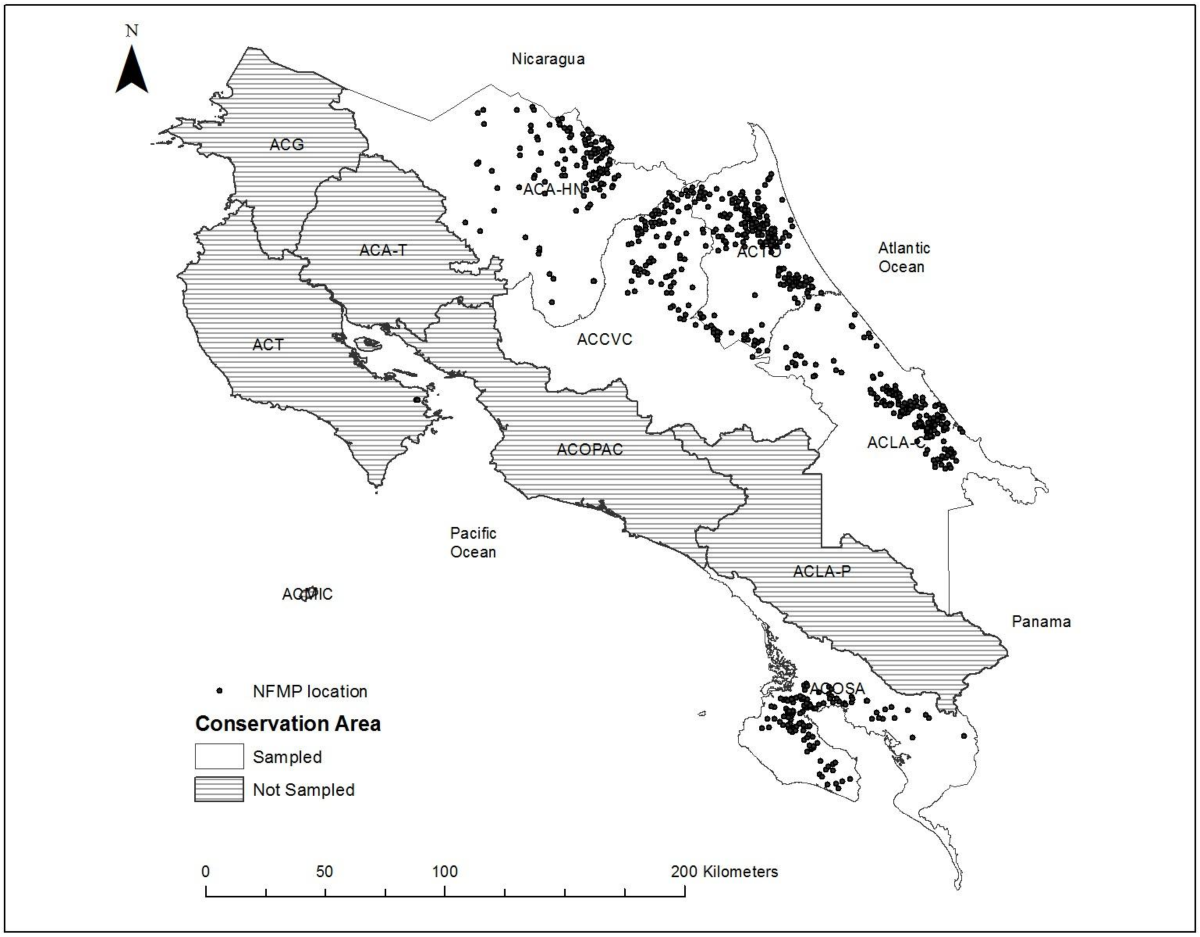

Figure 3.1. Map of Costa Rica and the conservation area system. Conservation areas where NFMPs were recorded are shown in white. The dark dots represent the NFMPs for which location data was available.

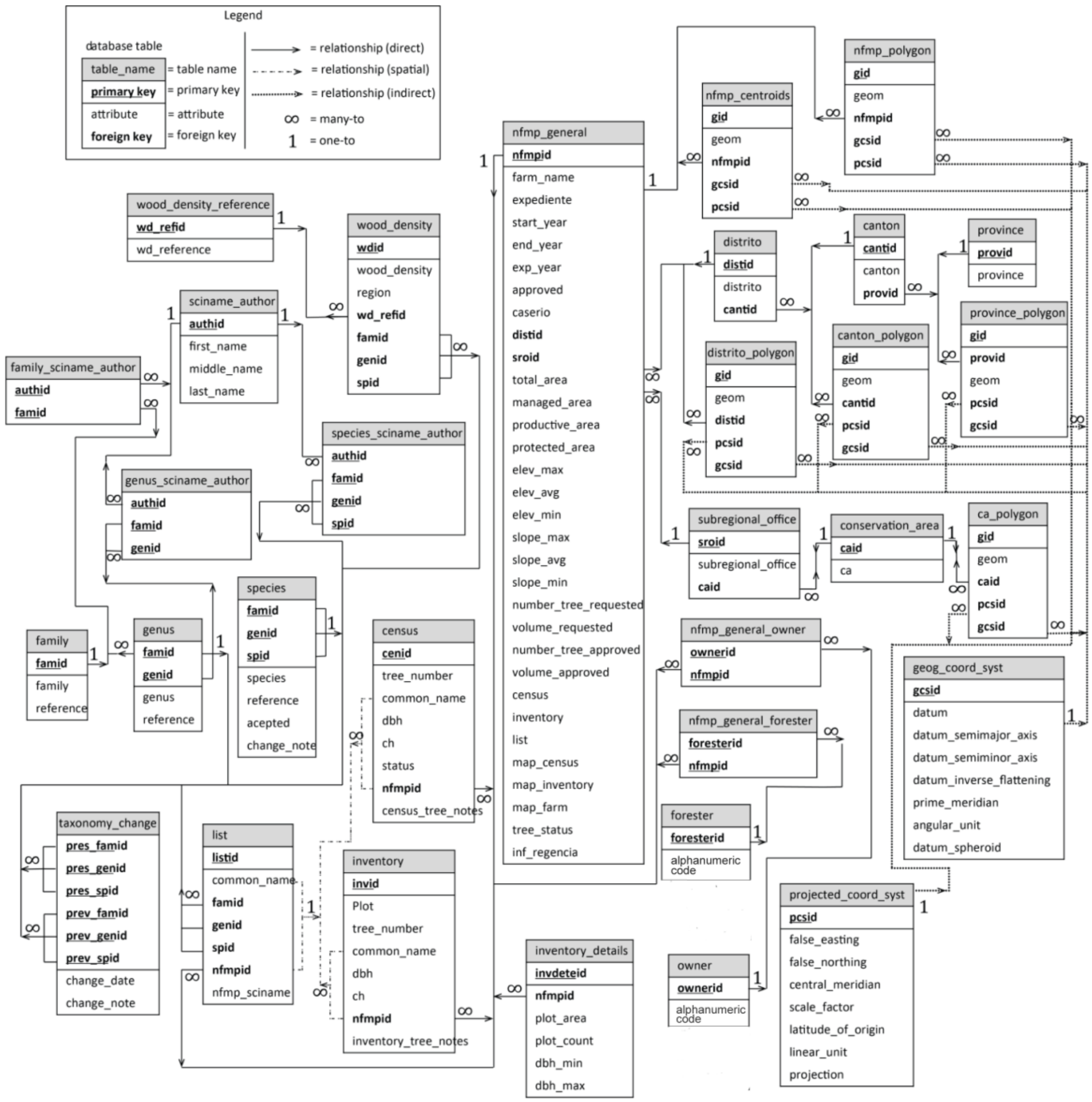

Figure 3.2. Schema of the relational structure of the FGIS geodatabase. The **primary key(s)** of each table is in bold and underlined. Attributes only in bold represent **foreign keys**. The symbol '1' represents a 'one to' relationship while the '∞' symbol represents a 'many-to' relationship. All solid lines depict a direct relationship between tables via foreign and primary keys. The black dashed lines linking the *census* and *inventory* tables to the *list* table represents an indirect relationship (i.e. the relationship is not through the use of primary/foreign key). The black dotted lines indicate the link between tables containing spatial data and the projected and geographic coordinate systems tables.

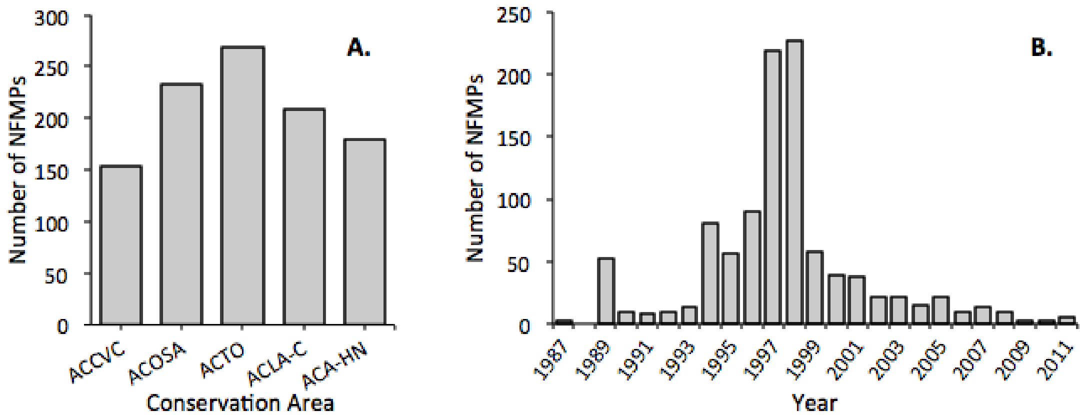

Figure 3.3. The general trends in quantities of NFMPs collected per (A) conservation area and (B) year. The years represent the year the NFMP was submitted to the sub-regional office. It is important to note that the NFMPs counted in this figure include those with missing data components (e.g. a NFMP without a census would have been included in the count).

4. A wood density and aboveground biomass variability assessment using pre-felling inventory data in Costa Rica

Citation: Svob, S., Arroyo-Mora, J.P. and Kalacska, M. 2014. A wood density and aboveground biomass variability assessment using pre-felling inventory data in Costa Rica. *submitted.*

4.1. Context within thesis

In this chapter, I use the FGIS database developed and explored in Chapter 3 to assess the variability of estimated aboveground biomass both among and within five Costa Rican conservation areas. The FGIS database provides the backbone dataset and tool for this study, which would not have been possible without the standardization, quality control, and logical structuring of the large forest management dataset. This study compares the mean and variance of estimated aboveground biomass among conservation areas, assesses the spatial variability of AGB across regions, investigates the relationship between estimated aboveground biomass and wood density, and finally, evaluates the uncertainty associated with the estimation of aboveground biomass using natural forest management data. Here, I utilize published pantropical allometric models to relate a measurement of diameter at breast height to an estimate of aboveground biomass. Further, by relating standardized scientific names to natural forest management tree data with the FGIS database, I am able to incorporate species and genus specific wood density values into my analysis of aboveground biomass.

4.2. Abstract

Background: The high spatio-temporal variability of aboveground biomass (AGB) in tropical forests is a large source of uncertainty in the estimation of forest carbon stocks. Due to their spatial distribution and sampling intensity, pre-felling inventories are a potential source of ground level data that could help reduce this uncertainty at larger spatial scales. Further, exploring the factors known to influence tropical forest biomass, such as wood density and large tree density, will improve our nowledge of biomass distribution across tropical regions. Here, we evaluate (1) the variability of wood density and (2) the variability of AGB across five ecosystems of Costa Rica.

Results: We found that the mean wood density of trees with a diameter at breast height (DBH) greater than 30 cm was highest, 0.623±0.182 g cm^{-3} (mean±SD) in the most northern region of Costa Rica studied, Huetar Norte. Using forest management or pre-felling inventories, we determined that the region with the highest estimated AGB (DBH≥30cm) was Costa Rica's Osa peninsula (173.47±60.23 Mg ha^{-1}). The density of large trees explained approximately 50% of the variability of estimated AGB across the five ecosystems studied. Comparing the AGB estimates of our study to published estimates reveals that, in the regions of Costa Rica where AGB has been previously sampled, our forest management data based approach produced a similar range of values.

Conclusions: This study presents the most spatially rich analysis of ground level AGB data in Costa Rica to date. Using data from pre-felling inventories from forest management plans, we found that the estimated AGB within and among five Costa Rican ecosystems is highly variable. Combining commercial logging inventories with ecological plots will provide a more representative ground level dataset for the calibration of the models and remotely sensed data used to estimate AGB at regional and national scales. Additionally, because it is the non-protected areas of the tropics that offer the greatest opportunity to reduce rates of deforestation and forest degradation, logging inventories offer a promising source of data to support mechanisms such as the United Nations REDD+ (Reducing Emissions from Tropical Deforestation and Degradation) program.

4.2.1 Keywords: Forest Management, Aboveground Biomass, Wood Density, Tropical Forest, Costa Rica

4.3 Introduction

Tropical forests play a vital role in regulating the Earth's climate through the processes of evapotranspiration and CO_2 uptake. While these areas represent only

7% of global land cover (Bradshaw et al., 2009), they store roughly 55% of the world's forest carbon stock (Pan et al., 2011). Tropical forests are among the most carbon dense ecosystems (242 Mg C ha^{-1}) in the world (Pan et al., 2011). Approximately 56% (193-223 Pg C) of their carbon is stored in the form of biomass alone (Pan et al., 2011; Saatchi et al., 2011; Baccini et al., 2012). During the 1990s and early 2000s, a substantial portion of this carbon stock suffered due to deforestation, which reached an estimated rate of $12.9 \cdot 10^6$ ha yr^{-1} (Malhi, 2000). The deforestation and degradation of tropical forests is also the second largest source of anthropogenic CO_2 emissions (Van de Werf et al., 2009), releasing carbon at an estimated net rate of 1.0 Pg yr^{-1}between 2000 and 2010 (Baccini et al., 2012).

The United Nations REDD+ (Reducing Emissions from Deforestation and Forest Degradation) program is an innovative global mechanism that aims to provide monetary benefits to developing tropical countries that can show an increase in forest carbon stocks from an established national baseline (Gibbs et al., 2007). In the past decade, the number of studies seeking to improve the methods and data used to accurately estimate the spatio-temporal variation of tropical forest carbon stocks, supporting REDD+, have substantially increased (Petrokofsky et al., 2011). Today, much of this research relies upon the relationship between aboveground biomass (AGB) and forest carbon. Despite recent efforts to estimate AGB in the tropics (and in turn carbon) (refer to Saatchi et al., 2001; Baccini et al., 2012; Asner et al., 2013), a large degree of uncertainty in the spatial distribution and accuracy of these estimates remains (Clark and Kellner, 2012; Mitchard et al., 2013). One of the key factors in reducing uncertainty in AGB estimates is using a spatial scale fine enough to capture variability across the landscape.

Remote sensing and ground data (i.e. forest inventories) are two techniques that have been proposed for the production of reliable carbon estimates (e.g. Hill et al., 2013). Remote sensing is an advantageous approach as it can provide wall-to-wall coverage of an entire country. Remotely sensed data, however, must be calibrated/validated with ground truth measurements (Drake et al., 2002; Patenaude et al., 2005). Additionally, remote sensing instruments may not be sensitive enough to detect the variability of biomass within and across the high-density forest stands

(Patenaude et al., 2005; Goetz et al., 2009) typical of tropical moist and tropical wet ecosystems. Ground data collected for scientific research (ecological data), is the most common data source employed to estimate AGB due to its high level of detail and systematic nature. Nevertheless, ecological data has its own weaknesses when estimating AGB such as: (1) the standard plot size of 0.1 ha (Brown and Lugo, 1992) is too small to capture AGB variability (Clark and Clark, 2000); (2) plots are sometimes biased towards high density (ideal) forest locations (Brown and Lugo, 1992); and (3) plots cover only a small fraction of a country's total forested area (Maniatis et al., 2011). Commercial logging inventories may provide a solution to these problems due to their spatial distribution and sampling intensity (Couteron et al., 2003; Maniatis et al., 2011). Logging inventories are common to tropical countries and represent a large source of data on forest structure and composition (Putz et al., 2011). With their success in measuring diversity on large spatial scales (Réjou-Méchain et al., 2011a), determining ecological factors that influence forest structure (Couteron et al., 2003), and estimating emission factors under REDD+ (Maniatis et al., 2011), logging inventories may provide a suitable source of forest data, complementing ecological datasets and helping to estimate baseline carbon stocks.

In Costa Rica, selective logging inventories or pre-felling inventories (we use this terminology in this paper indistinctively) are developed under a Natural Forest Management Plan (NFMP) framework (Arroyo-Mora et al., 2009). NFMP data is available for most of the country's ecosystems below an elevation of 300 meters, accurately representing the heterogeneity of the Costa Rica's lowland landscape The country is divided into 11 conservation areas (CAs) (Boza, 1993), each encompassing distinct forest ecosystems. Despite the country's small size (approximately 51,000 km^2), it contains a rich diversity of tropical ecosystems ranging from dry to wet forests.

A NFMP is required before the forest stand of a privately owned property can be selectively logged. In order to be legally approved to log, the owner must hire a certified forester to conduct an inventory and census of the forest stand (Arroyo-Mora et al., 2009). In an inventory, every tree with a diameter at breast height (DBH) equal to or greater than 30 cm is measured and identified in plots of 0.3 ha. For the same forest stand, a census is carried out to measure and identify every tree with a DBH equal to or greater than 60 cm. For this study we use a standardized relational geodatabase encompassing Costa Rican pre-felling inventory data (Svob et al. 2014a *in press*).

The main objective of this study is to assess to variability of wood density and estimated AGB across five ecosystems in Costa Rica. Wood density is an important predictive variable when estimating AGB (Brown and Lugo, 1992; Chave et al., 2005; Keeling and Phillips, 2007; Baker et al., 2009). As wood density is known to vary among different forest communities (Baker et al., 2004; Muller-Landau, 2004; Chave et al., 2009; Zhang et al., 2011), this variable is also critical to studying the differences in AGB across a landscape. Despite these findings, wood density has yet to be studied or implemented when estimating AGB across Costa Rica. Further, although our analysis is based on medium to large trees (30 and 60 cm DBH), studies have shown that large trees constitute a disproportionate fraction of AGB and drive the variations in biomass across the tropics (Slik et al., 2013). Therefore, despite a lack of tree data below the 30 cm DBH range, patterns of AGB variability may be discernible from our NFMP dataset. A standard method to estimate a tree's biomass employs an allometric equation to relate measurements on DBH to units of biomass. The choice of allometric model is critical and should be based upon both the aim of the study (Baker et al., 2004) and the characteristics of the dataset (Chave et al., 2004). Allometric models should be representative of the DBH range and ecosystem being studied (Chave et al. 2004). Additionally, to allow for regional scale comparisons, AGB estimates must be based on a consistent regression approach to avoid the confounding of results by variations inherent in different models (Baker et al., 2004).

In recent years, the pantropical allometric models developed by Chave et al. (2005) have been widely applied across the globe to estimate AGB. In Costa Rica, however, numerous studies of biomass have employed Brown’s (1997) equation for wet forests (e.g. Clark and Clark, 2000; Letcher et al., 2009; Clark et al., 2011; Saatchi et al., 2001) as it was calibrated with data collected at Costa Rica's La Selva

Biological Station. For this study, we believe Brown's (1997) equation has many disadvantages when compared to those developed by Chave et al. (2005). These shortcomings include: (1) the representation of a smaller DBH range; (2) the development of the equation from a smaller sample size; (3) the limited application of the equation outside of Costa Rica (making the comparison of AGB estimates with other countries/studies more complex); and (4) the absence of wood density as a parameter which is an aspect of forest structure that varies significantly at regional scales (Baker et al., 2004; Chave et al., 2009).

With the use of allometric models (Chave et al., 2005), the Global Wood Density database (Chave et al., 2009; Zanne et al., 2009), the pre-felling inventory database, and national measurements of wood density found in the scientific literature (e.g. Williamson and Wiemann, 2010), this study will first evaluate the variability in wood density and second assess the variability in estimated AGB across five ecosystems in Costa Rica. Specifically, our study uses a NFMP database for five conservation areas to address the following questions: (1) What are the patterns of wood density variability at the CA-level and between data produced by the census and inventory (i.e. sampling protocols)?; (2) What is the variability of estimated AGB within and among CAs?; (3) Do estimated AGB values differ between the two sampling protocols? (4) What is the uncertainty associated with AGB estimated using natural forest management data? As ground level data from pre- felling inventories covers a greater area than ecological plots within the five ecosystems being studied, our study will better capture the spatial heterogeneity of wood density and estimated AGB across the landscape. Through this analysis, we can enhance our understanding of the spatial distribution of estimated AGB and, in combination with both ecological and remotely sensed data, more reliably map and estimate national forest carbon stocks.

4.4 Methods

4.4.1 Study area and data

This study used a database of NFMPs from five conservation areas: ACLA-C (Caribbean La Amistad Conservation Area), ACAHN (Arenal Huetar Norte

Conservation Area), ACTO (Tortuguero Conservation Area), ACCVC (Central Volcanic Conservation Area), and ACOSA (Osa Conservation Area) (Figure 4.1). These conservation areas cover the country's Atlantic lowland forests, northern lowlands, and central and south Pacific forests, encompassing the regions where selective logging has been most heavily practiced. All five conservation areas include natural forest management plans that fall within the tropical wet (4000-8000 mm precipitation yr^{-1}) and/or rain (>8000 mm precipitation yr^{-1}) forest lifezones (defined by Holdridge 1979). Only ACLA-C and ACAHN include natural forest management plans that represent the tropical moist (2000-4000 mm precipitation yr^{-1}) forest lifezone. The management plans sampled largely represent a lowland ecosystem (0-500 m a.s.l.[3]) although a small subset of the data falls within the transition zone from lowland to premontane (500-1500 m a.s.l.) forest. Natural forest management plan forest type was classified using the Life Zone System Map from the Atlas Costa Rica 2008 (Instituto Tecnológico de Costa Rica, 2008). All management plans were carried out in primary forest.

4.4.2 The variability of wood specific gravity among CAs and sampling protocols

The variability of wood density across Costa Rica can be illustrated by differences found in wood density between conservation areas, NFMPs, and the different sampling protocols (census and inventory). The greater the variability in wood density, the more important it becomes to include this parameter when producing AGB estimates comparable at a landscape-scale. To carry out the analysis, we excluded data from NFMPs with less than 80% of their trees identified to the species or genus level. The wood density value for each tree in a NFMP was selected in decreasing order of preference from (1) a species-level average, (2) a genus-level average, and (3) a NFMP-level average. Mean NFMP wood density was calculated separately for each census and inventory. These NFMP averages were determined by summing the wood density of all stems with a species or genus level value. Differences between the wood density of CAs and the two different sampling protocols were tested with a one-way ANOVA and a multiple comparisons procedure

[3] In text, a.s.l. refers to above sea level.

using the grouping variables conservation area and sampling protocol. All conservation area level statistical analyses only included tree data that met the aforementioned NFMP taxonomic identification conditions. All of the analyses were carried out in MATLAB version R2013b (The MathWorks Inc., Natick, MA, USA) unless otherwise specified.

4.4.3 Estimating tree-level AGB

In this study, we denote estimated AGB as estimated aboveground biomass (EAGB), following Clark and Kellner (2012). To determine EAGB per tree, Chave et al.'s (2005) allometric models for wet and moist forests were applied. The EAGB of wet and rain forest lifezones was estimated by applying the wet forest equation:

$$EAGB = \rho\ x \exp\left(-1.239 + 1.980\ x\ \ln(\mathrm{DBH}) + 0.207\ x\ \ln(DBH)^2 - 0.0281\ x\ \ln(DBH)^3\right) \quad \text{(Equation 4.1)}$$

Correspondingly, the EAGB of forests within the tropical moist lifezones was estimated using the moist forest equation:

$$EAGB = \rho\ x \exp\left(-1.499 + 2.148\ x\ \ln(\mathrm{DBH}) + 0.207\ x\ \ln(DBH)^2 - 0.0281\ x\ \ln(DBH)^3\right) \quad \text{(Equation 4.2)}$$

where ρ is wood density in g cm^{-3}, DBH is in cm, and EAGB is given in kg $tree^{-1}$. Wood density values were selected in decreasing order of preference from (1) a species-level average, (2) a genus-level average, (3) a NFMP-level average, and (4) a conservation area level average. A large portion of the variation in wood density is captured at the genus-level, making mean genus wood density the second best option when estimating EAGB (Baker et al., 2004; Chave et al., 2009). We use the mean NFMP wood density for a tree that was present in a NFMP with at least 80% of its trees identified to the genus or species level and was missing a corresponding species or genus level wood density. For a tree that was reported in a NFMP with less than 80% of its trees identified to the species or genus level and lacked wood density information, we used a conservation area level mean wood density.

4.4.4 Estimating census and inventory AGB

To calculate EAGB per unit area of a census, the EAGB values per tree with a DBH greater than or equal to 60 cm were summed and then divided by the NFMP's productive area of the forest stand. The productive area is the total area of the forest stand sampled by a census. To calculate EAGB per unit area of an inventory, the EAGB values per tree with a DBH greater than or equal to 30 cm were summed across each plot and then divided by 0.3 ha (area of the plot). Finally, EAGB was averaged across all plots within a given NFMP.

An outlier analysis of all of the resulting EAGB values was performed, pinpointing cases where EAGB did not fall between the 1.5 lower and 1.5 upper interquartile range. All outliers were cross-checked with the original hardcopy versions of the NFMPs. If the values were the result of uncorrectable errors present in the original NFMPs, they were excluded from any further analyses.

To assess the amount of spatial autocorrelation among the NFMPs sampled, we evaluated the distribution of Moran's I with a spatial correlogram as applied in SAM version 4.0. Spatial correlograms indicate the correlation between pairs of spatial observations as the distance between them is increased (Rangel et al., 2006). As the values of Moran's I were relatively small, ranging between 0.154 and -0.209, we did not include any additional methodological approaches to account for spatial autocorrelation in later analyses.

4.4.5 Comparing the EAGB of the sampling protocols and CAs

Differences in EAGB between conservation areas and between sampling protocols were tested with a one-way ANOVA and a multiple comparisons procedure using CA and sampling protocol as the grouping variables. We also evaluated differences in the density of large trees (DBH≥70cm) between CAs and sampling protocols, as the recent study of Slik et al. (2013) found that large trees explained up to 70% of the variation in EAGB across the tropics.

In order to compare the data captured by inventories and censuses more directly, EAGB was recalculated for each inventory including only the trees that would be sampled during a census (DBH≥60cm). To assess whether there was a significant difference in this data, a paired t-test was applied. As multiple NFMPs only included census data (DBH≥60cm), we attempted to develop a model to estimate the EAGB of stems with a with a DBH≥30cm and a DBH<60 cm. We attempted to develop a model by relating, for each pre-felling inventory, the EAGB of trees with a DBH≥60cm to the EAGB of trees with a DBH≥30cm and a DBH<60cm. We compared the ability of a number of regression models (linear, exponential, logistic, and polynomial) to capture a relationship between EAGB (DBH≥60cm) and EAGB (60cm<DBH≤30cm).

4.4.6 Evaluating the uncertainty of AGB estimates

Uncertainty can be introduced to a tree's EAGB through DBH measurement errors (σMDBH), wood density measurement errors (σMρ), and errors inherent in the allometric model itself (σA) (Chave et al., 2004). In this study, we evaluated the uncertainty of tree-level AGB estimates introduced by error in wood density measurements (σMρ) at the four following levels: species (σMρ:sp), genus (σMρ:gen), NFMP (σMρ:NFMP), and CA (σMρ:CA). We hypothesized that the measurement error will increase as the source of wood density increases in taxonomic scale. To evaluate errors introduced by the allometric models themselves, we reiterated the calculation of a tree's EAGB while varying ε (Equations 4.3 and 4.4 before) based on the residual standard error (RSE). Monte Carlo simulations were run in MATLAB version R2013b (The MathWorks Inc., Natick, MA, USA) to simulate the parameters (wood density and residual standard error of the allometric model (ε)) and determine both wood density measurement error and allometric model error.

As wood density values vary at the tree level within NFMPs, conservation areas, species, and genera (Chave et al., 2009), using mean wood density values to estimate AGB will introduce measurement error. Further, in forest management inventories, trees are identified in the field by common names and later related to scientific names. This methodology can result in the misidentification of species (Lacerda and Nimmo, 2010) and therefore, additional wood density measurement errors. We evaluate wood density errors under the assumption that the errors have a centered normal distribution. The distribution of errors for each tree uses the calculated mean and standard deviation of the appropriate species, genus, NFMP, or conservation area. We randomly selected 10,000 trees (5,000 from the census and 5,000 from the inventory) to evaluate the uncertainty at each level. For each tree, we calculated EAGB 1,000 times while varying the wood density parameter by a random normal distribution.

As allometric models are typically created using a regression on log-transformed variables, there is inherent error in them. This uncertainty is the result of trees departing from the exact allometry described by the models (Chave et al., 2004). Errors in tree EAGB estimates due to the allometric model were assessed by varying ε (Equation 4.3 and 4.4) following the methodology of Maniatis et al. (2011). We assumed that ε followed a centered normal distribution with a mean of 0 and a standard deviation of 0.356 (residual standard error reported for the models in Chave et al., 2005). ε was incorporated into the EAGB equations using the same structure as Maniatis et al. (2011). For wet forests the model became:

$$EAGB = \rho\ x \exp(-1.302 + 1.980\ x \ln(DBH) + 0.207\ x\ \ln(DBH)^2 - 0.0281\ x \ln DBH3\ x \exp \varepsilon \quad \text{(Equation 4.3)}$$

while for moist forests is became:

$$EAGB = \rho\ x \exp(-1.562 + 2.148 \text{ x} \ln(DBH) + 0.207\ x\ \ln(DBH)^2 - 0.0281 \text{ x} \ln DBH3 \text{ x} \exp \varepsilon \quad \text{(Equation 4.4)}$$

Following the previously described methodology, 10 000 were randomly selected from the census and inventory and, for each tree, EAGB was calculated 1000 times while varying the ε parameter by a random normal distribution.

To evaluate the uncertainty of EAGB at the census and inventory levels, EAGB was simulated 1,000 times for every tree of 100 randomly selected censuses and 100 randomly selected inventories. For each simulation, we varied both wood density and ε simultaneously following the above sampling methodology. Simulated EAGB values were compared, revealing the uncertainty and precision of the reported EAGB values.

4.5 Results

4.5.1 The variability of wood density

Among all conservation areas, ACLA-C had a significantly lower wood density (inventory: 0.528±0.161 g cm^{-3}, $p<0.01$; census: 0.530±0.1520 g cm^{-3}, $p<0.01$) (mean±standard deviation unless otherwise specified) (Table 4.1). No difference in wood density was detected in the census data among ACOSA, ACCVC, and ACAHN. Based on the inventory data set (either DBH≥30cm or only including DBH≥60cm), however, ACAHN had a significantly higher wood density than the other four conservation areas (0.623±0.182 g cm^{-3} and 0.636±0.197 g cm^{-3} respectively). The greater mean wood density found in ACAHN is due to a larger fraction of trees within the 0.8 to 0.9 g cm^{-3} range (Figures 4.2a and 4.2b). This is a result of the high density of the *Dialium guianensis* in ACAHN (Svob et al., 2014b *submitted*). In ACCVC and ACTO, a prominent peak in the percent of stems within the 0.6 to 0.7 g cm^{-3} range can be attributed to the high relative frequency of *Pentaclethra macroloba*. The range of wood density values sampled was similar in all five CAs. Additionally, in all conservation areas, the mean wood density sampled by the inventory and that sampled by the census did not significantly differ.

4.5.2 The variability of EAGB

Based on inventory data, the estimated AGB (EAGB) (DBH≥30cm) found in ACAHN and ACCVC was significantly lower than that in ACLA-C and ACOSA ($p<0.05$) (Figure 4.3 and Table 4.2). While ACAHN, ACCVC, ACTO, and ACLA-C all shared similar inventory EAGB values with at least one other conservation area, only in ACOSA did EAGB differ significantly from all other CAs. In fact, ACOSA presented the highest mean inventory EAGB (173.47±60.23 Mg ha^{-1}, $p<0.05$).

Based on census data, ACAHN had the lowest EAGB, significantly differing from ACLA-C, ACTO, and ACOSA ($p<0.05$) (Figure 4.4 and Table 4.2). Simply ranking conservation areas in decreasing order of mean EAGB (Table 4.2), reveals that the overall trends are similar between sampling protocols. For example, ACAHN had one of the lowest mean EAGB values in both the census and inventory data (39.77±23.48 Mg ha^{-1} and 136.63±60.08 Mg ha^{-1}). A paired t-test comparing the EAGB of trees with a DBH≥60cm from the census and inventory detected a significant difference between the two sampling protocols (n=366, $p<0.01$). Across all five conservation areas, inventories generally produced higher EAGB values than censuses of the same forest stand (Figure 4.5).

Our attempt to relate the EAGB of trees with a DBH≥60cm to the EAGB of trees with a DBH≥30cm but <60cm did not indicate a strong relationship between the two variables (e.g. linear regression results adj R^2: 0.043, F=19.9, $p<0.01$, n=422). Although we were unable to find a relationship between the two variables, this analysis demonstrates the amount of variance in the structure of Costa Rican forests (Figure 4.6a).

4.5.3 The density of large trees (DBH≥70cm)

We found a significant correlation between the density of large trees (DBH≥70cm) and EAGB (DBH≥30cm) (correlation coefficient: 0.728, n: 422, adj R^2: 0.533, F= 470.6, $p<0.01$) (Figure 4.6b). Our results show that 53.3% of the variation in EAGB across five Costa Rican conservation areas was explained by the density of large trees. Additionally, trends in the density of large trees per conservation area match trends in EAGB per conservation area. For example, in ACOSA, the EAGB and density of large trees (16.48±8.08 tree ha^{-1}) were both significantly higher than in the four other CAs (Table 4.1). Furthermore, the two conservation areas with the lowest mean EAGB (ACAHN and ACCVC) also had the lowest mean density of large trees (6.20±3.54 and 6.73±4.00 tree ha^{-1} respectively). Across all conservation areas, a pairwise t-test indicates a significantly higher number of large trees ha^{-1} was recorded by the inventory than the census of the same forest stand (n=366, $p<0.01$).

4.5.4 Uncertainty analysis

At the tree-level, when moving from species wood density ($\sigma_{M\rho:sp}=0.110\langle EAGB\rangle$) to genus wood density ($\sigma_{M\rho:gen}=0.151\langle EAGB\rangle$), we found a 4% increase in the uncertainty of EAGB due to wood density measurement error ($\sigma_{M\rho}$). We found that an even greater amount of EAGB uncertainty resulted from using

NFMP ($\sigma_{M\rho:NFMP}$ =0.271⟨EAGB⟩) or conservation area ($\sigma_{M\rho:CA}$ =0.281⟨EAGB⟩) level wood density values. The uncertainty due to allometric model error (σA) for each tree was 0.371⟨EAGB⟩. Hence, based on our Monte Carlo uncertainty analysis, the uncertainty of a tree's EAGB can range from 48% to 65% of its EAGB depending on the level of wood density used.

At the stand level, random measurement and allometric model errors counteract one another, decreasing their impact on EAGB uncertainty and increasing the overall precision of EAGB (Figure 4.7). The uncertainty of EAGB from a NFMP's inventory ranged from 0.021⟨EAGB⟩ to 0.171⟨EAGB⟩. At the census level, the uncertainty of EAGB for each NFMP ranged from 0.011⟨EAGB⟩ to 0.101⟨EAGB⟩. We observed that the uncertainty of a given NFMP's EAGB was principally dependent on the number of trees sampled and the total area sampled (Figure 4.7). In Figure 4.7, we also observed that the uncertainty of EAGB decreases as the number of trees (Figure 4.7a) and the total area sampled increases (Figure 4.7b and 4.7c) following a power function.

4.6 Discussion

4.6.1 The variability of wood density

Our study demonstrates for the first time the variability of wood density across five Costa Rican conservation areas based on pre-felling inventory data. We found the most northern forests included in our study, located in ACAHN, typically contain trees of higher wood density than those located in the other conservation areas sampled. In contrast, our results show that the southeastern lowland forests of ACLA-C house trees that tend to have lower wood density values. The regional differences between wood density values detected by our study indicate the importance of including this variable for the production of AGB estimates that are comparable at regional scales across Costa Rica. This variation also suggests that using more general country or pantropical scale wood density values when estimating AGB may lead to inaccurate results, underestimating the variability of EAGB across tropical regions (Baker et al., 2004; Muller-Landau, 2004). The regional wood density values found express the similarities and differences in species composition between the five conservation areas studied.

Beyond species composition, it is also known that wood density is closely linked to a forest's functional composition as light-demanding fast-growing species commonly have lower wood densities than shade-tolerant ones (King, 1991; Whitmore, 1998). Building upon this idea, we believe that both natural and human disturbance regimes may play a key role in shaping the variation of EAGB and wood density among the forest stands and conservation areas studied.

4.6.2 The variability of EAGB

The variation of EAGB among NFMPs within any given conservation area reveals the heterogeneity of EAGB across the five conservation areas (Figure 4.1). We found, based on the pre- felling data, that the forest stands of ACOSA are some of the most biomass rich areas of Costa Rica while, those of ACAHN are some of the most biomass poor. Supporting the findings of Stegen et al. (2009), a comparison of conservation area level wood density and EAGB trends suggests that wood density alone cannot explain regional EAGB variability. For example, despite having one of the highest mean wood density values, ACAHN has one of the lowest mean EAGB values. The variation of EAGB between NFMPs was very high, as indicated by the large standard deviation of EAGB among conservation areas (Table 4.2). Our findings highlight the need for a better understanding of both the environmental and human variables influencing the distribution of EAGB across spatial scales. For example, studies have found that forest fragmentation has a strong negative impact on AGB and AGC (aboveground carbon) due to a significant increase in the mortality of large trees near forest edges (Laurance et al., 2000; Nacimento and Laurance, 2004). A greater comprehension of the factors controlling EAGB distribution will allow for the production of more reliable EAGB maps at local, regional, and national scales. Comparing the AGB estimates of our study to published estimates reveals that, in the regions of Costa Rica where EAGB has been previously sampled, our NFMP based approach produced a similar range of values. A study conducted at the La Gamba biological station in ACOSA reported the EAGB of trees with a DBH≥30cm was 218.46±29.01 Mg ha^{-1} (Hofhansl et al., 2012). After considering one standard deviation from the mean, the EAGB determined from NFMPs (173.47±60.23 Mg ha^{-1})

overlaps with the published estimates of Hofhansl et al. (2012). At ACCVC's La Selva biological station, Clark and Clark (2000) found the density of large trees ranged from 4.7 to 10.1 stems ha^{-1} and the EAGB of large trees (DBH≥70cm) ranged from 22.6 to 55.4 Mg ha^{-1}. The analysis of ACCVC NFMPs found values comparable to those reported by Clark and Clark (2000), with a large tree density of 5.27±3.39 tree ha^{-1} (census) and 6.71±4.02 tree ha^{-1} (inventory) and a large tree (DBH≥70cm) EAGB of 27.53±18.06 Mg ha^{-1} (census) and 34.84±22.43 Mg ha^{-1} (inventory). This indicates a positive aspect in using NFMP data for assessing biomass and carbon (Maniatis et al., 2011).

Over half of the variation in EAGB across the five conservation areas sampled in this study was explained by the density of large trees. Although the strength of this predictive variable was approximately 20% less than the value reported by Slik et al. (2013), our study supports the conclusion that large tree density accounts for a significant portion of EAGB variability across tropical regions. Additionally, we found that the patterns of EAGB and large tree density matched among conservation areas, demonstrating the importance of large trees as drivers of regional EAGB differences across Costa Rica.

Despite such a great amount of EAGB variability across NFMPs, a weak relationship was found between the EAGB of trees with a DBH<60cm but ≥30cm and the EAGB of trees with a DBH≥60cm. If these results are consistent throughout other components of forest biomass, they indicate that models developed to estimate unmeasured portions of forest biomass based solely on the EAGB of measured forest components (e.g. Maniatis et al., 2011) may lead to an underestimation of the variability of forest biomass across the tropics. Future studies aiming to identify key variables that best explain how EAGB is distributed throughout different DBH classes and other forest stand components (e.g. lianas, coarse woody debris) could greatly improve the accuracy of AGB estimates (particularly in smaller trees) and in turn, the effort required to conduct large scale studies.

Our results show that the plot based sampling methodology of NFMPs (i.e. the inventory) tends to overestimate EAGB when compared to EAGB values calculated from the census of an entire forest stand. Houghton et al. (2001) reported a similar result, finding a weak negative correlation between area sampled and EAGB. Both our results and those of Houghton et al. (2001) indicate that the total area sampled may have important negative impacts on tropical AGB estimates. An additional source of the differences found may be explained by the location and distribution of inventory plots within forest stands. Although NFMP protocols specify that plots be randomly placed, we found a significant bias towards higher EAGB regions of the forest stand. As we do not have data to fully resolve the reason behind the bias, we hypothesize that it may be explained by a desire to achieve a higher economic outcome from the NFMP (i.e. to log a greater number of species and trees) and/or to reduce sampling effort (i.e. the placement of plots in more convenient areas of the forest stand). No matter the reason behind the bias, this finding brings to light the need to evaluate the sustainability of forest management practices in Costa Rica. If inventories are not only overestimating the number and EAGB of trees with a DBH≥60cm, but the number and EAGB of trees within the 30 cm to 60 cm DBH range, they may also be overestimating the capacity of forests to recover after a selective logging event (Blanc et al., 2009).

4.6.3 The uncertainty of EAGB

Our uncertainty analysis explored how incorporating wood density values at different scales in allometric models will introduce different amounts of uncertainty into a tree's estimated AGB. We found that more general stand level and regional wood densities can lead to uncertainties in the EAGB of a single tree between 27% and 28%. Further, we investigated how much the uncertainty of a tree's EAGB will increase when using a genus versus a species level wood density average. Although the 4% increase in uncertainty reflects the taxonomic composition of the forest stands sampled in this study, we believe that future work should consider this source of uncertainty when reporting EAGB estimates. Particularly, studies should pay greater attention to species and genera that exhibit high levels of wood density variability in the tropics (refer to Chave et al., 2006 for a list of genera). The impact of wood density variability on EAGB uncertainty will be the greatest when (1) the species and/or genera sampled are highly variable and (2) the highly variable species and/or genera compose a notable portion of a forest stems and the wood density values incorporated into allometric models. When moving from the tree level to the plot level, the uncertainties introduced by measurement errors (wood density or DBH) decrease as the number of trees sampled increases. Errors introduced by the allometric model, on the other hand, can be either an issue of accuracy or precision (Chave et al., 2004; Maniatis et al., 2011). If the allometric error is consistent, regardless of sample size, an accuracy error is present (i.e. the allometric model does not apply to or represent the given area). If the allometric error differs between trees and decreases with an increase in sample size, it is a precision error. As this study does not include direct tree biomass measurements collected via destructive sampling, the accuracy of the allometric model could not be addressed (Clark and Kellner, 2012). Therefore, in our uncertainty analysis we simulated the impact of *random* allometric model error on EAGB to evaluate the *precision* of the allometric models, finding that allometric model errors introduced 37% uncertainty to tree level EAGB values. It is important to note that the development of Chave et al.'s (2005) wet forest model included samples collected in Costa Rica (La Selva) while the moist forest equation did not. Future studies should evaluate the accuracy and applicability of these models in Costa Rican forests, particularly those that did not incorporate any Costa Rican data during their development.

Allometric model selection is an important source of error that was not directly addressed in this study. Several studies comparing the results of allometric models have shown they produce vastly differing results (e.g. Chave et al., 2004; Segura and Kanninem, 2005; Komiyama et al., 2008; Pelletier et al., 2012). Further, Pelletier et al. (2012) demonstrated that two different allometric models can result in estimated annual emissions from deforestation that differ by up to 48%. Given the potential impact of model selection on forest biomass estimates and in turn, the success of international mechanisms such as REDD+, future studies will need to consider this source of uncertainty. In this study we consciously selected the Chave et al. (2005) models as they fulfill several key criteria (the inclusion of wood density, representation of a large DBH range, and development from a large sample size).

Feldpausch et al. (2012, 2013) found that an additional source of error in AGB estimates results from the exclusion of height as a predictor variable in allometric models. Regardless of this finding, Feldpausch et al. (2013) also reported that the decrease in AGB estimation occurred only in smaller DBH classes (≥40cm) and not larger ones. As our study only includes trees with a DBH≥30cm, we believe that the exclusion of height in the models applied to estimate AGB will not greatly increase the error of the reported values.

4.7 Conclusions

This study presents the most spatially rich analysis of ground level EAGB data in Costa Rica to date. Using data from forest management inventories, we found that the EAGB within and among five Costa Rican conservation areas is highly variable. Further, we detected bias in the NFMPs towards biomass rich areas of the forest stand, demonstrating the need to assess the sustainability of Costa Rican forest management practices. Expanding upon this finding, if ecological plots are also being preferentially located in more easily accessible areas of forest stands, studies may not be accurately capturing the EAGB of tropical forests (Clark and Kellner, 2012). Despite the potential taxonomic issues and missing DBH classes within NFMPs, our EAGB values were comparable to those reported in the scientific literature, supporting their inclusion in future EAGB assessments. Currently, the ground level data used to produce large-scale AGB and aboveground carbon maps is predominantly collected from ecological studies. Although this data is detailed and systematic in nature, ecological plots tend to sample either protected regions of the landscape or areas subject to a *lesser* amount of human impact. Forest management data, on the other hand, covers a considerable portion of the tropics and represents forests that are being greatly impacted by anthropogenic activities. In fact, Asner et al. (2013) found that human activity was the greatest driver of AGB and aboveground carbon in the forests of Panama, highlighting the need measure the impact of humans on the variability of AGB and AGC across the tropics. Combining commercial logging inventories with ecological plots will provide a more representative ground level dataset for the calibration of the models and remotely sensed data used to estimate AGB and aboveground carbon at regional and national scales. Additionally, it is the non-protected areas of the tropics that offer the greatest opportunity to reduce rates of deforestation and forest degradation. Therefore, by improving our knowledge on the variability of aboveground carbon and AGB through forest management data, studies can better support the REDD+ mechanism and the sustainable management of the rich natural resources of the tropics.

4.8 Tables and Figures

Table 4.1. Wood density (g cm^{-3}) per CA and sampling protocol.

	Wood Density (g cm^{-3}) (Mean±Std)		
CA	Inventory (DBH≥30cm)	Inventory (DBH≥60cm)	Census (DBH≥60cm)
ACAHN	0.623±0.182	0.636±0.197	0.602±0.189
ACCVC	0.602±0.144	0.579±0.170	0.600±0.143
ACLA-C	0.528±0.161	0.579±0.171	0.530±0.152
ACOSA	0.574±0.165	0.579±0.169	0.604±0.166
ACTO	0.565±0.143	0.578±0.169	0.564±0.140

Table 4.2. EAGB (Mg ha^{-1}) and large tree density (tree ha^{-1}) per CA and sampling protocol.

	Inventory (Mean±Std)			Census (Mean±Std)	
CA	$EAGB_{DBH\geq 30cm}$ (Mg ha^{-1})	$EAGB_{DBH\geq 60cm}$ (Mg ha^{-1})	Lrg tree density (tree ha^{-1})	$EAGB_{DBH\geq 60cm}$ (Mg ha^{-1})	Lrg tree density (tree ha^{-1})
ACAHN	118.07±54.09	65.25±33.81	6.29±3.55	39.77±23.48	4.25±2.72
ACCVC	116.17±44.48	63.35±29.99	6.71±4.02	45.22±26.05	5.27±3.39
ACLA-C	143.38±62.18	79.83±48.29	9.21±5.98	53.60±27.89	6.15±2.99
ACOSA	173.39±60.64	123.27±53.56	16.30±7.97	58.35±23.24	7.65±2.79
ACTO	130.30±51.05	80.12±39.21	10.70±6.26	52.36±31.46	6.98±4.27

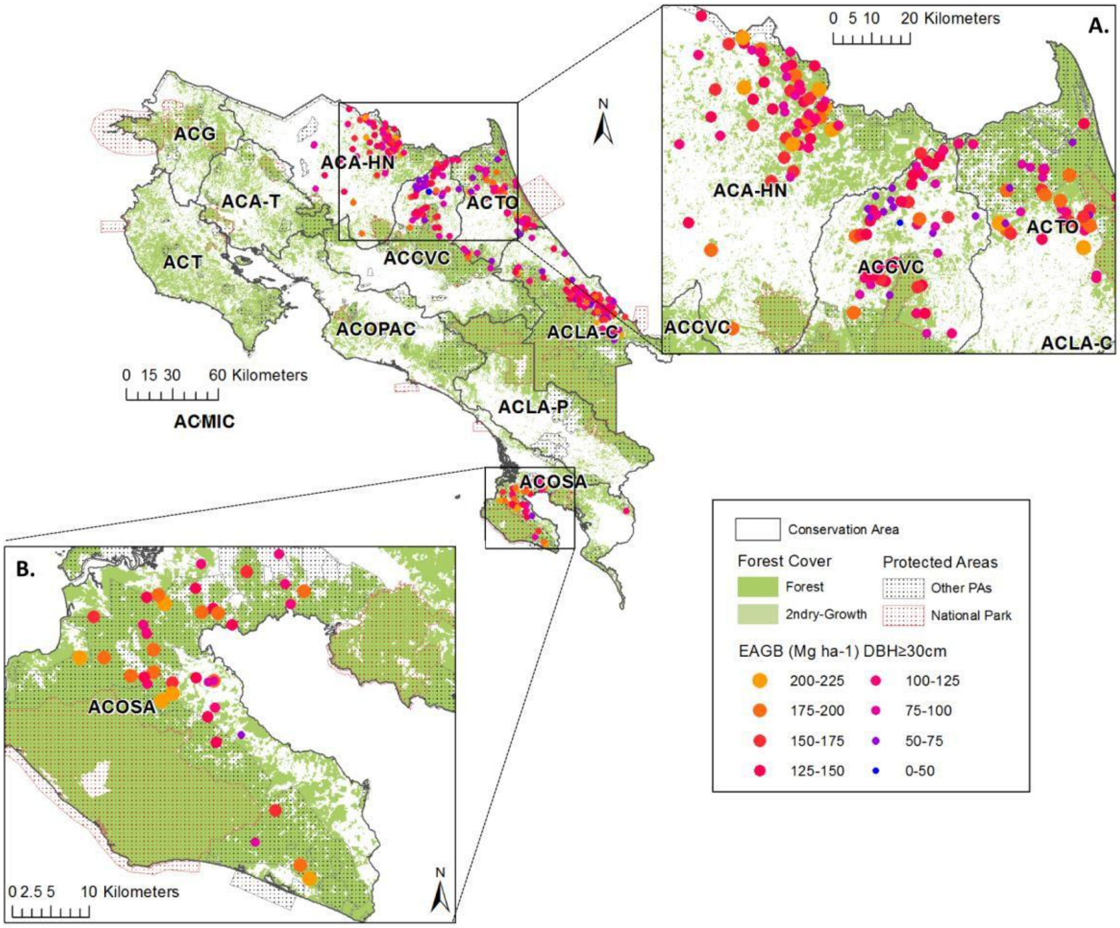

Figure 4.1. Map of the EAGB (Mg ha^{-1}) of trees with a DBH≥30cm per NFMP. (A) A view of the distribution NFMPs collected in ACAHN, ACCVC, and ACTO. (B) A view of NFMPs collected in ACOSA. Forest cover source: FONAFIFO 2005 Costa Rica forest cover assessment. Protected areas sources: Atlas of Costa Rica 2008 (Instituto Tecnológico de Costa Rica 2008).

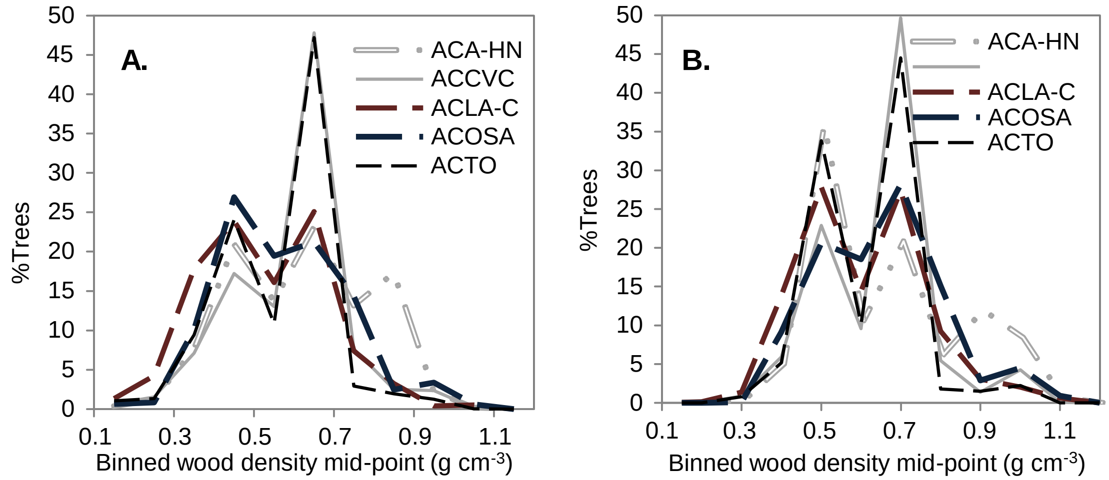

Figure 4.2. The distribution of wood density within five CAs. The percentage of trees that occupy consecutive 0.1 g cm^{-3} wood density bins in NFMP where (A) inventories (DBH≥30cm) and (B) censuses (DBH≥60cm) in the five CAs sampled.

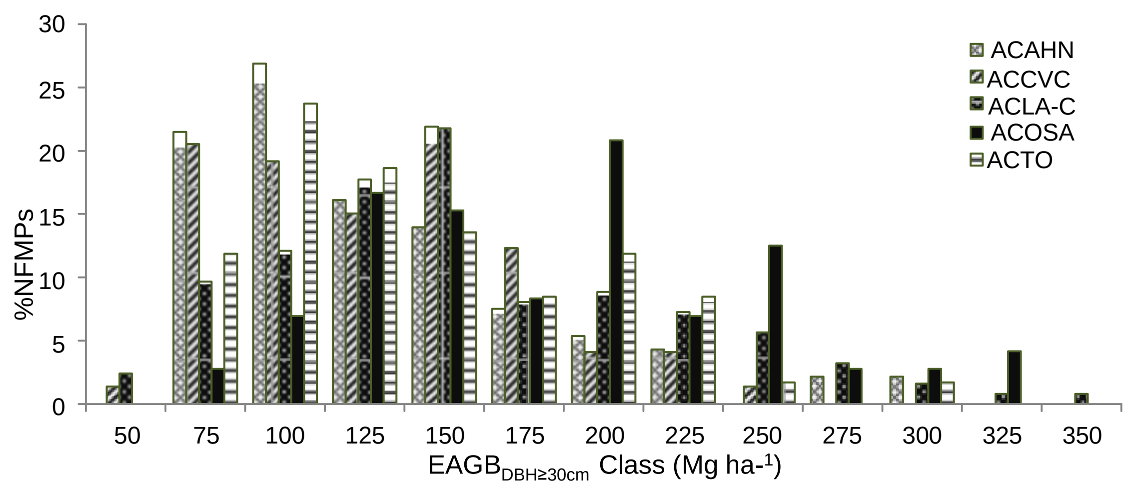

Figure 4.3. The distribution of inventory EAGB (DBH≥30cm) per CA. The distribution is shown by the percentage of NFMPs within 25 Mg ha^{-1} EAGB bins.

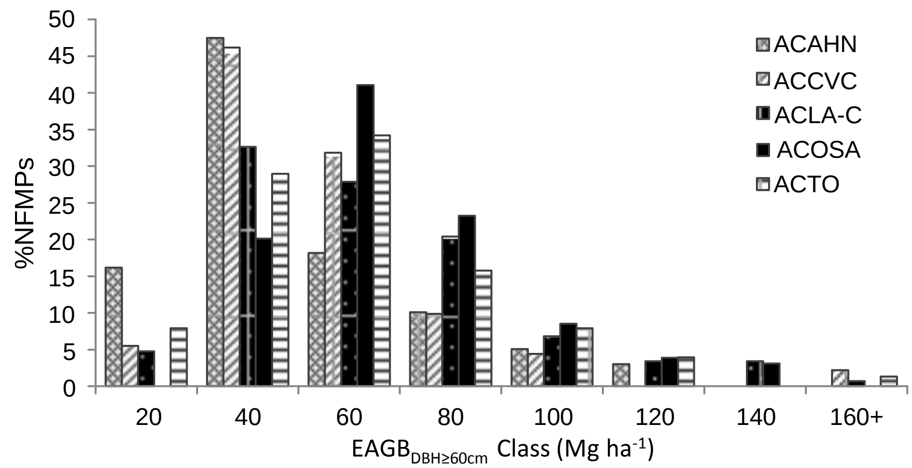

Figure 4.4. The distribution of census EAGB (DBH≥60cm) per CA. The distribution is shown by the percentage of NFMPs within 20 Mg ha^{-1} EAGB bins.

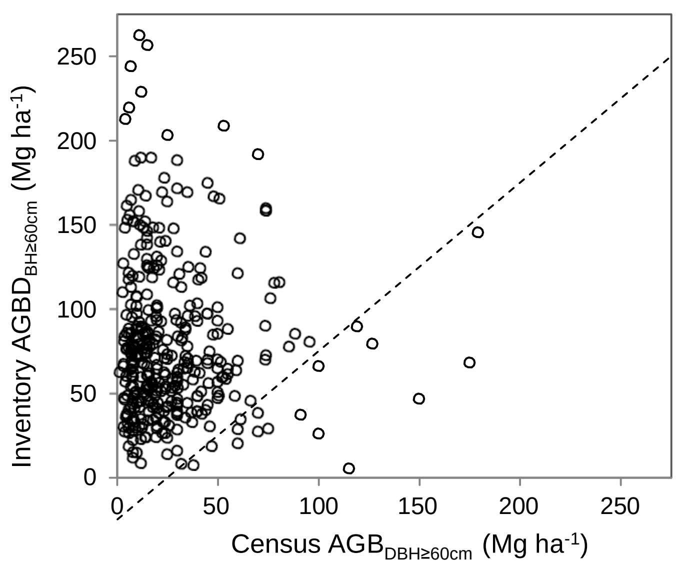

Figure 4.5. Scatter plot of census EAGB (DBH≥60cm) versus inventory EAGB (DBH≥60cm). The dashed diagonal line depicts a 1:1 relationship between census and inventory EAGB values.

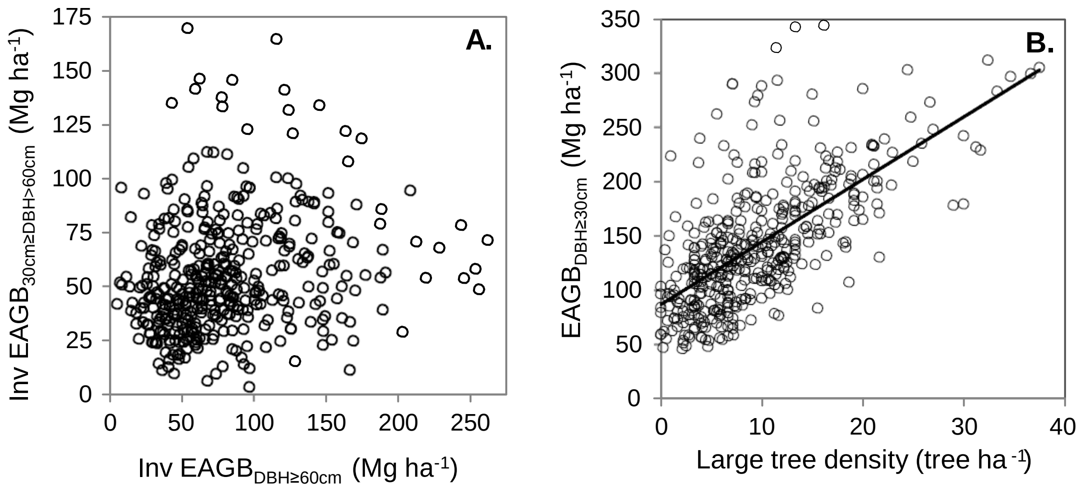

Figure 4.6. The relationship between the components of forest structure for trees of different sizes. (A) Scatter plot of inventory EAGB (60cm>DBH≥30cm) versus inventory EAGB (DBH≥60cm). The distribution of points within the scatter plot reveals the variation found between these two EAGB DBH classes; (B) The relationship between the density of large trees (DBH≥70cm) and EAGB (DBH≥30cm). The line fit through the data explains 53.3% of the variation in EAGB.

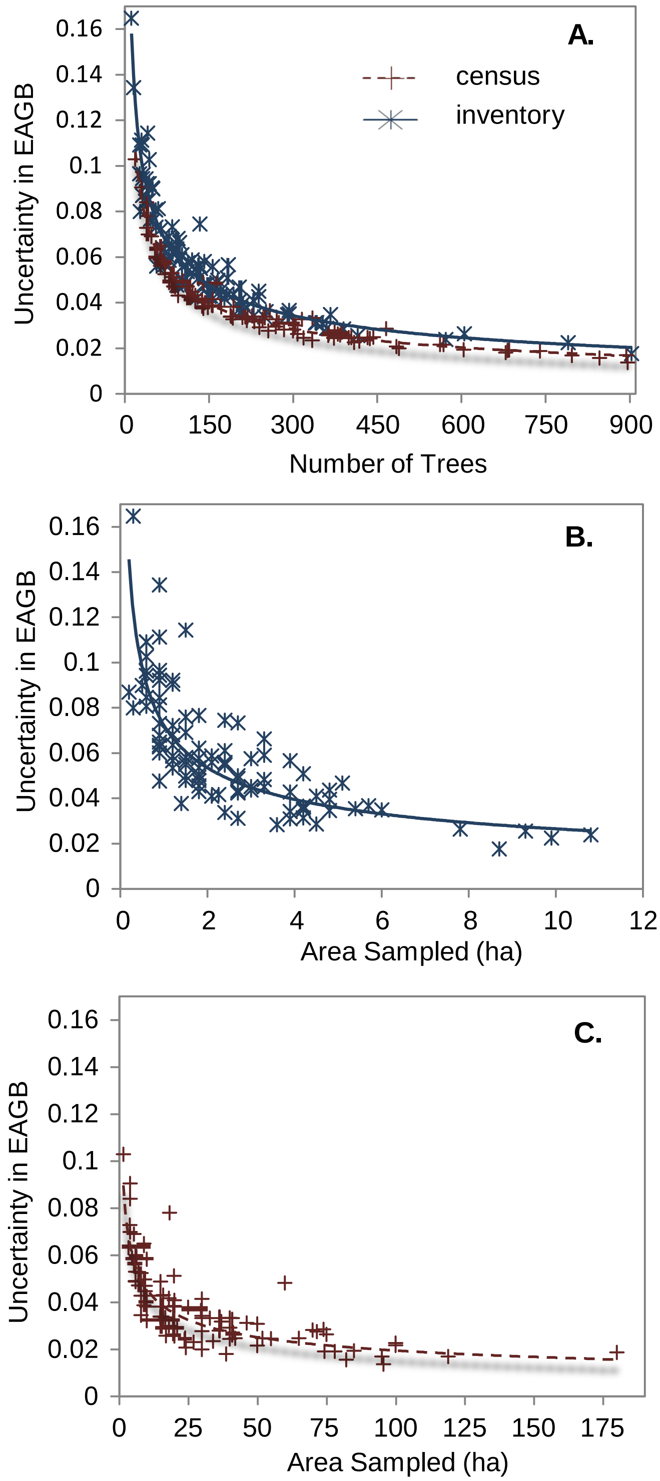

Figure 4.7. The relationship between the uncertainty in EAGB per NFMP and sampling effort. (A) The total number of trees sampled per NFMP; (B) the total area sampled per NFMP inventory; and (C) the total area sample per NFMP census. Uncertainty is given in unit-less value that varies from 0 to 1. The solid (inventory) and dashed (census) lines depict the best fit function (power).

5. Conclusions: Summary of findings and future research directions

5.1. Summary of findings

The distribution of aboveground biomass in tropical forests is highly heterogeneous. To best capture the variability of AGB across regions, the use of ground level data with a high sampling intensity (i.e. area sampled is large enough to capture the variation) and spatial distribution (i.e. sample points are well distributed across the landscape) is required. The majority of studies evaluating the variability of AGB across tropical regions rely on ecological data that tends to represent the areas of the landscape that are protected or subject to a *smaller* amount of human impact. Further, landscape scale forest inventories are rare among ecological studies due to their high cost and resource requirements (Greig-Smith 1983). In this study, I used forest management data retrieved from five Costa Rican ecosystems to study the variability of AGB across regions. Firstly, however, my Master's research aimed to combine and structure a large Costa Rican forest management dataset in a way that would allow for the efficient and repeatable analysis of the data.

My first objective was to develop a spatial/tabular database that integrates methods for standardization and quality control, ensuring forest management data is structured in a logical and accessible manner. Using PostgreSQL and PostGIS, I created a relational geodatabase, forming logical associations between the spatial and tabular components of the forest management data. With built-in standardization and quality control procedures, I was able to improve data integrity and reveal the magnitude, source, types, and implications of errors within the initial forest management dataset. The final NFMP tailored database permits flexible data access across 32 tables and over 250 000 tree records. Lastly, I was able to demonstrate the utility of the database by conducting an exploratory analysis of the standardized taxonomic data. The taxonomic analysis was conducted on the most complete digital record of natural forest management in Costa Rica to date. This analysis also revealed the strengths and weaknesses of the taxonomic components of the NFMPs including: (1) A genus level analysis is a favorable approach to take with the NFMP data due to the higher fraction of trees identified at the genus level; and (2) absolute species counts should be viewed with discretion due to the low frequency of a notable proportion of species within the NFMP tree records.

My second objective was to assess the variability of aboveground biomass among and within five Costa Rican conservation areas using forest management data. Based on NFMP data, I found that the most northern region of Costa Rica studied (ACAHN) had a significantly higher wood density for trees with a DBH≥30cm. Further, my results showed that Osa peninsula (ACOSA) houses a significantly higher AGB than the four other regions studied. The standard deviation and distributions of AGB within each conservation area supports studies that have found high levels AGB variability within the tropics. The density of large trees explained over 50% of the variability in estimated AGB across the five conservation areas considered in this thesis. Comparing the estimated AGB in this study to published estimates revealed that, in regions of Costa Rica where AGB has previously been sampled, my forest management based approach produced comparable values. Lastly, the AGB analysis indicated a bias in the NFMPs towards sampling biomass rich areas of forest stands, highlighting the need to assess the sustainability of forest management practices in Costa Rica.

5.2 Future research directions

In this thesis, I present the FGIS database which stores the largest forest management dataset in Costa Rica to date. This tool will make the analyses needed to address issues of conservation and sustainable resource management in Costa Rica possible. Further, future work could collect field samples from the NFMPs to refine our understanding of the issues related to taxonomic identification in Costa Rican NFMPs. Building upon previous studies (e.g. Couteron et al. 2003; Arroyo-Mora et al. 2009), the FGIS database's tree data could be used to assess species distribution, biodiversity, and species environment interactions from the local to landscape level. Further, the database could be used to expand the study of Arroyo-

Mora et al. (2009), evaluating the historical trends of natural forest management in Costa Rica.

Based on the AGB analysis, I detected bias in the NFMPs towards biomass rich areas of the forest stand. I hypothesize that the bias may be explained by a desire to achieve a higher economic outcome (i.e. log a greater number of trees and species) and/or to reduce sampling effort (i.e. sample more convenient areas of the forest stand). Future work should attempt to determine the cause for this bias and, in turn, assess the sustainability of the forest management plans. Expanding upon the uncertainty analysis, future work should assess the accuracy of the Chave et al.'s (2005) allometric models in Costa Rican forests. Recent studies of allometric models in the tropics suggest that height is a key variable, preventing the overestimation of AGB in lower DBH classes (DBH<40cm) (Feldpausch et al. 2011, 2012). Therefore, height-DBH models applicable in Costa Rica should be created to improve the accuracy of AGB estimates. Given the high level of variability in AGB across the five conservation areas studied, the next step is to evaluate the environmental and anthropogenic factors controlling the spatial distribution of AGB. I hypothesize that forest fragmentation characteristics will be an important factor influencing AGB as the NFMPs are largely located within the fragmented forests of Costa Rica. This hypothesis is also based on my finding that 'large tree' density explained over 50% of the AGB variability in NFMPs and the findings of previous studies that fragmentation increases the mortality of large trees near forest edges (Laurance et al. 2000; Nascimento & Laurance 2004). Following this study, AGB values could be converted to carbon and combined with ecological plots to form the ground level dataset required to produce a reliable and spatially explicit national carbon baseline in Costa Rica. Such an effort would improve the quality of the models and remotely sensed data used to create wall-to-wall coverage maps of forest carbon stocks, supporting the REDD+ mechanism in Costa Rica.

References

Alves, L.F., Viera, S.A., Scaranello, M.A., Camargo, P.B., Santos, F.A.M., Joly, C.A. & Martinelli, L.A. 2010. Forest structure and live aboveground biomass variation along an elevation gradient of tropical Atlantic moist forest (Brazil). *Forest Ecology and Management* 260: 679-691.

Andreae, M.O., Artaxo, P., Brandão, C., Carswell, F.E., Ciccioli, P., da Costa, A.L., Culf, A.D., Esteves, J.L., Gash, J.H.C., Grace, J., Kabat, P., Lelieveld, J., Malhi, Y., Manzi, A.O., Meixner, F.X., Nobre, A.D., Ruivo, M., Silva-Dias, M.A., Stefani, P., Valentini, R., von Jouanne, J. & Waterloo, M.J. 2002. Biogeochemical cycling of carbon, water, energy, trace gases, and aerosols in Amazonia: the LBA-EUSTACH experiments. *Journal of Geophysical Research-Atmospheres* 107(D20): 8066.

APG III. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnaean Society* 161: 105–121.

Arroyo-Mora, J.P. 2008. Natural forest management plans in Costa Rica: A potential framework for assessing tree biodiversity. Doctoral Dissertations. Paper AAI3329114. URL: http://digitalcommons.uconn.edu/dissertations/AAI3329114.

Arroyo-Mora, J. P., Chazdon, R.L., Kalacska, M., Obando, G., Aguilar, L. & Salas, L.F. Development of a forest management GIS for Costa Rica, a case study for the Central Volcanic Cordillera conservation area: management trends, lessons and potential uses in ecological research and conservation planning. 2009. *Proceedings of the XXIII World Forestry Congress.* 18-25 October, Buenos Aires, Argentina.

Asner, G.P. 2011. Painting the world REDD: addressing scientific barriers to monitoring emissions from tropical forests. *Environment Resource Letters* 6: 1-3.

Asner, G.P., Hughes, R.F., Varga, T.A., Knapp, D.E. & Kennedy-Bowdoin, T. 2009. Environmental and biotic controls over aboveground biomass throughout a tropical rain forest. *Ecosystems*12(2): 261-278.

Asner, G.P., Powell, G.V.N., Mascaro, J., Knapp, D.E., Clark, J.K., Jacobson, J., Kennedy-Bowdoin, T., Balaji, A., Paez-Acosta, G., Victoria, E., Secada, L., Valqui, M. & Hughes, R.F. 2010. High-resolution forest carbon stocks and emissions in the Amazon. *Proceedings of the National Academy of Sciences* 107(38): 16738-16742.

Asner, G.P., Mascaro, J., Anderson, C., Knapp, D.E., Martin, R.E., Kennedy-Bowdoin, T., van Breugel, M., Davies, S., Hall, J.S., Muller-Landau, H.C., Potvin, C., Sousa, W., Wright, J. & Bermingham E. 2013. High-fidelity national carbon mapping for resource management and REDD+. *Carbon Balance and Management* 8: 7.

Baccini, A., Goetz, S.J., Walker, W.S., Laporte, N.T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P.S.A., Dubayah, R., Friedl, M.A., Samanta, S. & Houghton, R.A. 2012. Estimating carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change* 2: 182-185.

Bagan, H., Wang, Q., Watanabe, M., Yang, Y. & Ma, J. 2005. Land cover classification from MODIS EVI time-series data using SOM neural network. *International Journal of Remote Sensing* 26(22): 4999-5012.
Baker, T.R., Philips, O.L., Malhi, Y., Almeidas, S., Arroyo, L., Di Fiore, A., Erwin, T., Killeen, T.J., Laurance, S.G., Laurance, W.F., Lewis, S.L., Lloyd, J., Monteagudos, A., Neill, D.A., Patiño, S., Pitman, N.C.A., Silva, J.N.M. & Martínez, R.V. 2004. Variation in wood density determines spatial patterns in Amazonian forest biomass. *Global Change Biology* 10: 545-562.
Baker, T.R., Phillips, O.L., Laurance, W.F., Pitman, N.C.A., Almeida, S., Arroyo, L., DiFiore, A., Erwin, T., Higuchi, N., Killen, T.J., Laurance, S.G., Nascimento, H., Monteagudo, A., Neil, D.A., Silva, J.N.M., Malhi, Y., Gonzalez, G.L., Peacock, J., Quesada, C.A., Lewis, S.L. & Lloyd, J. 2009. Do species traits determine patterns of wood production in Amazonian forests? *Biogeosciences* 6: 297-307.
Blanc, L., Echard, M., Herault, B., Bonal, D., Marcon, E., Chave, J. & Baraloto, C. 2009. Dynamics of aboveground carbon stocks in a selectively logged tropical forest. *Ecological Applications* 19(6): 1397-1404.
Bonan, G. 2008. Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science* 320: 1444-1449.
Boyle, B.L. 2006. *TaxonScrubber Version 2.0.* The SALVIAS Project. URL: http://www.salvias.net/pages/taxonscrubber.html. [Accessed March 10, 2013].
Boyle, B.L., Hopkins, N., Lu, Z., Garay, J.A.R., Mozzherin, D., Rees, T., Matasci, N., Narro, M.L., Piel, W.H., McKay, S.J., Lowry, S., Freeland, C., Peet, R.K. & Enquist, B.J. 2013. The taxonomic name resolution service: an online tool for automated standardization of plant names. *BMC Bioinformatics* 14:16.
Boza, M.A. 1993. Conservation in action: past, present, and future of the national park system of Costa Rica. *Conservation Biology 7*(2): 239-247.
Bradshaw, C.J.A., Sodhi, N.S., Peh, K.S.H. & Brook, B.W. 2007. Global evidence that deforestation amplifies flood risk and severity in the developing world. *Global Change Biology* 13: 2379-2395.
Bradshaw, C.J.A., Sodhi, N.S. & Brook, B.W. 2009. Tropical turmoil: a biodiversity tragedy in progress. *Frontiers in Ecology and the Environment* 7(2): 79-87.
Brown, S. 1997. *Estimating biomass and biomass change of tropical forests: a primer.* UN FAO Forestry Paper 134. FAO Rome: 55.
Brown, S. & Gaston, G. 1995. Use of forest inventories and geographic information systems to estimate biomass density of tropical forests: Application to tropical Africa. *Environmental Monitoring and Assessment* 38: 157-168.
Brown, S. & Lugo, A.E. 1992. Aboveground biomass estimates for tropical moist forests of the Brazilian Amazon. *Interciencia* 17(1): 8-18.
Brown, S., Gillepsie, A.J.R. & Lugo, A.E. 1989. Biomass estimation methods for tropical forests with applications to forest inventory data. *Forest Science* 35(4): 881-902.
Cain, J. 2001. Planning improvements in natural resources management: guidelines for using Bayesian Networks to support the planning and management of

development programmes in the water sector and beyond*. Centre for Ecology and Hydrology*. Wallingford, UK.

Campbell, B.M. 2009. Beyond Copenhagen: REDD+, agriculture, adaptation strategies and poverty. *Global Environmental Change* 19: 397-399.

Castilho, C.V. de, Magnusson, W.E., de Araujo, R.N.O., Luizao, R.C.C., Luizao, F.J., Lima, A.P. & Higuchi, N. 2006. Variation in aboveground tree live biomass in central Amazonian forest: Effects of soil and topography. *Forest Ecology and Management* 234(1-3): 85-86.

Chambers, J.Q., Higuchi, N., Schimel, J.P., Ferreira, L.V. & Melack, J.M. 2000. Decomposition and carbon cycling of dead trees in tropical forests of the central Amazon. *Oecologia* 122*:* 380-388.

Chambers, J.Q., Schimel, J.P. & Nobre, A.D. 2001. Respiration from coarse wood litter in central Amazon forests. *Biogeochemistry* 52: 115–131.

Chambers, J. Q., Higuchi, N., Teixeira, L.M., dos Santos, J., Laurance, S.G. & Trumbore, S.E. 2004. Response of tree biomass and wood litter to disturbance in a central Amazon forest. *Oecologia* 141: 596–611.

Chave, J. 2008. Spatial variation in tree species composition across tropical forests: pattern and process. In: Carson, W., & Schnitzer, S. (eds.) *Tropical Forest Community Ecology*, pp. 11-30. Blackwell.

Chave, J., Chust, G., Condit, R., Aguilar, S., Hernandez, A., Lao, S. & Perez, R. 2004. Error propagation and scaling for tropical forest biomass estimates. In *Tropical forests and global atmospheric change.* Edited by Malhi Y, Philips O. London: Oxford University Press: 155-166.

Chave, J., Andalo, C., Brown, S., Cairns, M.A., Chamber, J.Q., Eamus, D., Fölster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J.-P., Nelson, B.W., Ogawa, H., Puig, H., Riéra, B. & Yamakura, T. 2005. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia* 145: 87-99.

Chave, J., Muller-Landau, H.C., Baker, T.R., Easdale, T.A., Ter Steege, H. & Webb, C.O. 2006. Regional and phylogenetic variation of wood density across 2456 neotropical tree species. *Ecological Applications* 16(6): 2356-2367.

Chave, J., Olivier, J., Bongers, F., Chatelet, P., Forget, P.-M., Meer, P.V.D., Norden, N., Riera, B. & Charles-Dominique, P. 2008. Above-ground biomass and productivity in a rain forest of eastern South America. *Journal of Tropical Ecology* 24: 355-366.

Chave, J., Coomes, D., Jansen, S., Lewis, S.L., Swenson, N.G. & Zane, A.E. 2009. Towards a worldwide wood economics spectrum. *Ecology Letters* 12: 351-366.

Chen, P.P.S. 1976. The entity-relationship model—toward a unified view of data. *ACM Transactions on Database Systems (TODS)* 1(1): 9-36.

Clark, D.B. & Clark, D.A. 2000. Landscape-scale variation in forest structure and biomass in a tropical rain forest. *Forest Ecology and Management* 137: 185-198.

Clark, D.B. & Kellner, J. 2012. Tropical forest biomass estimation and the fallacy of misplaced concreteness. *Journal of Vegetation Science* 23: 1191-1196.

Clark, D.B., Clark, D.A. & Read, J.M. 1998. Edaphic variation and the mesoscale distribution of tree species in a neotropical rain forest. *Journal of Ecology* 86: 101-112.

Clark, D.B., Palmer, M.W. & Clark, D.A. 1999. Edaphic factors and the landscape-scale distributions of tropical rain forest trees. *Ecology* 80: 2662-2675.

Clark, M.L., Roberts, D.A., Ewel, J.J. & Clark, D.B. 2011. Estimation of tropical rain forest aboveground biomass with small-footprint lidar and hyperspectral sensors. *Remote Sensing of Environment* 115(11): 2931-2942.

Cleaveland, C.C. & Townsend, A.R. 2006. Nutrient additions to a tropical rain forest drive substantial soil carbon dioxide losses to the atmosphere. *Proceedings of the National Academy of Sciences* 103(27): 10316-10321.

Cleaveland, C., Townsend, A.R. & Constance, B.C. 2004. Soil microbial dynamics in Costa Rica: Seasonal and biogeochemical constraints. *Biotropica* 36(2): 184-1995.

Codd, E.F. 1970. A relational model of data for large shared data banks. *Communications of the ACM* 13(6): 377-387.

Codd, E.F. 1979. Extending the database relational model to capture more meaning. *ACM Transactions on Database Systems* 4(4): 397-434.

Condit, R., Lao, S., Singh, A., Esufali, S. & Dolins, S. 2013. Data and database standards for permanent forest plots in a global network. *Forest Ecology and Management:* http://dx.doi.org/10.1016/j.foreco.2013.09.011. Couteron, P., Pélissier, R., Mapaga, D., Molino, J.F. & Tellier, L. 2003. Drawing ecological insights from a management-oriented forest inventory in French Guiana. *Forest Ecology and Management* 172: 89-108.

DeWalt, S. & Chave, J. 2004. Structure and biomass of four lowland neotropical forests. *Biotropica* 36(1): 7-19.

Dirzo, R. & Raven, P.H. 2003. Global state of biodiversity and loss. *Annual Review of Environment Resources* 28: 137-167.

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

Drake, J.B., Knox, R.G., Dubayah, R.O., Clark, D.B., Condit, R., Blair, J.B. & Hofton, M. 2003. Above-ground biomass estimation in closed canopy Neotropical forests using lidar remote sensing: factors affecting the generality of relationships. *Global Ecology & Biogeography* 12: 147-159.

Dubayah, R.O., Sheldon, S.L., Clark, D.B., Hofton, M.A., Blair, J.B., Hurti, G.C. & Chazdon, R.L. 2010. Estimation of tropical forest height and biomass dynamics using lidar remote sensing at La Selva, Costa Rica. *Journal of Geophysical Research: Biogeosciences* 115: G00E09.

Elias, M. & Potvin, C. 2003. Assessing inter- and intra-specific variation in trunk carbon concentration for 32 neotropical tree species. *Canadian Journal of Forest Research* 33: 1039-1045.

Evans, J.S. & Cushman, S.A. 2009. Gradient modeling of conifer species using random forests. *Landscape Ecology* 24*:* 673-683.

FAO (Food and Agricultural Organization). 2006. Global forest resources assessment 2005: progress towards sustainable forest management. *FAO Forestry Paper*

No. 147. Food and Agriculture Organization of the United Nations (FAO), Rome, Italy, 320 p.
Fayolle, A., Engelbrecht, B., Freycon, V., Mortier, F., Swaine, M., Réjou-Méchain, M., Doucet, J.-L., Fauvet, N., Cornu, G. & Gourlet-Fleury, S. 2012. Geologic substrates shape tree species and trait distributions in African moist forests. *PLoS ONE* 7(8): e42381.
Feldpausch, T.R., Banin, L., Phillips, O.L., Baker, T.R., Lewis, S.L., Quesada, C.A., Affum-Baffoe, K., Arets, E.J.M.M., Berry, N.J., Bird, M., Brondizio, E.S., de Camardo, P. de, Chave, J., Djagbletey, G., Domingues, T.F., Drescher, M., Fearnside, P.M., França, Fyllas, N.M., Lopez-Gonzalez, G., Hladik, A., Higuchi, N., Hunter, M.O., Iida, Y., Salim, K.A., Kassim, A.R., Keller, M., Kemp, J., King, D.A., Lovet, J.C., Marimon, B.S., Marimon-Junior, B.H., Lenza, E., Marshall, A.R., Metcalfe, D.J., Mitchard, E.T.A., Moran, E.F., Nelson, B.W., Nilus, R., Nogueira, E.M., Palace, M., Patiño, S., Peh, K.S.-H., Raventos, M.T., Reitsma, J.M., Saiz, G., Schrodt, F., Sonké, B., Taedoung, H.E., Tan, S., White, L., Wöll, H. & Lloyd, J. 2011. Height-diameter allometry of tropical forest trees. *Biogeosciences* 8: 1081-1106.
Feldpausch, T.R., Lloyd, J., Lewis, S.L., Brienen, R.J.W., Gloor, M., Monteagudo Mendoza, A., Lopez-Gonzalez, G., Banin, L., Abu Salim, K., Affum-Baffoe, K., Alexiades, M., Almeida, S., Amaral, I., Andrade, A., Aragão, Araujo Murakami, A., Arets, E.J.M.M., Arroyo, L., Aymard C., 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., Erwin, T.L., Fearnside, P.M., França, M.B., Freitas, M.A., Higuchi, N., Honorio C., E., Iida, Y., Jiménez, E., Kassim, A.R., Killeen, T.J., Laurance, W.F. Lovett, J.C., Malhi, Y., Marimon, B.S., Marimon-Junior, B.H., Lenza E., Marshall, A.R., Mendoza, C., Metcalfe, D.J., Mitchard, E.T.A., Neill, D.A., Nelson, B.W., Nilus, R., Nogueira, E.M., Parada, A., Peh, K.S.-H., Pena Cruz, A., Peñuela, Pitman, N.C.A., Prieto, A., Quesada, C.A., Ramírez, F., Ramírez-Angulo, H., Reitsma, J.M., Rudas, A., Saiz, G., Salomão, R.P., Schwarz, M., Silva, N., Silva-Espejo, J.E., Silverira, M., Sonké, B., Stopp, J., Vásquez, R., Vilanova, E., Vos, V.A., White, L., Willcock, S., Woell, H. & Phillips, O.L. 2012. Tree height integrated into pantropical forest biomass estimates. *Biogeosciences* 9: 3381-3403.
Foley, J.A., DeFries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, C., Gibbs, H.K., Helkowski, J.H., Holloway, T., Howard, E.A., Kucharik, C.J., Monfreda, C., Patz, J.A., Prentice, C., Ramankutty, N. & Snyder, P.K. 2005. Global consequences of land use. *Science* 309(5734): 570-574.
Foody, G.M. 1999. Applications of the self-organizing feature map neural network in community data analysis. *Ecological Modeling* 120: 97-107.
Fonseca, W., Benayas, J.M.R. & Alice, F.E. 2011. Carbon accumulation in the biomass and soil of different aged secondary forests in the humid tropics of Costa Rica. *Forest Ecology and Management* 262: 1400-1408.
FUNDECOR, 2008. Methodology for calculating carbon and CO_2 from biomass in the forest located in the area of interest, using FUNDECOR permanent plot measurements: 1-18. *Unpublished.*

Gardner, T.A., Barlow, J., Chazdon, R., Ewers, R.M., Harvey, C.A., Peres, C.A. & Sodhi, N.S. 2009. Prospects for tropical forest biodiversity in a human-modified world. *Ecology Letters* 12: 561-582.
Gaston, K.J. & Williams, P.H. 1993. Mapping the world's species - the higher taxon approach. *Biodiversity Letters* 1: 2-8.
Ghazoul, J., Butler, R.A., Mateo-Vega, J. & Kon, L.P. 2010. REDD: a reckoning of environment and development implications. *Trends in Ecology and Evolution* 25(7): 396-402.
Gibbs, H.K., Brown, S., Niles, J.O. & Foley, J.A. 2007. Monitoring and estimating tropical forest carbon stocks: making REDD a reality. *Environmental Research Letters* 2: 1-13.
Girardin, C.A.J., Malhi, Y., Aragao, L.E.O.C., Mamani, M., Huaraca Huasco, W., Durand, L., Feeley, K.J., Rapp, J., Silva-Espejo, J.E., Silman, M., Salinas, N. & Whittaker, R.J. 2010. Net primary productivity allocation and cycling of carbon along a tropical forest elevational transect in the Peruvian Andes. *Global Change Biology* 16(12): 3176-3192.
Giraudel, J.L. & Lek, S. 2001. A comparison of self-organizing map algorithm and some conventional statistical methods for ecological community ordination. *Ecological Modelling* 146(1-3): 329-339.
Goetz, S., Baccini, A., Laporte, N., Johns, T., Walker, W., Kellndorfer, J., Houghton, R. & Sun, M. 2009. Mapping and monitoring carbon stock with satellite observations: a comparison of methods. *Carbon Balance and Management* 4(1): 1184-1182.
GOFC-GOLD. 2011. A sourcebook of methods and procedures for monitoring and reporting anthropogenic greenhouse gas emissions and removals caused by deforestation, gains and losses of carbon stocks in forests remaining forests, and forestation. Alberta, Canada: Natural Resources Canada. *GOFCGOLD Report version COP17-1.*
Gond, V., Fayolle, A., Pennec, A., Cornu, G., Mayaux, P., Camberlin, P., Doumenge, C., Fauvet, N. & Gourlet-Fleury, S. 2013. Vegetation structure and greenness in Central Africa from Modis multi-temporal data. *Philosophical Transactions of the Royal Society of London Biological Sciences* 368: 20120309.
Gourlet-Fleury, S., Rossi, V., Réjou-Méchain, M., Freycon, V., Fayolle, A., Saint-André, L., Cornu, G., Gerard, J., Sarrailh, J.M., Flores, O., Baya, F., Billand, A., Fauvet, N., Gally, M., Henry, M., Hubert, D., Pasquier, A. & Picard, N. 2011. Environmental filtering of dense-wooded species controls above-ground biomass stored in African moist forests. *Journal of Ecology* 99: 981-990.
Greig-Smith, P. 1983. *Quantitative plant ecology*. 3rd. ed. Blackwell, Oxford, UK.
Griscom, B., Shoch, D., Stanley, B., Cortez, R. & Virgilio, N. 2009. Sensitivity of amounts and distribution of tropical forest carbon credits depending on baseline rules. *Environmental Science & Policy* 12: 897-911.
Hanazaki, N., Mazzo, R., Duarte, A.R., Souza, V.C. & Rodrigues, R.R. 2007. Ecologic salience and agreement on the identification of tree species from Brazilian Atlantic forest. *Biota Neotropica* 10(1): 77-84.
Hansen, M.C., Stehman, S.V., Potapov, P.V., Loveland, T.R., Townshend, J.R.G., DeFries, R.S., Pittman, K.W., Arunarwatl, B., Stolle, F., Steininger, M.K., Corroll, M. &

Dimiceli, C. 2008. Humid tropical forest clearing from 2000 to 2005 quantified by using multitemporal and multiresolution remotely sensed data. *Proceeding of the National Academy of Sciences* 105(27): 9349-9444.
Herrera, B. & Finegan, B. 1997. Substrate conditions, foliar nutrients and the distribution of two canopy tree species in a Costa Rican secondary rain forest. *Plant and Soil* 191: 259-267.
Herwitz, S.R. 1981. Regeneration of selected tropical species in Corcovado National Park, Costa Rica. *University of California Publications in Geography* 24 (vii-xii): 1-109.
Hill, T.H., William, M., Bloom, A.A., Mitchard, E.T.A. & Ryan, C.M. 2013. Are inventory based and remotely sensed above-ground biomass estimates consistent? *PLoS ONE* 8(8): e74170.
Hofhansl, F., Wanek, W., Drage, S., Huber, W., Weissenhofer, A. & Richter, A. 2012. Controls of hydrochemical fluxes via stemflow in tropical lowland rainforests: Effects of meteorology and vegetation characteristics. *Journal of Hydrology* 452-453: 247-258.
Holdridge, L. R. 1979. *Life zone ecology*. IICA Press, San Jose, Costa Rica.
Houghton, R.A. 2005a. Tropical deforestation as a source of greenhouse gas emissions. Mountinho, P., Shwartzman, S. (Eds.) *Tropical Deforestation and Climate Change* IPAM: Balem, Brazil, and Environmental Defense, Washington DC 2005: pp. 13-21.
Houghton, R.A. 2005b. Aboveground forest biomass and the global carbon balance. *Global Change Biology* 11: 945-958.
Houghton, R.A., Lawrence, K.T., Hackler, J.L. & Brown, S. 2001. The spatial distribution of forest biomass in the Brazilian Amazon: a comparison of estimates. *Global Change Biology* 7: 731-746.
Houghton, R.A., Hall, F. & Goetz, S.J. 2009. Importance of biomass in the global carbon cycle. *Journal of Geophysical Research* 114: 1-13.
Huber, W., Weissenhofer, A. Zamora, N. & Weber, A. 2008. Plant diversity and biogeography of the Gulfo Dulce region, Costa Rica. In: Weissenhofer, A., W. Huber, V. Mayer, S. Pamperl, A. Weber. Aubrecht, G. (eds.) *Natural and Cultural History of the Gulfo Dulce Region, Costa Rica - Linz: Staphnia 88*. pp. 97-104.
Instituto Tecnológico de Costa Rica. 2008. *Atlas Costa Rica 2008.* Cartago, Costa Rica. 1 CD.
Jensen, F.V. & Nielsen, T.D. 2001. Bayesian networks and decision graphs. 2nd ed. New York: Springer.
Jordan, M.I. 1999. *Learning in graphical model.* Cambridge, MA: MIT Press, 1999.
Kalacska, M., Sanchez-Azofeifa, A., Caelli, T., Rivard, B. & Boerlage, B. 2005. Estimating leaf area index from satellite imagery using Bayesian Networks. *IEEE Transactions of Geoscience and Remote Sensing* 42(8): 1866-1873.
Keeling, H.C. & Phillips, O.L. 2007. The global relationship between forest productivity and biomass. *Global Ecology and Biogeography* 16: 618-631.
Kenfack, D., Duncan, W.T., Chuyong, G. & Condit, R. 2007. Rarity and abundance in a diverse African forest. *Biodiversity and Conservation* 16: 2045-2074.

King, D.A. 1991. Correlations between biomass allocation, relative growth rate and light environment in tropical forest saplings. *Functional Ecology* 5(4): 485-492.

Komiyama, A., Eong Ong, J. & Poungparn, S. 2008. Allometry, biomass, and productivity of mangrove forests: A review. *Aquatic Botany* 89(2): 128-137.

Lacerda, A.E.B. & Nimmo, E.R. 2010. Can we really manage tropical forests without knowing the species within? Getting back to the basics of forest management through taxonomy. *Forest Ecology and Management* 259: 995-1002.

Laumonier, Y., Edin, A., Kanninen, M. & Munandar, A.W. 2010. Landscape-scale variation in the structure and biomass of the hill dipterocarp forest of Sumatra: Implications for carbon stock assessments. *Forest Ecology and Management* 259: 505-513.

Laurence, W.F., P.F. Fearnside, P.F., Laurance, S.G., Delamonica, P., Lovejoy, T.E., Rankin-de Merona, J.M., Chambers, J.Q. & Gascon, C. 1999. Relationship between soils and Amazon forest biomass: a landscape-scale study. *Forest Ecology and Management* 118: 127-138.

Laurance, W.F., Delamônica, R., Laurance, S.G., Vasconcelos, H.L. & Lovejoy, T.E. 2000. Rainforest fragmentation kills big trees. *Nature* 404: 836.

La Torre-Cuadros, M., Herrando-Perez, S. & Young, K.R. 2007. Diversity and structural patterns for tropical montane and premontane forests of central Peru with an assessment of the use of higher-taxon surrogacy. *Biodiversity and Conservation* 16: 2965-2988.

Lippitt, C.D., Rogan, J., Li, Z., Eastman, R. & Jones, T.G. 2008. Mapping selective logging in mixed deciduous forest: A comparison of machine learning algorithms. *Photogrammetric Engineering & Remote Sensing* 14(10): 1201-1211.

Le Duc, M.G., Yang, L. & Marrs, R.H. 2007. A database application for long-term ecological field experiments. *Journal of vegetation science* 18: 509-516.

Letcher, S.G. & Chazdon, R.L. 2009. Rapid recovery of biomass, species richness, and species composition in a forest chronosequence in Northeastern Costa Rica. *Biotropica* 41(5): 608-617.

Lewis, S.L. 2006. Tropical forests and changing earth system. *Philosophical Transactions of the Royal Society of London Biological Sciences* 361(1465): 195-210.

Loh, D.K., Hsieh, Y.-T.C., Choo, Y.K. & Holtfrerich, D.R. 1994. Integration of a rule-based expert system with GIS through a relational database management system for forest resource management. *Computers and Electronics in Agriculture* 11: 215-228.

Lloyd, J. 2004. The above-ground wood productivity and net primary productivity of 100 Neotropical forest plots. *Global Change* Biology 10: 563-591.

Malhi, Y. & Grace, J. 2000. Tropical forests and atmospheric carbon dioxide. *TREE* 15(8): 332-337.

Malhi, Y., Wood, D., Baker, T.R., Wright, J., Phillips, O.L., Cochrane, T., Meir, P., Chave, J., Almeida, S., Arroyo, L., Higuchi, N., Killeen, T.J., Laurence, S.G., Laurance, W.F., Lewis, S.L., Monteagudo, A., Neill, D.A., Vargas, R.N., Pitman, N.C.A., Quesada, C.A., Salomao, R., Silva, J.N.M., Lezama, A.T., Terborgh, J., Martinez,

R.V. & Vincet, B. 2006. The regional variation of aboveground live biomass in old-growth Amazonian forests. *Global Change Biology* 12: 1107-1138.
Malhi, Y., Doughty, C. & Galbraith, D. 2011. The allocation of ecosystem net primary productivity in tropical forests. *Philosophical Transactions of the Royal Society of London Biological Sciences* 1366(1582): 3225-3245.
Maniatis, D., Malhi, Y., André, L.S., Mollicone, D., Barbier, N., Saatchi, S., Henry, M., Tellier, L., Schwartzenberg, M. & White, L. 2011. Evaluation the potential of commercial forest inventory data to report on forest carbon stock and forest carbon stock changes for REDD+ under the UNFCCC. *International Journal of Forestry Research* 2011: 1-14.
Martin, A.R. & Thomas, S.C. 2011. A Reassessment of Carbon Content in Tropical Trees. *PLoS ONE* 6(8): e23533.
Mas, J.-F., Valázquez, A., Diáz-Gallegos, J.A., Mayorga-Saucedo, R., Alcántaram, C., Bocco, G., Castro, R., Fernández, T. & Pérez-Vega, A. 2004. Assessing land use/cover changes: a nationwide multidate spatial database for Mexico. *International Journal of Applied Earth Observation and Geoinformation* 5: 249-261.
McCann, R.K., Marcot, B.G. & Ellis, R. 2006. Bayesian belief networks: applications in ecology and natural resource management. *Canadian Journal of Forest Research* 36: 3053-3062.
McGinley, K. & Finegan, B. 2003. The ecological sustainability of tropical forest management: evaluation of the national forest management standards of Costa Rica and Nicaragua, with emphasis on the need for adaptive management. *Forest Policy and Economics* 5: 421-431.
McIntosh, A.C.S., Cushing, J.B, Nadkarni, N.M. & Zeman, L. 2007. Database design for ecologists: Composing core entities with observations. *Ecological Informatics* 2: 224-236.
McMinn, J.W. & Crossley, D.A. 1993. Biodiversity and coarse woody debris in southern forests. *USDA Forest Service Report* SE-94.
Melilo, J.M., McGuire, A.D., Kicklighter, D.W., Moore, B., Vorosmarty, C.J. & Schloss, A.L. 1993. Global climate change and terrestrial net primary productivity. *Nature* 363: 234-240.
Miller, M.J. 2011. Persistent illegal logging in Costa Rica: The role of corruption among forestry regulators. *The Journal of Environment & Development* 20(1): 50-58.
Miller, S.D., Goulden, M.L., Hutyra, L.R., Saleska, S.R., Wofsky, S.C., Figueira, A.M.S., da Rocha, H.R. & de Camargo, P.B. 2011. Reduced impact logging minimally alters tropical rainforest carbon and energy exchange. *Proceedings of the National Academy of Sciences* 108(48): 19431-19435.
MINAE. 2008. Article No. 115 Decree No. 34559. Estándares de sostenibilidad para manejo de bosques naturales en Costa Rica: principios, criterios e indicadores, código de prácticas, manual de procedimientos.
Missouri Botanical Garden. 2013. TROPICOS database. URL: http://www.tropicos.org/ [Accessed October 13, 2013].
Mitchard, E.T.A., Saatchi, S.S., Baccini, A., Asner, G.P., Goetz, S.J., Harris, N.L. & Brown, S. 2013. Uncertainty in the spatial distribution of tropical forest biomass: a comparison of pan-tropical maps. *Carbon Balance and Management* 8: 10.

Muller-Landau, H.C. 2004. Interspecific and inter-site variation in wood specific gravity of tropical trees. *Biotropica* 36(1): 20-32.
Mustafa, Y.T., Stein, A. & Tolpekin, V. 2011. Improving forest growth estimates using a Bayesian Network Approach. *ASPRS 2011 Annual Conference* Milwaukee, Wisconsin May 1-5 2011.
Nasciemento, H.E.M. & Laurance, W.F. 2004. Biomass dynamics in Amazon forest fragments. *Ecological Applications* 14(4) supplement: S127-S138.
Norsys Software Corporation. Tutorial basic Netica operation. 2012. URL: http://www.norsys.com/tutorials/netica/secB/tut_B3.htm. [Accessed March 16, 2012].
Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L., Canadell, J.G., Ciais, P., Jackson, R.B., Pacala, S.W., McGuire, A.D., Piao, S., Rautiainen, A., Sitch, S. & Hayes, D. A. 2011. Large and persistent carbon sink in the world's forests. *Science* 333: 988-993.
Paoli, G.D., Curran, L.M. & Slik, J.W.F. 2008. Soil nutrients affect spatial patterns of aboveground biomass and emergent tree density in southwestern Borneo. *Oecologia* 155: 287-299.
Patenaude, G.R., Milne, R. & Dawson, T.P. 2005. Synthesis of remote sensing approached for forest carbon estimation: reporting on the Kyoto Protocol. *Environmental Science & Policy* 8(2): 161-178.
Peacock, J., Baker, T.R., Lewis, S., Lopez-Gonzalez, G. & Phillips, O.L. 2007. The RAINFOR database: monitoring forest biomass and dynamics. *Journal of Vegetation Science* 18: 535-542.
Pelletier, J., Kirby, R. & Povin, C. 2012. Significance of carbon stock uncertainties on emission reductions from deforestation and forest degadation in developing countries. *Forest Policy and Economics* 24: 3-11.
Petrokofsky, G., Homgren, P. & Brown, N.D. 2011. Reliable forest carbon monitoring — systematic reviews as a tool for validating the knowledge base. *International Forestry Review* 13(1): 56-66.
Pielke, R.A., Marland, G., Betts, R.A., Chase, T.N., Eastman, J.L., Niles, J.O., Niyogi, D.D.S. & Running, S.W. 2002. The influence of land-use change and landscape dynamics on the climate system: relevance to climate-change policy beyond the radiative effect of greenhouse gases. *Philosophical Transactions of the Royal Society London A* 360: 1705-1719.
Potter, C.S., Randerson, J.T., Field, C.B., Matson, P.A., Vitousek, P.M., Mooney, H.A. & Klooster, S.A. 1993. Terrestrial ecosystem production: a process model based on global satellite and surface data. *Global Biogeochemical Cycles* 7: 811-841.
Potvin, C. & Bovarnik, A. 2008. Reducing emissions from deforestation and forest degradation in developing countries: Key actors, negotiations and actions. *Carbon & Climate Law Review3.*
Powers, J.S. 2004. Changes in soil carbon and nitrogen after contrasting land-use transitions in Northeastern Costa Rica. *Ecosystems* 7: 134-146.
Prinzing, A., Klotz, S., Stadler, J. & Brandl, R. 2003. Woody plants in Kenya: expanding the higher-taxon approach. *Biological Conservation* 110: 207-314.
Putz, R.E., Blate, G.M., Redford, K.H., Fimbel, R. & Robinson, J. 2001. Tropical forest

management and conservation of biodiversity: an overview. *Conservation Biology* 15: 7-20.
Rangel, T. F. L. V. B., Diniz-Filho, J. A. F. & Bini, L. M. 2006. Towards an integrated computational tool for spatial analysis in macroecology and biogeography. *Global Ecology and Biogeography* 15: 321–327.
Réjou-Méchain, M., Fayolle, A., Nasi, R., Gourlet-Fleury, S., Doucet, J.-L., Gally, M., Hubert, D., Pasquier, A. & Billand, A. 2011a. Detecting large-scale diversity patterns in tropical trees: Can we trust commercial forest inventories? *Forest Ecology and Management* 261: 187-194.
Réjou-Méchain, M., Flores, O., Bourland, N., Doucet, J.-L., Fétéké, R.F., Pasquier, A. & Hardy, O.J. 2011b. Spatial aggregation of tropical trees at multiple scales. *Journal of Ecology* 99: 1373-1381.
Saatchi, S.S., Houghton, R.A., Alvala, R.C.D.S., Soares, J.V. & Yu, Y. 2007. Distribution of aboveground live biomass in the Amazon Basin. *Global Change Biology* 13(4): 816-837.
Saatchi, S., Marlier, M., Chazdon, R.L., Clark, D.B. & Russel, A.E. 2011. Impact of spatial variability of tropical forest structure on radar estimation of aboveground biomass. *Remote Sensing of Environment* 115(11): 2836-2849.
Saatchi, S.S., Harris, N.L., Brown, S., Lefsky, M., Mitchard, E.T.A., Salas, W., Zutta, B.R., Buermann, W., Lewis, S.L., Hagen, S., Petrova, S., White, L., Silman, M. & Morel, A. 2011. Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences* 108: 9899-9904.
Salimon, C.I., Putz, F.E., Menezes-Filho, L., Anderson, A., Silveira, M., Foster Brown, I. & Oliveira, L.C. 2011. Estimating state-wide biomass carbon stocks for a REDD plan in Acre, Brazil. *Forest Ecology and Management* 262: 555-560.
Segura M. & Kanninen, M. 2005. Allometric models for tree volume and total aboveground biomass in a humid tropical forest in Costa Rica. *Biotropica* 37(1): 2-8.
Sivia, D.S. 1996. *Data analysis. A Bayesian tutorial*. Oxford Science Publications, ISBN 0-19-851889-7.
SIREFOR. 2012. List of the scientific names of Costa Rican trees. *Unpublished.*
Slik, J.W., Paoli, G., McGuire, K., Amaral, I., Barroso, J., Bastian, M., Blanc, L., Bongers, F., Boundja, P., Clark, C., Collins, M., Dauby, G., Ding, Y., Doucet, J.-L., Eler, E., Ferreira, L., Forshed, O., Fredriksson, G, Gillet, J.-F., Harris, D., Leal, M., Laumonier, Y., Malhi, Y., Mansor, A., Martin, E., Miyamoto, K., Araujo-Murakami, A., Nagamasu, H., Nilus, R., Nurtjahya, E., Oliveira, Á., Onrizal, O., Parada-Gutierrez, A., Permana, A., Poorter, L., Poulsen, J., Ramirez-Angulo, H., Reitsma, J., Rovero, F., Rozak, A., Sheil, D., Silva-Espejo, J., Silveira, M., Spironelo, W., Steege, H.T., Stevart, T., Navarro-Aguilar, G.E., Sunderland, T., Suzuki, E., Tang, J., Theilade, I., Heijden, G.V.D., Valkenburg, J.V., Do, T.V., Vilanova, E., Vos, V., Wich, S., Wöll, H., Yoneda, T., Zang, R., Zhang, M.-G. & Zweifel, N. 2013. Large trees drive forest aboveground biomass variation in moist lowland forests across the tropics. *Global Ecology and Biogeography* 22(12): 1261-1271.
Smith, C.S., Howes, A.L., Price, B. & McALpine, C.A. 2007. Using a Bayesian Belief Network to predict suitable habitat of an endangered mammal - the Julia

Creek dunnart (*Smithopsis douglasi*). *Biological Conservation*: 333-347.
Sodhi, N.S., Brook, B.W. & Bradshaw, C.J.A. 2007. *Tropical conservation biology*. Oxford, UK: Blackwell Publishing.
Stanley, W.G. & Montagnini, F. 1999. Biomass and nutrient accumulation in pure and mixed plantations of indigenous tree species grown on poor soils in the humid tropics of Costa Rica. *Forest Ecology and Management* 113(1): 91-103.
Stegen, J.C., Swenson, N.G., Valencia, R., Enquist, B.J. & Thompson, J. 2009. Above-ground forest biomass is not consistently related to wood density in tropical forests. *Global Ecology and Biogeography* 18: 617-625.
Stevens, P.F. 2001. *Angiosperm Phylogeny Website Version 12*, *July 2012.* URL: http://www.mobot.org/MOBOT/research/APweb/welcome.html. [Accessed March 10, 2013].
Svob, S., Arroyo-Mora, J.P. & Kalacska, M. 2014a. The development of a forestry geodatabase for natural forest management plans in Costa Rica. *Submitted*.
Svob, S., Arroyo-Mora, J.P. & Kalacska, M. 2014b. A wood density and aboveground biomass variability assessment using pre-felling inventory data in Costa Rica. *Submitted.*
Swift, M.J., Heal, O.W. & Anderson, J. 1979. *Decomposition in terrestrial ecosystems.* Berkeley: University of California Press.
UNFCCC. 2005. Draft conclusions for Agenda Item 6: Reducing emissions from deforestation in developing countries UNFCCC/COP-11 draft decision. URL: http:/unfccc.int/resource/docs/2005/cop11/eng/102.pdf.
UNFCCC. 2007. Reducing emissions from deforestain in developing countries: approaches to stimulate action. FCCC/SBSTA/2007/L.23?Add.1/Rev.1 URL: http://unfccc.int/resource/docs/2007/sbsta/eng/123a01r01.pdf. [Accessed May 10, 2014].
UNFCCC. 2008. Reducing emission from deforestation in developing countries: approaches to stimulate action. FCCC/SBSTA/2008/L.23 URL: http://unfccc.int/resource/docs/2008/sbsta/eng/123.pdf. [Accessed May 11, 2014].
UNFCCC. 2009. Report on the expert meeting on methodological issues relating to reference emission levels and reference levels. FCCC/SBSTA/2009/2. Subisidiary Body for Scientific and Technological Advice, Thirtieth Session, Bonn June 1-10 2009. URL:http//unfccc.int/resource/docs/2009/sbsta/eng/02.pdf. [Accessed May 11, 2014].
Urquiza-Haas, T., Dolman, P.M. & Peres, C.A. 2007. Regional scale variation in forest structure and biomass in the Yucatan Peninsula, Mexico: Effects of forest disturbance. *Forest Ecology and Management* 247*:* 80-90.
Uusitalo, L. 2007. Advantages and challenges of Bayesian networks in environmental modelling. *Ecological Modelling* 203: 312-318.
Van der Werf, G.R., Morton, D.C., DeFries, R.S., Olivier, J.G.J., Kasibhatla, P.S., Jackson, R.B., Collatz, G.J. & Randerson, J.T. 1998. CO_2 emissions from forest loss. *Nature Geoscience* 2: 737-738.
Vansanto, J., Himberg, J., Helsinki, E. & Parhankangs, J. 1999. Self-organizing map in Matlab: the SOM Toolbox. *Proceedings of the Matlab DSP Conference Espo*

Finland November 16-17: 35-40.
Villaseñor, J.L., Ibarra-Manríquez, G., Meave, J.A. & Ortíz, E. 2005. Higher taxa as surrogates of plant diversity in a megadiverse country. *Conservation Biology* 19(1): 232-238.
Vitousek, P.M. & Sanford, R.L. 1986. Nutrient cycling in moist tropical forest. *Annual Review of Ecological Systems* 17: 137-167.
Wagner, F., Rutishauser, E., Blanc, L. & Herault, B. 2010. Effects of plot size and census interval on descriptors of forest structure and dynamics. *Biotropica* 42(6): 664-671.
Walton, M.E., Vay, L.L., Lebata, J.H., Binas, J. & Primavera, J.H. 2006. Seasonal abundance, distribution and recruitment of mud crabs (*Scylla spp.*) in replanted mangroves. *Estuarine, Coastal and Shelf Science* 66(3-4): 493-500.
Webb, E.L. 1999. Growth and ecology of *Carapa nicaraguensis* Aublet. (Meliaceae): Implications for natural forest management. *Biotropica* 31(1): 102-110.
Webb, E.L. & Peralta, R. 1998. Tree community diversity of lowland swamp forest in Northeast Costa Rica, and changes associated with controlled selective logging. *Biodiversity and Conservation* 7: 565-583.
Williamson, G.B. & Wiemann, M.C. 2010. Age-Dependent Increases in Wood Specific Gravity of Tropical Pioneers in Costa Rica. *Biotropica* 42(5): 590-597.
Whitmore, T.C. 1998. *An Introduction to Tropical Rain Forests.* Oxford University Press, Oxford.
Zanne, A.E., Lopez-Gonzalez, G., Coomes, D.A., Ilic, J., Jansen, S., Lewis, S.L., Miller, R.B., Swenson, N.G., Wiemann, M.C. & Chave, J. 2009. Data from: Towards a worldwide wood economics spectrum. *Dryad Digital Repository.* doi:10.5061/dryad.234.
Zeiler, M. & E.S R. Institute. 1999. Modeling our world: the ESRI guide to geodatabase design. ESRI Press.
Zhang, S.-B., Slik, J.W.F., Zhang, J.-L. & Cao, K.-F. 2011. Spatial patterns of wood traits in China are controlled by phylogeny and the environment. *Global Ecology and Biogeography* 20: 241-250.
Zolkos, S.G., Goetz, S.J. & Dubayah, R. 2013. A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. *Remote Sensing of Environment* 128: 289-298.

Appendix 1

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

Dec 08, 2013

This is a License Agreement between Sienna AP Svob ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3284280260593
License date	Dec 08, 2013
Licensed content publisher	John Wiley and Sons
Licensed content publication	Journal of Vegetation Science
Licensed content title	The RAINFOR database: monitoring forest biomass and dynamics
Licensed copyright line	2007 IAVS - the International Association of Vegetation Science
Licensed content author	J. Peacock,T.R. Baker,S.L. Lewis,G. Lopez-Gonzalez,O.L. Phillips
Licensed content date	Feb 24, 2009
Start page	535
End page	542
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	Figure 1
Will you be translating?	No
Total	0.00 USD
Terms and Conditions	

Permit for the use of Figure 1 from Peacock et al. (2007) in thesis of Sienna Svob.

Appendix 2. Natural forest management plan terminology.

Table S2. Natural forest management terminology. All terms are defined according to Forestry Law 7575 and Article No.115, Decree No. 34559 (MINAE 2008).

Term	Definition
Management unit	An area of forest belonging to one or more estates that shares a set of technical standards for the regulation of any future forest management activities
Management plan	A set of technical standards contained in a document to govern the management activities in one or more management units in order to use, conserve, and develop the existing timber resources, according to the principles of the sustainable use of renewable natural resources.
Natural forest	A native or indigenous ecosystem intervened or not, regenerated by natural succession or other forestry practices, which occupies an area ≥ 2 ha, characterized by the presence of mature trees of different ages, species, and sizes, with one or more canopies covering > 70% of the surface and > 60 trees per ha of a DBH ≥ 15 cm.
Total area	The total area of a property as defined in a cadastral map of the owner's land parcel(s). Total area may include areas under different forms of land use such as forest management or agriculture.
Managed area	The total area of natural forest being governed under a forest management plan.
Protected area	The area within the managed area classified as protected based on a set of legal regulations including areas near water bodies and on steep slopes (table x)
Productive area	The area outside of the protected area and within the managed area where timber harvest occurs.
Inventory	A diagnostic sampling of every tree with a DBH ≥ 30 cm in plots of 0.3 ha (30 m by 100 m) within the productive area. The inventory aims to understand the silvicultural variables (species, basal area, number of trees) in the forest and the proportion of trees of each species and functional group, especially within the short lived shade-intolerant group. The inventory attempts to sample trees below the minimum cutting DBH to permit the sustainable use and regeneration of the commercial species.
Census	Along transects separated by 50 m to 100 m every tree with a DBH ≥ 60 cm is recorded within the entire productive area. The census aims to identify the commercial timber stock. The results of the census are used to plan future logging operations.
Baquiano	A local expert knowledgeable of the tree community composition responsible for identifying trees by common names during the census and inventory of the productive forest area.
Parent tree	A commercially viable species that is selected based on a set of phenotypic characteristics for seed production. These characteristics include being physiologically mature, phytosanitary (i.e. no pests or pathogens), and an adult and having a good form, well-developed crown, and straight trunk.
Remnant tree	A tree left in the forest after harvest and having the necessary features (mostly free of damage) to remain part of the residual mass for the next harvest.
Low frequency species	A species with an abundance of less than 0.3 trees per hectare of trees with a DBH ≥ 30 cm based on the inventory. Low frequency species cannot be logged.