CORDILLERAN FOREST SCALING DYNAMICS AND DISTURBANCE REGIMES

QUANTIFIED BY AERIAL LIDAR

by

Tyson L. Swetnam

_____

A Dissertation Submitted to the Faculty of the

SCHOOL OF NATURAL RESOURCES AND THE ENVIRONMENT

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2013

UMI Number: 3605921

UMI
Dissertation Publishing

UMI 3605921

ProQuest

THE UNIVERSITY OF ARIZONA

GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Tyson L. Swetnam titled, "Cordilleran forest scaling dynamics and disturbance regimes quantified by aerial LiDAR" and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

_____ Date: 08/23/2013

Donald A. Falk

_____ Date: 08/23/2013

Brian J. Enquist

_____ Date: 08/23/2013

Ann M. Lynch

_____ Date: 08/23/2013

D. Phillip Guertin

_____ Date: 08/23/2013

Stephen R. Yool

_____ Date: _____

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

_____ Date: _____

Dissertation Director: Donald A. Falk

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Tyson L. Swetnam

DEDICATION

For my eldest child, Alayna Rose Swetnam

Contents

ABSTRACT

Semi-arid forests are in a period of rapid transition as a result of unprecedented landscape scale fires, insect outbreaks, drought, and anthropogenic land use practices. Understanding how historically episodic disturbances led to coherent forest structural and spatial patterns that promoted resilience and resistance is a critical part of addressing change. Here my coauthors and I apply metabolic scaling theory (MST) to examine scaling behavior and structural patterns of semi-arid conifer forests in Arizona and New Mexico. We conceptualize a linkage to mechanistic drivers of forest assembly that incorporates the effects of low-intensity disturbance, and physiologic and resource limitations as an extension of MST. We use both aerial LiDAR data and field observations to quantify changes in forest structure from the sub-meter to landscape scales. We found: (1) semi-arid forest structure exhibits MST-predicted behaviors regardless of disturbance and that MST can help to quantitatively measure the level of disturbance intensity in a forest, (2) the application of a power law to a forest overstory frequency distribution can help predict understory presence/absence, (3) local indicators of spatial association can help to define first order effects (e.g. topographic changes) and map where recent disturbances (e.g. logging and fire) have altered forest structure. Lastly, we produced a comprehensive set of above-ground biomass and carbon models for five distinct forest types and ten common species of the southwestern US that are meant for use in aerial LiDAR forest inventory projects. This dissertation presents both a conceptual framework and applications for investigating local scales (stands of trees) up to entire ecosystems for diagnosis of current carbon balances, levels of departure from historical norms, and ecological stability. These tools and models will become more important as we prepare our ecosystems for a future characterized by increased climatic variability with an associated increase in frequency and severity of ecological disturbances.

1. INTRODUCTION

1.1 Background

1.1.1 The Carbon Cycle

The Earth system's carbon cycle has been fundamentally altered by the reintroduction of large amounts of fossil carbon into the environment by humans (Prentice et al. 2001, Hansen et al. 2008, Parry et al. 2008). Carbon in the atmosphere, as carbon dioxide (CO_2) or methane (CH_4) acts as a greenhouse gas (GHG), where it absorbs infrared radiation during the day and reemits that radiation at night. This heating leads to decreased atmospheric cooling at night, and increased air temperatures during the day (Solomon 2007, Stocker et al. 2013). The dominant GHGs in the atmosphere are: water vapor (H_2O), nitrous oxide (N_2O), ozone (O_3), methane (CH_4), and carbon dioxide (CO_2). Atmospheric concentrations of CO_2 surpassed 400 parts per million (ppm) in early 2013, and are continuing to increase annually at $\sim 2 \frac{ppm}{yr^{-1}}$ (Solomon 2007, Füssel 2009, Stocker et al. 2013). The 400 ppm value is substantially higher than the historical average of 280 ppm over at least the last 400 thousand years (Kyr^{-1})(Neftel et al. 1985, 1994, Barnola et al. 2003, Lüthi et al. 2008), with levels not seen as high for at least the last 3.5 million years (Myr^{-1}) (Brigham-Grette et al. 2013). I refer to the consequence of humans altering the Earth's climate by changing the dynamics of the carbon, water, and nitrogen cycles as 'anthropogenic climate change'.

The Earth's dominant carbon sinks are (1) the oceans and (2) the terrestrial biosphere. Historically these sinks sequestered a sustained amount of the annual emitted CO_2 in synchrony with biogenic and geologic emission sources. At present the CO_2 sequestration capacitance of the oceans are in decline due to an increase in the carbonic-acid content of sea water ($CO_2 + H_2O \rightleftharpoons H_2CO_3$). Troublingly, as the oceans' pH falls their capacity to sequester carbon is further reduced (Caldeira and Wickett 2003). Simultaneous to the degradation of the sequestration rate and capacitance of the oceans the terrestrial biospheres' sequestration rate and potential are also being reduced. Much of the Earth's arable land surface area has been converted to agriculture and pastoral uses, with the total

area modified by humans now making up a significant portion (~34%) of the total surface (Foley et al. 2005, Ramankutty et al. 2008). Forests and soils are the two largest terrestrial carbon sinks; these include longer term depositions of carbon which leave the active carbon cycle [e.g. the formation of peat and coal]. Also, as part of the water cycle dissolved organic carbon from forests and forest fires travel down large rivers to the ocean where it is deposited in sediment (Jaffé et al. 2013, Masiello and Louchouarn 2013). Importantly, forests can act as both sinks and sources depending on age, stand density, and type of disturbance regime. Recently, boreal, temperate, and semi-arid forests have increasingly been impacted by wildfires (Westerling et al. 2006), drought (van Mantgem et al. 2009, Allen et al. 2010, Williams et al. 2012), and insect pests (Raffa et al. 2008, Negron et al. 2009), the result being many recent carbon sinks have turned into sources with large amounts of CH_4, and CO_2 released as woody decomposition occurs. Forest fires are also believed to contribute to a positive feedback loop (Ramanathan and Carmichael 2008) related to increasing temperatures of the atmosphere (Stocks et al. 1998, Cox et al. 2000, Friedlingstein et al. 2006).

Scientists have repeatedly and in near consensus warned continued emissions levels will lead to increased warming and the catastrophic collapse of the Earth's cryosphere (Pachauri and Reisinger 2007, Hansen et al. 2008). The ensuing rise in sea surface levels will inundate coastal cities and flood large portions of agricultural lands along the coasts; meanwhile increased air temperatures will have profound effects on ecosystem services and human health (Hansen et al. 2008, Parry et al. 2008, Burkett and Davidson 2012). To slow down or reverse anthropogenic climate change humanity must move toward a sustainable CO_2 emissions scenario where the atmospheric CO_2 concentration decreases below 350 ppm (Hansen et al. 2008). Reductions in the use of fossil fuel for energy generation, infrastructure development, industry, and transportation have been suggested. Adoption of sustainable agriculture and forestry practices that sequester more carbon are also a critical first step toward lowering humanity's C footprint.

1.1.2 Forest management

Forest management and ecosystem restoration in the 21st century are major challenges for managers and stake-holders in the face of anthropogenic climate change. Anthropogenic climate change is suspected to be altering the frequency and intensity of weather events (Solomon 2007, Pachauri and Reisinger 2007, Füssel 2009), increasing the length and severity of (1) wildfire seasons (Westerling et al. 2006), (2) insect outbreaks (Logan et al. 2001, 2003, Kurz et al. 2008, Anderegg and Calloway 2012), and (3) drought induced mortality in forests (Breshears et al. 2005, 2008, Allen and Breshears 2007, Adams et al. 2009, McDowell 2011, Anderegg et al. 2012, Willams et al. 2012). In western North America a century of logging, rural and urban development, and policy toward forest management [e.g. the 10AM policy of fire suppression (Swetnam 1990)] have further sent forests away from their historical ranges of variability (HRV) (Morgan et al. 1994) that were once maintained in many geographic regions by frequent low-intensity wildfires. The continued development of fossil energy (Lyon and Anderson 2003, Sawyer et al. 2006) and increased building of residential homes in the rural and ex-urban interface (Cohen 2000, Radeloff et al. 2005) has further fragmented forests, putting humans at risk, and altering the management strategies involved in the event of disaster. As disturbances continue to modify and shift ecosystems away from their HRV the potential for complete type-conversions across 'tipping-points' (Lenton et al. 2008, Adams 2013) are moving these systems toward configurations with lower-complexity and decreased diversity.

Recent management policies [i.e. fire suppression and wide-spread logging] are partly responsible for the observed change in fire number and size of fires, specifically in western North America; however the increase in the timing and intensity of fires is directly related to an increase in mean temperatures and decreases in the level of fuel moistures during the fire season and those effects can be directly linked to climate (Westerling et al. 2006).

1.1.3 Monitoring landscape change

In the present study I evaluate the condition of semi-arid forests in Arizona (AZ) and New Mexico (NM) through a combination of plot observations and remotely sensed data. I also present allometric scaling models for estimating total above ground biomass (AGB) across five distinct forest types and ten common species that dominate across AZ and NM.

Monitoring ecosystem scale patterns and processes requires very large datasets. Forest population studies by convention rely upon plot based field surveys because to inventory an entire forest is both economically and physically impractical. Measurement of individual trees at ecosystem scale, an impossible task a decade ago, can now be completed by aerial Light Detection and Ranging (LiDAR), in only a few days (Van Leeuwen and Nieuwenhuis 2010, Ustin and Gamon 2010). Aerial LiDAR provides inventory-scale information at costs that become fractional to those of field surveys when considered at the landscape scale.

Monitoring the pools and fluxes of aboveground-biomass (AGB), below-ground biomass (BGB), and above ground carbon (AGC) will become a vital component of managing ecosystems in the future (Turner et al. 2004). Globally, forests make up approximately 18% of the terrestrial carbon sink and in the United States forests sequester 12-19% of annual total carbon output (Ryan et al. 2010); in the Southwestern US semi-arid forests account for 20% of the total land area (61.1 million hectares), contributing to the bulk of the region's carbon sink and storage capacity.

Conservation efforts to reduce greenhouse gas emissions and encourage sustainable forests practices include the United Nations Reducing Emissions from Deforestation and Forest Degradation (UN-REDD) and nationally-led REDD+ initiatives (Gibbs et al. 2007, Goetz and Dubayah 2011). Along with efforts in the United States (Bachelet et al. 2001, Ryan et al. 2010, McKinley et al. 2011) such as the Forest Inventory Analysis (FIA) program (Birdesy et al. 2004, Smith et al. 2006, Woodbury et al. 2007) scientists have already begun to report on biomass and C flux at the regional and national scale. The FIA serves as a baseline for the continual monitoring of changes

in sequestered carbon by biomass over time and space (Blackard et al. 2008, Sundquist et al. 2009, Ryan et al. 2010, McKinley et al. 2011). However, the spatial resolutions of these projects are coarse and their estimates of sequestered C and C flux vary significantly. If forested countries, including the United States, intend to monitor and manage carbon effectively at the local scale, better estimates and models are needed.

We must also consider sequestered carbon in temperate semi-arid forests remains for centuries longer as biomass and in soils than in tropical forests. These slower growing carbon sinks are at a disproportionately increased risk from the increasing rates of disturbance i.e. wildfire, drought, and pathogens driven by anthropogenic climate change. It is time critical that we identify which forest types are most at risk, and where old growth stands are located so that we can mitigate these effects.

1.2 Analysis techniques

Power-law spectra in nature result when a steady-state disintegrates [from minor perturbation] toward a more intermediate number of phenomena under observation (Bak et al. 1987). Self-organized critical systems (SOC), power laws, chaos, and fractals share a common property of self-similarity where there is invariance in the scale or size of the phenomena under observation (Mandelbrot 1983, Bak et al. 1987). Below I briefly define the concepts of scale, SOC, and state space in-between SOC with random distributions.

Semantically, the word 'forest' is used to describe any area that has trees, regardless of cover percentage or total tree height. I considered the 'forest' level to be the next order above the 'stand' level, which is above the 'patch' level, and is resolved at the fundamental unit of the 'individual' (Table 1). Anecdotally, I also include a scale reference for time in Table 1, suggesting that as area increases the time it takes to completely change or modify the scalar measure also increases.

Table 1: Defining spatial and temporal scales of forests, inspired by Pianka (2011, Fig.1).

	Smallest	→	Largest				
Spatial (m^2)	Individual	→	Patch	→	Stand	→	Forest
Scalar	$10^{-3} - 10^2$		$10^0 - 10^4$		$10^3 - 10^6$		$10^5 - 10^{11}$
Temporal (yr^{-1})	Individual	→	Cohort	→	Community	→	Population

Unoccupied state space represents a possible Hutchinsonian niche (Hutchinson 1959, Holt 2009) for a species to occupy. The niche concept is supported by other ecological theories such as diversity-disturbance relationships and the intermediate disturbance hypothesis (Connell 1978, Dial and Roughgarden 1988, Wilkinson 1999, Miller et al. 2011, Hall et al. 2012) suggesting that systems with some disturbance may actually increase their relative diversity by creating a spectrum of structure sizes and patch configurations that support more niches. Species area relationships (Harte et al. 1999) and diversity (Morse et al. 1985) have been shown to follow a self-similar distribution that mirrors the fractal dimension of the surfaces on which the species exist. Kellner and Asner (2009) also reported the frequency of canopy gap openings in a forest follow a power law equivalent to a zeta function. This suggests a forest with complex structure and variability across space [e.g. a natural landscape with multiple disturbance histories] may in fact harbor more species than one with a uniform size frequency distribution [i.e. a tree plantation], when other factors are controlled.

In forests with multiple disturbance types [e.g. drought, fire, and insects] a single SOC state may not be achieved, and if it is it will be broken down into 'heterogeneous' [i.e. incoherent or random] patterns over successive events. If disturbance is rare and the rate of the system to recover fast than the processes that dominate the forest will likely involve individuals undergoing density dependent competition (Enquist et al. 1998, 2009, Enquist and Niklas 2001, Niklas et al. 2003, West et al. 2009). Solé and Manrubia (1995) reported that the canopy gaps in a tropical rain forest canopy are indistinguishable from the gap openings generated by the output of a SOC 'forest-game' automata program. They also proposed that rain forests self-maintain in the same way as the sand pile model,

and thus exhibits scaling by SOC as suggested by Bak et al. (1987). Enquist et al. (2009) showed that the SFD of diameter in tropical rain forests exhibit power-law behavior and that these conditions are maintained over time (Enquist et al. 1998, 1999, 2009, Brown et al. 2004, West et al. 2009). Solé and Manrubia's (1995) study, and the West, Enquist, and Brown (WBE) rainforest work were all done on the Barro Colorado Island Experimental forest in Panama where detailed observations of tropical forest structure are uniquely available.

Spatial heterogeneity is a common denominator for western US semi-arid forest stand structure (Fulé 1997, Mast et al. 1999, Sánchez-Meador et al. 2011, Larson and Churchill 2012). The mechanisms by which this forest structure develops are still poorly defined in the literature. I incorporate terms that universally describe mathematical and physical properties of self-similarity and self-organized phenomena and relate these to the configurations of trees [that form 'forests'] in space.

A system that exhibits SOC dynamically maintains itself in a 'critical state' between incremental developmental stage increases and discrete event decreases (Bak et al. 1987, Newman et al. 2005). In dynamical systems these critical points, or attractors, display scale-invariance in both space and time and have been explained as a function of $1/f^{\beta} \equiv f^{-\beta}$ 'noises' (Bak et al. 1987, Mandelbrot 2002). Generally, SOC is typified by small perturbations or avalanches that maintain a log-log slope, such as the famous 'sand pile' model, where the dynamic exponent β is constant over time (Bak et al. 1987). SOC is the corollary of a 'steady-state'. Steady state and SOC are typically considered independent of one another as SOC arises from a condition of the steady-state (Bak et al. 1987, Hwa and Kardar 1989). Frequently used examples of SOC in nature include earthquakes, forest fires, and financial markets (Bak et al. 1987, 1991, Drossel and Schwabl 1992, 1994, Reed and McKelvey 2002, Newman 2005, Schoenberg and Patel 2012). Importantly, developmental increases in a SOC system lead to a corresponding increase in the magnitude of its disturbances (Newman 2005). For example, forest growth simulations simplified as 'percolation' models (Drossel and Schwabl 1992, 1994, Newman 2005) if allowed to synchronize across the entire model extent will pass a

critical point or phase transition where pixels become inter-connected (i.e. fuels become closely oriented in space allowing a contagion to spread); the consequence is a fire that burns the entire model (Newman 2005). Fire sizes in nature, however, are scale-dependent and limited to the area of fuels – i.e. an event cannot grow larger than the total vegetated area, and as scale increases the observed number of very large fires begins to decrease, creating a divergence from a heavy-tailed Pareto distribution (Newman 2005, Schoenberg and Patel 2012). This is also the case of numerous other phenomena, including earthquakes, and importantly to my work, tree size.

'Synchrony' is another important property of forests and is explained as the simultaneous similarity of a developmental stage across a cohort or community [in time] and patch or stand [in space]. Synchrony can describe a $1/f^\beta$ noise temporally at short scale with high frequency [e.g. annual seedling recruitment], and at long scale and low frequency [e.g. an old growth grove of trees that occupy the side of a mountain for millennia]. Importantly, synchrony is a continuous function (from small to large size), and is spatially independent, though it is likely to be entrained by first order spatial effects such as topographic relief and type of parent material.

'Heterogeneity' is a broadly defined term that explains differences across more than two measures. A forest is considered to be 'heterogeneous' if it has different types of trees or different configurations of patch sizes with different sized trees within a defined area. Thus, heterogeneity is an intrinsically scaled concept. Heterogeneity is in a sense equivalent to fractal lacunarity, from the Latin *lacuna* for 'pool' or 'gap', also described as the 'porosity' of an object (Mandelbrot 1983, 2002, Halley et al. 2004). When considered at a defined scale [e.g. the size frequency distribution of overstory vs. understory] the presence or absence of understory has no effect on the distribution of trees in the overstory. Lacunarity is manifest at large scale in natural openings between stands of trees [i.e. meadows or rock-slides] and at small scales where a gap forms after a canopy tree falls and even within trees where gaps allow sun flecks to pass through a leafy canopy and illuminate a patch of forest floor.

1.2.1 Formulae and derivation of the Pareto distributions

SOC phenomena can be described using power laws. In 1896 Vilfredo Pareto described the behavior of power-laws with a probability distribution that takes his name.

The Pareto shape parameter α is defined as any positive integer that can undergo factorial multiplication of a real number x, thus x^α is defined as x multiplied α times of itself. When α is negative, i.e. $x^{-\alpha}$, the equation can also be written: $\frac{1}{x^\alpha}$. The scale parameter β determines the scale or statistical dispersion of the probability distribution. If β is large, then the range of the function is large and vice versa when it is small.

Newman (2005) gives a full derivation for determining the exponent α from a power-law distribution. The derivation, or rate of change in the dependent variable (y) over the span of the independent variable x (possibly space or time), can be written as a function (f) of the two variables $y = f(x)$.

When $y = \alpha$, the slope of the relationship is 0:

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} 0 = 0$$

where the derivative of $f(x) = 0$; thus the application of a correlation coefficient (a constant) will have no effect on the derivative of the function. For x^α if $y = f(x) = \alpha x$, than the slope of the function α is written:

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{\alpha \Delta x}{\Delta x} = \alpha$$

Thus the derivative of $f(x) = \alpha x$ is α.

In the case of a function with more than one independent variable, as is the case for the tapered and truncated Pareto: *partial derivatives* must be applied. The total derivative $\frac{dy}{dx}$, is broken into partial derivatives, where ∂ replaces d. The total derivative is

evaluated in parts for each individual variable while the other independent variables are held constant:

$$\frac{dy}{dx} = \left(\frac{\partial y}{\partial x}\right)_{\gamma,\tau} dx + \left(\frac{\partial y}{\partial \gamma}\right)_{x,\tau} d\gamma + \left(\frac{\partial y}{\partial \tau}\right)_{x,\gamma} d\tau$$

here γ and τ are additional independent variables.

For a probability density function (PDF) represented by $f(x)$, where P is the probability of a certain number of observed events, and is $0 \leq P \leq 1$, the derivation can be defined as the interval of values of the independent variable where the total probability is:

$$P = \int_{x_m}^{\beta} f(x)dx$$

where x_m, and β represent the limits of the interval and $f(x)$ is the PDF.

The likelihood function explains the parameters of a statistical model and is the same as the 'probability' in statistical inference. The likelihood function of a Pareto distribution for the parameters α and x_m, given a sample $x = x_1, x_2, x_3, \ldots, x_n$ is:

$$L(\alpha, x_m) = \prod_{i=1}^{n} \alpha \frac{x_m^{\alpha}}{x_i^{\alpha+1}} = \alpha^n x_m^{n\alpha} \prod_{i=1}^{n} \frac{1}{x_i^{\alpha+1}}.$$

The estimator for α by its partial derivative becomes zero at:

$$\frac{\partial l}{\partial \alpha} = \frac{n}{\alpha} + n \ln x_m - \sum_{i=1}^{n} \ln x_i = 0.$$

The Pareto is also related to the exponential distribution if $F(x)$ with a minimum β and exponent α equals $Y = \log(\frac{f(x)}{\beta})$ where α is the rate parameter. Equivalently, if Y is an exponential distribution: βe^{Y} is Pareto with a minimum β and exponent α.

For an independent identically distributed (i.i.d) set of observations the log-likelihood function of the tapered Pareto takes the form (from Kagan and Schoenberg 2001):

$$\log L(\beta, \theta) = \sum_{i=1}^{n} \log\left(\frac{\alpha}{x_i} + \frac{1}{\theta}\right) + \alpha n \log \beta - \alpha \sum_{i=1}^{n} \log x_i + \frac{\alpha n}{\theta} - \frac{1}{\theta}\sum_{i=1}^{n} x_i.$$

Setting the α parameter to zero for the derivative of $\log L$ in respect to θ gives:

$$\frac{\theta}{n}\sum_{i=1}^{n}\frac{x_i}{\alpha\theta + x_i} = \bar{x} - \beta$$

and

$$\theta\sum_{i=1}^{n}\frac{x_i}{\alpha\theta + x_i} = \sum_{i=1}^{n}\log x_i - n\log\beta$$

Where $\bar{x}$ is the sample mean: $(x_1 + x_2 + \cdots + x_n)/n$.
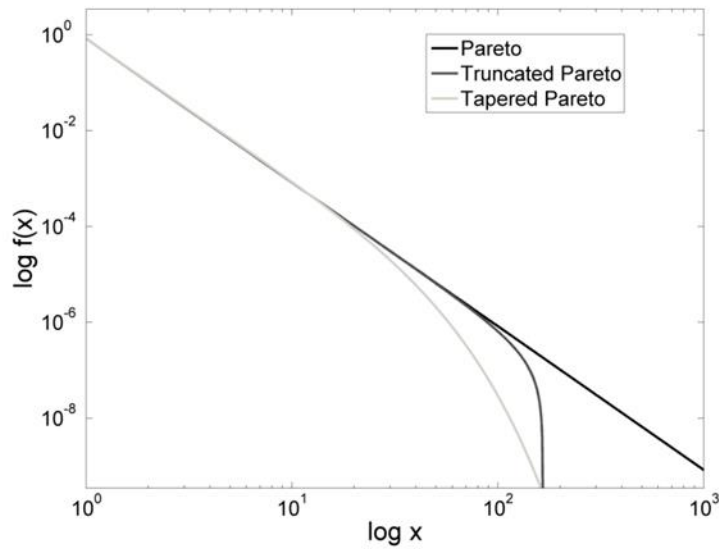
Figure 1: Probability Density Functions of the Pareto, Tapered, and Truncated Pareto distributions.
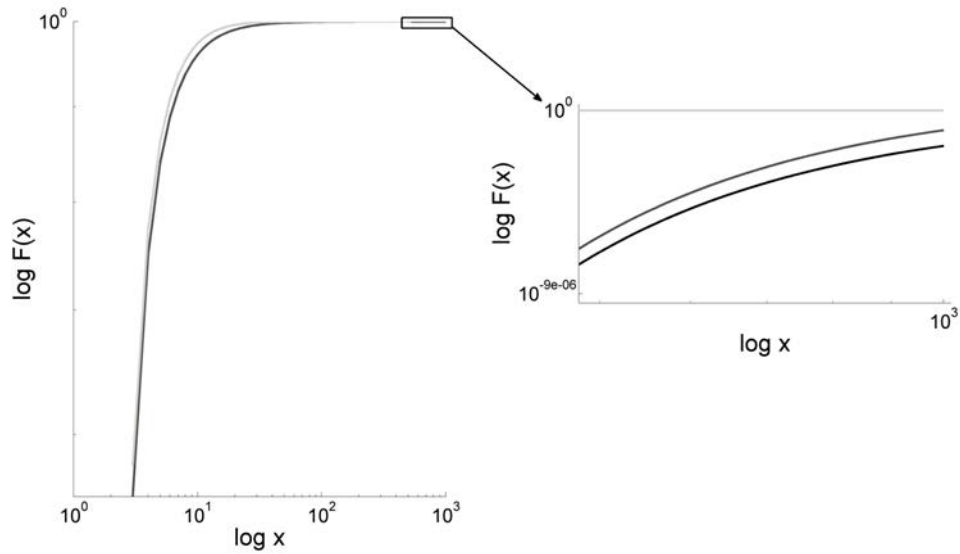
Figure 2: Cumulative distribution function of the Pareto, Tapered, and Truncated Pareto distribution.

1.2.2 Least-squares regression

Least-squares regression is a common technique for determining the best-fit of a function using an "explanatory" or "independent" (input or cause) variable and the "response" or "dependent" (output or effect) variable. My coauthors and I used least-squares regression analysis to determine the parameters of both power law and Pareto distributions in our study data in MATLAB 2012B Curve Fitting Tool (The Mathworks 2012). For the most part we left the default settings unchanged when estimating parameters in Curve Fitting Tool. In some cases the Curve Fitting Tool could not find a fit and we were forced to modify the default parameters. When we did so, we set the power law or Pareto shape parameter α to values that were predicted by MST, and the scale parameter β we set *a priori* based on the point at which the distribution experienced a cross-over (e.g. from positive to negative trend), in the case of the truncated Pareto distribution's upper truncation parameter ω we set the value to the observed maximum size.

To obtain a correlation coefficient least-squares uses the residual of the ith data point r_i, defined as the difference between the observed value y_i and the fitted value $\hat{y}_i$ which is identified as the error within the data:

$$r_i = y_i - \hat{y}_i$$

The summed square of the residuals are given by:

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

where n is the number of data points included in the fit and S is the sum of squares estimated error. The assumptions made about the error in the data include error in only the response, not the predictor, and that they are normally (Gaussian) distributed with zero mean and constant variance:

$$error \rightarrow N(0, \sigma^2)$$

where σ^2 is the constant variance. The matrix form of the non-linear model is given by the formula:

$$y = f(\boldsymbol{X}, \boldsymbol{\beta}) + \varepsilon$$

Where y is an $n - by - 1$ vector of responses, f is a function of β and $\boldsymbol{X}$, $\boldsymbol{\beta}$ is an $m - by - 1$ vector of coefficients, $\boldsymbol{X}$ is the $n - by - m$ design matrix for the model, and ε is an $n - by - 1$ vector of errors. MATLAB's CFT uses a 'Trust-region' algorithm that first starts with an initial estimate for each coefficient; model coefficients were set to reasonable starting values and constrained boundaries to ensure a fit.

1.2.3 Maximum Likelihood Estimation

White et al. (2008) found that by fitting a slope to continuous data that has been linearly binned the wrong exponent may be reported and suggest that maximum likelihood estimation (MLE) should be used in place of ordinary least squares for

determining dynamic exponents. We fit our data's continuous distributions by MLE in MATLAB with the algorithms developed by White et al. (2008).

For a Pareto distribution the MLE of the dynamic exponent α is:

$$\alpha = -1 - \left[\frac{1}{n \sum_{i=1}^{n} \log(k_i)} \right]^{-1}$$

The MLE of α for an upper-truncated-Pareto is:

$$\overline{\ln k} = -\frac{1}{\alpha + 1} + \frac{b^{\alpha+1} \ln b - \gamma^{\alpha+1} \ln \alpha}{b^{\alpha+1} - \gamma^{\alpha+1}}$$

for $\tilde{\alpha}$: where γ and v are known:

$$\frac{n}{\tilde{\alpha}} + \frac{n \left(\frac{\gamma}{v} \right)^{\tilde{\alpha}} \ln \left(\frac{\gamma}{v} \right)}{1 - \left(\frac{\gamma}{v} \right)^{\tilde{\alpha}}} - \sum_{i=1}^{n} \left[\ln K_{(i)} - \ln \gamma \right] = 0$$

Details on the algorithms used in the appendix manuscripts can be found in White et al. (2008) and online at http://whitelab.weecology.org/software-powerlaw-distributions.

Schoenberg and Patel (2012) suggest that MLE is unstable when β is unknown and because the θ parameter is dependent upon β it too would be unstable. Because we impose the threshold parameters from the data the estimation of β is not necessary.

1.3 Research Objectives

The objective of this dissertation is to contribute to a growing methodology for measuring forests in arid and semi-arid regions with new technology and to incorporate recently developed theories on individual to ecosystem assembly and function.

Aerial LiDAR presents us with several game-changing advancements for forestry applications: (1) in scale it surpasses any kind of field inventory and (2) it gives us the ability to quantify landscapes like never before. Within a geographic information system (GIS) we can quantify across a vast virtual space the measurements taken from thousands of trees at hundreds of measured forest plots and extrapolate those observations to tens of millions of individuals [i.e. the LiDAR segmented tree inventory] across entire landscapes.

Metabolic Scaling Theory (MST) (Enquist et al. 1998, 1999, 2003, 2009, Brown et al. 2004, West et al. 1999, 2009) gives us the ability to quantitatively measure the fractal-like structure and metabolic behavior of trees and their spatial assembly in forests. By incorporating scale invariance into our analyses we quantitatively diagnose the health of forests from the individual to forest scale.

This research aimed to answer four broad questions:

1) How can we address disturbance caused divergent, i.e. non log-log linear, scaling behavior when measuring forests at ecosystem scale?

To address this question my coauthors and I first conduct a brief overview of MST, general tree allometry, and the forest self-thinning rule. Next we discuss the ways in which disturbance can modify steady-state forest structure [operating with a self-organized criticality (SOC)] and in what way these changes are expressed in the size frequency distribution (SFD) of the population at different scales.

2) Can airborne laser swath mapping derived overstory structure help us predict the composition of understory in semi-arid forests?

To address this question my coauthors and I first developed a variable area local maxima segmentation algorithm that incorporates MST predictions of canopy allometry to isolate tall-to-medium height trees in dense stands and all trees in open canopy architectures. Because semi-arid forests are mostly made up of a single overstory canopy layer with conical shaped canopy geometries the algorithm does an excellent job of differentiating individuals from one another. The next step was to fit the distribution functions of both the overstory and the understory from observed field inventories to a scale invariant tapered Pareto distribution to see if there was a positive correlation between the two.

3) Can we use local indicators of spatial association (LISA) for individual trees to discriminate past disturbances that may or may not have been recorded in the historical record?

To address this question my coauthors and I took the results of the tree segmentation algorithm that we describe in Appendix B and conducted a LISA analysis on the points [i.e. trees]. The local statistics tell us information about the neighboring individual trees and the individual in relation to its neighbors. My coauthors and I demonstrate that in sections of forest that were either logged or burned recently the legacy of disturbance on the forest remains distinct from other sections that did not experience the same disturbance.

4) How much biomass and carbon are on the forest landscapes in the aerial LiDAR datasets?

To answer this question my coauthors and I developed allometric equations from an extensive forest plot network in Arizona and New Mexico, defining 5 distinct forest types with eight of the most common tree species. Because aerial LiDAR can measure both stand scale characteristics and individual trees we created grid-scale and individual tree-scale models for estimating the above ground biomass (AGB) and carbon density (ACD). My coauthors and I include an in-depth review of historical observations of AGB and ACD in western forests, as well as a discussion of how to apply these observations to aerial LiDAR. We also include a brief discussion on estimating below ground biomass (BGB), although we did not physically measure BGB directly in the study. Current LiDAR cannot directly discriminate between species. However, future technology appears poised to make this leap with hyper-spectral sensing inventory of plant 'spectranomics'. Other uses include discriminating leaves from bark or branches, and grasses from dead litter or other ground features. In anticipation of this technological leap my coauthors and I also report individual species allometric scaling models for AGB and ACD.

1.4 Dissertation Format

Following the INTRODUCTION is the "PRESENT STUDY" which provides a summary of the analyses and conclusions of the five papers prepared for peer-review publication. The five articles included as appendices are as follows:

Appendix A: "The generality of metabolic scaling theory in differing forest disturbance regimes." In this paper my coauthors and I conceptualize and summarize the scaling behavior of forests and discuss how disturbance, and physiological and resource limitation affect power-law scaling behavior in relation to metabolic scaling theory.

Appendix B: "Application of metabolic scaling theory to reduce error in local maxima segmentation from aerial LiDAR". In this paper my coauthors and I utilize predictions of the MST to generate virtual inventories of forest overstory populations and test those against observational inventory data. My coauthors and I also compare other published segmentation models and discuss the strengths and weaknesses of those studies in regard to the MST.

Appendix C: "Predicting rank-size distributions of forest mid and understory in two semi-arid conifer forests with an aerial LiDAR derived forest inventory." In this paper my coauthors and I demonstrate a way to predict the number of trees in forest understory with an airborne LiDAR data set by fitting a power law function to only the overstory data, my coauthors and I also suggest why density-dependence and scale-invariance of forests described in Appendix A makes this possible.

Appendix D: "Characterizing forest structure and disturbance legacies with aerial LiDAR using Getis-Ord G_i^* and Anselin Moran's I_i local indicators of spatial association." In this paper my coauthors and I demonstrate how local-statistics reveal where forests change structurally, emphasizing the effect of disturbance and first-order [topographic variability] controls on vegetation spatial pattern.

Appendix E: "Biomass and carbon models for Aerial Laser Swath Mapping inventories in Southwestern North American conifer forests." In this paper my coauthors and I synthesize twenty two above ground biomass and carbon models at various levels of complexity [universal generalized, study area generalized, forest

type, and species specific] and scalars [grid and point scale] from field measured tree allometry and aerial LiDAR data over three study areas in Arizona and New Mexico.

2. PRESENT STUDY

The methods, results, and conclusions of this study are presented in the papers appended to this dissertation. The following is a summary of the most important findings of each of the papers.

2.1 Appendix A: THE GENERALITY OF METABOLIC SCALING THEORY IN DIFFERING FOREST DISTURBANCE REGIMES

My coauthors and I were interested in whether the size frequency distribution (SFD) of trees in semi-arid forests that experience episodic disturbance, i.e. wildfires, exhibit scaling in size and frequency that is predicted by the MST (West et al. 1999, 2009, Enquist et al. 1998, 2009, Kerkhoff and Enquist 2007). We based our null assumption on data from tropical forests (Condit et al. 1999, West et al. 2009, Enquist et al. 2009) that do not experience similar exogenous disturbances, and assumed the controls on tree and ecosystem growth rates in semi-arid forests are set by MST. We looked at both plot data and aerial LiDAR derived SFD to determine whether MST-type scaling exists in the study systems. Our questions were: (1) do semi-arid forests exhibit any sort of power-law SFD for their primary size measures and are those similar or identical to MST; and (2) if semi-arid forests do exhibit power-law SFD, when and where does this occur; or (3) if semi-arid forests do not exhibit power-law SFD how can we describe their distribution?

We found: (1) semi-arid fire-frequent forest types in Arizona and New Mexico do exhibit MST scaling, albeit at different spatial scale. The scaling is limited across scalar measures: small areas that have not recent experienced disturbance exhibit MST in the smallest observed size classes, but when larger spatial extents are considered MST scaling is only apparent in larger size classes. This suggests disturbance eliminates small individuals up to a minimum size; whereafter individuals large enough to be physically insulated from low-intensity fire with thick bark or high canopy base height survive and fill space in MST fashion. (2) Semi-arid forest SFD exhibit divergence in the upper-tail as the consequence of exogenous stochastic noise [i.e. disturbance], and physiological

and resource limitation [i.e. maximization of soil available water in the tree's exclusive root-footprint limits maximum amount physiologically determined sustainable biomass]. The divergence is addressed by fitting a truncated Pareto function to the distribution versus the heavy-tailed Pareto distribution which would tend to over-predict the number of individuals at large size. (3) Divergence in the lower-tail is a consequence of increasing the lacunarity (gap spacing) in the canopy and can be described when accounting for unoccupied space in the maximum density dependence of the stand.

These results suggest semi-arid forests are continually moving toward resource steady-state and that disturbance effects the SFD non-linearly based on its intensity and the existing forest structure. The effect of low severity disturbances can be modeled as either a 'contamination' of power-law behavior in the system or as a change in the normalization constant [i.e. the maximum density of the possible SFD] with increasing lacunarity. The reduced value of the constant therefore becomes a quantitative measure of how 'disturbed' the forest is, versus its maximum density at steady-state. Notably, semi-arid forests have likely evolved around semi-disturbed states; when they reach maximum density dependence across space they have a much higher probability of experiencing catastrophic disturbance, as seen recently in the exceptionally large wildfires ravaging western states.

2.2 Appendix B: APPLICATION OF METABOLIC SCALING THEORY TO REDUCE ERROR IN LOCAL MAXIMA TREE SEGMENTATION FROM AERIAL LIDAR

We developed a variable-area local maxima (VLM) algorithm that incorporates predictions of the MST to reduce the frequency of commission error in a local maxima individual tree inventory derived from aerial LiDAR. We also evaluated why variations in other reported aerial LiDAR segmentation models may or may not be appropriate in certain forest types. By comparing the MST prediction to 663 species of North American champion-sized trees [which include the tallest and the largest trees on the planet], and 610 measured trees in semi-arid conifer forests in Arizona and New Mexico we showed the MST model $r_{can} = \beta h^\alpha$ where β is the normalization constant, h is height, and α is a

dynamic exponent predicted by MST to be $\alpha = 1$, can be applied as a general model in any forest.

We found that the MST model accurately characterizes semi-arid conifer tree allometry in a consistent way, and that this consistency can be applied to the local-maximum derived inventory in a way that reduces the error of commission in the inventory segmentation.

2.3 Appendix C: PREDICTING RANK-SIZE DISTRIBUTIONS OF FOREST MID- AND UNDERSTORY IN TWO SEMI-ARID CONIFER FORESTS WITH AN AERIAL LIDAR DERIVED FOREST INVENTORY

Building from the results of Appendices A & B, we asked whether an aerial LiDAR derived forest overstory inventory [from a variable area local maxima (VLM) algorithm] can predict the frequency of presence/absence in the understory. The LiDAR does a good job of inventorying the overstory but decreases in accuracy when discriminating trees in the mid- and understory due to (1) the two-dimensionality of the data hiding understory trees beneath canopy trees, and (2) understory trees tend to have interconnected canopies that cannot be differentiated by aerial LiDAR. We test whether a power law tapered Pareto distribution, discussed briefly in Appendix A and expanded upon here, fit by least squares to the overstory LiDAR predicts the SFD of individuals in the understory.

We found (1) the tree height SFD of two observed forests had the same probability density function (PDF) parameters when fit with a tapered Pareto distribution, and (2) the shape parameters for the function do not change regardless of canopy cover. What does change, however, is the scale parameter (a normalization constant) based on the size of the area under observation. (3) The frequency distribution of the observed height inventory was indistinguishable from the inventory derived by the VLM in the forested plots [by a two-sample Kolmogorov-Smirnov test]. (4) In stands with < 50% cover, the VLM has a low rate of omission of understory, making understory estimation unnecessary, as the stems are visible to the sensor. (5) When canopy cover >50% the

observed distribution fit a tapered Pareto distribution with less variance between linear bins as cover % continues to increase.

Biometricians and foresters interested in generating virtual inventories based on aerial LiDAR should consider first parameterizing their LiDAR to field observations based on the local species allometry. Aerial LiDAR derived inventories of forest stands could help in monitoring forest conditions and processes without expending valuable resources on monitoring teams that can measure only tiny portions of complex landscapes with variable disturbance histories or local microclimates. For ecologists, the benefit of having landscape level inventories opens up entirely new questions about pattern and process in forests. The scale of landscape LiDAR inventory removes the need for statistical extrapolation from small plot areas and provides the capability to quantify structure and biomass distributions at both large spatial extents and fine spatial resolution. For both managers and ecologists interested in studying Earth-system scale processes, scale and pattern are critical measures toward quantifying the health of the environment as it exists today, to monitor it, and to detect change.

2.4 Appendix D: CHARACTERIZING FOREST STRUCTURE AND DISTURBANCE LEGACIES WITH AERIAL LIDAR USING GETIS-ORD G_I^* AND ANSELIN'S MORAN I LOCAL INDICATORS OF SPATIAL ASSOCIATION

Here my coauthors and I use an aerial LiDAR individual tree overstory inventory [Appendix B] from multiple aerial LiDAR flights over semi-arid coniferous forests in Arizona and New Mexico to identify structural legacies of known historical disturbance. Parts of these landscapes were either preserved from logging and wildfires or have been logged and burned, in some cases extensively, in the recent past through the present. We obtained two different local-statistics for the individual tree inventory: the Getis-Ord G_i^*, and the Anselin Moran's I_i in a GIS. Our results show clearly where human-caused impacts as well as natural disturbance of forested stands can be delineated across young to old-growth semi-arid conifer forest.

The main points of this chapter are: (1) Both Anselin Moran's I_i and Getis-Ord G_i^* reveal change in forest structure as a result of differences in first order effects and

disturbance legacies that are not readily apparent in other forms of remotely sensed data. (2) Anselin Moran's I_i provides additional information over Getis-Ord G_i^* because it acknowledges both the z-score of the neighbors (as Getis-Ord G_i^* does) but also reports the condition of the point of reference (in this case, a tree). (3) Cross validation with observed disturbance histories revealed significant impacts that were to-date not properly mapped or quantified explicitly. This new information is likely to be useful to forest managers and ecologists interested in tracking or discovering historic events that may be lost or poorly recorded in the written record.

2.5 Appendix E: BIOMASS AND CARBON MODELS FOR AERIAL LASER SWATH MAPPING INVENTORIES IN SOUTHWESTERN NORTH AMERICAN SEMI-ARID CONIFER FORESTS

Models of total above ground biomass (AGB) and above ground carbon (AGC) using individual tree height and aerial LiDAR mean canopy height (MCH) profiles as independent variables are created using inventory data from three forests in Arizona and New Mexico by species and mixed-species forest type [by dominant species plant functional types (PFT): Ponderosa pine, White-fir dominated, Mixed-Conifer, Spruce-fir, and seral Aspen forest]. The PFTs are taxonomically and climatically similar to other forests across the Southwestern United States, making these models potentially generalizable to a larger geographic area. We developed the general and PFT-specific models for both individual tree height and MCH profiles, as well as ten common species models for individual tree height. The general models using MCH estimated AGB in 500 m^2 and 1,000 m^2 plots to within ± 18% of the observed values at the 95% confidence interval. The accuracy of the MCH PFT models was higher ($r^2 = 0.770 - 0.840$) than the MCH general model ($r^2 = 0.630$). Four of the five individual tree PFT models exhibited the same or higher accuracy ($r^2 = 0.685 - 0.760$) [except Aspen ($r^2 = 0.492$)] than the general model ($r^2 = 0.685$). Estimation of AGB at the individual tree scale was based on the MST. Typically the independent variable for estimating AGB is bole radius: $AGB = \beta r^\alpha$ where β is the normalization constant and α is the scale parameter found by least squares regression $\alpha = 2.47 \pm 0.02$. This α was

indistinguishable from published tropical tree allometry reported elsewhere to be $\alpha = 2.5$ and is a rejection of the MST prediction of $\alpha = 2.66\bar{6}$. However, AGB models generated by individual species found α varied between 2.22 ± 0.05 and 2.65 ± 0.04. Further, when health was considered scaling of individual bole radius to height $HT = \beta r^{\alpha}$, found the unhealthiest trees had $\alpha = 0.599 \pm 0.030$ and the healthiest $\alpha = 0.675 \pm 0.015$. The plot observed AGB was similar to reported forest biomass in the southwestern US, however some plots far exceeded what is predicted using conventional models, for example, in mature (+200 year old) mixed-conifer stands left undisturbed over the last century a maximum of $1,495\ Mg\ ha^{-1}$ was observed this far exceeds a model 99th percentile estimate ($\sim 640\ Mg\ ha^{-1}$).

Two novel findings of this research are (1) when the condition of trees are taken into account the MST predicted scaling exponent is observed in the healthiest trees, however, MST predictions were found to fail when all trees regardless of condition were used in the least-squares regression. In those cases the observed estimates were indistinguishable from the Jenkins et al. (2004) and Chave et al. (2005) models; and (2) AGB derived from height based observations (individual tree height or MCH) is shown to significantly under predict biomass in old-growth semi-arid conifer forests.

We suggest two mechanisms are at work which account for the greater than expected level of biomass with shorter than expected tree height and MCH profiles: (1) a truncation of maximum height based on the combination of species physiology and locally limiting factors [i.e. precipitation, temperature, and total plant available water in the soil profile], and (2) the particularly long lives [up to 450 years (Grissino-Mayer et al. 1995)] of trees in the observed old growth stands. A semi-arid conifer tree may reach a vertical height near its local eco-physiological limit early in its life cycle [50-150 years], after which and over the proceeding decades to centuries the tree continues to increase in cambial diameter and maximizing its canopy volume while maintaining the same vertical height relative to the amount of water it is able to obtain. Such a tree eventually is shorter than predicted using general models based on its diameter. Importantly, the age of the

stand becomes a critical factor in determining whether the stand has an exceptional amount of AGB.

2.6 Future Research Opportunities

- An expansion of allometric scaling models to include other common woodland and forest types in the Southwestern US, e.g. Piñon and Juniper woodland, Limber and Lodgepole pine forest, and Madrean pine-oak forests.

- Exploration of fractal dimensionality across forests, related to self-affinity and space-filling and area preserving density dependence [i.e. their Riemann Zeta Function].

- Exploring the impact of disturbance on forest species assemblages, and resilience.

- Finding 'tipping point' behaviors, e.g. positive or negative autocorrelation where the Hurst exponent relates to fractal dimension and $\frac{1}{f^\beta}$ noise with SOC behavior in the size frequency distribution of both disturbances and forest patch sizes.

- Advanced point cloud segmentation algorithms that measure individual tree canopy features of the point cloud to better estimate physical parameters [e.g. Leaf Area Index, canopy bulk density, etc] and species specific characteristics [e.g. crown geometry, spectral reflectance].

3. REFERENCES

Adams, H. D., Guardiola-Claramonte, M., Barron-Gafford, G. A., Villegas, J. C., Breshears, D. D., Zou, C. B., ..., Huxman, T. E. (2009). Temperature sensitivity of drought-induced tree mortality portends increased regional die-off under global-change-type drought. *Proceedings of the National Academy of Sciences*, *106*(17), 7063-7066.

Adams, M. A. (2013). Mega-fires, tipping points and ecosystem services: Managing forests and woodlands in an uncertain future. *Forest Ecology and Management*. 294 (15), 250-261.

Anderegg, W. R., Berry, J. A., Smith, D. D., Sperry, J. S., Anderegg, L. D., Field, C. B. (2012). The roles of hydraulic and carbon stress in a widespread climate-induced forest die-off. *Proceedings of the National Academy of Sciences*, *109*(1), 233-237.

Anderegg, W. R., Callaway, E. S. (2012). Infestation and hydraulic consequences of induced carbon starvation. *Plant physiology*, *159*(4), 1866-1874.

Allen, C. D., Breshears, D. D. (2007). Climate-induced forest dieback as an emergent global phenomenon. *Eos, Transactions American Geophysical Union*, *88*(47), 504.

Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., ... ,Cobb, N. (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management*, *259*(4), 660-684.

Bachelet D., R.P. Nelson, J.M. Lenihan, R.J. Drapek. (2001) Climate Change Effects on Vegetation Distribution and Carbon Budget in the United States. *Ecosystems*. *Forest Biodiversity Under Global Change, 4*, 164–185.

Bak P., C. Tang, K. Wiesenfeld. (1987) Self-organized criticality: An explanation of the 1/f noise. *Physical Review Letters*, *59*, 381-384.

Bak, P., Chen, K. and Tang, C. (1990) A forest-fire model and some thoughts on turbulence. *Physics Letters A, 147*, 297–300.

Bak P., K. Chan. (1991). Self-organized criticality. *Scientific American, 264*, 46–54.

Barnola, J.-M., D. Raynaud, C. Lorius, Barkov. N.I. (2003). Historical CO_2 record from the Vostok ice core. In Trends: A Compendium of Data on Global Change. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.

Birdsey R., R. Kolka, M.L. Smith, M. Ryan, D. Hollinger, L. Health, C. Hoover (2004) Landscape Carbon Monitoring and Analysis at the Experimental Forest Network. In *Proceedings, 3rd Annual Conference on Carbon Sequestration,* 3-6.

Blackard, J.A., Finco, M.V., Helmer, E.H., Holden, G.R., Hoppus, M.L., Jacobs, D.M., Lister, A.J., Moisen, G.G., Nelson, M.D., Riemann, R., Ruefenacht, B., …, Tymcio, R.P. (2008). Mapping U.S. forest biomass using nationwide forest inventory data and moderate resolution information: *Remote Sensing of Environment 112*, 1658–1677.

Breshears, D. D., Cobb, N. S., Rich, P. M., Price, K. P., Allen, C. D., Balice, R. G., ... Meyer, C. W. (2005). Regional vegetation die-off in response to global-change-type drought. *Proceedings of the National Academy of Sciences of the United States of America*, *102* (42), 15144-15148.

Breshears, D. D., Myers, O. B., Meyer, C. W., Barnes, F. J., Zou, C. B., Allen, C. D., ... Pockman, W. T. (2008). Tree die-off in response to global change-type drought: mortality insights from a decade of plant water potential measurements. *Frontiers in Ecology and the Environment*, *7* (4), 185-189.

Brigham-Grette, J., M. Melles, P. Minyuk., A. Andreev, et al. 2013. Pliocene Warmth, Polar Amplification, and Stepped Pliocence Cooling Recorded in NE Arctic Russia. *Science* 1233137 Published online 9 May 2013

Brown JH, Gillooly JF, Allen AP, Savage VM, West GB (2004) Toward a metabolic theory of ecology. *Ecology* ,*85*(7), 1771-1789.

Burkett, V.R. and Davidson, M.A. [Eds.]. (2012). Coastal Impacts, Adaptation and Vulnerability: A Technical Input to the 2012 National Climate Assessment. Cooperative Report to the 2013 National Climate Assessment., 150 p.

Caldeira, K., M. E.Wickett. (2003). Anthropogenic carbon and ocean
pH. *Nature,* 425(6956), 365–365.

Chave, J., Andalo, C., Brown, S., Cairns, M. A., Chambers, J. Q., Eamus, D., ...
Yamakura, T. 2005. Tree allometry and improved estimation of carbon stocks and
balance in tropical forests. *Oecologia*, *145*(1), 87-99.

Cohen, J. D. (2000). Preventing disaster: home ignitability in the wildland-urban
interface. *Journal of forestry*, *98*(3), 15-21.

Condit, R. 1998. Tropical Forest Census Plots. Springer-Verlag and R. G. Landes
Company, Berlin, Germany, and Georgetown, Texas.

Connell, J. H. 1978. Diversity in tropical rain forests and coral reefs. *Science, 199*, 1302-
1310.

Cox, P. M., Betts, R. A., Jones, C. D., Spall, S. A., Totterdell, I. J. (2000). Acceleration
of global warming due to carbon-cycle feedbacks in a coupled climate
model. *Nature*, *408*(6809), 184-187.

Dial, R., Roughgarden, J. (1988). Theory of marine communities: the intermediate
disturbance hypothesis. *Ecology, 79*, 1412–1424.

Dale, V. H., Joyce, L. A., McNulty, S., Neilson, R. P., Ayres, M. P., Flannigan, M. D., ...
Michael Wotton, B. (2001). Climate Change and Forest Disturbances: Climate
change can affect forests by altering the frequency, intensity, duration, and timing
of fire, drought, introduced species, insect and pathogen outbreaks, hurricanes,
windstorms, ice storms, or landslides. *BioScience*, *51*(9), 723-734.

Drossel, B., and F. Schwabl (1992), Self-organized critical forest-fire model, *Phys. Rev.
Lett., 69*, 1629–1632.

Drossel, B., Schwabl, F. (1994). Formation of space-time structure in a forest-fire
model. *Physica A: Statistical Mechanics and its Applications*, *204*(1), 212-229.

Enquist B., J. Brown, G. West. 1998. Allometric scaling of plant energetics and
population density. *Nature*, 395, 163-165.

Enquist B., G. West, E. Charnov, J. Brown. 1999. Allometric scaling of production and
life-history variation in vascular plants. *Nature*, *401*, 907-911.

Enquist, B.J., K.J. Niklas 2001. Invariant scaling relations across tree-dominated communities. *Nature*, *410*, 655-660.

Enquist, B. J., Economo, E. P., Huxman, T. E., Allen, A. P., Ignace, D. D., Gillooly, J. F. (2003). Scaling metabolism from organisms to ecosystems. *Nature*, *423*(6940), 639-642.

Enquist B.J., G.B. West, J.H. Brown. (2009) Extensions and evaluations of a general quantitative theory of forest structure and dynamics. *Proceedings of the National Academy of Sciences, 106(17), 7046-7051.*

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., Von Bloh, W., Brovkin, V., ... Zeng, N. (2006). Climate-carbon cycle feedback analysis: Results from the C4MIP model intercomparison. *Journal of Climate*, *19*(14), 3337-3353.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., ... Snyder, P. K. (2005). Global consequences of land use. *Science*, *309*(5734), 570-574.

Fulé, P. Z., Covington, W. W., Moore, M. M. (1997). Determining reference conditions for ecosystem management of southwestern ponderosa pine forests. *Ecological Applications*, *7*(3), 895-908.

Füssel, H. M. (2009). An updated assessment of the risks from climate change based on research published since the IPCC Fourth Assessment Report. *Climatic Change*, *97*(3-4), 469-482.

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*(4), 045023.

Goetz, S., Dubayah, R. (2011). Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. *Carbon Management*, *2*(3), 231–244.

Grissino-Mayer, H.D., Baisan, C.H. and Swetnam, T.W. 1995. Fire history in the Pinaleño Mountains of southeastern Arizona: Effects of human related disturbances. In Debano, L.F., Gottfried, G.J., Hamre, R.H., Edminster, C.B.,

Ffolliott, P.F. and Ortega-Rubio, A., editors, Biodiversity and management of the Madrean Archipelago: the Sky Islands of Southwestern United States and Northwestern Mexico, US Department of Agriculture Forest Service, General Technical Report RM-GTR-264, 399–407.

Hall, A. R., Miller, A. D., Leggett, H. C., Roxburgh, S. H., Buckling, A., Shea, K. (2012). Diversity–disturbance relationships: frequency and intensity interact. *Biology Letters*, *8*(5), 768-771.

Hansen J., M. Sato, P. Kharecha, D. Beerling, R. Berner, V. Masson-Delmotte, M. Pagani, M. Raymo, D.L. Royer, J.C. Zachos. 2008. Target Atmospheric CO_2: Where Should Humanity Aim? *arXiv preprint arXiv:0804.1126*.

Harte, J., Kinzig, A., Green, J. (1999). Self-similarity in the distribution and abundance of species. *Science*, *284* (5412), 334-336.

Holt, R. D. (2009). Bringing the Hutchinsonian niche into the 21st century: ecological and evolutionary perspectives. *Proceedings of the National Academy of Sciences*, *106* (Supplement 2), 19659-19665.

Hutchinson, G. E. (1959). Homage to Santa Rosalia or why are there so many kinds of animals? *The American Naturalist*, *93*(870), 145-159.

Jaffé, R., Ding, Y., Niggemann, J., Vähätalo, A. V., Stubbins, A., Spencer, R. G., ... Dittmar, T. (2013). Global charcoal mobilization from soils via dissolution and riverine transport to the oceans. *Science*, *340*(6130), 345-347.

Jenkins, J.C.; Chojnacky, D.C.; Heath, L.S.; Birdsey, R.A. 2004. A comprehensive database of biomass regressions for North American tree species. Gen. Tech. Rep. NE-319. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station. 45 p.

Kellner, J. R., Asner, G. P. 2009. Convergent structural responses of tropical forests to diverse disturbance regimes. *Ecology letters*, *12*(9), 887-897.

Kempes CP, West GB, Crowell K, Girvan M (2011) Predicting Maximum Tree Heights and Other Traits from Allometric Scaling and Resource Limitations. *PLoS ONE, 6*(6), e20551.

Koch G.W., S.C. Sillett, G.M. Jennings, S.D. Davis (2004) The limits of tree height. *Nature, 428*, 851-854 (22 April 2004)

Kurz, W. A., Dymond, C. C., Stinson, G., Rampley, G. J., Neilson, E. T., Carroll, A. L., ... Safranyik, L. (2008). Mountain pine beetle and forest carbon feedback to climate change. *Nature*, *452*(7190), 987-990.

Larson, A. J., Churchill, D. (2012). Tree spatial patterns in fire-frequent forests of western North America, including mechanisms of pattern formation and implications for designing fuel reduction and restoration treatments. *Forest Ecology and Management*, *267*, 74-92.

Lenton, T. M., Held, H., Kriegler, E., Hall, J. W., Lucht, W., Rahmstorf, S., Schellnhuber, H. J. (2008). Tipping elements in the Earth's climate system. *Proceedings of the National Academy of Sciences*, *105*(6), 1786-1793.

Logan, J. A., Regniere, J., Powell, J. A. (2003). Assessing the impacts of global warming on forest pest dynamics. *Frontiers in Ecology and the Environment*, *1*(3), 130-137.

Logan, J. A., Powell, J. A. (2001). Ghost forests, global warming, and the mountain pine beetle (Coleoptera: Scolytidae). *American Entomologist*, *47*(3), 160.

Lüthi, D., M. Le Floch, B. Bereiter, T. Blunier, J.-M. Barnola, et al. 2008. High-resolution carbon dioxide concentration record 650,000-800,000 years before present. *Nature,* 453: 379-382.

Lyon, A. G., Anderson, S. H. (2003). Potential gas development impacts on sage grouse nest initiation and movement. *Wildlife Society Bulletin*, 486-491.

Mandelbrot B.B. 1983 The Fractal Geometry of Nature Macmillan 1983

Mandelbrot, B.B., 2002. Gaussian Self-Affinity and Fractals: Globality, the earth, 1/f noise, and R/S. Springer, New York. 654 p.

Masiello, C. A., Louchouarn, P. (2013). Fire in the Ocean. *Science*, *340*(6130), 287-288.

Mast, J. N., Fule, P. Z., Moore, M. M., Covington, W. W., Waltz, A. E. (1999). Restoration of presettlement age structure of an Arizona ponderosa pine forest. *Ecological applications*, *9*(1), 228-239.

McDowell, N. G. (2011). Mechanisms linking drought, hydraulics, carbon metabolism, and vegetation mortality. *Plant physiology*, *155*(3), 1051-1059.

McKinley, D.C., M.G. Ryan, R.A. Birdsey, C.P. Giardina, M.E. Harmon, L.S. Heath, R.A. Houghton, R.B. Jackson, J.F. Morrison, B.C. Murray, D.E. Pataki, K.E. Skog. 2011. A synthesis of current knowledge on forests and carbon storage in the United States. *Ecological Applications, 21*, 1902–1924.

Miller, A. D., Roxburgh, S. H., Shea, K. (2011). How frequency and intensity shape diversity–disturbance relationships. *Proceedings of the National Academy of Sciences*, *108*(14), 5643-5648.

Morgan, P., Aplet, G. H., Haufler, J. B., Humphries, H. C., Moore, M. M., Wilson, W. D. (1994). Historical range of variability: a useful tool for evaluating ecosystem change. *Journal of Sustainable forestry*, *2*(1-2), 87-111.

Morse, D. R., Lawton, J. H., Dodson, M. M., Williamson, M. H. (1985). Fractal dimension of vegetation and the distribution of arthropod body lengths. *Nature*, *314*(6013), 731-733.

Neftel, A., E. Moor, H. Oeschger, Stauffer. B. (1985). Evidence from polar ice cores for the increase in atmospheric CO_2 in the past two centuries. *Nature, 315*:45-47.

Neftel, A., H. Friedli, E. Moor, H. Lötscher, H. Oeschger, U. Siegenthaler, and B. Stauffer. (1994). Historical CO_2 record from the Siple Station ice core. In Trends: A Compendium of Data on Global Change. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.

Negron, J.F., McMillin, J.D., Anhold, J.A., Coulson, D. (2009). Bark beetle-caused mortality in a drought-affected ponderosa pine landscape in Arizona, USA. *Forest Ecology and Management,* 257, 1353–1362.

Newman, M.E.J. (2005). Power-laws, Pareto distributions, and Zipf's law. *Contemporary Physics,* 46, 323-351.

Niklas, K. J., Midgley, J. J., Rand, R. H. (2003). Tree size frequency distributions, plant density, age and community disturbance. *Ecology letters*, *6*(5), 405-411.

Pachauri, R. K., Reisinger, A. (2007). Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. *Intergovernmental Panel on Climate Change*.

Radeloff, V. C., Hammer, R. B., Stewart, S. I., Fried, J. S., Holcomb, S. S., McKeefry, J. F. (2005). The wildland-urban interface in the United States. *Ecological applications*, *15*(3), 799-805.

Raffa, K.F., Aukema, B.H., Bentz, B.J., Carroll, A.L., Hicke, J.A., Turner, M.G., Romme, W.H. (2008). Cross-scale drivers of natural disturbances prone to anthropogenic amplification: the dynamics of bark beetle eruptions. *Bioscience*, *58*, 501–517.

Ramankutty et al. (2008), Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000, *Global Biogeochemical Cycles*, *22*, GB1003.

Ramanathan, V., Carmichael, G. (2008). Global and regional climate changes due to black carbon. *Nature geoscience*, *1*(4), 221-227.

Reed, W.J., and K.S. McKelvey. (2002). Power-law behavior and parametric models for the size distribution of forest fires. *Ecological Modeling*, *150*, 239-254

Ryan M.G., Harmon M.E., Birdsey R.A., et al. (2010) A Synthesis of the Science on Forests and Carbon for U.S. Forests. Issues in Ecology. Vol. 13 Washington, DC: Ecological Society of America.1–16.

Parry M, Paluyokof J, Hanson C, Lowe J. (2008). Squaring up to reality. *Nature Reports Climate Change* 2008; 2: 68–70.

Pianka, E. R. (2011). *Evolutionary ecology*, 7th edition. Erik Pianka, Harper & Row, University of Michigan, 416 p.

Prentice, I.C., Farquhar, G.D., Fasham, M.J.R., Goulden, M.L., …,Wallace, D.W.R. (2001). The carbon cycle and atmospheric carbon dioxide. In: Houghton, J.T., Ding, Y., Griggs, D.J., Noguer, M., …, Johnson, C.A. Eds., Climate Change 2001: The Scientific Basis. Contribution of Working Group 1 to the Third

Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, 183–237.

Sánchez Meador, A. J., Parysow, P. F., Moore, M. M. (2011). A new method for delineating tree patches and assessing spatial reference conditions of ponderosa pine forests in northern Arizona. *Restoration Ecology*, *19*(4), 490-499.

Savage, V. M., Bentley, L. P., Enquist, B. J., Sperry, J. S., Smith, D. D., Reich, P. B., von Allmen, E. I. (2010). Hydraulic trade-offs and space filling enable better predictions of vascular structure and function in plants. *Proceedings of the National Academy of Sciences*, *107*(52), 22722-22727.

Shinozaki K., Yoda K., Hozumi K., Kira T. (1964a) A quantitative analysis of plant form – the pipe model theory. I. Basic analyses *Japanese Journal of Ecology, 14*, 97–105

Shinozaki K., Yoda K., Hozumi K., Kira T. (1964b) A quantitative analysis of plant form – the pipe model theory. II. Further evidence of the theory and its application in forest ecology. *Japanese Journal of Ecology, 14*, 133–139

Smith J.E., L.S. Heath, K.E. Skog, R.A. Birdsey. 2006. Methods for Calculating Forest Ecosystem and Harvested Carbon With Standard Estimates for Forest Types of the United States. USDA Forest Service GTR-NE-343. Newtown Square, PA.

Solé, R.V., Manrubia, S.C. (1995). Are rainforests self-organized in a critical state? *Journal of Theoretical Biology*, *173*(1), 31-40.

Solomon, S. (Ed.). (2007). Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC (Vol. 4). Cambridge University Press.

Stocker, T. F., Qin, D., & Platner, G. K. (2013). Climate Change 2013: The Physical Science Basis. Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Summary for Policymakers (IPCC, 2013).

Stocks, B. J., Fosberg, M. A., Lynham, T. J., Mearns, L., Wotton, B. M., Yang, Q., ..., McKenney, D. W. (1998). Climate change and forest fire potential in Russian and Canadian boreal forests. *Climatic Change*, *38*(1), 1-13.

Sundquist, E. T., Ackerman, K. V., Bliss, N. B., Kellndorfer, J. M., Reeves, M. C., Rollins, M. G. (2009). Rapid assessment of US forest and soil organic carbon storage and forest biomass carbon sequestration capacity. *US Geological Survey Open-File Report*, *1283*, 15.

Swetnam, T. W. (1990). Fire history and climate in the southwestern United States. In Proceedings of Symposium on Effects on Fire in Management of Southwestern Natural Resources, GTR-RM-191. US Forest Service, 6-17.

Ustin, S., J.A. Gamon (2010) Remote sensing of plant functional types. *New Phytologist*, *186*(4), 795-816 June 2010

Van Leeuwen, M., Nieuwenhuis, M. (2010). Retrieval of forest structural parameters using LiDAR remote sensing. *European Journal of Forest Research*, *129*(4), 749-770.

Van Mantgem, P. J., Stephenson, N. L., Byrne, J. C., Daniels, L. D., Franklin, J. F., Fulé, P. Z., ..., Veblen, T. T. (2009). Widespread increase of tree mortality rates in the western United States. *Science*, *323*(5913), 521-524.

West, G. B., Brown, J. H., Enquist, B. J. (1999). The fourth dimension of life: fractal geometry and allometric scaling of organisms. *Science*, *284*(5420), 1677-1679.West et al. 2009

West G.B., B.J. Enquist, J.H. Brown (2009) A general quantitative theory of forest structure and dynamics. *Proceedings of the National Academy of Sciences, 106(17), 7040-7045.*

Wilkinson, D.M. (1999). The disturbing history of intermediate disturbance. *Oikos* 84 (1): 145–147

Williams, A. P., Allen, C. D., Macalady, A. K., Griffin, D., Woodhouse, C. A., Meko, D. M., ..., McDowell, N. G. (2012). Temperature as a potent driver of regional forest drought stress and tree mortality. *Nature Climate Change,* 3, 292-297.

Woodbury P.B., Smith J.E., Heath L.S. (2007). Carbon sequestration in the US forest sector from 1990 to 2010. *Forest Ecology and Mana*gement. *241*(1), 14-27.

APPENDIX A: THE GENERALITY OF METABOLIC SCALING THEORY IN DIFFERING FOREST DISTURBANCE REGIMES

Tyson L. Swetnam[1,3], Donald A. Falk[1] , Brian J. Enquist[2]

Format of Submission: The American Naturalist

[1] School of Natural Resources and Environment, The University of Arizona Tucson, AZ 85721, USA. Email: tswetnam@email.arizona.edu, phone: (520) 247-2293

[2] Department of Ecology and Evolutionary Biology, The University of Arizona Tucson, AZ 85721, USA. Email: benquist@email.arizona.edu, phone: (520) 626-3329

[3] Corresponding Author

Abstract

Mortality of trees not caused by competitive self-thinning [e.g. external disturbance, such as fire or herbivory] can disrupt the sorting and size frequency distribution (SFD) of trees in forests non-linearly across both spatial and temporal scale. The quantitative theory of forest structure, based on the Metabolic Scaling Theory (MST), explains why steady-state forest SFD follow power-law distributions based on primary size measures, but does not explicitly account for the role of external disturbance impacts on the SFD. In the present study we asked: Do semi-arid forests, which experience a wide range of disturbances, exhibit spacing and packing which scale similarly with tropical forests considered to be at a MST steady-state? To assess this expanded version of MST we compared three different semi-arid forest inventory datasets from Arizona and New Mexico, USA to a long-term old-growth tropical rain forest site in Panama. Specifically, we test (1) if semi-arid forests exhibit self-similarity for diameter and height SFD; and (2) if these relationships are consistent with the predictions of the MST. We extend the conceptual foundation of MST to include the role of episodic disturbance and limits to maximum tree size. As predicted by MST, despite differences in sample size and area measured, empirical diameter and height distributions at all sites strongly reflected the predicted self-thinning trajectories with either a -2 or -3 scaling exponent, respectively. Our findings suggest despite differences in disturbance regimes, the sizes and packing of trees in both semi-arid and tropical forests are similarly constrained by MST mechanisms.

Introduction

Density dependence in forested ecosystems

Forest population size-frequency distributions (SFD) are known to follow power-law scaling relationships as a result of density-dependent competition where seedlings grow to adult canopy trees while competing for a limiting resource(s) [typically light and/or nutrients] (Enquist et al. 1998, 1999, 2009, Enquist and Niklas 2001, Coomes et al. 2003, 2007, Purves et al. 2007, Strigul et al. 2008, West et al. 2009, Deng et al. 2012). The need to accurately describe the size frequency distribution (SFD) of forest populations across space has become more important as modern remote-sensing (Ustin and Gamon 2010) and computer tools enable enumeration over large areas. Because the rate of density-dependent growth and mortality scale with individual size (White et al. 2007, Enquist et al. 2009, Deng et al. 2012) the scaling of how less competitive trees die leaving survivors [i.e. winners] to grow and eventually fill the losers' space and assimilate more of a limiting resource is predicted to be alike across all forests (Enquist et al. 2009). This metabolic 'packing-rule' builds upon earlier geometric scaling arguments for self-thinning in forests (Reineke 1933, Yoda et al. 1963). We assess the generality of MST predictions by comparing the frequency of primary size-measures [e.g. diameter at breast height (DBH), or maximum tree height] in forests with episodic disturbance, as well as how physiological and resource limits affect tree size distribution. Our approach allows us to quantify how differing disturbance regimes modify the spacing and packing of canopies – and ultimately the functioning of forests. Our approach builds from work by Kerkhoff and Enquist (2007), Enquist et al. (2009), West et al. (2009), and Deng et al. (2012).

There has been much debate concerning the appropriateness of Metabolic Scaling Theory (MST) (West et al. 1999, Brown et al. 2004) in understanding forest structure and functioning. On the one hand, several studies have claimed to support metabolic self-thinning predictions (West et al. 2009, Enquist et al. 2009, Deng et al. 2012). On the other hand, several studies have claimed there is variation in the scaling behavior of individuals and forests (Kolokotrones et al. 2010, Mori et al. 2010) and mechanisms

invoked within MST to explain the spacing and packing of forests canopies are incomplete (Purves et al. 2007, Strigul et al 2008, Sperry et al. 2012, Anfodillo et al. 2013, Lin et al. 2013) or wrong (Agutter and Wheatly 2004, Muller-Landau et al. 2006, Coomes and Allen 2009). An important question that has emerged within the debate is: What is the role of disturbance in modifying forest structure? For example, the approximate steady-state behavior predicted by MST may be generally limited to the most productive forests that experience continuous recruitment only after minor disturbance events such as tree-fall (Enquist and Niklas 2001, Kerkhoff and Enquist 2007, West et al. 2009, Enquist et al. 2009). Semi-arid forests, on the other hand, experience episodic fire, herbivory, insect pests, and drought; they may also lack continuous recruitment after the disturbance event has ceased. Such forests are likely to be disproportionately impacted within the smaller size classes, leaving open spaces or gaps in recruitment for prolonged periods.

The origins of spatial patterns in forests are a consequence of first-order effects across space: the variation in annual temperature and precipitation, the amount of plant available soil water, and strength of the atmospheric vapor pressure deficit. These effects are themselves controlled by differences in elevation, topographic position, slope, aspect; and the parent material depth and age of the soil. The range of conditions under which a forest may grow fluctuates widely across semi-arid ecosystems which typically are water limited at low elevations and temperature limited at high elevation. Importantly past disturbance events are what initialized the present conditions, modifying the chaotic trajectories (Hastings et al. 1993) of any given stand. For example, low-intensity fires and herbivory kill most understory plants and sapling trees but leave overstory trees intact; high intensity fire kills all trees but often only in topographically aligned, wind driven strips, or in clustered stands that enabled the contagion to spread.

Seuront and Mitchell (2008) described how one or multiple process noises can 'contaminate' a power law and impact the emergence of log-log linear behavior in the tails of a SFD. The Seuront and Mitchell (2008) study focused mainly on marine ecosystems [i.e. phytoplankton blooms which undergo increased copepod predation

pressure]. This is conceptually the same in terrestrial ecosystems where increased spatial packing facilitates contagion [e.g. fire, insect outbreaks, and parasitic infestations]. Mechanistically, low intensity disturbance like herbivory or episodic surface fire is likely to have a greater impact on small individuals in the population SFD.

Kellner and Asner (2009) reported five tropical forests with differing canopy gap frequencies [e.g. disturbance levels] all exhibited power law scaling equivalent to a Zeta function [a complex function that continues the sum of a power-law series]. Shier and Bourke (2013), working with geometric and fractal shapes [conceptually identical to fractal-like branching trees], reported any randomly placed shape filling an empty space resulted in a distribution of sizes that follow a Zeta function. These two examples are comparing space from opposite positions: one is looking at absence (e.g. canopy gaps, Kellner and Asner 2009) the other at presence (e.g. the fractal shapes, Shier and Bourke 2013); importantly both studies found the distribution exhibits a Zeta function. The space filling nature of individual trees results in what is conventionally referred to as a forest and is fundamentally linked to the Zeta function. The value of the Zeta function varies by the scaling exponent [equivalent to the fractal dimension] based on the level of presence/absence in the stand.

Theoretical predictions described by MST do not well explain how disturbance fits into spatial and temporal effects of forest growth, recruitment, and mortality. In Figure 1 we conceptualize how we expect disturbance to create divergence from space-filling power law behavior. First, external processes like low severity disturbances [e.g. surface fires] kill mostly small individuals and open up gap space amongst larger individuals; this is expressed as a divergence in the lower tail (Figure 1). Second, other external processes such as stochastic mortality over time which randomly remove individuals from the population regardless of size are expressed as a divergence in the upper tail where the loss of a single individual is significant on log-log scale. Last, physiological size limits [e.g. maximum tree height] and resource limitation [e.g. light, water, or nutrients] truncate the maximum possible size trees can reach in a lifetime and are expressed as a tapering or truncation in the upper tail (Figure 1).

Metabolic scaling theory

MST provides a quantitative and mechanistic framework from which numerous *a priori* predictions for the scaling of organism's resource utilization, growth, reproduction, and mortality rate are made (Brown et al. 2004, Enquist et al. 2009, West et al. 2009). Conceptually, as trees grow from seedlings to adult canopy trees they compete with other trees for limiting resources and finite space within the forest. Every cohort of trees experiences mortality and a reduction of numbers over time as individuals grow to require more space to sustain their metabolism. This *"packing-rule"* results in fewer larger trees from a once numerous cohort having undergone a '*self-thinning*' process (Reineke 1933, Yoda et al. 1963, Enquist et al. 1998, 2009, Niklas et al. 2003, West et al. 2009, Deng et al. 2012). If recruitment is not limited and there is no extrinsic disturbance then all mortality across the different sizes of trees are due to self-thinning (Enquist et al. 2009). As a result, various models of self-thinning (Enquist et al. 1998, Enquist and Niklas 2001) predict in general the resultant SFD for primary measures of trees follow a power-law:

$$N(x) \propto x^{-\alpha}, 0 < \alpha, \qquad \text{Eq. 1}$$

where the number $N(x)$ of individuals of size larger than x decreases as a power of α, the scaling exponent (Enquist et al. 1998, 1999, 2009, Enquist and Niklas 2001, Niklas and Enquist 2001, Niklas et al. 2003). Typically, α changes with the size measure being considered (i.e. mass, diameter, length, or leaf area) and can be normalized by a constant C as: $Cx^{-\alpha}$ to fit any distribution size (West et al. 2009, Enquist et al. 2009).

At steady-state the SFD is expected to self-thin with $\alpha = -3/2$, a Euclidean geometric exponent advanced by Reineke (1933) and Yoda et al. (1963). The explanation for the Euclidean exponent is simply in a volume-filled space the packing density of individuals per unit volume (l^3) over the unit area (l^2) becomes $V_k \propto N_k^{-\frac{3}{2}}$, reduced to the geometric '-3/2 thinning law' (Reineke 1933, Yoda et al. 1963). The WBE framework predicts the mass distribution amongst individual size classes k as a function of population size distribution $N_k \propto M_k^{-\frac{3}{4}}$, where M is organism mass. Re-writing this

equation with mass as the dependent variable, the SFD of a population of kth size becomes $M_k \propto N_k^{-4/3}$, which reveals the -4/3 self-thinning exponent. It is important to understand the volume-area occupied versus mass-area occupied are subtly different measures, and it is this difference which accounts for the geometric and fractal relationships. In this paper we consider either self-thinning rule to be correct, but stress they should be properly defined when establishing the scaling exponents for a SFD.

In their general quantitative theory of forest structure West et al. (2009) suggest 'the forest is the tree,' implying the distribution of tree sizes and flux among trees within a stand scale in the same way as the distribution of branch sizes and flux through a single branching network (West et al. 1999, 2009, Enquist et al. 1998, 1999, 2009). Thus, as trees fill space their SFD for primary size measures becomes:

$$N_k \propto r_k^{-2} \propto h_k^{-3} \qquad\qquad \text{Equation 2}$$

where N_k, the number of individuals of kth size. Importantly the theory also predicts the distribution to become truncated at large sizes due to physiologic limitations (Enquist et al. 2009, Eq. 10):

$$N_k \approx \frac{\dot{R}}{\dfrac{K+1}{b_0}} r_k^{-2} \ or \ f(r) = \frac{dn}{dr} = \frac{\dot{R}}{r_{max} b_0} r^{-2} \qquad\qquad \text{Equation 3}$$

where $\dot{R}$ is the rate of a limiting resource, K is the total number of size classes, r_{max} is the maximum size, and b_0 is a normalization constant.

Pareto distributions

The Pareto distribution (Pareto 1896) has been used to describe the frequency of many different natural phenomena including sand particles, the populations of cities, landslides, earthquakes, and forest fires (Bak et al. 1987, 1990, Stark and Hovius 2001, Newman 2005). The Pareto has the notable feature of being 'heavy-tailed' in the right side of its distribution and its PDF is not exponentially bound (Newman 2005). The Pareto's heavy-tail tends to over predict frequency distribution of large sizes in nature (Kagan and Schoenberg 2001, Schoenberg and Patel 2012). In order to account for physiologic and resource limitations to maximum tree size we incorporated the tapered

and upper-truncated Pareto distribution for our fitting of the observed data (Kagan and Schoenberg 2001, White et al. 2008, Patel 2011, Schoenberg and Patel 2012). Anecdotally, MST predicts there to be a maximum size to trees [Eq. 3] (Enquist et al. 2009), equivalent to the truncated Pareto.

Since we are working in natural systems with density-dependence and finite size limits the Pareto may not be a good fit to our data in the right tail at large population sizes (Schoenberg and Patel 2012). We are however still interested in plotting the Pareto along our observed distribution to determine if and where power-law behavior may be evident. The change d, in tree size x, for the number of individuals in each size class $n(x)$ is:

$$n(x)dx = N f(x) \, dx \qquad \text{Equation 4}$$

where N is the maximum population size and $f(x)$ is the PDF. The Pareto distribution also assumes *a priori* there is a lower limit β to the distribution, where $0 < \beta < x$. The Pareto's cumulative distribution function (CDF) is written as:

$$F(x) = 1 - \left(\frac{\beta}{x}\right)^{\alpha} = \beta^{\alpha} x^{-\alpha} \qquad \text{Equation 5}$$

where $\alpha > 0$ is the shape parameter, and β is the scale parameter for $\beta \leq x$. The Pareto's PDF is written as:

$$f(x) = \frac{\alpha \beta^{\alpha}}{x^{\alpha+1}} \equiv \alpha \beta^{\alpha} x^{-\alpha-1}. \qquad \text{Equation 6}$$

If an observed SFD perfectly fits this function it has a Pareto distribution. The absence of heavy-tailed Pareto behaviors in natural phenomena led to the introduction of the tapered or truncated Pareto distributions whose parameters are used to accurately fit the upper-tail of the observed power-law distribution (Pareto 1896, Kagan and Schoenberg 2001, Patel 2011, Schoenberg and Patel 2012, Sornette et al. 2009). Prior to the development of the upper-truncated distribution, Pareto himself reported on a power-law with a tapered tail, now known as the tapered Pareto (Pareto 1896, Kagan and Schoenberg 2001, Patel 2011, Schoenberg and Patel 2012) whose CDF is:

$$F(x) = 1 - \left(\frac{\beta}{x}\right)^{\alpha} \exp\left(\frac{\beta - x}{\theta}\right). \qquad \text{Equation 7}$$

where α is the shape parameter, β is the scale known *a priori* as the lower bound, $\beta \leq x < \infty$, and the parameter θ governs the shape of the upper tail: as $\theta \to \infty$ the $F_{taper}(x) \to F_{Pareto}(x)$, and as $\alpha \to 0$ $F_{taper}(x)$ approaches an exponential distribution (Patel 2011). The tapered Pareto's PDF is given by:

$$f(x) = \left(\frac{\alpha}{x} + \frac{1}{\theta}\right)\left(\frac{\beta}{x}\right)^{\alpha} \exp\left(\frac{\beta - x}{\theta}\right). \qquad \text{Equation 8}$$

Importantly the tapered Pareto is not heavy tailed like the Pareto, thus it exhibits a lower frequency of extreme size events or objects (Schoenberg and Patel 2012).

The truncated-Pareto differs from the tapered-Pareto in it more slowly deviates from the Pareto distribution then more quickly diverges to become bounded at its maximum value. The truncated Pareto's CDF is:

$$1 - F_X(x) = P(X > x) = \frac{\beta^{\alpha}(x^{-\alpha} - \omega^{-\alpha})}{1 - \left(\frac{\beta}{\omega}\right)^{\alpha}} = \frac{1 - \left(\frac{\beta}{x}\right)^{\alpha}}{1 - \left(\frac{\omega}{\beta}\right)^{\alpha}} \qquad \text{Equation 9}$$

where ω is the upper-truncation limit, and $0 < \beta \leq x \leq \omega \leq \infty$, and $\beta < \omega$. The truncated Pareto's PDF is written as:

$$f(x) = \frac{\alpha\beta^{\alpha}x^{-(\alpha-1)}}{1 - \left(\frac{\beta}{\omega}\right)^{\alpha}}. \qquad \text{Equation 10}$$

MST predicts there to be a maximum limit to size (Enquist et al. 2009, Eq. 10), and the distribution is equivalent to the truncated Pareto. The θ and ω parameters of the tapered (Eq. 7 & 8) and upper-truncated Pareto (Eq. 9 & 10) become most important when evaluating large forest populations where the heavy tail of a Pareto distribution would suggest sizes physiologically impossible in nature. The *a priori* understanding of θ and ω comes first from observation: ω is unlikely to exceed 50 m in our study area forests [this value will vary for the study region of interest and biometricians should be cautious in determining ω without observed data]. The θ parameter changes perhaps the most of any of the observed parameters based on sample size: as sample sizes increases the point at which the population begins an exponential decay will decrease θ until it could theoretically be equal to or less than the β parameter.

Hypotheses

We sought to determine whether power law scaling exists in our study systems by looking for log-log linear scaling in the SFD for primary size measures while applying the tapered and upper-truncated Pareto fits to the data to account for physiological limitation. If the semi-arid forests do exhibit scaling in their SFD, we are interested at what scale does this occur, considering these forests are thought to be disturbed; and if they do not exhibit power-law SFD, how then can we best describe their SFD?

MST predicts the number of individuals in any area scale as $\frac{dn}{dt} \approx r^{-2}$ [Eq. 2] for bole radius. As a null hypothesis we assumed the SFD of trees in semi-arid conifer forests are the result of self-thinning [scaling as -2 for diameter and -3 for height], based on the predictions of MST (Enquist et al. 1998, 2009, West et al. 1997, 1999a, 1999b, 2009, White et al. 2007). This follows from research in tropical forests with continuous recruitment where canopy gaps are refilled quickly (Kohyama 1993, 1994, Enquist et al. 1998, 2009, West et al. 2009). As an alternate hypothesis we assume external processes influence the distribution by altering the shape parameter of the SFD, resulting in either a flatter or steeper SFD.

Methods

Study areas

We examined three different semi-arid forest inventory datasets from Arizona and New Mexico, USA (Figure 2). We chose these forest systems based on (1) their well understood disturbance regimes, and (2) their density of standing trees being greater than what is generally considered to be savannah-like but not so dense as closed canopy forests. Two of the data sets: the Pinaleño Mountains in Arizona, and Jemez Mountains in New Mexico, originated from plot-to-forest assessments established to validate measures of aeriel LiDAR collections (Laes et al. 2008,2009, Mitchell et al. 2012, Swetnam and Falk *in review*, Swetnam et al. *in review*). The other data set are a competition demography project in the Monument Canyon Research Natural Area (Falk et al. unpublished data). An independent comparison forest: the 1995 Barro Colorado Island

Experimental forest inventory monitoring data from Panama (Figure 1)(Condit 1999) are shown as an example of a forest at steady-state (Enquist et al. 2009, West et al. 2009).

Arizona

Pinaleño Mountains

The Pinaleño Mountains (PM) are located south of Safford, Arizona at 32.7° N, 109.9° W (Figure 2). The Pinaleño reach 3,267 m above mean sea level (amsl) atop Mount Graham. The mountains are characteristic of the Basin and Range province, and are a complex of steeply sided canyons with relatively gentle high elevation uplands above 2,700 m amsl (Figure 2). The lower elevation forests of the Pinaleño are typical of Madrean Sky Islands (1,830-2,440 m amsl) consisting of Madrean-type oaks (*Quercus hypoleucoides, Q. rugosa*) and *Pinus* (*P. ponderosa, P. strobiformis*). The pines coexist above 2,130 m amsl with Gambel oak (*Q. gambellii*), Douglas-fir (*Pseudotsuga menziesii*) and white fir (*Abies concolor*). In the highest elevation forests (>2,740 m amsl) Engelmann spruce (*Picea engelmannii*), and corkbark fir (*A. lasiocarpa var. arizonica*) coexist (Whittaker and Niering 1975, Niering and Lowe 1984). Post-fire seral communities of quaking aspen (*Populus tremuloides*) are present throughout the upper elevations of the PM. The Pinaleño inventory data come from a tree population demographics and disturbance study (O'Connor et al 2010). Seventy nine forest inventory 0.05 ha radial plots were collected across a systematic grid above 2,300 amsl; a total of 2,862 trees were measured for DBH and height.

New Mexico

Valles Caldera

The Valles Caldera National Preserve (VCNP) is located in the Jemez Mountain range west of Santa Fe, New Mexico at 35.9° N, 106.5° W (Figure 2). Elevations range from 2,300m amsl in Redondo Meadow to 3,431 m amsl atop Redondo Peak. The Valles Caldera is a collapsed volcanic caldera with a rim approximately 19km wide; within the caldera are resurgent domes over 200 m high, locally referred to as *Cerros*. The Valles Caldera shares many of the same species as at the Pinaleño, with notable absences of Madrean species. *P. ponderosa* is common in the lowest elevations of the Preserve

(2,100-2,400 m amsl), with some limber pine (*P. flexilis*) and Douglas-fir on mesic sites. Gambel oak is common in post-fire seral stands amongst ponderosa and quaking aspen. North aspects tend to be dominated by Douglas-fir and white fir, as well as by sub-alpine fir and Engelmann spruce. The highest elevations of the Valles Caldera are dominated by spruce-fir. Prior to the extensive logging of the Valles Caldera beginning in the 1930's (Balmat and Kupfer 2004) there appears to have been evidence of more extensive Douglas-fir on north aspects now dominated by spruce-fir types. Forty-eight 0.1 ha radial plots were collected in 2011 by members of the Santa Catalina – Jemez Mountains Critical Zone Observatory, a total of 1,520 live and dead trees were measured, and 3,952 trees were counted (including seedlings and saplings > 15cm tall).

Monument Canyon

A second dataset within the Valles Caldera aerial LiDAR flight area is the Monument Canyon (MCN) Research Natural Area, 35.8° N, 106.6° W also located on the Santa Fe National Forest (Figure 2). The Monument Canyon data come from a long term study of fire history, stand dynamics, and competition (Falk 2004, Marshall et al., in prep). For the last century fire disturbances have been suppressed in a 255 ha (1^2 mi) area. A 90 year old cohort of ponderosa pine has in-filled the overstory forest in most areas (Falk 2004). Ten demography plots were established around large trees and within a variable radius (15.5 m -26.0 m) all neighbor trees were measured by diameter. The ten radial plot areas are equivalent to ~1.462 ha, and a total of 6,695 trees were measured in this extent. Falk et al. (unpublished data) estimate the maximum density of 1.0 cm DBH trees to be ~12,500 per hectare. This is the maximum density-dependent value we apply to the other Arizona and New Mexico plot data for estimating correlations from the PDF. We also included an aerial LiDAR segmentation of tree height (See Swetnam Dissertation 2013 and Swetnam and Falk *in review* for full details) for the entire MCN (289 Ha [1^2 mi], n = 106,529).

Barro Colorado Island, Panama

The last dataset we use is the 1995 survey of the 50 Ha Barro Colorado Island (BCI) Experimental Forest in Lake Gatun, Panama, 9.15° N, 79.85° W. The species

composition of BCI shares no species in common with the semi-arid forests in Arizona and New Mexico. We included the BCI data here because of the relevance to the MST quantitative theory of forest structure (West et al. 2009, Enquist et al. 2009), and to show scaling of SFD is invariant regardless of forest type. All trees >1cm were stem mapped and measured, resulting in a dataset of 229,068 trees (Condit et al. 1999). The BCI dataset has a measured maximum density of ~6,000 per hectare for trees 1.0 cm DBH. Previous research has shown BCI exhibits a -2 scaling exponent for its diameter SFD (West et al. 2009, Enquist et al. 2009).

Disturbances in semi-arid forest ecosystems

Most forests undergo continuous and/or periodic disturbances, e.g. wildfires, trampling, herbivory, wildfire, tree fall from wind-throw, root rot, fatal attacks by insect pests, lightning, or mechanical damage by large mammals. Disturbance from steady state can affect a forest SFD in several ways: (1) it can kill all of the individuals in an area; (2) it can have a directional mortality factor based on the individual species, its size, age, and metabolic demands; or (3) it can be randomly distributed across species, sizes, and age classes. We know *a priori* certain kinds of disturbance affect semi-arid forest structure in specific ways: high intensity fire can kill all the individuals in an entire area (having 'high severity'); low-intensity fire may only kill understory trees and have little effect on overstory trees which have thick bark and high canopy base heights (having 'low severity'). Notably, the concept of a 'mixed-severity' fire is really just a combination of low and high severity measured together without accounting for the spatial scale changes across a burned area. Other types of disturbance, such as from insect attack or drought, may impact specific size classes of trees differently depending on their spatial configuration and competitive status with their neighbors.

Globally most forests experience disturbance as an initiation and/or episodic part of their life cycle and possibly as the agent of their death. We were interested in the generality of the origin and dynamics of size distributions in forests with episodic recruitment and disturbance. In Arizona and New Mexico, where our study is focused, the historical and recent land use histories are fairly well known (Woolsey 1911, Shreve

1915, Whittaker and Niering 1965, 1975, Niering and Lowe 1984, Hoffmann and Istock 1995, Allen et al. 2002, Falk and Swetnam 2002, Falk et al. 2004, Brown and Wu 2005, Lynch 2009, Sanderson and Koprowski 2009, Swetnam et al. 2009, Swetnam et al. 2011). Semi-arid forests in western North America exhibit overstory population densities which are thought to be well below their potential self-thinning limit. For example, prior to European settlement in the southwestern US, region wide ponderosa pine (*Pinus ponderosa var. brachyptera*) forests were observed to be predominantly made up of open stands of large-diameter trees (Woolsey 1911, Shreve 1915, Allen et al. 2002). After settlement pervasive logging, grazing, and fire suppression led to major conversions in structure and ecosystem function (Allen et al. 2002). The mechanism responsible for such open conditions and lack of understory is attributed widely to frequent low-severity surface fire regimes which would have been lethal to the vast majority of small trees in the stand but beneficial to the larger trees through the release of nutrients and a subsequent lack of competing understory vegetation (Swetnam and Betancourt 1990, Grissino-Mayer et al. 1995, Allen et al. 2002, Falk et al. 2004, Swetnam et al. 2009). Less frequently fires achieved high severity and burned smaller areas, importantly, these events occurred on a longer interval than wide spread low-severity fires based on the observed ages of old-growth stands of forests across the western US (Allen et al. 2002, Falk et al. 2004, Brown and Wu 2005, Margolis et al. 2011). Other less common disturbances (relative to surface fires) include: semi- to multi-decadal droughts (Williams et al. 2012), wind and topography driven crown fires (Margolis et al. 2011), and periodic insect outbreaks (Swetnam and Lynch 1993). The legacies of lower frequency high impact events, such as drought, insect outbreaks, and catastrophic fires should not be overlooked or discounted as they likely played significant roles in deriving existing forest structure.

Analytical methods

Linear Regression and Maximum Likelihood Estimation

We found the shape parameter of the Pareto and upper-truncated Pareto by maximum likelihood estimation (MLE) (White et al. 2008) in MATLAB 2012b (The

Mathworks 2012). MLE is considered superior to least squares for estimating the scaling exponent of power-laws because of the biases of using linear binning on continuous data (White et al. 2008). Both the Pareto and truncated-Pareto MLE algorithm (White et al. 2008) use the continuous distribution of data without binning.

An argument against fitting the tapered Pareto by MLE is made by Schoenberg and Patel (2011): the estimates of the parameter θ are dependent on the lower truncation point β, and may be unstable in cases where β is estimated from data. Here we only present the Pareto and truncated-Pareto distributions fit by MLE.

In the cases of the semi-arid forest data we had to establish where we felt the β value was no longer divergent [e.g. increasing in frequency with size]. We estimated β based on the point at which the frequency of tree size stopped increasing and began a decline toward the upper tail; for the upper-truncation limit ω we used the largest observed tree size in the dataset. In Table 1 we report the range $\beta < x < \omega$ from which we fit the probability distribution by MLE. The lower-tail of the observed SFDs of the semi-arid forests [with the exception of the Monument Canyon competition data] were unlikely to exhibit density dependence across small size scale, as they were observed to have 'open structure' [i.e. the lack of an understory].

To visualize the data we linearly binned in 1 cm width bins for diameter or 0.5 m bins for height. The choice in bin width was set at a scale which the variance between bins was small. Because binning is subjective we did not attempt to fit the exponent by least-squares regression.

Results

Size Frequency Distribution: Diameter

Despite differences in sample size and area of measurement the scaling exponent with a decline in abundance in DBH, as estimated by MLE, at all sites ranged between $-2.172 < \alpha < -1.94$ for the Pareto distribution and $-2.005 < \alpha < -1.893$ for the upper-truncated Pareto (Table 1a); these values support the self-thinning prediction of MST $\alpha = -2$, and result in a failure to reject our null hypothesis.

The linearly binned SFD appear to begin declining in Monument Canyon > 1.75 cm, Pinaleno > 7 cm, Valles Caldera > 6 cm, and BCI > 1 cm (Figure 3). For the truncated Pareto we set ω (Eq. 9, 10) as the largest observed DBH in each dataset: Monument Canyon = 80.0 cm, Pinaleno = 124.4 cm, Valles Caldera = 113.0 cm, and BCI = 246.0 cm (Table 1a).

In Figure 3 we plot the three semi-arid forests and the BCI DBH distributions (in 1 cm bin width) on a per-hectare (Table 1a) frequency basis. The three semi-arid forests all share a similar SFD scaling parameter at the hectare scale. The BCI SFD, on the other hand, is lower than the other three semi-arid forest datasets. This is because: (1) the presence of very large trees in the BCI dataset mean a few large trees fill a significantly larger area resulting in less area for smaller trees to pack into, and (2) the canopy diameter allometry relative to bole diameter of broad-leaved tropical trees in the BCI are likely much wider than conifer trees in the three semi-arid forest data sets.

Size Frequency Distribution: Height

The field data from BCI and Monument Canyon competition plots did not include height measurement so these data are not included. Locations with height data are the Pinaleño and Valles Caldera; we also include an airborne laser swath mapping inventory from Monument Canyon (data from Swetnam Dissertation Appendix C) (Figure 3).

The 0.5 m wide linearly binned SFD for height begins declining in the Pinaleño at > 5 m and the Valles Caldera > 5.5 m (Figure 4). We also plot an aerial LiDAR stem segmentation from Swetnam and Falk (*in review*) and Swetnam et al. (*in review*) for the entire Monument Canyon research natural area; we know *a priori* the segmentation is increasingly inaccurate for trees > 11 m in height, so we set the scale parameter at 11 m (Figure 3). For the truncated Pareto we set ω (Eq. 9, 10) as the tallest tree in each dataset: Pinaleño = 35.9 m, Valles Caldera = 38.4 m, Monument Canyon = 48.2 m (Table 1a).

The observed range of exponents as estimated by MLE, at the three sites ranged between $-3.358 < \alpha < -3.094$ for the Pareto distribution and $-2.983 < \alpha < -2.93$ for the upper-truncated Pareto (Table 1b); these values support the self-thinning prediction of MST $\alpha = -3$.

The frequency distribution of very large trees appears to diverge exponentially in the upper tail; we interpret this upper-truncation or taper in the distribution as due to the physiological limitation of vertical height (Koch et al. 2004, Kempes et al. 2011), and not as a violation of MST.

Discussion

At least two mechanisms modify the emergent power law scaling in the semi-arid forest data: (1) external processes [e.g. disturbances create open spaces and kill trees across all size classes], and (2) internal processes [e.g. physiological and resource limits lead to a truncation in the size distribution at the tallest tree height]. Regardless of these process noises we found the different semi-arid forest study areas still exhibit MST predicted power-law behavior [i.e. r^{-2} and h^{-3}] for the MLE derived shape parameter α [Eq. 2]. Because process noise is also present and the distribution diverges no one of the semi-arid study area forest was considered to be at steady-state across its entire SFD in the same way as the Barro Colorado Island data (Figure 3). In the semi-arid forests the observed frequency distribution tended to be below the absolute packing-rule density dependence, anecdotally the absence is the result of episodic disturbance and spatially is expressed as an increase in the lacunarity of the stand and a decrease in the fractal dimension of the Zeta function (Kellner and Asner 2009, Shier and Bourke 2013). A lack of continuous recruitment after disturbance has left these gaps unoccupied in the present time in more locations than would be expected under MST-predicted self-thinning. Even if these gaps are eventually colonized the cohort of new individuals would be expected to exhibit self-thinning amongst like-size seedlings, and not the older overstory. A second point is that divergence from a heavy tail Pareto power-law is expected for the upper tail where stochastic mortality [e.g. wind-throw, root-rot, lightning, fire, insect attack, etc], physiological limitation [e.g. maximum tree height due to hydraulic cavitation], and resource limitation [e.g. total soil available water and nutrients] truncate the maximum size trees can attain in a lifetime. This effect is most notable in the semi-arid forest maximum tree height which does not exceed 49 m.

Observed density dependence in semi-arid forests

To exhibit power-law behavior in the lower tail of its size distribution a forest floor must support a carpet of seedlings and saplings amongst the boles of larger trees. Remarkably this was the observed case for Monument Canyon. Prior to a mastication thinning project in 2008 Falk et al. (unpublished data) observed an approximately 90-year old cohort of *P. ponderosa* had stem densities as high as 6,000 stems per hectare. Most of the trees in the cohort were stagnated and did not exhibit typical healthy growth forms [i.e. vertical and straight standing boles, and a large amount of canopy needles relative to tree mass], suggesting the stand was indeed undergoing density-dependent self-thinning. The competitive strategy of these trees had changed from an open environment where saplings may not be directly competing with their neighbors to a closed environment encouraging fast-growing individuals who had less mass, and subsequently a higher mortality rate, analogous to what Smith et al. (2013) proposed for animal populations and Enquist et al. (2009) predict in forests. Conversely, on similar environmental conditions an open stand of 90-year old *P. ponderosa* averaged 35 cm DBH in the Pinaleño Mountains (O'Connor et al. 2010) whereas the dense stand of Monument Canyon trees were all <10 cm DBH. In 2009, a year following the mastication treatment in Monument Canyon, field observers found by measurement and extrapolation a maximum density of ~26,000 seedlings/ha smaller than 1cm diameter above root crown in the newly masticated openings.

The slow rate of decomposition in semi-arid forests means dead wood can remain on the surface for decades to centuries (White 1986, Hart 1992). Without a faster source of biogenic decomposition downed woody fuels do not return their nutrients to the soil and can infill the surface area needed for new trees to establish. The presence of woody debris decreases the potential of a forest to completely fill all of the available space. Such was the case in Monument Canyon until the mastication project removed the tangled dog-hair thicket, after which an immediate re-initiation of seedlings took place.

Effects of physiology on the size frequency distribution

The results in Figure 3 and Table 1a show beyond the β estimate the diameter SFD of the semi-arid forests exhibit an MST predicted decline in abundance regardless of process noise. We infer that external and internal process noise in these systems are what result in the divergence of the upper tail, and a lack of complete space filling [e.g. open stand structure] result in the divergence in the lower tail of the distribution.

For most tree species the limit to maximum tree height are thought to be caused by (1) the species mechanistic support potential, e.g. its elastic buckling strength (McMahon 1973, McMahon and Kronauer 1976); (2) the maximum energy and resource flux capacity of the organism in its exclusive footprint (Enquist and Niklas 2001, Savage et al. 2010, Kempes et al. 2011); and (3) the tension on the vertical water column with tracheid fluid conductance from gravity and the negative water potential of the atmosphere (Koch et al. 2004, Kempes et al. 2011). Prior to 20th century logging the three semi-arid forests had more mature tall and massive trees based both on historical accounts and physical measurements taken from remnant stumps in the observed forest plots. Today it is rare, if at all possible, to find an undisturbed stand of trees in unprotected areas. The complex topography and associated variability of soil depth, and the lack of old growth stands left from which to base assumptions of maximum size, complicates the application of the tapered and truncated Pareto for estimating maximum tree size in forests still recovering from mid- to late-20th century logging.

For the observed tree height data the MST -3 exponent for the Pareto distribution was found when β was determined *a posteriori*: we selected β where the SFD stopped increasing and began to decline toward the upper tail. In Table 1b we note the range of heights at which the SFD exhibited the reported MLE slopes. If the β parameter is increased the scaling exponent α rapidly becomes more negative. Importantly, the truncation height ω of a location co-varies with first order spatial affects (such as soil depth and water availability). Without accounting for an upper truncation in height, the theoretical -3 self-thinning exponent breaks down and can result in more negative α exponents.

Effect of disturbance on the size frequency distributions

Our results are equivalent to what Westoby (1981, 1984) called the Competition–Density (C–D) effect: at initiation states communities increase until they reach a theoretical self-thinning line, but at intermediate time periods only portions of their SFD are self-thinning. Every forest cohort scales to the gap or patch scale it occupies. Pulsed age-size structure is therefore not disruptive of emergent self-organization within the stand, and follows the metastable state hypothesis (Reed and McKelvey 2002, Moritz et al. 2011). In fire-frequent systems open canopy gaps may experience repeated recruitment and disturbance events before a tree establishes and eventually enters the overstory canopy. Forests with episodic disturbances are therefore expected to have gaps where either continuous or pulsed recruitment of new cohorts that may soon recolonize or be killed (Falk et al. 2007, McKenzie et al. 2011).

Other studies that discuss spatial patterns of disturbance in forests (White 1985, Lorimer 1989, Foster 1997, 1998, Franklin et al. 2002, Larson and Churchill 2008, 2012) do not mention scale explicitly in quantifying the level of disturbance in the system. Recent work by Anfodillo et al. (2013) shows how simple allometric relationships, at the individual scale, shape the structure of the whole forest community and suggest the difference in the shape parameter α and the observed SFD provide diagnostic tools to access the impact of disturbance on the forest community. Our findings are further support of this concept, while staying in good agreement with the MST.

Today's ecologists must grapple with how process-scale ecological events become emergent features on landscapes, and how large-scale top-down processes, such as disturbances, influence mechanisms at the stand to individual level. Measuring successional states in forests, by analyzing their patch size distribution and the SFD of trees within similarly constrained spatial units [e.g. the 'stand-level'] will allow ecologists to better gauge the resilience and resistance of forests to potential future disturbance events and increased climatic variability [i.e. more intense droughts or pluvials (Woodhouse et al. 2005)]. We believe the results of this paper lay the foundation for further analyzing forest disturbances spatially and for delineating the effects of

disturbance by other factors such as geographic location and local climate conditions. We attempted here to establish a framework for analyzing ecosystem pattern with field and aerial LiDAR datasets, and show how these data can be used to assess both forest structure, and forest health.

Understanding ecosystem complexity will be instrumental in managing landscapes for future ecosystem service and ecosystem stability monitoring. Specifically, as global warming causes more frequent and intense weather events the complete reorganization of extant ecosystems are predicted to occur (Dale et al. 2001, Westerling et al. 2006, Chapin et al. 2011, Williams et al. 2012). Predicting ecosystem vulnerability to catastrophic disturbance will be a critical component of management and monitoring in the arid western United States and indeed throughout the Earth System in the near future.

Conclusions

Our conceptual figures suggest how episodic recruitment and disturbance influence the size distribution of a forest. The overall effect of disturbance in the stand is characterized by the degree of departure from an MST predicted self-thinning line. For example, low intensity disturbances [e.g. surface fire] kills most of the small individuals in the stand, making the SFD divergent below a specific size, after which the SFD may begin to experience self-thinning. Gap sizes larger than the largest individual lead to a decrease in the normalization constant used to estimate the maximum density of the stand.

This work extends the MST to allow integration of more detailed processes in a predictive model of forest form and function. Our findings suggest MST mechanisms are conserved across physical scale in both understory trees and overstory trees in the example semi-arid systems. If the window of observation is carefully attenuated the scaling behavior of the forest is clear to see: at small spatial scale saplings compete in a density-dependent fashion until they are over-run at some point by a disturbance, the survivors then grow without inter-tree competition until they once again crowd their neighbors [either with their canopies or roots] for a limiting resource, in accordance with the MST.

References

Allen, C. D., M. Savage, D. A. Falk, K. F. Suckling, T. W. Swetnam, T. Schulke, P. B. Stacey, P. Morgan, M. Hoffman, and J. T. Klingel. 2002. Ecological restoration of southwestern ponderosa pine ecosystems: a broad perspective. Ecological Applications 12:1418-1433.

Anfodillo, T., M. Carrer, F. Simini, I. Popa, J. R. Banavar, and A. Maritan. 2013. An allometry-based approach for understanding forest structure, predicting tree-size distribution and assessing the degree of disturbance. Proceedings of the Royal Society B: Biological Sciences 280(1751). 22 p.

Bak, P., K. Chen, and C. Tang. 1990. A forest-fire model and some thoughts on turbulence. Physics letters A 147:297-300.

Bak, P., C. Tang, and K. Wiesenfeld. 1987. Self-organized criticality: An explanation of 1/f noise. Physical Review Letters 59:381-384.

Balmat, J., and J. Kupfer. 2004. Assessment of timber resources and logging history of the Valles Caldera National Preserve. University of Arizona Technical Report VCT04011 for Valles Caldera Trust, Tucson, AZ.

Brown, J. H., J. F. Gillooly, A. P. Allen, V. M. Savage, and G. B. West. 2004. Toward a metabolic theory of ecology. Ecology 85:1771-1789.

Brown, P. M., and R. Wu. 2005. Climate and disturbance forcing of episodic tree recruitment in a southwestern ponderosa pine landscape. Ecology 86:3030-3038.

Chapin III, F. S., and P. P. A. Matson. 2011. Principles of terrestrial ecosystem ecology. Springer.

Chen, K., P. Bak, and M. H. Jensen. 1990. A deterministic critical forest fire model. Physics Letters A 149:207-210.

Condit, R. 1998. Tropical forest census plots: methods and results from Barro Colorado Island, Panama and a comparison with other plots. Springer. 217 p.

Coomes, D. A., and R. B. Allen. 2009. Testing the metabolic scaling theory of tree growth. Journal of Ecology 97:1369-1373.

Coomes, D. A., and R. B. Allen. 2007. Mortality and tree-size distributions in natural mixed-age forests. Journal of Ecology 95:27-40.

Coomes, D. A., R. P. Duncan, R. B. Allen, and J. Truscott. 2003. Disturbances prevent stem size-density distributions in natural forests from following scaling relationships. Ecology Letters 6:980-989.

Coomes, D. A., E. R. Lines, and R. B. Allen. 2011. Moving on from Metabolic Scaling Theory: hierarchical models of tree growth and asymmetric competition for light. Journal of Ecology 99:748-756.

Dale, V. H., L. A. Joyce, S. McNulty, R. P. Neilson, M. P. Ayres, M. D. Flannigan, P. J. Hanson, L. C. Irland, A. E. Lugo, and C. J. Peterson. 2001. Climate Change and Forest Disturbances: Climate change can affect forests by altering the frequency, intensity, duration, and timing of fire, drought, introduced species, insect and pathogen outbreaks, hurricanes, windstorms, ice storms, or landslides. Bioscience 51:723-734.

Deng, J., W. Zuo, Z. Wang, Z. Fan, M. Ji, G. Wang, J. Ran, C. Zhao, J. Liu, and K. J. Niklas. 2012. Insights into plant size-density relationships from models and agricultural crops. Proceedings of the National Academy of Sciences 109:8600-8605.

Enquist, B. J., J. H. Brown, and G. B. West. 1998. Allometric scaling of plant energetics and population density. Nature 395:163-165.

Enquist, B. J., E. P. Economo, T. E. Huxman, A. P. Allen, D. D. Ignace, and J. F. Gillooly. 2003. Scaling metabolism from organisms to ecosystems. Nature 423:639-642.

Enquist, B. J., and K. J. Niklas. 2002. Global allocation rules for patterns of biomass partitioning in seed plants. Science 295:1517-1520.

Enquist, B. J., and K. J. Niklas. 2001. Invariant scaling relations across tree-dominated communities. Nature 410:655-660.

Enquist, B. J., G. B. West, and J. H. Brown. 2009. Extensions and evaluations of a general quantitative theory of forest structure and dynamics. Proceedings of the National Academy of Sciences 106:7046-7051.

Enquist, B. J., G. B. West, E. L. Charnov, and J. H. Brown. 1999. Allometric scaling of production and life-history variation in vascular plants. Nature 401:907-911.

Falk, D. A. 2004. Scaling rules for fire regimes. Ph.D. Dissertation. Department of Ecology & Evolutionary Biology. University of Arizona, Tucson AZ.

Foster, D. R., J. D. Aber, J. M. Melillo, R. D. Bowden, and F. A. Bazzaz. 1997. Forest response to disturbance and anthropogenic stress. Bioscience 47:437-445.

Foster, D. R., D. H. Knight, and J. F. Franklin. 1998. Landscape patterns and legacies resulting from large, infrequent forest disturbances. Ecosystems 1:497-510.

Franklin, J. F., T. A. Spies, R. V. Pelt, A. B. Carey, D. A. Thornburgh, D. R. Berg, D. B. Lindenmayer, M. E. Harmon, W. S. Keeton, and D. C. Shaw. 2002. Disturbances and structural development of natural forest ecosystems with silvicultural implications, using Douglas-fir forests as an example. Forest Ecology and Management 155:399-423.

Grissino-Mayer, H.D., Baisan, C.H. and Swetnam, T.W. 1995. Fire history in the Pinaleño Mountains of southeastern Arizona: Effects of human related disturbances. In Debano, L.F., Gottfried, G.J., Hamre, R.H., Edminster, C.B., Ffolliott, P.F. and Ortega-Rubio, A., editors, Biodiversity and management of the Madrean Archipelago: the Sky Islands of Southwestern United States and Northwestern Mexico, US Department of Agriculture Forest Service, General Technical Report RM-GTR-264, 399–407.

Hart, S. C., M. K. Firestone, and E. A. Paul. 1992. Decomposition and nutrient dynamics of ponderosa pine needles in a Mediterranean-type climate. Canadian Journal of Forest Research 22:306-314.

Hastings, A., C. L. Hom, S. Ellner, P. Turchin, and H. C. J. Godfray. 1993. Chaos in ecology: is mother nature a strange attractor? Annual Review of Ecology and Systematics :1-33.

Kagan, Y. Y., and F. Schoenberg. 2001. Estimation of the upper cutoff parameter for the tapered Pareto distribution. Journal of Applied Probability 38:158-175.

Kellner, J. R., and G. P. Asner. 2009. Convergent structural responses of tropical forests to diverse disturbance regimes. Ecology Letters 12:887-897.

Kempes, C. P., G. B. West, K. Crowell, and M. Girvan. 2011. Predicting maximum tree heights and other traits from allometric scaling and resource limitations. PloS one 6:e20551.

Kerkhoff, A. J., and B. J. Enquist. 2007. The implications of scaling approaches for understanding resilience and reorganization in ecosystems. Bioscience 57:489-499.

Koch, G. W., S. C. Sillett, G. M. Jennings, and S. D. Davis. 2004. The limits to tree height. Nature 428:851-854.

Kohyama, T. 1994. Size-structure-based models of forest dynamics to interpret population-and community-level mechanisms. Journal of Plant Research 107: 107-116.

Kohyama, T. 1993. Size-structured tree populations in gap-dynamic forest--the forest architecture hypothesis for the stable coexistence of species. Journal of Ecology 131-143.

Kolokotrones, T., E. J. Van Savage, and W. Fontana. 2010. Curvature in metabolic scaling. Nature 464:753-756.

Larson, A. J., and D. Churchill. 2012. Tree spatial patterns in fire-frequent forests of western North America, including mechanisms of pattern formation and implications for designing fuel reduction and restoration treatments. Forest Ecology and Management 267:74-92.

Larson, A. J., and D. Churchill. 2008. Spatial patterns of overstory trees in late-successional conifer forests. Canadian Journal of Forest Research 38:2814-2825.

Lin, Y., U. Berger, V. Grimm, F. Huth, and J. Weiner. 2013. Plant Interactions Alter the Predictions of Metabolic Scaling Theory. PloS one 8:e57612.

Lorimer, C. G. 1989. Relative effects of small and large disturbances on temperate hardwood forest structure. Ecology 70:565-567.

Lynch AM (2009) Spruce Aphid, Elatobium abietinum (Walker) Life History and Damage to Engelmann Spruce in the Pinaleño Mountains, Arizona. In: Sanderson, H.R.; Koprowski, J.L., eds. The last refuge of the Mt. Graham red squirrel: Ecology of endangerment. Tucson, AZ: University of Arizona Press. p. 318-338.

Margolis, E., T. Swetnam, and C. Allen. 2011. Historical stand-replacing fire in upper montane forests of the Madrean Sky Islands and Mogollon Plateau, southwestern USA. Fire Ecol 7:88-107.

McGaughey, R. 2012. FUSION/LDV: Software for LIDAR Data Analysis and Visualization, Version 3.01. US Department of Agriculture, Forest Service, Pacific Northwest Research Station, University of Washington. Available online at: http://forsys.cfr.washington.edu/fusion/fusionlatest.html (last accessed 24 August 2012) .

McMahon, T. 1973. Size and shape in biology. Science 179:1201-1204.

McMahon, T. A., and R. E. Kronauer. 1976. Tree structures: deducing the principle of mechanical design. Journal of theoretical biology 59:443-466.

Mori, S., K. Yamaji, A. Ishida, S. G. Prokushkin, O. V. Masyagina, A. Hagihara, A. R. Hoque, R. Suwa, A. Osawa, and T. Nishizono. 2010. Mixed-power scaling of whole-plant respiration from seedlings to giant trees. Proceedings of the National Academy of Sciences 107:1447-1451.

Moritz, M. A., P. F. Hessburg, and N. A. Povak. 2011. Native fire regimes and landscape resilience. Pages 51-86 *in* Anonymous The Landscape Ecology of Fire. Springer.

Muller-Landau, H. C., R. S. Condit, J. Chave, S. C. Thomas, S. A. Bohlman, S. Bunyavejchewin, S. Davies, R. Foster, S. Gunatilleke, and N. Gunatilleke. 2006. Testing metabolic ecology theory for allometric scaling of tree size, growth and mortality in tropical forests. Ecology Letters 9:575-588.

Newman, M. E. 2005. Power laws, Pareto distributions and Zipf's law. Contemporary physics 46:323-351.

Niering, W. A., and C. H. Lowe. 1984. Vegetation of the Santa Catalina Mountains: community types and dynamics. Vegetatio 58:3-28.

Niklas, K. J., and B. J. Enquist. 2001. Invariant scaling relationships for interspecific plant biomass production rates and body size. Proceedings of the National Academy of Sciences 98:2922-2927.

Niklas, K. J., J. J. Midgley, and R. H. Rand. 2003. Tree size frequency distributions, plant density, age and community disturbance. Ecology Letters 6:405-411.

O'Connor, C., D.A. Falk, A.M. Lynch, C.P. Wilcox, T.W. Swetnam, T.L. Swetnam. 2010. Growth and demography of Pinaleño high elevation forests. RJVA 07-JV-11221615317 Performance Report. Tucson, AZ: University of Arizona, Laboratory of Tree-Ring Research; School of Natural Resources and the Environment. 21 p.

Olberding, S. D., and M. M. Moore. 2008. Fort Valley Experimental Forest-A Century of Research 1908-2008 (P-53).

Pareto, V. 1896. Cours d'économie Politique, reprinted as a volume of Oeuvres Completes. Droz, Geneva 1965.

Patel, R. D. 2011. Testing Local Self-Similarity in Univariate Heavy-Tailed Data. Doctoral dissertation, University of California Los Angeles.

Patel, R. D., and F. P. Schoenberg. 2011. A graphical test for local self-similarity in univariate data. Journal of Applied Statistics 38:2547-2562.

Perry, D. A. 1995. Self-organizing systems across scales. Trends in Ecology & Evolution 10:241-244.

Popescu, S. C., R. H. Wynne, and R. F. Nelson. 2002. Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window size. Computers and Electronics in Agriculture 37:71-95.

Purves, D. W., J. W. Lichstein, N. Strigul, and S. W. Pacala. 2008. Predicting and understanding forest dynamics using a simple tractable model. Proceedings of the National Academy of Sciences 105:17018-17022.

Reed, W. J., and K. S. McKelvey. 2002. Power-law behaviour and parametric models for the size-distribution of forest fires. Ecological Modelling 150:239-254.

Reineke, L. H. 1933. Perfecting a stand-density index for even-aged forests. US Government Printing Office.

Sanderson, H. R., and J. L. Koprowski. 2009. The Last Refuge of the Mt. Graham Red Squirrel: Ecology of Endangerment. University of Arizona Press. 429 p.

Schoenberg, F. P., and R. D. Patel. 2012. Comparison of Pareto and tapered Pareto distributions for environmental phenomena. The European Physical Journal Special Topics 205:159-166.

Seuront, L., and J. G. Mitchell. 2008. Towards a seascape typology. I. Zipf versus Pareto laws. Journal of Marine Systems 69:310-327.

Shier, J., and P. Bourke. 2013. An Algorithm for Random Fractal Filling of Space. .

Smith, J. A., L. J. Baumgartner, I. M. Suthers, D. S. Fielder, and M. D. Taylor. 2013. Density-Dependent Energy Use Contributes to the Self-Thinning Relationship of Cohorts. The American Naturalist 181:331-343.

Sornette, D. 2009. Dragon-kings, black swans and the prediction of crises. .

Sperry, J. S., D. D. Smith, V. M. Savage, B. J. Enquist, K. A. McCulloh, P. B. Reich, L. P. Bentley, and E. I. Allmen. 2012. A species-level model for metabolic scaling in trees I. Exploring boundaries to scaling space within and across species. Functional Ecology 26:1054-1065.

Stark, C. P., and N. Hovius. 2001. The characterization of landslide size distributions. Geophysical Research Letters 28:1091-1094.

Storch, D., and K. J. Gaston. 2004. Untangling ecological complexity on different scales of space and time. Basic and Applied Ecology 5:389-400.

Strigul, N., D. Pristinski, D. Purves, J. Dushoff, and S. Pacala. 2008. Scaling from trees to forests: tractable macroscopic equations for forest dynamics. Ecological Monographs 78:523-545.

Swetnam, T.L., Falk, D.A., *in review*. A variable-area local maxima tool for segmenting individual trees from aerial LiDAR: allometric scaling rules to reduce error in individual tree segmentation. Forest Ecology and Management.

Swetnam, T.L., Falk, D. A., Lynch, A. M., Guertin, D. P., Yool, S.R., *in review*. A Variable-area Local Maxima tool for segmenting individual trees from aerial LiDAR: Predicting rank-size distributions of forest mid- and understory in two semi-arid conifer forests. Forest Ecology and Management.

Swetnam, T. W., C. Baisan, and H. Grissino-Mayer. 2009. Tree-ring perspectives on fire regimes and forest dynamics in mixed conifer and spruce-fir forests on Mt. Graham. Sanderson HR, Koprowski JL (tech eds) The last refuge of the Mt.Graham red squirrel: ecology of endangerment, University of Arizona Press, Tucson .

Swetnam, T. W., and J. L. Betancourt. 1990. Fire-southern oscillation relations in the southwestern United States. Science 249:1017-1020.

Ustin, S. L., and J. A. Gamon. 2010. Remote sensing of plant functional types. New Phytologist 186:795-816.

Van Leeuwen, M., and M. Nieuwenhuis. 2010. Retrieval of forest structural parameters using LiDAR remote sensing. European Journal of Forest Research 129:749-770.

Walker, B., C. S. Holling, S. R. Carpenter, and A. Kinzig. 2004. Resilience, adaptability and transformability in social--ecological systems. Ecology and society 9:5.

West, G. B., and J. H. Brown. 2005. The origin of allometric scaling laws in biology from genomes to ecosystems: towards a quantitative unifying theory of biological structure and organization. Journal of Experimental Biology 208:1575-1592.

West, G. B., J. H. Brown, and B. J. Enquist. 1997. A general model for the origin of allometric scaling laws in biology. Science 276:122-126.

West, G. B., J. H. Brown, and B. J. Enquist. 1999a. The fourth dimension of life: fractal geometry and allometric scaling of organisms. Science 284:1677-1679.

West, G. B., J. H. Brown, and B. J. Enquist. 1999b. A general model for the structure and allometry of plant vascular systems. Nature 400:664-667.

West, G. B., B. J. Enquist, and J. H. Brown. 2009. A general quantitative theory of forest structure and dynamics. Proceedings of the National Academy of Sciences 106:7040-7045.

West, G. B., B. J. Enquist, and J. H. Brown. 2009. A general quantitative theory of forest structure and dynamics. Proceedings of the National Academy of Sciences 106:7040-7045.

Westerling, A. L., H. G. Hidalgo, D. R. Cayan, and T. W. Swetnam. 2006. Warming and earlier spring increase western US forest wildfire activity. Science 313:940-943.

Westoby, M. 1984. The self-thinning rule. Advances in Ecological Research 14:167-226.

Westoby, M. 1981. The place of the self-thinning rule in population dynamics. The American Naturalist 118:581-587.

White, A. S. 1985. Presettlement regeneration patterns in a southwestern ponderosa pine stand. Ecology 66:589-594.

White, C. 1986. Effects of prescribed fire on rates of decomposition and nitrogen mineralization in a ponderosa pine ecosystem. Biology and Fertility of Soils 2:87-95.

White, E. P., S. Ernest, A. J. Kerkhoff, and B. J. Enquist. 2007. Relationships between body size and abundance in ecology. Trends in ecology & evolution 22:323-330.

White, E.P., Enquist, B.J., Green, J.L., 2008. On estimating the exponent of power-law frequency distributions. Ecology 89, 905-912.

Whittaker, R. H., and W. A. Niering. 1975. Vegetation of the Santa Catalina Mountains, Arizona. V. Biomass, production, and diversity along the elevation gradient. Ecology 771-790.

Williams, A. P., C. D. Allen, A. K. Macalady, D. Griffin, C. A. Woodhouse, D. M. Meko, T. W. Swetnam, S. A. Rauscher, R. Seager, and H. D. Grissino-Mayer. 2012. Temperature as a potent driver of regional forest drought stress and tree mortality. Nature Climate Change 3, 292-297.

Woodhouse, C. A., K. E. Kunkel, D. R. Easterling, and E. R. Cook. 2005. The twentieth-century pluvial in the western United States. Geophysical Research Letters 32:L07701.

Woolsey, T. S. 1911. Western yellow pine in Arizona and New Mexico. US Department of Agriculture, Forest Service.

Yoda, K. 1963. Self-thinning in over-crowded pure stands under cultivated and natural conditions.(In-traspecific competition among higher plants. XI.). J Biol Osaka City Univ 14:107-129.

Figure Captions

Figure 1: Theoretical size frequency distribution where β_0 is the maximum density of individual trees, β_x is the point at which log-log linear power-law behavior begins; α is the slope [scaling exponent] where the self-thinning imposed limit by maximum density dependence is observed; θ is imposed by physiological or resource limitation causing an exponential decline from log-log linear behavior, θ increases to ω at small area scale, and decreases to β at large area scale. Λ is the lacunarity function $f(\lambda_k)$ of the individual gaps λ_k and changes across scale, e.g. from sun-flecks up to inter-canopy gaps $< \omega$. Notably for gaps $> \omega$ the self-thinning line is reduced log-log linearly, reflecting a decrease in β_0.

Figure 2: (A) Location of the study areas, (B) aerial LiDAR coverage and plot locations, (C) Plot design 500 m^2 for Pinaleno, and 1,000 m^2 for Valles Caldera. The Barro Colorado (50 Ha) was an inventoried 50 ha rectangular area (no plots).

Figure 3: Sample size distribution for tree diameter [in 1cm wide linear bins] from the data in the 5 study areas (left) normalized to represent the frequency of trees per hectare. A reference -2 slope (black line) is the predicted MST exponent for tree diameter.

Figure 4: Sampled size distribution for tree height [in 0.5 m wide linear bins] from the data in the three semi-arid forest datasets with height measure observations (left panel). The frequency distribution normalized for area (estimated by the maximum density dependence per hectare) along with the truncated Pareto (Eq. 10) set to the maximum observed height (45 m) (thick black line, right panel). The data are from the field observed tree height in the Pinaleño, Valles Caldera; and aerial LiDAR segmentation of the entire Monument Canyon RNA (Swetnam and Falk *in review*). Linear axes (left) and log-log axes (right).

Figures
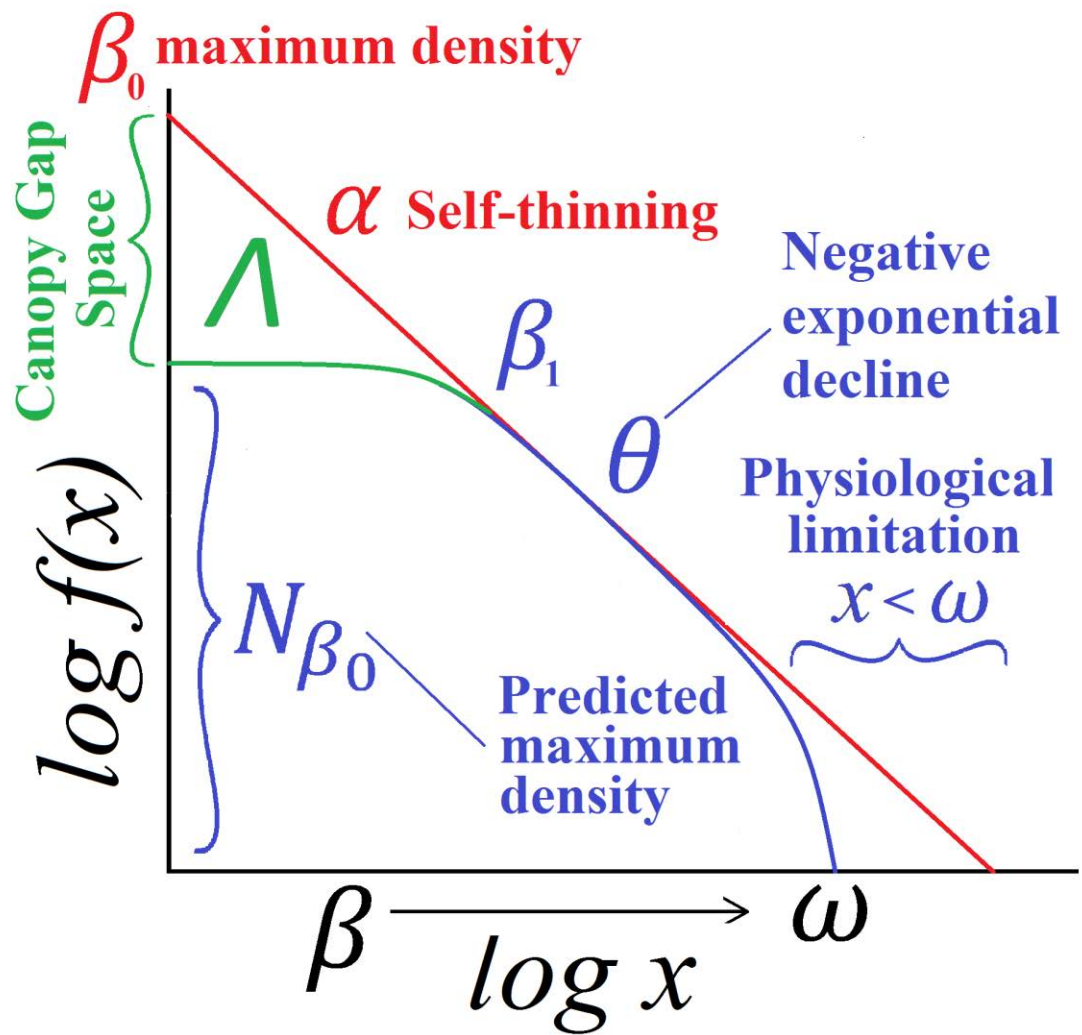
Figure 1.
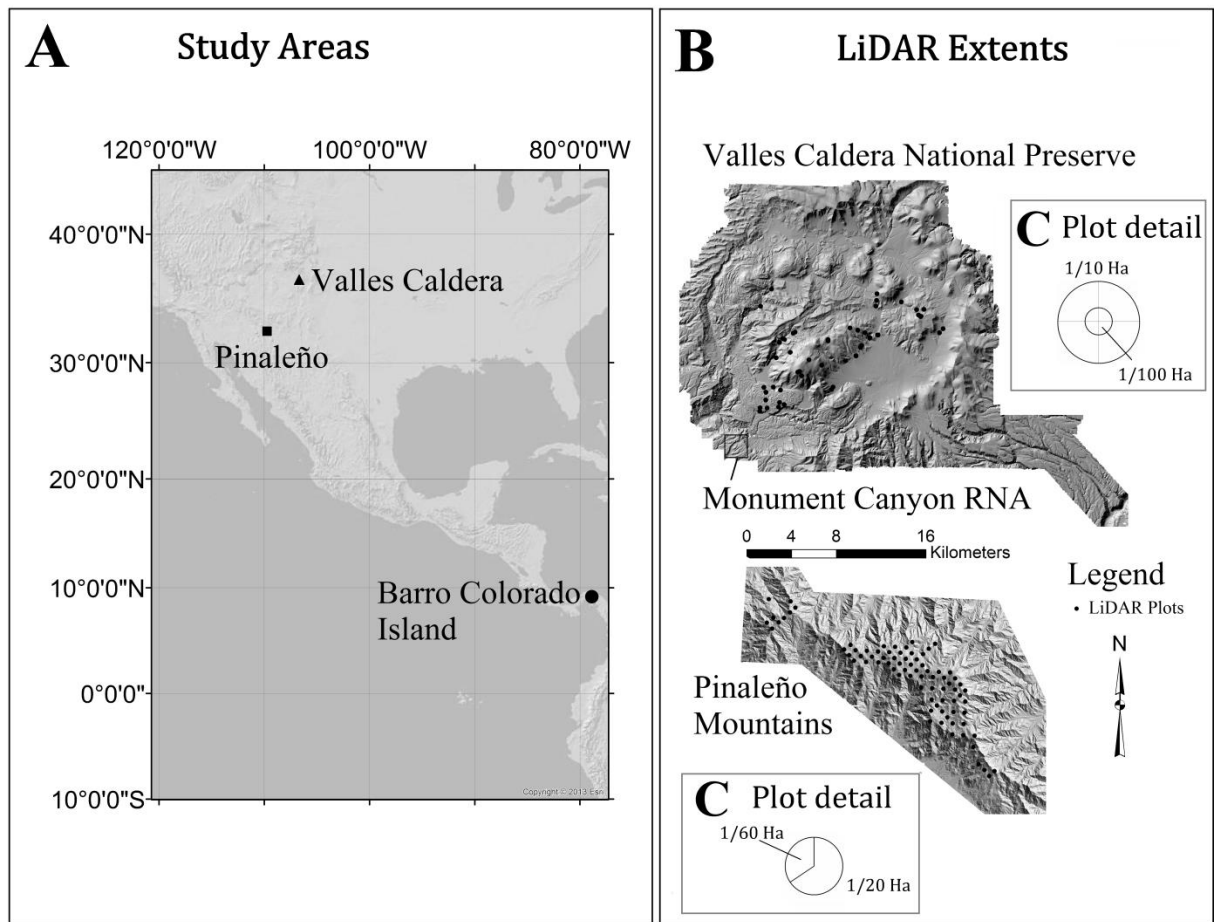
Figure 2.
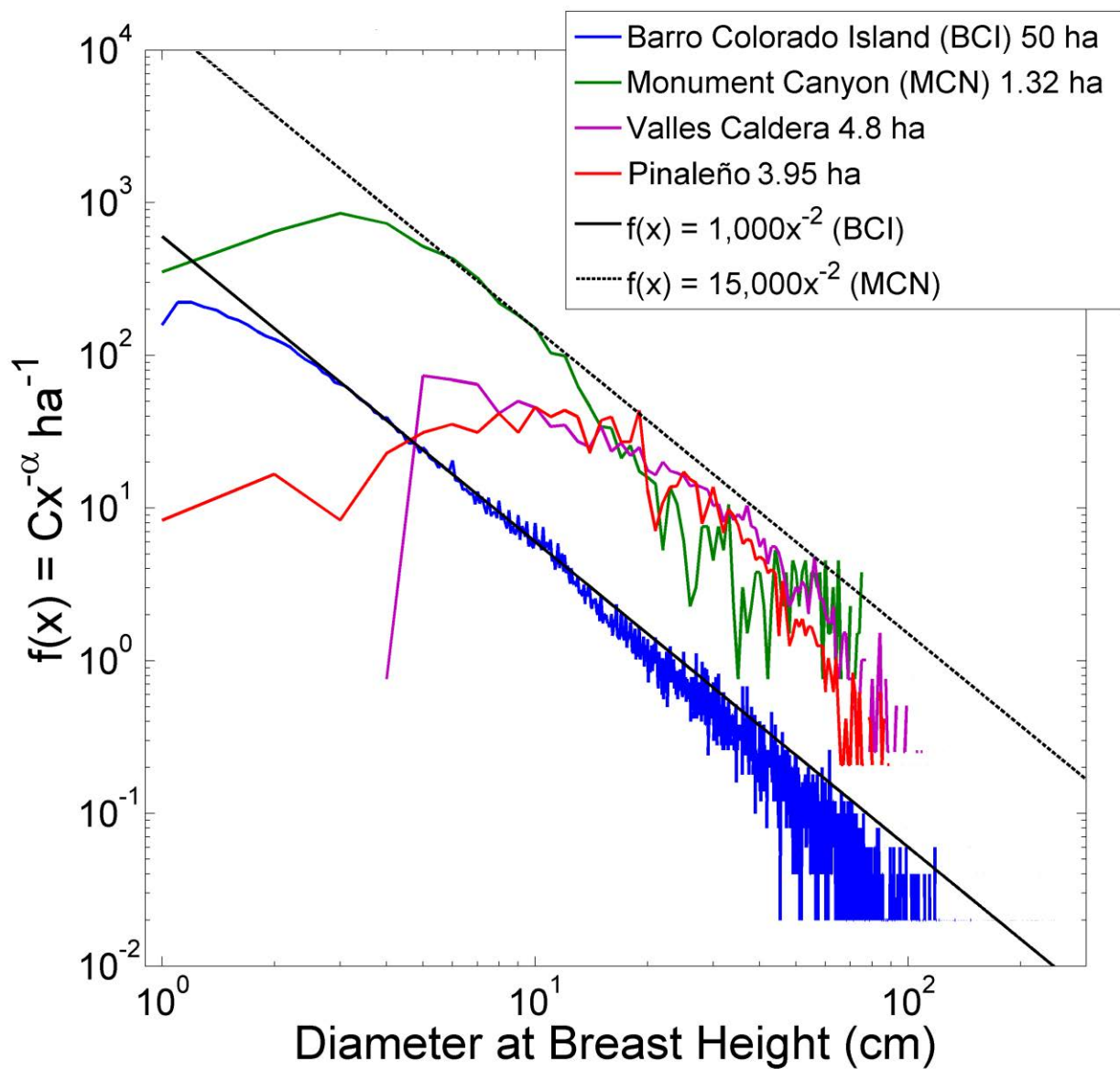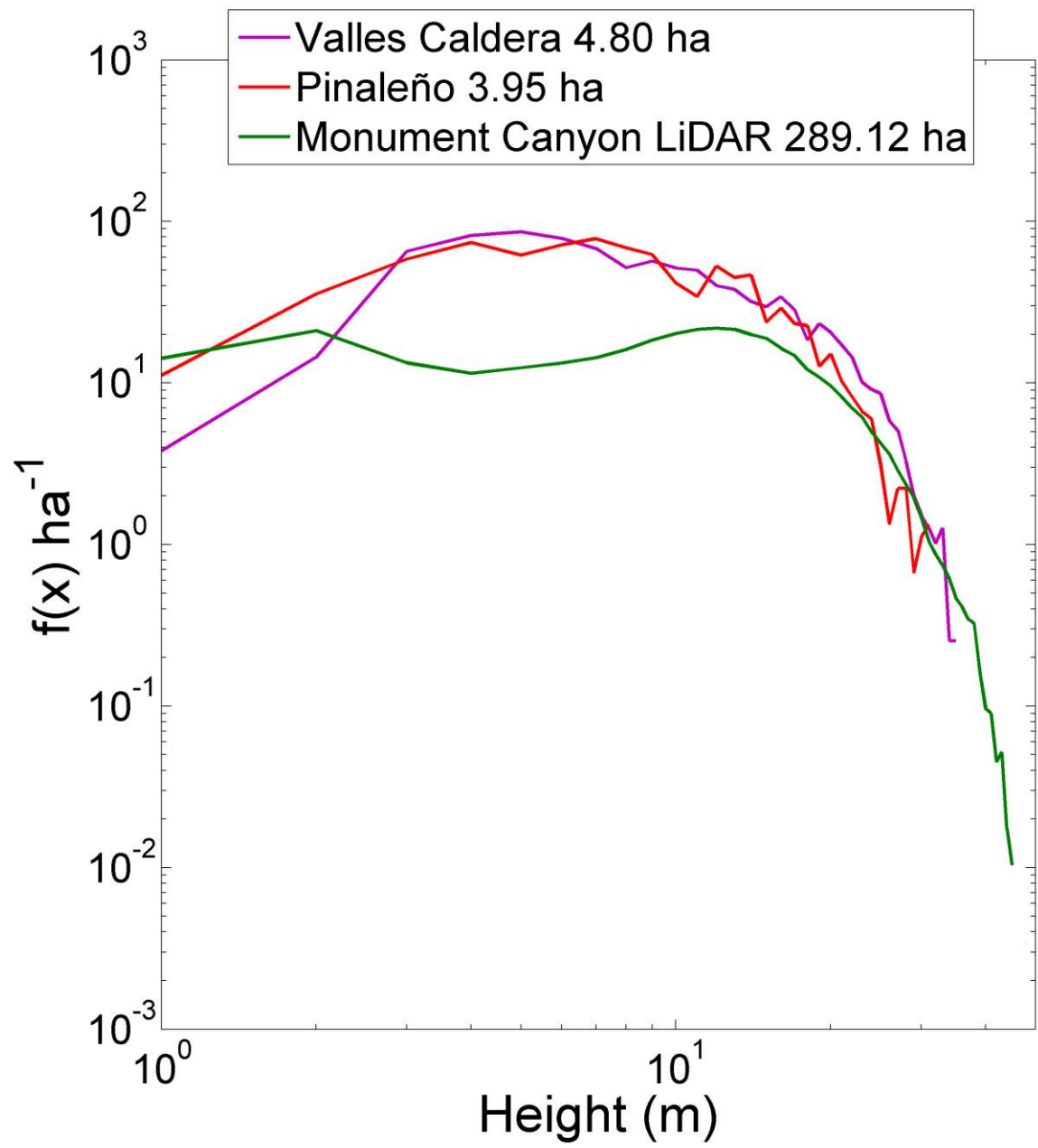
Figure 3.

Figure 4.

Tables

Table 1a. Pareto and truncated Pareto distribution fit by MLE (White et al. 2008) to diameter at breast height (cm).

Study area	Sample size		Diameter (cm)	Pareto	truncated Pareto
	ha	n	$\beta \leq x < \omega$	$MLE\ \hat{\alpha}$	$MLE\ \hat{\alpha}$
PM Plots	3.95	3,765	$7.0 < x < 124.4$	-2.11	-2.005
VCNP Plots	4.8	4,108	$6.0 < x < 113.0$	-2.172	-1.963
MCN Competition	1.32	6,691	$2.0 < x < 80.0$	-2.168	-2.080
BCI Inventory	50	229,068	$1.0 < x < 246.0$	-1.940	-1.908

Table 1b. Power-law and truncated Pareto fit by MLE [the same as Table 1a], but for tree height (m). BCI and MCN competition data are not included because height was not measured. The MCN height data are from an aerial LiDAR segmentation of the entire Research Natural Area (289.0 ha).

Study area	Sample size		Height (m)	Pareto	truncated Pareto
	ha	n	$\beta \leq x < \omega$	$MLE\ \hat{\alpha}$	$MLE\ \hat{\alpha}$
PM Plots	3.95	3,765	$5.5 < x < 35.9$	-3.185	-2.983
VCNP Plots	4.8	4,108	$5.0 < x < 38.4$	-3.094	-2.930
MCN VLM	289.0	106,529	$11.0 < x < 48.24$	-3.358	-2.953

Supplemental Materials A: Conceptual impacts on power-law size frequency distribution

The so-called 'packing-rule' (Rieneke 1933) suggests the frequency of individuals filling space follows a power law (Figure SM1 Panel A). Process-noise [e.g. stochastic mortality] is most easily detected in the upper tail as a divergence from the log-log linear scaling at small sample size (Figure SM1 Panel B). If the frequency of large individuals increases [e.g. an old growth forest] or there are large open spaces [e.g. meadows or clearings] the total number of individuals that fit into the remaining occupied space decreases the SFD non-linearly (Figure SM1 Panel C). The proportion of the area filled follows the packing rule as:

$$\delta = 1 - \left(\frac{\sum_i^\omega \lambda_i}{A} \right)$$
Equation SM1

where $\sum_i^\omega \lambda_i$ is the sum of all empty spaces λ_i from the area an ith size tree occupies up to ω [e.g. the area occupied by the largest single tree], and A is the total area (Figure SM1 Panel C). For gaps larger than the largest tree [$x_\omega < \lambda_\omega$], a log-log linear decrease in the population size results (Figure SM1 Panel C). The change in lacunarity across sizes becomes Λ a function of the absence in the distribution that sums for progressively smaller sizes as:

$$\Lambda = f(\lambda) = \sum_i^\omega \lambda_{i \to \omega}$$
Equation SM2

where the total number of individual trees that could fit into any canopy gap λ_x, whose area equals the primary size measure x as $\lambda_x = Ax^\alpha$ between $x_{min} < x_i < x_{max}$, and diverges in the lower tail (Figure SM1, Panel D). When we incorporate both the process noise ε along with Λ we see the SFD is divergent in both tails (Figure SM1, Panel E). When we incorporate all of the terms: the lacunarity gap distribution Λ, the proportion of filled space δ, the process noise ε, and the maximum density β we arrive at a function that best describes the observed size frequency distribution of a natural stand (Figure SM1, Panel F).

From these known states we have to make several assumptions: (1) occupied or empty spaces larger that are than the largest tree footprint decrease the absolute number

of individuals in the population log-log linearly by reducing the total area occupied, (2) the distribution of trees are following the -4/3 self-thinning rule for organism mass (Enquist et al. 1998, 2009, Deng et al. 2012] (Figure SM1, Panel C), and (3) in gaps smaller than the largest tree footprint the frequency of trees in the lower tail becomes divergent but does not change the self-thinning rule (Figure SM2).

Supplemental Materials B: Modeling stochastic noise

We defined 'disturbance' as any decline on the frequency of individuals in the population not associated with self-thinning. 'Heterogeneity' created in space after a disturbance is scale dependent and moves a system away from its steady-state attractor (Walker et al. 2004). In the period following the disturbance the system will begin to move back toward its strongest basin of attraction as new individuals are recruited into the newly open space (Westoby 1981,1984). The existing SFD of a forest is therefore a key measure of forest development relative to its potential steady-state basin of attraction.

A low-intensity disturbance event, i.e. a surface fire, affects a forest SFD differently based on the level of exposure and the continuity and density of the population in which the disturbance spreads. Surface fires in particular kill smaller trees that they come into contact with whereas the larger trees may be unaffected. Other size-dependent disturbances include herbivore browsing pressure, insect pests, and drought. These disturbances vary from having a stochastic event or directional effect on the SFD.

Seuront and Mitchell (2008) refer to disturbance effects on power-law distribution functions as noise 'contaminations'. The contamination of a power-law (Eq. A1) by external noise (and observation error) or internal noise is written as:

$$Y_n = kN^\alpha - \varepsilon \qquad \text{Equation SM3}$$

where ε is a stochastic process whose amplitude is a percentage of the maximum value of Y_n. Seuront and Mitchell (2008) also give examples of internal process noise within communities, such as oceanic plankton communities undergo greater herbivory pressures by copepods at higher population density. We consider Seuront and Mitchell's 'ingestion

function' from the perspective of a fire and its impact on the overall synchrony of the SFD in a forest stand. This so called ingestion function, was written as:

$$S = f(x) - 10^{\frac{f(x)}{k}}$$ Equation SM4

where k is a constant and the function represents the increased impact of disturbance on higher concentrations of individuals. Decreasing the value of k increases the impact of disturbance in the lower tail (Figure SM3). Anecdotally, the impact of a low severity fire would increase in areas of high population density; e.g. a continuous cohort of recently recruited seedlings and saplings, such as what was observed in Monument Canyon, where the fire could both spread and kill the smaller individuals in the space.

Applying Seuront and Mitchell's (2008) ingestion function S (Eq. A2) causes a divergence in a the Pareto distribution with variable levels of intensity; the larger the function the earlier the Pareto diverges from the limit, the smaller the longer it remains near the limit (Figure SM3, left panel). We modified k to the cross-over point in the SFD where the lower-tail ends its divergence from the power law. When we add in negative process noise ε (Figure SM3, right panel) the uncertainty in the simulated tail is close to what we observe in the actual data.

Supplemental Materials C: Data from other forest demography studies with disturbance

Following our initial analysis, reported in the main text, we sought to include other data from the published literature regarding density of stems per hectare. We incorporated Figure 1B from Coomes et al. (2003) who report with 2 cm linear binning tree diameter from 2 cm to 92 cm (22.6 ha) for forest plots in New Zealand; and the Gentry data from Enquist and Niklas (2001): 2 cm to 122 cm (61.8 ha). We also plot data from the HJ Andrews Experimental Forest (Oregon) from a 2007 inventory for 130 250 m^2 plots (Halpern and Dyrness 2010). We have re-plotted the three semi-arid forest and HJ Andrews data sets at 2 cm linear bins to match the scaling of the Coomes et al. (2003) data (Figure SM4).

Notably, at the frequency per hectare scale all of the datasets appear to be declining in abundance with size with similar maximum densities.

Halpern, C.; Dyrness, C. 2010. Plant succession and biomass dynamics following logging and burning in the Andrews Experimental Forest Watersheds 1 and 3, 1962-Present. Long-Term Ecological Research. Forest Science Data Bank, Corvallis, OR. [Database]. Available: http://andrewsforest.oregonstate.edu/data/abstract.cfm?dbcode=TP073 (31 October 2013).

Supplemental Figure Captions

Figure SM1: The different contaminations of a power-law size distribution. (A) power law with a negative scaling exponent, (B) the power law + negative process noise, (C) the power law multiplied by a constant representing the total area occupied ≤ 1, (D) the power law multiplied by a function of the increasing lacunarity Λ smaller than the area of the largest tree, (E) the power law with lacunarity and process noise, (F) the power law with lacunarity, area occupied, and process noise.

Figure SM2: Conceptual model of the effect of increased lacunarity on the size frequency distribution. The MST predicted negative scaling exponent does not change (orange color lines all have the same slope [e.g. scaling exponent]); what is modified is the total maximum density of the forest with an increasing frequency of gaps equal to or smaller than the individuals of each size class. The black line becomes the best-fit distribution, if we were to consider applying least-squares regression.

Figure SM3: The Pareto PDF with a -2 exponent minus density dependent $10^{f(x)/k}$ (left), and minus negative stochastic noise ε (right) for two Pareto PDF. The second distribution (gray) appears very similar to those that are observed in the semi-arid forest data.

Figure SM4: Example of the same semi-arid forest data from Arizona and New Mexico used in the main text, as well as data from two other separate studies: Coomes et al. (2003) and Halpern and Dyrness (2010) (left panel). When the data are plotted on the same frequency per hectare [f(x) ha ^ -1] (right panel) the distributions all appear to lock together along a similar decline in abundance with size which reflects MST $-$ 2 scaling.

Supplemental Figures
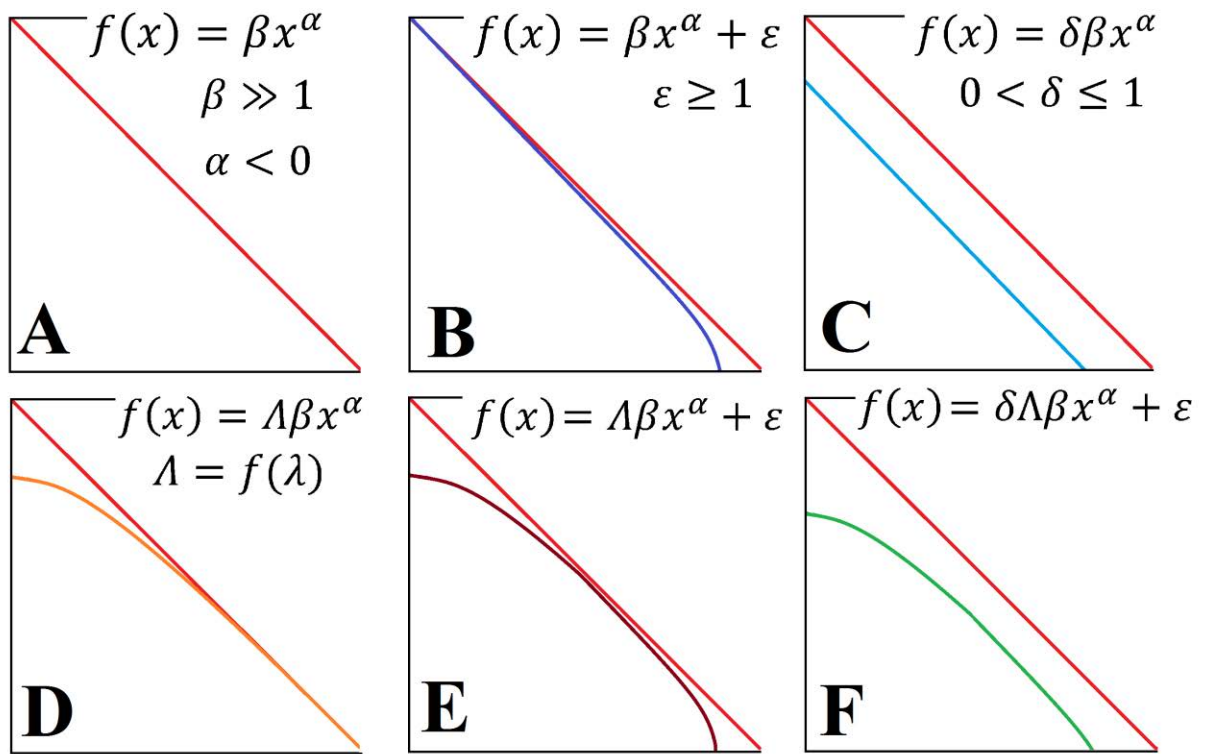
Figure SM1
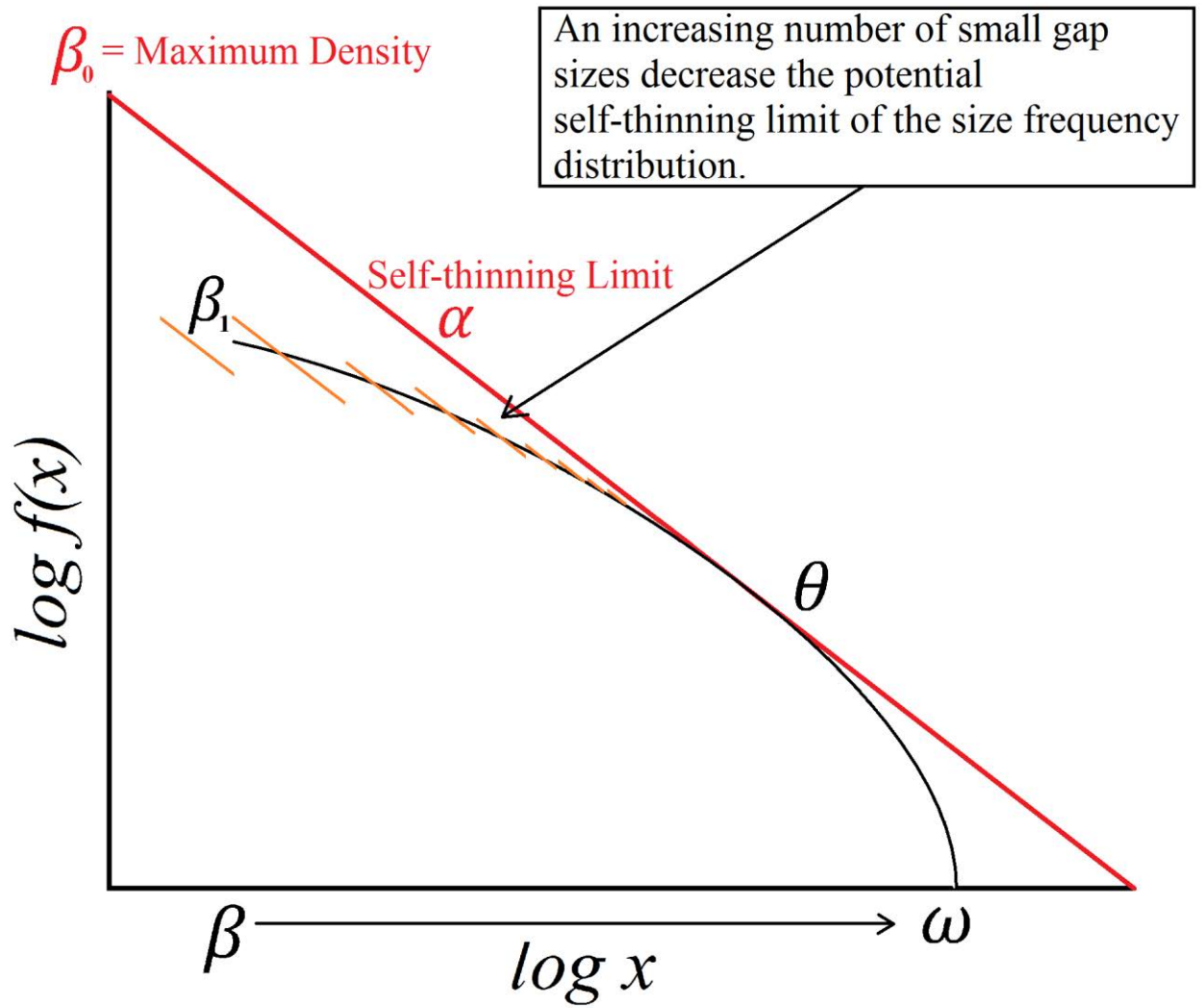
Figure SM2.
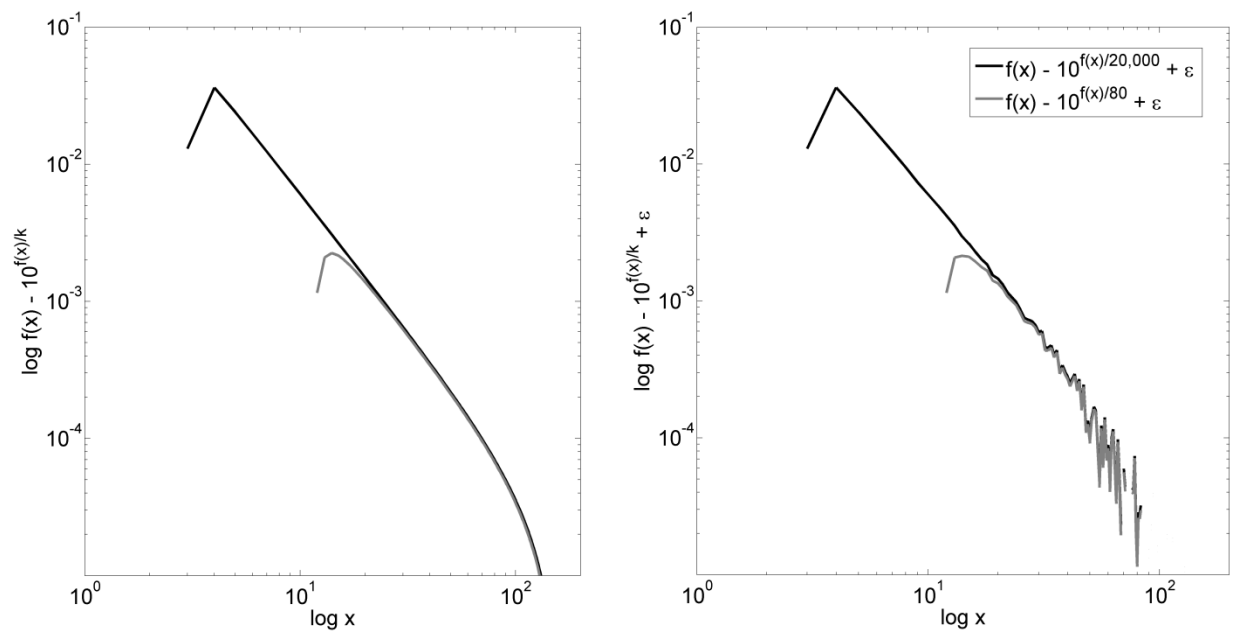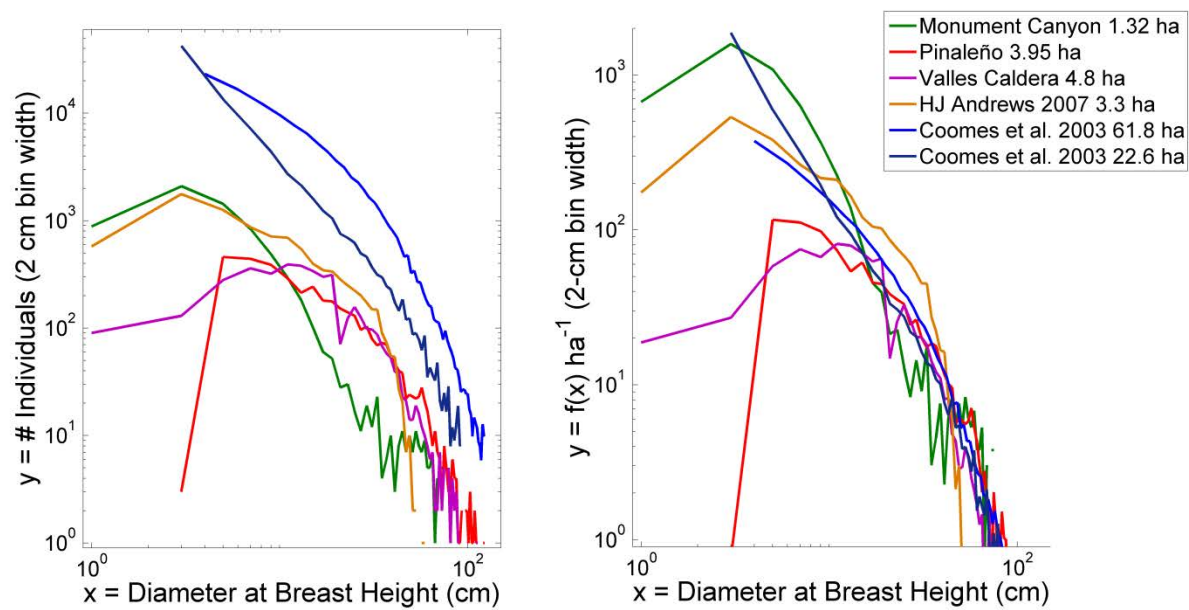
Figure SM3

Figure SM4.

APPENDIX B: APPLICATION OF METABOLIC SCALING THEORY TO REDUCE ERROR IN LOCAL MAXIMA TREE SEGMENTATION FROM AERIAL LIDAR

Tyson L. Swetnam[1,2] Donald A. Falk[1]

Format of Submission: Forest Ecology and Management

[1] The University of Arizona School of Natural Resources and Environment

[2] Corresponding Author: email: tswetnam@email.arizona.edu, phone: (520) 247-2293

School of Natural Resources and Environment

Biological Sciences East

Tucson, AZ 85721, USA

Abstract

Identifying individual trees across large landscapes is one of the major potential benefits of aerial LiDAR. However, current approaches to individual tree segmentation of aerial LiDAR data does not always reflect how the allometry of tree canopies change with height, age, or competition for limiting space and resources. We developed a variable-area local maxima (VLM) algorithm that incorporates predictions of the Metabolic Scaling Theory (MST) to reduce the frequency of commission error in a local maxima individual tree inventory derived from aerial LiDAR. We also evaluate why variations in other reported aerial LiDAR segmentation models may or may not be appropriate in certain forest types. By comparing the MST prediction to 663 species of North American champion-sized trees [which include the tallest and the largest trees on the planet], and 610 measured trees in semi-arid conifer forests in Arizona and New Mexico we show that the MST model $r_{can} = \beta h^\alpha$ where β is the normalization constant, h is height, and α is a dynamic exponent predicted by MST to be $\alpha = 1$, can be applied as a general model in any forest. MST also informs the estimate of individual tree bole diameter d_{bole} [which aerial LiDAR does not measure directly] based on two primary size measures easily obtained from the aerial LiDAR: height h and canopy diameter d_{can}. A two parameter model $\beta h \sqrt{d_{can}}$ is shown to better predict bole diameter (r^2 = 0.811, RMSE = 7.66 cm) than a single parameter model of either canopy diameter or height alone: βd_{can}^α (r^2 = 0.51 RMSE = 12. 4 cm) or βh^α (r^2 = 0.753, RMSE = 8.94 cm). These methods provide improved accuracy in estimates of total forest stand density, structural diversity, above ground biomass and carbon.

Highlights

- A Metabolic Scaling Theory (MST) general model of tree canopy and height allometry
- A variable area local maximum algorithm is demonstrated for conifer tree segmentation
- MST rule increases aerial LiDAR tree segmentation precision and accuracy

1 Introduction

Aerial LiDAR has the potential to revolutionize studies of forest macrostructure across entire landscapes with unprecedented accuracy and detail compared to conventional aerial surveys (Ustin and Gamon 2010, Van Leeuwen and Nieuwenhuis 2010). An important application of aerial LiDAR for forestry involves the segmentation of individual trees [both for forest inventory and biophysical tree metrics] from either the three dimensional point cloud (Li et al. 2012, Yao et al. 2012) or canopy height model (CHM) (Lefsky et al. 1999) that can be produced from the point cloud. Demonstrated techniques for segmenting trees from a CHM include: (1) the local maxima [with filtering] (Dralle and Rudemo 1996, Hyyppä et al. 2001, 2005, Persson et al., 2002, Monnet et al. 2010), (2) variable area window (Popescu et al. 2002, Popescu and Wynne 2004), (3) hierarchical inverse watersheds (Chen et al. 2006, Zhao and Popescu 2007), and (4) spatial wavelet (Falkowski et al. 2006, 2008). While some of those authors discuss allometric scaling of height and canopy radii as they relate to canopy spacing [for improving segmentation accuracy] (Table 1) they do not explain how allometry relates to more fundamental principles of forest structure (Enquist and Niklas 2001, Niklas and Enquist 2001, Niklas et al. 2003, West et al. 2009, Enquist et al. 2009, Deng et al. 2012). Thus, a universal or general model for segmenting individual tree canopies from aerial LiDAR cannot be made based on those assessments. Here we develop a variable-area local maxima (VLM) algorithm incorporating allometric scaling rules predicted by the Metabolic Scaling Theory (MST) (Enquist et al. 1998, 2003, West et al. 1999a, 1999b) to reduce errors of commission from a local maximum segmentation.

A stated problem in the existing literature (above) is the presence of false-positives (error of commission), whereby individual branch tips or multiple leader branches in an individual tree are wrongly identified as maxima and this incorrectly scored as separate trees. One solution to the false-positive problem has been to apply a Gaussian smoothing (Perona and Mallick 1990) to the CHM (Chen et al. 2006, Zhao et al. 2009, Monnet et al. 2010) to reduce the number of false-positives in the reported inventory. The problem with applying such a smoothing filter is that distinct variations in

shorter or smaller trees are lost. Mid-development height trees become 'blurred' together, leading to an increased error of omission.

The algorithm we created (Supplemental Materials Appendix B) incorporates a simple logic-rule based on MST which dis-allows maxima within a variable radius distance of trees that are taller. This segmentation technique is similar to that of Hyyppä et al. (2005), who split large tree crowns based on allometric knowledge of individual tree canopies failing to exceed a certain size. In our approach we take an inverse perspective, using segmentation rules to eliminate commission error trees, rather than split trees that are creating omission error.

MST provides a framework based on energy flow through connected networks, providing a mechanistic basis for density dependence and 'self-thinning' (Yoda 1963, Enquist et al. 1998, 2009, Niklas et al. 2003, West et al. 2009, Deng et al. 2012). MST also incorporates critical buckling tolerance (McMahon 1973, McMahon and Kronauer 1976)], the minimum bole radius to support its height, and a minimum canopy volume to support its total biomass based on that height (Savage et al. 2010, Kempes et al. 2011). These properties provide a more accurate predictive model for tree allometry, and inform our segmentation approach. MST predictions also improve the accuracy of a local maxima-derived forest inventory by allowing us to use an allometric variable search radius for isolating errors of commission, while keeping the minimum mapping unit as small as possible to pick up fine scale variations in smaller tree canopies. As a further result, applications of MST-based predictions are shown to improve the estimation of other allometric measures of trees not made by aerial LiDAR such as basal area and total above ground biomass.

2 Methods

2.1 Metabolic Scaling Theory in trees and forests

The 'pipe model' of tree architecture (Shinozaki et al. 1964a, 1964b) predicts that the primary size measures in vascular plants follow fractal-like branching rules. The predictions of MST extend these original predictions based on more detailed observations of how matter and energy are distributed in living organisms. First, energy

flow in trees is conducted by a continuous network of vessel elements from roots to xylem to foliage (Shinozaki et al. 1964a, 1964b, West et al. 1999b, Savage et al. 2010, Bentley et al. 2013). Second, vessels are both space filling and area preserving (in other words, there is constant cross-sectional area at all branching levels) (West et al. 1999b, 2009, Enquist et al. 1999). Finally, this vessel architecture maximizes efficiency by transporting water and nutrients from the roots to the petioles through capillary action driven by the more negative water potential of the atmosphere versus the water potential of the soil (Savage et al. 2010, Kempes et al. 2011). Variations in observed canopy geometry are life history adaptations that alter path length and the angle of the branches off the primary bole (Savage et al. 2010, Bentely et al. 2013) to maximize efficiency in resource use.

The geometry of trees varies widely across the Earth system. Canopy geometry is often described in Euclidean terms (e.g., spherical, elliptical, cone, or inverse cone canopies). In these terms, gymnosperm canopies tend to be conical or elliptical with apical leaders located near the vertical center of the standing bole, whereas angiosperms tend to have elliptical, spherical, or inverse cone geometry. However, the underlying physiology of trees is not Euclidean; trees are in-fact fractal-like objects (Enquist et al. 1998, 2009, West et al. 1997, 2009). It is only after fractal-like branching terminates that tree assume an approximation of Euclidean geometry.

Changes in canopy geometry as trees grow larger and taller are thought to be due to the physiological limits of (1) gravity working against capillary action [inducing a vertical height limit in very tall trees (Koch et al. 2004, Kempes et al. 2011)], (2) the size, shape and position of crowns of other trees in the existing forest canopy (Purves et al. 2007), and (3) the critical buckling strength of woody tissue with branch length and weight (McMahon 1973, McMahon and Kronauer 1976).

Because branches are area preserving (Savage et al. 2010, Bentely et al. 2013) the aboveground volume of a tree is equivalent to a cylinder, $V = \pi r_{bole}^2 l = \frac{\pi d_{bole}^2}{4} l$ [Eq. 1], where r_{bole} is the bole radius, bole diameter $d_{bole} \equiv 2r_{bole}$, and l is the length of the trunk proportional to maximum height h if the tree is vertically upright [as it the case for

most conifers]. The diameter of the bole is typically measured at breast height (1.3 m above ground level) or above the butt swell of the roots. Measures of bole diameter by convention are considered to be the strongest independent variable for predicting above ground biomass (Enquist et al. 1998, Jenkins et al. 2004, Savage et al. 2010, Bentley et al. 2013): $M = \beta r_{bole}{}^{\alpha}$ [Eq. 2], where M is the aboveground biomass of the tree, β is a normalization constant, and α is the dynamic exponent.

The MST quantitative theory of forest structure (Enquist et al. 2009, West et al. 2009) predicts the scaling of canopy diameter [or radius] with height as $r_{can} = \beta_0 h^{\alpha}$ [Eq. 3] where $\alpha = 1$. Equivalently, bole radius to canopy radius is predicted as: $r_{can} = \beta_1 r_{bole}^{\alpha}$ [Eq. 4] where $\alpha = 2/3$. For Eq. 3 and Eq.4 the normalization constants β_0 and β_1 determine how the individual species canopy geometry fill space (Enquist et al. 2009). For example, $\beta_0 = 1$ results in a tree whose height is equal to its canopy radius [e.g. spherical geometry]; $\beta_0 < 1$ results in a tree that is taller than it is wide, and $\beta_0 > 1$ results in a tree wider than it is tall. Since we are interested in the prediction of other measures [primarily bole diameter which is not observed by aerial LiDAR], we take the inverse function $r_{bole} = \beta_2 r_{can}^{\frac{3}{2}}$ [Eq. 5], and $r_{bole} = \beta_3 h^{3/2}$ [Eq. 6] to predict the diameter of the bole or of tree biomass.

2.2 Tree Data

2.2.1 Semi-arid conifer forests in Arizona and New Mexico

Field observations of conifers from the semi-arid southwestern USA came from two field campaigns in the mountains of Arizona and New Mexico, both part of the Santa Catalina Mountains and Jemez River Basin (SCM-JRB) Critical Zone Observatory (CZO) and nearby areas (Mitra et al. in prep). The Santa Catalina Mountains are located north of Tucson, Arizona, at 32.4° N, 110.7° W and are managed by the Coronado National Forest (NF), United States Forest Service (USFS). The Jemez Mountains are west of Santa Fe, New Mexico at 35.8° N, 106.5° W, and are managed by the Santa Fe NF, USFS and the Valles Caldera National Preserve under the supervision of USFS.

Field measures of trees included diameter at breast height (DBH), tree height, canopy base height, and canopy radius. Observations in the Santa Catalina Mountains

were taken for 203 trees at two locations in Marshall Gulch [named 'Granite site' and 'Schist site' because of their different parent materials] in spring of 2012. Observed species included Ponderosa pine (*Pinus ponderosa*), Arizona white pine (*Pinus strobiformis*), White-fir (*Abies concolor*), and Douglas-fir (*Pseudotsuga menziesii*). The other set of observations from the Valles Caldera National Preserve in the Jemez River Basin are for two eddy covariance tower sites (named 'Lower Ponderosa' and 'Upper Spruce-fir'). Those observations were collected by field technicians of the SCM-JRB CZO in summer 2012. A total of 100 trees were observed at the Lower site, all ponderosa pine; and 94 trees at the upper site, all Engelmann spruce (*Picea engelmannii*) (Table 2).

An additional 238 conifer trees were measured from the aerial LiDAR in FUSION/LDV tree modeler (McGaughey 2012). These observations only include height and canopy diameter (Table 2). A total of 133 old growth Engelmann spruce and sub-alpine-fir (*Abies lasiocarpa*) trees were measured on a northeast aspect of Redondo Peak in the Valles Caldera National Preserve. A total of 105 old growth Ponderosa were measured in the Monument Canyon Research Natural Area, adjacent to the Valles Caldera National Preserve, on the Santa Fe National Forest. The two old growth stands were selected to determine the upper limit of size to trees in the semi-arid conifer forests, as the field plots were both from young- or mid-development stands.

To determine how well Eq. 3 predicts the distribution of the data we fit the data for both the field and aerial LiDAR measured semi-arid conifer trees [n = 610] by least-squares regression in MATLAB 2013a Curve Fitting Tool (Mathworks 2013).

Because the semi-arid conifer forest data represent only fraction of the potential range of sizes and geometries exhibited by trees, and because other reported aerial LiDAR studies are from forests with potentially very different canopy geometries we decided to include the North American Champion Tree List from the National Register of Big Trees (data are available from: www.americanforests.org). The 2013 list has 761 champion or co-champion trees from 663 species including both angiosperms and gymnosperms. These include the largest trees in the world, giant sequoia (*Sequoiadendron giganteum*), and the tallest, the coast redwood (*Sequoia sempervirens*).

The benefit of the champion tree data is that it represents the range of physiological maximum potential for tree geometry across North America. We reproduced tree size data (n = 185) from Chen et al. (2006) by capturing the observations in a figure-to-data conversion tool (Frantz 2000). The lone species in Chen et al. (2006) is the California Blue Oak (*Quercus douglasii*); which was observed to exhibit spherical to elliptical canopy geometry unlike the conifer species from Arizona and New Mexico. We analyzed three other reported canopy radius-to-height models (Popescu et al. 2002, Chen et al. 2006, Falkowski et al. 2006). Most of these models were originally published with canopy diameter relationships, which we normalized to represent canopy radius. The data that Popescu et al. (2002) used to fit their model were not available; however we plot their model against our data. The data from Chen et al. (2006) were plotted and the reported model fit. The species of tree in Chen et al. (2006) are California Blue oak which are notable for their spherical canopy architecture. The Falkowski et al. (2006) data were not available; Many of the tree species in the Falkowski et al. (2006) study are the same as the semi-arid forest data from Arizona and New Mexico.

2.3 Aerial LiDAR

The aerial LiDAR data used in this study were from a combination of two snow-off and snow-on acquisitions in 2010 over the Valles Caldera National Preserve by the SCM-JRB CZO (data available from OpenTopography.org: http://www.opentopography.org/index.php), and a 2012 acquisition for the Southwest Jemez Mountains Landscape Restoration by the USFS. Both flights were collected with high density [>12 pulses per square meter (ppsm)], and were parameterized based on recommendations similar to those in Laes et al. (2008).

The CHM was generated in USFS FUSION/LDV (McGaughey 2012) after a Digital Terrain Model (DTM) of the bare earth returns was calculated. The CHM raster grid was created at 33.3 cm pixel size by subtracting the maximum elevation value of laser hits within the pixel from the DTM elevation. The density of the pulse return cloud, in the case of Monument Canyon, exceeded 45 ppsm when both the 2010 and 2012 acquisitions were loaded together; the average density of the point cloud exceeded 25

ppsm in the old growth spruce-fir stand in the Valles Caldera when the 2010 snow-on and snow-off data were loaded together.

For trees with field-measured heights versus heights measured in FUSION LDV tree modeler (McGaughey 2012) the Pearson's correlation coefficient for n = 394 trees was $r = 0.9837, MSE = 0.8097\ m$.

2.4 Variable-Area Local Maxima Algorithm

We wrote a local maximum algorithm in MATLAB 2013a (The Mathworks 2013) to isolate local maxima pixels from the high resolution (0.333 m) CHM [See Supplemental Materials for full code details]. We named our local maxima algorithm a 'variable-area local maxima' or 'VLM' for short. The VLM was based on the variable area window technique (Popescu et al. 2002, Popescu and Wynne 2004). The VLM also builds on other local maxima techniques which use Gaussian smoothing (Chen et al. 2006, Zhao et al. 2009, Monnet et al. 2010).

The VLM tree list in many cases will have errors of commission [e.g. too many local maxima points that are actually parts of another tree crown]. To flag these as potential errors of commission we use a spatial packing rule which states (1) there cannot be more than one maximum stem within the canopy radius determined by that tree's height, and (2) if there are other stems inside the radius, and those stems are shorter, they are flagged for removal.

In the companion paper to this work, the average number of pulses per square meter (ppsm) within the inventory plots (Pinaleño n = 78 500 m^2, and Valles Caldera n = 48 1,000 m^2) ranged from $9.3 < \mu < 25.1\ ppsm$ (Swetnam et al. *in review*), where the average distance (m) between pulses is measured as $d_\mu = \sqrt{\dfrac{1}{\mu}}$ (Chen et al. 2006). To ensure a minimum of one pulse per pixel and limit under-sampled 'NoData' pixels (Figure SM2) a $0.333m^2$ [$d_\mu = \sqrt{\dfrac{1}{9}} = 0.33\overline{3}$] pixel size was used for all plots. Salt and pepper voids are expected because the aerial LiDAR pulse data are not uniformly distributed. Rather than decrease the minimum mapping unit, we instead apply an anisotropic diffusion (Perona and Mallick 1990) to the surface image. The benefit of

Gaussian smoothing [versus inverse distance weighting or Kriging] is anisotropic transforms retain the edge detail between canopies while reducing the gaps within canopies.

We intentionally limit the number of smoothing iterations applied to our CHM to ensure that the fine scale variability of the midstory component is not lost. The output from the VLM includes a unique tree identifier number, the geo-position of the tree [in Universal Transverse Mercator coordinates] by the X- and Y-axes, the tree height on the Z-axis, and an estimated canopy radius which we derive using Eq. 3. The parameter for estimating canopy radius was based on the results of our least squares regression of the observed canopy radius-to-height scaling.

The VLM uses a structural element [the 'strel' function in MATLAB] – and a disk shaped dilation [the 'imdilate' function in MATLAB] that reduces no-data pixels. The disk-shaped dilation eliminates other maxima within a n-distance; if the maximum is the highest point it is retained while all other points are overlapped by taller pixel heights.

The allometric rule of the VLM (Supplemental Materials Appendix B) uses a kNN 'exhaustive searcher' to identify neighbor trees within 21 m of the target tree. This was based on our observation that the largest trees in the study area had a maximum canopy radii ≤ 7 m, so we arbitrarily increased the search distance to twice that width to ensure we were identifying enough neighbor trees of equal or greater size. The script uses an 'IF' statement to flag any local maxima within the predicted minimum canopy radius [from Eq.3] of the target maximum: If the distance to the neighbor maximum is less than the predicted canopy radius the neighbor maximum is deleted. The radius of the target tree was determined by setting the parameter β_0 [Eq. 3] to the minimum 95% prediction interval canopy diameter of the observed trees $\beta_0 = 0.075$ in the semi-arid conifer forest data (Figure 1, Table 1), as determined by least-squares regression.

3 Results

3.1 Height to Canopy Allometry

The least squares regression of height to canopy diameter for all of the observed semi-arid conifer forest data (n=610) was $d_{can} = 0.63 \pm 0.185 h^{0.7202 \pm 0.056}$ $(r^2 =$

0.569, $RMSE = 160\ cm$) (Table 1). For the MST predicted $\alpha = 1$ parameter in Eq. 3 the least-squares regression was: $d_{can} = 0.258 \pm 0.011h^1$ ($r^2 = 0.505$, $RMSE = 172\ cm$) (Table 1).

When we plot the Champion Tree canopy radius and height allometry along with the semi-arid conifer tree allometry (Figure 1), the observed range of the semi-arid conifer studies occupy the lower portion of the over-all potential distribution of canopy radius-to-height scaling of trees across North America. Remarkably, for the semi-arid conifer data a normalization constant of $\beta_0 = 0.126$ [Eq. 3] aligns with the canopy radius-to-height scaling of the tallest of all trees in the dataset, the coast redwoods, which exceed 116 m in height. The widest canopy species tend to exhibit little more than 1:1 ratio canopy radius-to-height [e.g. spherical geometry] ($\beta_0 = 1$), while the narrowest scale with 1:13 canopy radius-to-height ($\beta_0 = 0.075$). Above 40 m height canopy radius-to-height does not exceed 1:2 ($\beta = 0.5$), and at 85 m canopy radius-to-height does not exceed 1:5 ($\beta = 0.2$). This is potentially a consequence of critical buckling tolerance in branches as they grow away from the bole (McMahon 1973, McMahon and Kronauer 1976). Notably, no trees in the Champions or semi-arid conifer list that exceed 20 m in height have a 1:1 canopy radii to height ratio (Figure 1).

In general, angiosperms such as Chen et al. (2006) California Blue oak (Figure 2) have larger normalization constants ($\beta_0 > 0.25$) than do gymnosperm such as the cone shaped conifers in the Arizona and New Mexico data ($\beta_0 > 0.129$) (Figure 2). The models reported in Popescu et al. (2002), Chen et al. (2006), and Falkowski et al. (2006) (Figure 2, Table 1) vary for both their slope and intercept points – but do so within the globally observed range of tree height to canopy allometry. The Falkowski et al. (2006) model closely fits the observed 95% minimum confidence interval of the Arizona and New Mexico fit model, with a similar normalization constant and identical shape parameter $\alpha = 1$.

3.2 Basal area prediction

The necessary field measurements (canopy diameter and tree height) from which to compare aerial LiDAR measures to field measured diameter came from three semi-arid

forests: the Santa Catalina mixed-conifers (n=203), Valles Caldera upper flux tower Engelmann spruce (n=94), and Valles Caldera lower flux tower Ponderosa pine (n=100) (Figure 3). The two sets of fitted lines (solid and dashed) are shown for the least-squares regression of the entire data set (n =397). The solid line represents the best fit without any *a priori* predictions of the scaling exponent or normalization constant; the dashed line represents the theoretical MST shape parameter [Eq.4]. For prediction of bole diameter, based on observed canopy diameter (Figure 3, left panel), the observed variance for height with canopy diameter (Table 1) was relatively wide for both models (RMSE = ±12.4 cm $r^2 = 0.51$ and RMSE = ±14.9 cm $r^2 = 0.28$, respectively). The predicted scaling exponent $\alpha = 0.82 \pm 0.09$ was lower than the MST prediction of $\alpha = 1.5$. Tree height as a predictor of bole diameter (center panel) produced a lower RMSE and higher coefficient of determination for both models (RMSE = ±8.94 cm $r^2 = 0.75$ and RMSE = ±9.29 $r^2 = 0.73$cm, respectively) and a scaling parameter closer to theory ($\alpha = 1.26 \pm 0.08$). The strongest model of predicting bole diameter used both canopy diameter and tree height, based on the Hyyppä et al., 2005 (Table 1) (Figure 3, right panel): $d_{bole} = \beta\left(h * \sqrt{d_{can}}\right)^1$, had the smallest RMSE and largest COD of all the models (RMSE = ±7.66 cm, $r^2 = 0.81$).

3.3 VLM

Error of omission for the VLM increases with the number of iterations the anisotropic diffusion is conducted on the CHM (Figure SM2 and SM3). Functionally, the greater number of anisotropic diffusion iterations the fewer maxima are preserved due to the smoothing of the small differences between points that have interconnected canopies. We observed anisotropic diffusion with more than three iterations (SM Figure 4) resulted in an increasing error of omission in tall trees, which without smoothing are well represented in the overstory local maxima inventory with few or no errors of omission. The VLM does a poor job of inventorying trees in the mid- and understory where canopies tend to be interconnected or located beneath overstory canopy trees (Swetnam et al. *in review*). Swetnam et al. (in review) report that the VLM becomes divergent from

the observed inventory between 12 m and 16 m tree height for open to closed stands in the same study forests in Arizona and New Mexico.

On the other hand, error of commission in the VLM resulted from at least two observed phenomena: (1) The presence of 'corn-row' or 'salt and pepper' effects in the CHM which create gaps between individual pixels resulted in numerous maxima within a single canopy (SM Figure 2), and (2) the actual presence of multiple apical leaders or daughter branches within a single tree canopy resulted in error of commission (Figure 4). Corn-rows and salt and peppering occurs when the CHM is produced at too high of a resolution compared to the underlying aerial LiDAR point cloud. These effects were not common in the validation plots used by Swetnam et al. (in review), based on the observed pulse return density in all plots exceeding 9 ppsm with a 3x3 neighborhood per meter square (0.3334 m pixel size) (Figure 4).

Based on our findings reported in section 3.1 we assumed the scaling of canopy radii with height follows the MST model. Thus we set the minimum canopy radii of a single local maximum to $\beta_0 = 0.075$. [Eq. 3]; based on the observed interval from our data (Figures 2 and 4, Table 1). After each local maximum was identified (Figure 4) the VLM allometric rule self-limits other maxima within a k-nearest neighborhood. The allometric rule resulted in fewer errors of commission observed inside each individual tree canopy (Figure 4).

There were however errors of omission generated by the application of the allometric rule. These errors tended to be located in and around: (1) stands of mid development or midstory trees that are tightly packed in space, (2) trees with a significant lean may have a maximum near the edge of their canopy while most of the canopy area is located to one side of the isolated maximum [this resulted in the elimination of small trees that were located on the near side of the maximum], and (3) snags and standing dead poles are not defined and may result in small trees of a new cohort being removed around the standing pole. When we increased the normalization constant $[0.075 < \beta_0]$ for the predicted canopy radius the risk of increasing omission error appears to also appears to increase.

4 Discussion and Conclusions

The main objective of this paper was to determine whether MST's prediction of canopy-height scaling is generalizable first across all North American trees, using a novel variable-area local maxima (VLM) method for segmenting individual trees from an ASLM point cloud. We compared other published aerial LiDAR studies which reported alternate canopy-height scaling (Popescu et al. 2002, Chen et al. 2006, Falkowski et al. 2006) to the scaling of trees in New Mexico and Arizona (Figure 2). We found good agreement between our data and the Falkowksi et al. (2006) model; this is not surprising considering that both of these studies were in conifer forests in the western US. They are however on opposite ends of the distribution of these species north-south distribution. Other existing aerial LiDAR models for estimating canopy radius with height (Popescu et al. 2002, Chen et al. 2006, Falkowski et al. 2006) varied widely, although within the observed range of variation for all trees (Figure 2). Collectively, these results suggest that canopy scaling of conifers across the western US may share common scaling parameters as reported here.

The accuracy of the VLM varied by forest type and by cover percentage. The authors report the accuracy of the VLM for two independent forest inventories with aerial LiDAR and use the VLM inventory to predict mid- and understory distributions in the companion paper (Swetnam et al. *in review*). In summary, plots with high cover percentage (>60%) the VLM was less accurate for trees > 16 m in height; in plots with low cover percentage (<50%) the VLM was in general accurate for all heights (Swetnam et al. *in review*). Midstory trees are likely to be obscured by overstory trees, or have interconnected canopies with their neighbor trees, making local maxima undifferentiable. Those results are supportive of the MST prediction that trees grow with interconnected canopies when they are small and eventually separate due to physical and physiological mechanisms (Enquist et al. 2009. West et al. 2009).

The output of the self_thin.m algorithm with $\beta_0 = 0.129$ successfully segregated individual trees from the ASLM point cloud (example in Figure 4). Importantly, the β_0 normalization constant is expected to change both amongst trees of the same species and

across different forst types. In our example, the VLM.m reports all of the observed maxima; the self_thin.m will uses a logic rule to eliminate the shorter of the two maxima because of their relative distance to one another being allometrically too close to suggest they are two separate trees. This inference can be used to estimate the dominance class of the largest trees and help determine whether a maximum is an error of commission.

The scaling of canopy geometry with height [in both semi-arid conifers and in champion size American trees] appears to broadly support the MST prediction across the entire range of possible tree height in North America (Figure 1). Because of the plasticity of canopies within a single species and across all species due to life history variation [e.g. the effect of competitive self-thinning versus open growing conditions] (Purves et al. 2007, Strigul et al. 2008), the proportionality constant used to generate estimates of population size, density, and size-frequency distributions may vary in ecological space. For example, trees growing in competitive self-thinning conditions may have narrower canopy radii than trees of the same species that grow in open spaces (Purves et al. 2007).

Unlike the other aerial LiDAR studies, we present a general model based on mechanistic predictions made by the MST (Enquist et al. 2009, West et al. 2009). The VLM and self-thinning algorithm developed here can be modified based on the observed canopy architecture of any forest, allowing the user to parameterize their model differently in forests with varying canopy geometry. We suggest that local-scale β values (Eq. 3) be determined with local inventory data. The two parameter model decreases the RMSE by ~1.63 cm, suggesting the multiple parameter model yields stronger predictions for estimating above ground biomass and carbon density, versus a single parameter model alone.

The general quantitative theory of forest structure (Enquist et al. 2009) can also be used to predict the distance between tree boles and between canopies. In this manuscript we did not explore this relationship, as the applied local maxima logic rule only pertained to the minimum radius of a single tree and not its self-spacing with neighbor trees. Future research that is interested in the conditions of the stand should consider applying those MST predictions.

Acknowledgements

References

ASPRS 2011. LAS specification version 1.4 – R12. 10 June 2012. 27 p. Bethesda, Maryland. www.asprs.org.

Bentley, L.P., Stegen, J.C., Savage, V.M., Smith, D.D., Allmen, E.I., Sperry, J.S., Reich, P.B., Enquist, B.J., 2013. An empirical assessment of tree branching networks and implications for plant allometric scaling models. Ecol. Lett. 16(8):1069-78.

Chen, Q., Baldocchi, D., Gong, P., Kelly, M., 2006. Isolating individual trees in a savanna woodland using small footprint LiDAR data. Photogramm. Eng. Remote Sensing 72, 923-932.

Chen, Q., Gong, P., Baldocchi, D., & Tian, Y. Q. (2007). Estimating basal area and stem volume for individual trees from LiDAR data. Photogramm. Eng. Remote Sensing 73(12), 1355.

Deng, J., Zuo, W., Wang, Z., Fan, Z., Ji, M., Wang, G., Ran, J., Zhao, C., Liu, J., Niklas, K.J., 2012. Insights into plant size-density relationships from models and agricultural crops. Proceedings of the National Academy of Sciences 109, 8600-8605.

Dralle, K., Rudemo, M., 1996. Stem number estimation by kernel smoothing of aerial photos. Canadian Journal of Forest Research 26, 1228-1236.

Enquist, B.J., Brown, J.H., West, G.B., 1998. Allometric scaling of plant energetics and population density. Nature 395, 163-165.

Enquist, B.J., Economo, E.P., Huxman, T.E., Allen, A.P., Ignace, D.D., Gillooly, J.F., 2003. Scaling metabolism from organisms to ecosystems. Nature423, 639-642.

Enquist, B.J., Niklas, K.J., 2001. Invariant scaling relations across tree-dominated communities. Nature 410, 655-660.

Enquist, B.J., Niklas, K.J., 2002. Global allocation rules for patterns of biomass partitioning in eed plants. Science 295, 1517-1520.

Enquist, B.J., West, G.B., Brown, J.H., 2009. Extensions and evaluations of a general quantitative theory of forest structure and dynamics. Proceedings of the National Academy of Sciences 106, 7046-7051.

Enquist, B.J., West, G.B., Charnov, E.L., Brown, J.H., 1999. Allometric scaling of production and life-history variation in vascular plants. Nature 401, 907-911.

ESRI 2012. ArcGIS Desktop: Release 10.1. Redlands, CA: Environmental Systems Research Institute.

Falkowski, M.J., Smith, A.M., Gessler, P.E., Hudak, A.T., Vierling, L.A., Evans, J.S., 2008. The influence of conifer forest canopy cover on the accuracy of two individual tree measurement algorithms using lidar data. Canadian Journal of Remote Sensing 34, S338-S350.

Falkowski, M.J., Smith, A.M., Hudak, A.T., Gessler, P.E., Vierling, L.A., Crookston, N.L., 2006. Automated estimation of individual conifer tree height and crown diameter via two-dimensional spatial wavelet analysis of lidar data. Canadian Journal of Remote Sensing 32, 153-161.

Frantz, J., 2000. g3data.exe 1.5.1 GNU. GPL. https://github.com/pn2200/g3data

Friedman, J.H., Bentley, J.L., Finkel, R.A., 1977. An algorithm for finding best matches in logarithmic expected time. ACM Transactions on Mathematical Software 3, 209-226.

Hyyppä, J., Kelle, O., Lehikoinen, M., Inkinen, M., 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. Geoscience and Remote Sensing, IEEE Transactions on39, 969-975.

Hyyppä, J., Mielonen, T., Hyyppä, H., Maltamo, M., Yu, X., Honkavaara, E., Kaartinen, H., Using individual tree crown approach for forest volume extraction with aerial images and laser point clouds, 2005. In: Proceedings of ISPRS workshop laser scanning. Citeseer, 12-14.

Jenkins, J.C., Chojnacky, D.C., Heath, L.S., Birdsey, R.A., 2004. Comprehensive Database of Diameter-Based Biomass Regressions for North American Tree Species. USDA Forest Service General Technical Report NE, 319. Newtown Square, PA.

Koch, G. W., S. C. Sillett, G. M. Jennings, and S. D. Davis. 2004. The limits to tree height. Nature 428:851-854.

Kempes, C.P., West, G.B., Crowell, K., Girvan, M., 2011. Predicting maximum tree heights and other traits from allometric scaling and resource limitations. PloS one6, e20551.

Kerkhoff, A. J., and B. J. Enquist. 2007. The implications of scaling approaches for understanding resilience and reorganization in ecosystems. Bioscience 57:489-499.

Laes, D.; Reutebuch, S.; McGaughey, B.; Maus, P.; Mellin, T.; Wilcox, C.; Anhold. J.; Finco, M.; Brewer, K. 2008. Practical lidar acquisition considerations for forestry applications. RSAC-0111-BRIEF1. Salt Lake City, UT: U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center. 7 p.

Lefsky MA, Cohen WB, Acker SA, Parker GG, Spies TA, Harding D. 1999. Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. Remote Sens Environ 70:339–61.

Li, W., Guo, Q., Jakubowski, M.K., Kelly, M., 2012. A new method for segmenting individual trees from the lidar point cloud. Photogramm. Eng. Remote Sensing 78, 75-84.

MATLAB version 8.0.0783 (R2012B) Natick, Massachusetts: The Mathworks Inc. 2012.

McMahon, T., 1973. Size and shape in biology. Science 179, 1201-1204.

McMahon, T.A., Kronauer, R.E., 1976. Tree structures: deducing the principle of mechanical design. J. Theor. Biol. 59, 443-466.

Mitchell, B., Walterman, M., Mellin, T., Wilcox, C., Lynch, A.M., Anhold, J., Falk, D.A., Koprowski, J., Laes, D., Evans, D., 2012. Mapping vegetation structure in the Pinaleño Mountains using lidar-phase 3: Forest inventory modeling. RSAC-100007-RPT1. Salt Lake City, UT: U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center. 17 p.

Mitra, B., Papuga, S.A., Alexander, R., Swetnam, T.L., Abramson, N., *in prep*. Are allometric relationships for mixed-conifer species generalizable? Implications for up-scaling in sap flow applications.

Monnet, J., Mermin, É., Chanussot, J., Berger, F., 2010. Tree top detection using local maxima filtering: a parameter sensitivity analysis. Proceedings of Silvilaser 2010.

Newman, M. E. 2005. Power laws, Pareto distributions and Zipf's law. Contemporary physics 46:323-351.

Niklas, K.J., Enquist, B.J., 2001. Invariant scaling relationships for interspecific plant biomass production rates and body size. Proceedings of the National Academy of Sciences 98, 2922-2927.

Niklas, K.J., Midgley, J.J., Rand, R.H., 2003. Tree size frequency distributions, plant density, age and community disturbance. Ecol. Lett. 6, 405-411.

Perona, P., Malik, J., 1990. Scale-space and edge detection using anisotropic diffusion. Pattern Analysis and Machine Intelligence, IEEE Transactions on 12, 629-639.

Persson, A., Holmgren, J., Söderman, U., 2002. Detecting and measuring individual trees using an airborne laser scanner. Photogramm. Eng. Remote Sensing 68, 925-932.

Popescu, S.C., Wynne, R.H., 2004. Seeing the trees in the forest: using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height. Photogramm. Eng. Remote Sensing 70, 589-604.

Popescu, S.C., Wynne, R.H., Nelson, R.F., 2002. Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window size. Comput. Electron. Agric. 37, 71-95.

Popescu, S.C., Wynne, R.H., Nelson, R.F., 2003. Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. Canadian journal of remote sensing 29, 564-577.

Purves, D.W., Lichstein, J.W., Pacala, S.W., 2007. Crown plasticity and competition for canopy space: a new spatially implicit model parameterized for 250 North American tree species. PLoS One2, e870.

Reineke, L.H., 1933. Perfecting a Stand-Density Index for Even-Aged Forests. US Government Printing Office.

Samberg, A., An implementation of the ASPRS LAS standard, 2007. In: ISPRS Workshop on Laser Scanning and SilviLaser, 363-372.

Savage, V., Bentley, L., Enquist, B., Sperry, J., Smith, D., Reich, P., von Allmen, E., 2010. Hydraulic trade-offs and space filling enable better predictions of vascular

structure and function in plants. Proceedings of the National Academy of Sciences 107, 22722-22727.

Shinozaki K., Yoda K., Hozumi K., Kira T. 1964a. A quantitative analysis of plant form- the pipe model theory. I. Basic analyses. Jpn. Ecol. 14:97–105.

Shinozaki K., Yoda K., Hozumi K., Kira T. 1964b. A quantitative analysis of plant form- the pipe model theory: II. Further evidence of the theory and its application in forest ecology. Jpn. Ecol. 14:133–139.

Strigul, N., Pristinski, D., Purves, D., Dushoff, J., Pacala, S., 2008. Scaling from trees to forests: tractable macroscopic equations for forest dynamics. Ecol. Monogr. 78, 523-545.

Swetnam, T.L., Falk, D. A., Lynch, A. M., Guertin, D. P., Yool, S.R., *in review*. A Variable-area Local Maxima tool for segmenting individual trees from aerial LiDAR: Predicting rank-size distributions of forest mid- and understory in two semi-arid conifer forests. Forest Ecology and Management.

Takahashi, T., Awaya, Y., Hirata, Y., Furuya, N., Sakai, T., Sakai, A., 2010. Stand volume estimation by combining low laser-sampling density LiDAR data with QuickBird panchromatic imagery in closed-canopy Japanese cedar (Cryptomeria japonica) plantations. Int. J. Remote Sens. 31, 1281-1301.

Ustin, S. L., & Gamon, J. A. (2010). Remote sensing of plant functional types. New Phytologist, 186(4), 795-816.

Van Leeuwen, M., & Nieuwenhuis, M. (2010). Retrieval of forest structural parameters using LiDAR remote sensing. European Journal of Forest Research, 129(4), 749-770.

West, G.B., Brown, J.H., Enquist, B.J., 1997. A general model for the origin of allometric scaling laws in biology. Science 276, 122-126.

West, G.B., Brown, J.H., Enquist, B.J., 1999a. The fourth dimension of life: fractal geometry and allometric scaling of organisms. Science 284, 1677-1679.

West, G.B., Brown, J.H., Enquist, B.J., 1999b. A general model for the structure and allometry of plant vascular systems. Nature 400, 664-667.

West, G.B., Enquist, B.J., Brown, J.H., 2009. A general quantitative theory of forest structure and dynamics. Proceedings of the National Academy of Sciences 106, 7040-7045.

Yao, W., Krzystek, P., Heurich, M., 2012. Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. Remote Sens. Environ.123, 368-380.

Yoda, K., 1963. Self-thinning in over-crowded pure stands under cultivated and natural conditions.(In-traspecific competition among higher plants. XI.). J Biol Osaka City Univ 14, 107-129.

Zhao, K., Popescu, S., 2007. Hierarchical watershed segmentation of canopy height model for multi-scale forest inventory. Proceedings of the ISPRS working group Laser Scanning 436-442.

Figure Captions

Figure 1: Canopy radius and tree height relationships. On the x-axis are height measures (cm), on the y-axis are canopy radius measures (cm) for (1) the field measured trees (pink and orange), (2) FUSION/LDV measured trees (red and blue), (3) North American Champion Tree List (dark gray), and (4) data reproduced from Chen et al. (2006) (light gray). The model, $r_{can} = \beta h$, represents the MST scaling with various correlation coefficients. When $\beta_{max} = 1.0$ the tree has 1:1 scaling or spherical canopy geometry. The median value $\beta_{median} = 0.5$ is equivalent to 1:2 scaling. At the 95% prediction interval for the observed conifer species from Arizona and New Mexico: $\beta_{high} = 0.251$ approximately 1:4 scaling, $\beta_{low} = 0.075$ approximately 1:13 scaling [likely approaching the critical buckling strength (McMahon 1973)]. The least-squares best fit $\beta_{best} = 0.129$ is approximately 1:8 scaling, similar to the scaling of the Coast Redwood, the tallest trees on the planet [far right side of the figure].

Figure 2: Fit of other allometric models to tree size data. Same data shown in Figure 1 with the exception of the Chen et al. (2006) data now highlighted(blue). Three other published aerial LiDAR models for canopy radii to height: Popesecu et al. 2002 (*Pinus spp.* and *Quercus spp.*) (green), Chen et al. 2006 (*Quercus douglasii*) (blue), and Falkowski et al. 2006 (*Pinus spp., Pseudotsuga menziesii, Pica spp. Abies spp.*) (orange) are shown relative to this study's data. Model parameters are changed from the original publications to represent canopy radii where in the original text a model may have represented canopy diameter and units may have been in meters rather than centimeters.

Figure 3: Canopy and bole diameter. Relationship of canopy diameter to diameter at breast height (left), height to DBH (center), and a model of height multiplied by the square root of canopy diameter versus DBH (right). The multiple parameter model of both canopy diameter and height accounts for the highest amount of variation in DBH ($r^2 = 0.881$); there is only one best fit because there was no theoretical model.

Figure 4: Example of a spruce-fir forest plot (Valles Caldera #404). The VLM process, where a CHM (upper left) is anisotropically smoothed and dilated (upper left to right), the local maxima are isolated, and the allometric-rule for spatial sorting and

packing is applied to the k-nearest neighbors to remove errors of commission (bottom right to left). The black crosses are the local maxima, the black dots are the tallest maxima whose predicted canopy radii (empty black circles) eliminate the errors of commission in individual tree canopies. A histogram with 2.5 meter wide bins (lower left) of the observed inventory (gray bars) versus the local maxima and allometric-rule maxima show the allometric-rule maxima more accurately reflects the observed overstory inventory. There are several additional trees in the inventory along the perimeter of the plot boundary which are not included in the plot inventory; suggesting that the VLM still suffers from some error of omission.

Figure 5: Example of the VLM output over a section of Monument Canyon, Jemez Mountains, New Mexico. The canopy height model is 0.333m resolution and is colored from black-to-white for zero height to the tallest height. Individual maxima are shown with the predicted mean-canopy radius (yellow circles) which was used to reduce error of commission about individuals with multiple apical leaders or branches.
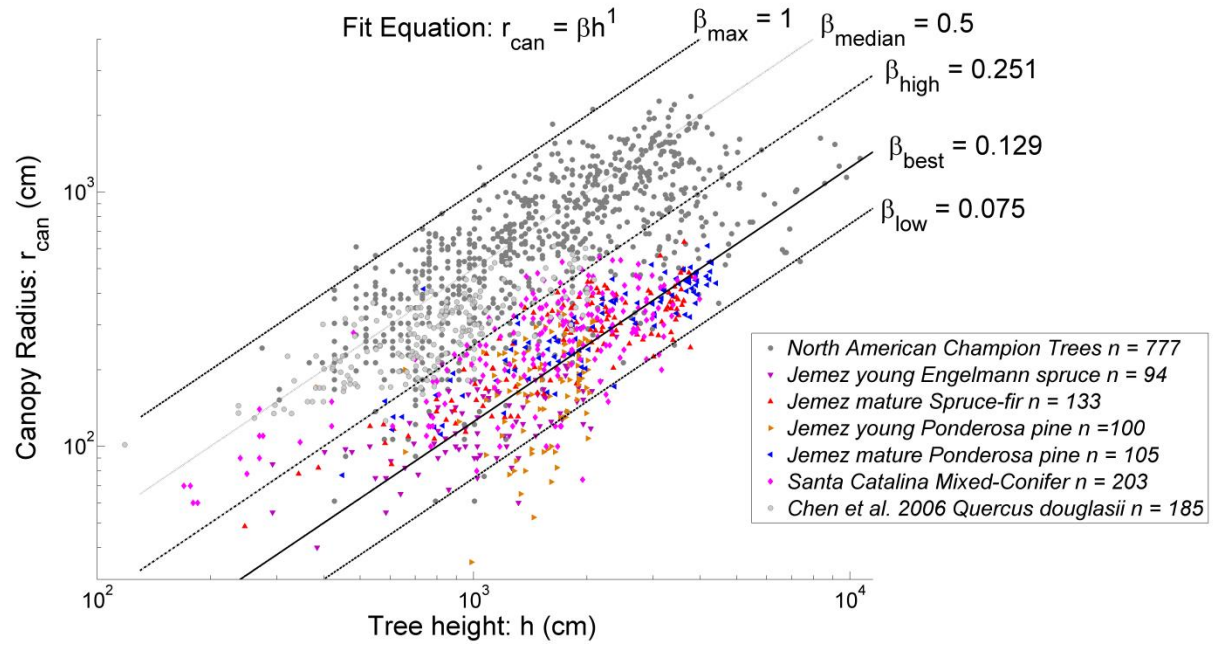
Figures & Tables

Figure 1.

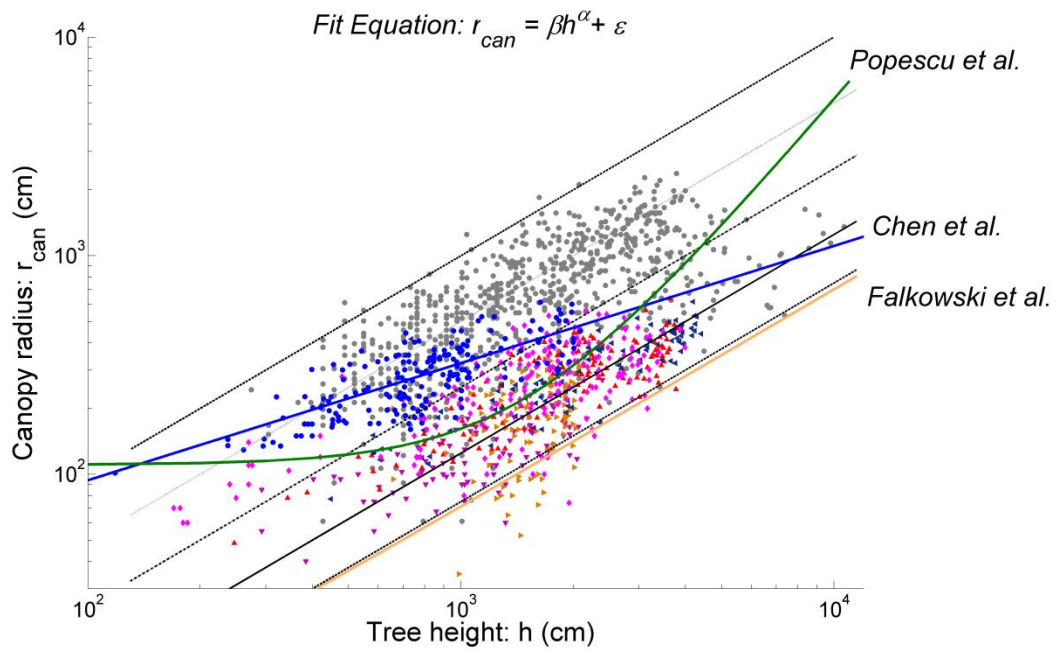*Fit Equation:* $r_{can} = \beta h^{\alpha} + \varepsilon$

Popescu et al.

Chen et al.

Falkowski et al.
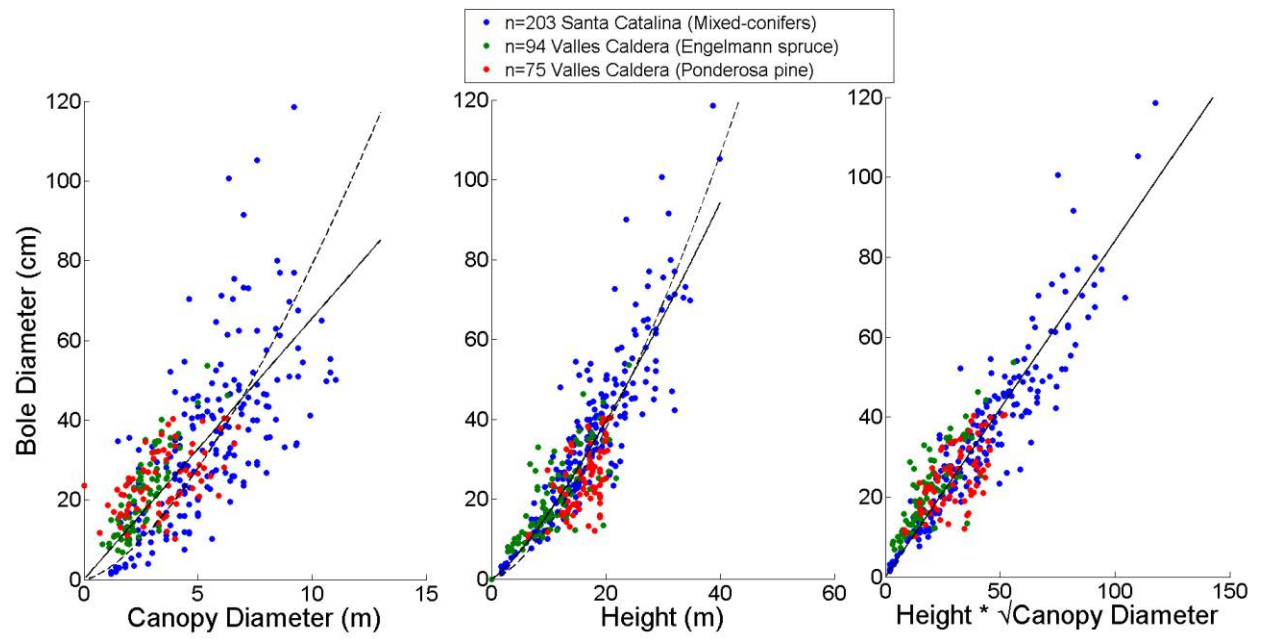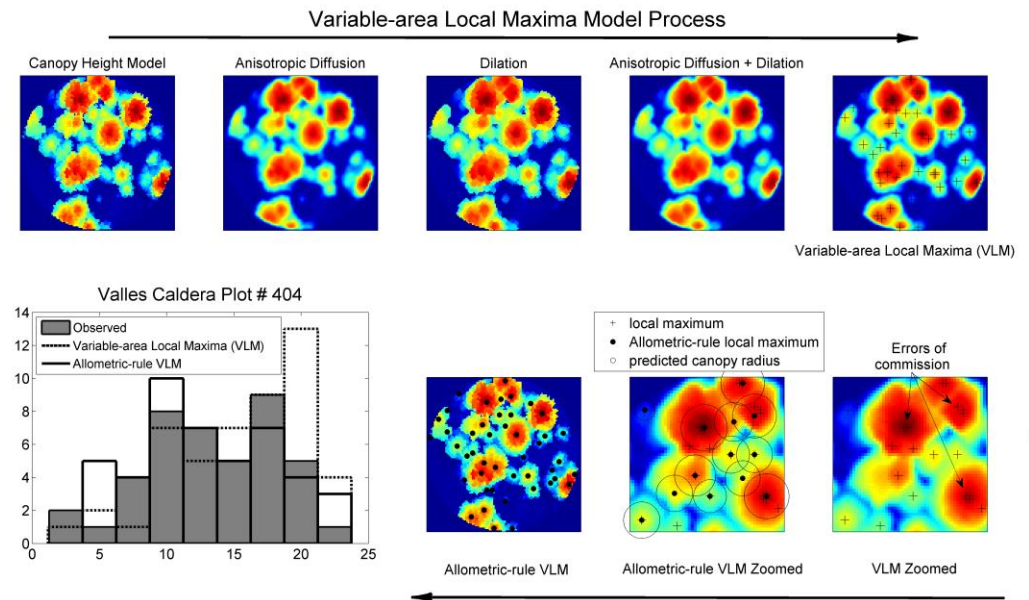
Figure 2.
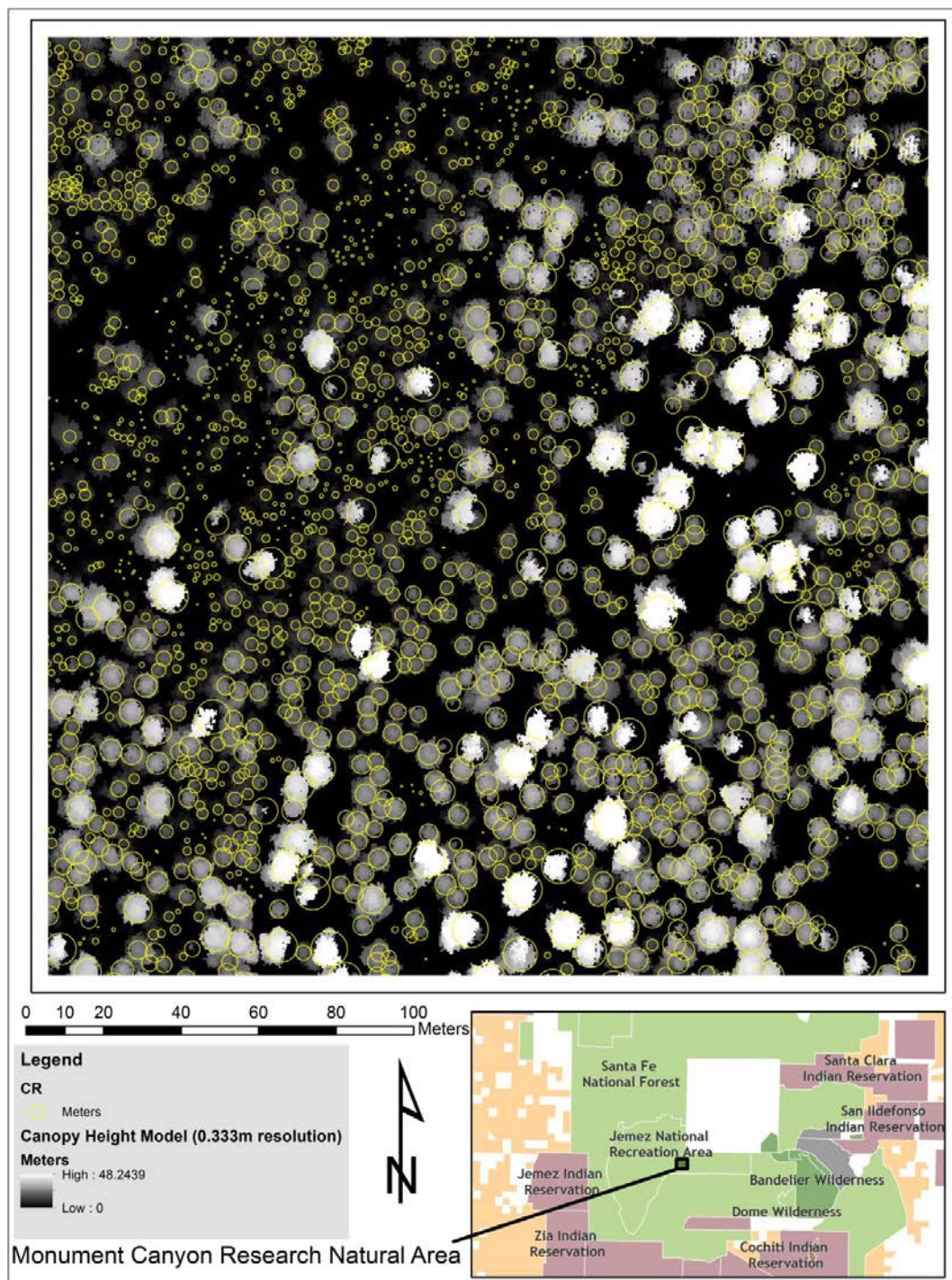
Figure 3.

Figure 4.

Figure 5.

Table 1: Allometric models from this study, and other aerial LiDAR tree segmentation studies in the current literature. The models from the present study were determined by least-squares regression in Matlab 2013a Curve Fitting Tool (Mathworks 2013); the first models are best fit without *a priori* assumptions of the α exponent; the second model uses the α value predicted by MST .

Reference	Location	Vegetation Types / Species	Model(s)	r^2 / RMSE
Present Study	New Mexico, Arizona, US	*Pinus* spp., *Pseudotsuga menziesii*, *Abies* spp., *Picea* spp.	$d_{can} = \beta \pm SE * h^{\alpha \pm SE}$: $d_{can} = 0.63 \pm 0.185 h^{0.7202 \pm 0.056}$ $d_{can} = 0.258 \pm 0.011 h^1$ (MST)	$r^2 = 0.569$, $RMSE = 160\ cm$ $r^2 = 0.505$, $RMSE = 172\ cm$
			$d_{bole} = \beta * d_{can}^{\alpha}$: $d_{bole} = 9.1\ d_{can}^{0.82}$ $d_{bole} = 2.5\ d_{can}^{1.5}$ (MST)	$r^2 = 0.51$, $RMSE = 12.4\ cm$ $r^2 = 0.28$, $RMSE = 14.9\ cm$
			$d_{bole} = \beta \pm SE * h^{\alpha \pm SE}$: $d_{bole} = 9.1\ h^{1.26 \pm 0.08}$ $d_{bole} = 2.5\ h^{1.5}$ (MST)	$r^2 = 0.753$, $RMSE = 8.94\ cm$ $r^2 = 0.73$, $RMSE = 9.29\ cm$
			$d_{bole} = \beta \pm SE * h * \sqrt{d_{can}}$: $d_{bole} = 0.82 \pm 0.09\ h * \sqrt{d_{can}}$	$r^2 = 0.811$, $RMSE = 7.66\ cm$
Popescu et al. 2002	Virginia, US	Pine plantation, mixed-hardwood semi-natural *Quercus spp.* *Pinus spp.*	$d_{can} = \varepsilon + \beta * h^{\alpha}$: $d_{can} = 2.21 + 0.01022 h^2$	$r^2 = 0.51$ / --
Hyyppä et al 2005	Kalkkinen, Finland	Semi-natural forest *Picea abies* *Pinus sylvestris* *Betula spp.*	$d_{bole} = -3.405 + 0.783 * \sqrt{h} + 1.28 * \sqrt{d_{can}}$	
			$\sqrt{d_{bole}} = 2.065 * \sqrt{d_{can}} + 0.538$	
Chen et al. 2006 Lower prediction	California, US	Savannah –Woodland *Quercus douglasii*	$d_{can} = 1.7425 * h^{0.5566}$	$r^2 = 0.59$ / --
			$\tilde{Y}_x = \hat{Y} - t(1 - \alpha; n - 2)\sqrt{s^2 + xSx'}$	
Chen et al. 2007	California, US	Savannah –Woodland *Quercus douglasii*	$d_{bole} = 2.29 + 1.6h + 0.26 C_a$ $C_a = \pi \left(\frac{d_{can}}{2}\right)^2$ = canopy area $BA = \pi \left(\frac{d_{bole}}{2}\right)^2$	$r^2 = 0.79$, $RMSE = 207\ cm^2$
Falkowski et al. 2006 Mexican hat wavelet	Montana, US	Mixed-conifer forest *Pinus spp.* *Pseudotsuga menziesii* *Abies grandis* *Picea engelmannii*	$d_{can} = \beta * h^{\alpha} + \varepsilon$: $d_{can} = 0.14 * h^1 + 2.56$	$r^2 = 0.34$ / $161 cm$
			$\psi(x, y) = (1 - x^2 - y^2) \exp[-(x^2 + y^2)/2]$	$r^2 = 0.86$ / $135 cm$
Takahashi et al. 2010	Ibaraki Prefecture, Japan	Plantation *Cryptomeria japonica*	$d_{can} = 0.099 d_{bole} + 0.49$	$r^2 = 0.83$ / --
			$V = \ln \beta_0 + \beta_1 h + \beta_2 \ln A_{can}$	$r^2 = 0.73$ / $0.178 m^3$

Table 2: Study area, species, and allometric measurements taken either by hand or FUSION/LDV tree model.

	Locations	*Picea eng.+Abies lasioscarpa*	*Picea engelmannii*	*Pinus ponderosa*	*Abies concolor*	*Pseudotsuga menziesii*	DBH (cm)	Canopy Diameter (m)	Height (m)	Field = 🖐 / LiDAR = ✳
New Mexico	Lower Flux			✓			✓	✓	✓	🖐 ✳
	Upper Flux		✓				✓	✓	✓	🖐 ✳
	Monument Canyon			✓				✓	✓	✳
	Redondo Peak	✓						✓	✓	✳
Arizona	Santa Catalina			✓	✓	✓	✓	✓	✓	🖐 ✳

APPENDIX C: PREDICTING RANK-SIZE DISTRIBUTIONS OF FOREST MID- AND UNDERSTORY IN TWO SEMI-ARID CONIFER FORESTS WITH AN AERIAL LIDAR DERIVED FOREST INVENTORY

Tyson L. Swetnam[1,5,6], Donald A. Falk[1,5], Ann M. Lynch[3], and Stephen R. Yool[4,5]

Format of Submission: Forest Ecology and Management

[1] School of Natural Resources and Environment

Biological Sciences East

Tucson, AZ 85721, USA

[2]Department of Ecology and Evolutionary Biology

[3]United States Forest Service, Rocky Mountain Research Station

[4]School of Geography and Development

[5]The University of Arizona

[6]Corresponding Author: email: tswetnam@email.arizona.edu, phone: (520) 247-2293

Abstract

Aerial LiDAR local maxima forest inventories cannot discriminate mid- and understory trees because of sampling limitations inherent to the technology. To address this issue, we take advantage of scale invariance [e.g. self-similarity] in natural phenomena rank-size distributions where any part of the distribution shares the same shape and scale parameters as the whole. We first tested whether the field observed data sets fit a tapered Pareto distribution with an exponential decline in rank-size of very large members. We next matched two aerial LiDAR segmented inventory plot networks to the observed mid- and overstory distribution [in our case: the heights of individual trees] fit to a tapered Pareto using linear binning and least squares regression. In plots where canopy cover exceeded 50% the observed height size distribution [1 meters (m) < x < 38 m] fit a tapered Pareto distribution [$r^2 = 0.88$], based on a two-sample Kolmogorov-Smirnov test [p-value ≥ 0.497, K-S statistic = 0.184]. In plots with less than 50% cover the inventory did not fit a tapered Pareto; however, the local maxima segmentation had a low rate of omission error for mid- and understory trees because trees were openly spaced, making estimation of mid- and understory inventory unnecessary. In plots with > 50% cover the local maxima segmentation of overstory was accurate for trees > 12 m height; and the rank-size distribution of the local maxima were indistinguishable from the observed rank-size tapered Pareto distribution using a two-sample K-S test [p-value ≥ 0.153, K-S statistic = 0.296]. These results suggest shape and scale parameters of an overstory distribution alone are sufficient to generate a tapered Pareto distribution prediction of understory trees. Foresters and ecologists interested in stand dynamics at landscape scale can apply these techniques to aerial LiDAR forest inventory stands to better estimate the rank-size and density of understory where field measurements are physically or economically impractical.

Highlights

- Forest rank-size distributions are self-similar and described by a tapered Pareto

- Scale invariance allows prediction of forest rank-size distributions

- Overstory tapered Pareto distribution predicts mid- and understory rank-size

Keywords: LiDAR, Scaling, Pareto distribution, understory trees, local maximum algorithm

1 Introduction

Creating a complete tree inventory of an entire forest, at landscape scale [e.g. 10^5 -10^8 individuals], would have been considered an impossible undertaking fifteen years ago. Forest inventory analyses have traditionally relied upon plot-based field surveys, but it is economically and physically impractical to inventory geographically large areas using plot methods. Such an inventory can today be surveyed by aerial Light Detection and Ranging (LiDAR), in a few days (van Leeuwen and Nieuwenhuis 2010, Ustin and Gamon 2010). However, one weakness is the inability to differentiate trees in the mid- and understory where canopies are interconnected or obscured beneath overstory trees from the aircraft sensor. As a consequence, despite its emerging importance, aerial LiDAR has been capable of producing only inventories of the larger trees and not complete forest stand structure, with attendant implications for estimates of standing biomass, carbon sequestration, and species composition (Chen et al. 2006, Falkowski et al. 2006, 2008, Frazer et al. 2011, Gatziolis et al. 2010, Hudak et al. 2008).

We were interested in whether this limitation of aerial LiDAR can be addressed by *a priori* understanding of density-dependence and scaling relationships of trees in forests, many of which follow power-laws for their rank-size distribution (Enquist et al. 1998, 1999, Niklas and Enquist 2001, Niklas et al. 2003, Deng et al. 2012). Thus we asked: can the overstory size distribution predict the rank-size frequency of the mid- and understory that heretofore is invisible in most aerial LiDAR inventories?

Techniques for segmenting aerial LiDAR data (Lefsky et al. 2002, van Leeuwen and Nieuwenhuis 2010, Ustin and Gamon 2010) from canopy height models [the vertical difference between the top of tree canopy height return, and the surface elevation] include (1) local maxima [with filtering] (Dralle and Rudemo 1996, Hyyppä et al. 2001, 2005, Persson et al., 2002, Monnet et al. 2010), (2) variable area window (Popescu et al. 2002, Popescu and Wynne 2004), (3) hierarchical inverse watersheds (Chen et al. 2006, Zhao and Popescu 2007), and (4) spatial wavelet (Falkowski et al. 2006, 2008). There are also more recent techniques to segment the discrete pulse return point cloud (Li et al. 2012, Yao et al. 2012) as well. Each of these techniques offer strengths and weaknesses for

inventorying in various forest types but none have demonstrated the ability to discriminate mid- or understory with a level of significance equivalent to that of a field inventory with the possible exception of the Yao et al. (2012) vector technique. While the Yao et al. (2012) approach appears to be a much needed step forward in inventorying forests with aerial LiDAR, many existing datasets have pulse densities that are too low to resolve features of the mid- to understory canopy architecture.

1.1 Self-similarity in rank-size distributions

Forests are shown to have power-law rank-size distributions (Niklas et al. 2003, Enquist et al. 2009, West et al. 2009)] which are variously approximated by Zipf's law or the Pareto distribution (Pareto 1896, Newman 2005). Importantly, rank-size distributions that exhibit scale invariance [e.g. self-similarity] share the same shape and scale parameters for all parts of the whole.

As trees grow to fill space, their canopies and roots become interconnected and compete for limiting resources (Enquist et al. 1998, 1999, Niklas and Enquist 2001, Niklas et al. 2003). In a resource-limited forest environment, at steady state, the rank-size distribution of individuals follow a power law $N_x = \beta x^\alpha$ [Eq. 1] where N_x is the number of individuals of x size measure, β is a normalization constant, x is a primary size measure [e.g. diameter, height, canopy volume], and α the dynamic exponent $\alpha < 0$ (Enquist et al. 1998, 1999, Niklas et al. 2003, Savage et al. 2010, Kempes et al. 2011, Deng et al. 2012). The general quantitative theory of forest structure (Enquist et al. 2009, West et al. 2009) predicts that as trees fill space the rank-size distribution of trees becomes: $N_x = \beta_1 r_x^{-2} = \beta_2 h_x^{-3}$, where r is bole radius, height is h, and β_1, β_2 are unique normalization constants.

Swetnam et al. (*in review*) show forest rank-size distributions become divergent from Eq. 1 in forests that experience episodic disturbance. The tails of the distribution are affected by process noise, and because of the metabolic and physiological limitations fall off at large member size from the power law scaling. Importantly, forests with episodic disturbance are not considered to be at a density dependent steady-state, and therefore their rank-size distribution is not predicted to follow Eq. 1.

Along with allometric scaling of canopy diameter (Purves et al. 2008, Strigul et al. 2008, West et al. 2009, Enquist et al. 2009), maximum tree height (Koch et al. 2004, Kempes et al. 2011), and the elastic buckling strength of a woody bole (McMahon 1973, McMahon and Kronauer 1976), trees must maintain a certain bole diameter to maintain themselves upright and occupy enough space for their canopy to support their metabolic demand. The minimum area required to sustain a tree of any given size can be estimated from the filled space of its individual canopy footprint or belowground root volume (West et al. 2009, Savage et al. 2010, Kempes et al. 2011). Allometric scaling rules can help determine the maximum number of individuals of any size that can occupy a given space. The so-called packing rule (Reineke 1933, Yoda et al. 1963, Enquist et al. 1998, 2009, Niklas et al. 2003, West et al. 2009, Deng et al. 2012) can also be used to establish an accurate stem count of overstory trees with local maxima from the aerial LiDAR (Swetnam and Falk *in review*). Here we suggest that the physiological limits to tree height (Koch et al. 2004, Kempes et al. 2011) are best expressed by a tapered Pareto rank-size distribution (Schoenberg and Patel 2012) in the upper-tail of the distribution (Swetnam et al. *in review*). We applied the tapered Pareto using least square regression to a set of linearly binned rank-size frequency data, based on the understanding that trees cannot continue to increase in size indefinitely and therefore diverge from a heavy tailed Pareto distribution (Newman 2005) in the upper tail with an exponential or hyper-exponential decline in abundance at large size.

2 Methods

2.1 Study Areas

The Pinaleño Mountains (PM) are located south of Safford, Arizona at 32.7° N, 109.9° W (Figure 1). The Pinaleño reach 3,267 m above mean sea level (amsl) atop Mount Graham. The mountains are characteristic of the Basin and Range province, and are a complex of steeply sided canyons with relatively gentle high elevation uplands above 2,700 m amsl (Figure 1). Lower elevation (1,830-2,440 m amsl) forests of the Pinaleño are typical of Madrean Sky Islands (Whittaker and Niering 1975, Niering and Lowe 1984). Common pine species include Ponderosa pine (*P. ponderosa*) and

Southwestern white pine (*P. strobiformis*) which coexist above 2,130 m amsl with Gambel oak (*Quercus gambellii*), Douglas-fir (*Pseudotsuga menziesii*) and white fir (*Abies concolor*). In the highest elevation forests (>2,740 m amsl) Engelmann spruce (*Picea engelmannii*), and corkbark fir (*A. lasiocarpa var. arizonica*) coexist (Whittaker and Niering 1975, Niering and Lowe 1984, O'Connor 2013). Post-fire seral communities of quaking aspen (*Populus tremuloides*) are present throughout the upper elevations of the Pinaleños. The Pinaleño inventory data come from a tree population demographic and disturbance study (O'Connor et al 2010). Seventy nine forest inventory 0.05 ha radial plots were collected across a systematic grid above 2,300 amsl; a total of 2,862 trees were measured for diameter at breast height (DBH, 1.3 m above ground level) and height. Within each plot there was a random 1/3 area in which all trees were measured; for the 0.05 ha plot only trees >20.0 cm DBH were measured.

The Valles Caldera National Preserve (VCNP) is located in the Jemez Mountain range west of Santa Fe, New Mexico at 35.9° N, 106.5° W (Figure 1). Elevations range from 2,300 m amsl in Redondo Meadow to 3,431 m amsl atop Redondo Peak. The Valles Caldera is a collapsed volcanic caldera with a rim approximately 19km wide; within the caldera are resurgent domes over 200 m high, locally referred to as *cerros*. The Valles Caldera shares many of the same species (Muldavin and Tonne 2003) as at the Pinaleño: *P. ponderosa* is common in the lowest elevations of the Preserve (2,100-2,400 m amsl), with some limber pine (*P. flexilis*) and Douglas-fir on mesic sites, Gambel oak is common in post-fire seral stands along with ponderosa and quaking aspen. North aspects tend to be dominated by Douglas-fir and white fir, as well as by sub-alpine fir and Engelmann spruce. The highest elevations of the Valles Caldera are dominated by spruce-fir. Prior to the extensive logging of the caldera beginning in the 1930's (Balmat and Kupfer 2004) there appears to have been evidence of more extensive Douglas-fir on north aspects now dominated by spruce-fir types. Forty-eight 0.1 ha radial plots were collected in 2011 by Santa Catalina Mountains – Jemez River basin Critical Zone Observatory staff (Swetnam 2013); a total of 1,520 live and dead trees were measured, and 3,952 trees were counted (including seedlings and saplings > 15cm tall). Within each

0.1 ha plot a 0.01 ha inner plot was inventoried for all trees; in the 0.1 ha plot all trees >18.5 cm DBH were measured.

2.2 Aerial LiDAR data acquisition

LiDAR flights for the Pinaleño and Valles Caldera were conducted by different vendors with different scanning units but used similar flight and scan parameters and both achieved similar pulse return densities (Table 1). The point cloud data are dense enough [>8 points per meter square (ppsm)] to conduct rigorous assessments of canopy structure in complex terrain (Laes et al. 2008, 2009, Gatziolis et al. 2010). The pulse returns have a footprint of ~20-50 cm at ground level assuming an average flight elevation of 1,000 m above ground level and laser beam divergence of 0.20 - 0.5 milliradians (Table 1). Because the data do not have 100% illumination we use a ratio of pulse returns to estimate the actual cover percentage (cover % = # returns > 2 m ground level / total # returns).

We processed the *.LAS files (ASPRS 2011) in FUSION/LDV (McGaughey 2012). We removed outlier points from the analysis [< -2 m below ground level, and > 60 m above ground level]. A digital terrain model (*.DTM) was created for the bare earth (ASPRS 2011). All study plots were clipped from the full *.LAS tiles with 'PolyClipData' extension in FUSION (McGaughey 2012) using circular polygon buffers equal to the area of the measured plots that were created in ArcGIS 10.1 (ESRI 2012) and exported as *.SHP files.

Canopy height models were generated at a consistent resolution greater than the minimum pulse density [equivalent to surface area illumination]. For all plots in both data sets the μ LiDAR pulse density was 9.3 $< \mu <$ 25.1 ppsm which allowed us to create 0.333 m pixel canopy height models. No smoothing or median filters were applied to the initial canopy height model. The plot extracted canopy height models included some NoData pixels, salt and pepper voids, and corn row spacing effects suggesting that the actual density of the pulse return data are not uniform. Canopy height model data were then converted to ASCII file format (*.ASC) for export to ArcGIS 10.1 (ESRI 2012) and MATLAB 2012b (Mathworks 2012).

The *.ASC canopy height model data were imported into MATLAB. Segmentation was done by local maxima from the canopy height model using a custom algorthim called the 'variable-area local maxima' or VLM (full algorithm details in Swetnam and Falk *in review*).

2.2.1 LiDAR Plot Observations

Surveying trees first in the plot and later in a GIS is labor intensive but essential to validate the corresponding local maxima of the aerial LiDAR. Tree observations included: species, DBH, maximum height, canopy base height, tree canopy condition (live good, fair, poor, or dead), and field photographs. Tree locations within the plot were taken using horizontal distance and compass bearing from the differential GPS-located plot center (Trimble 2005). Field-measured tree heights were re-measured in the FUSION LiDAR Data Viewer (LDV)(McGaughey 2012), a subtotal of 399 identified trees (Pearson's $r = 0.9837, MSE = 0.8097\ m$; Figure SM1) were positively identified from the plot data; if there was any doubt whether a tree position was stem mapped correctly the tree was dropped from the analysis. For the tree rank-size distribution we compared the frequency of the observed vs. local maxima distribution without directly identifying each tree, instead focusing on the distribution of tree sizes for each plot.

2.3 Modeling rank-size distributions

The authors created the two semi-arid forest inventory data sets with a local maxima segmentation algorithm (Swetnam and Falk *in review*). First, we determined by least-squares regression the tapered Pareto distribution (Pareto 1896, Schoenberg and Patel 2012) of the observed overstory and conducted a two-sample Kolmogorov-Smirnov test (Wang et al. 2003). Second, we examined whether an increase in canopy cover affected the shape, taper, or scale parameters of the tapered Pareto distribution across plots with increasing levels of total canopy cover > 2 m above ground level. Third, we compared the local maxima tapered Pareto fit to the observed inventory's distribution to determine whether the parameters and normalization constants were equivalent to one another.

A continuous real variable of any distribution [Eq. 1] has probability $p(x)\,dx$ in the interval $x + dx$. In our case, for a power law distribution the lowest value at which the power law is obeyed is x_{min} which we refer to as the scale parameter β for the cumulative distribution function (CDF): $F(x) = 1 - \left(\frac{\beta}{x}\right)^{\alpha}$ [Eq.2a] and Probability Density Function (PDF) $f(x) = \frac{\alpha}{\beta}\left(\frac{\beta}{x}\right)^{\alpha+1}$ [Eq. 2b] for the Pareto; and tapered Pareto distribution CDF: $F(x) = 1 - \left(\frac{\beta}{x}\right)^{\alpha} \exp\{\frac{\beta-x}{\theta}\}$ [Eq. 3a], and PDF: $f(x) = (\frac{\alpha}{x} + \frac{1}{\theta})\left(\frac{\beta}{x}\right)^{\alpha} \exp\{\frac{\beta-x}{\theta}\}$ [Eq. 3b]. Where the scale parameter β is taken *a priori* from the observed data, the shape parameter α [the dynamic exponent], and θ [the point at which $f(x)$ begins an exponential taper toward zero] (Schoenberg and Patel 2012) can be estimated either by maximum likelihood (White et al. 2008), or by least-square regression. The taper parameter theta is scale dependent – it decreases for very large population sizes because the upper limit to the distribution is finite. Distributions whose tails are finite with a hyper-exponential decline have a negative shape parameter; for exponentially decreasing distributions like the normal or generalized Pareto the shape parameter is equal to zero, and sub-exponentially decreasing distributions have a positive shape parameter (Figure 2). A tapered Pareto distribution changes with different shape parameters α from positive to negative; for negative α the lower tail diverges from log-log linear behavior (Figure 2). We estimated α and θ by least-squares in MATLAB 2012b Curve Fitting Tool (The Mathworks 2012).

Tapered Pareto PDF were fit in MATLAB 2012b Curve Fitting Tool (The Mathworks 2012) for $C * f(x)$ [Eq. 3], where C is a normalization constant applied to both the Pinaleño and Valles Caldera field observed data. For all data, at each study site, the height distribution was $\beta < x < \omega$ where ω is the tallest observed height. Because the Curve Fitting Tool was unlikely to find a fit without any constraints on the distribution we bound a range of the potential parameters based on *a priori* knowledge. The minimum height $\beta = 1\,m$ was established *a priori*; the α parameter value was set at

0 reflecting an expected exponential decline in size; and the θ parameter was initialized at 2 m with a range of 1 to 38 m [the range of the observed height distribution].

We grouped study area plots by canopy cover percentage and examined their observed rank-size distribution by study area. The plots are grouped by cover percentage into 10% categories, the observed range of plots are between <40% to 90% (Table 2).

2.4 Goodness-of-fit tests

The two-sample Kolmogorov-Smirnov (K-S) test (Massey 1951, Wang et al. 2003) is a nonparametric test for the equality of two sample distributions sharing the same empirical distribution function. If the p-value of the test is small we can assume that the samples come from different distributions. The K-S statistic is: $D = max|F_1(x) - F_2(x)|$

where F_1 and F_2 are the empirical distribution functions of the first and second sample respectively. We reject the null hypothesis (that the distributions are the same) if

$D > c(\alpha)\sqrt{\frac{n_1+n_2}{(n_1 n_2)}}$ where α is the level at which to reject (e.g. p = 0.1, 0.05, 0.01), and n

are the independent identically distributed (*iid*) observations of each distribution.

We used the two sample K-S test (Wang et al. 2003) to determine whether (1) the full observed inventories (Pinaleño and Valles Caldera) have tapered Pareto distributions, (2) the observed overstory has the same distribution as the VLM over the same range of height, and (3) the tapered Pareto distribution derived from the VLM overstory fit the observed distribution mid- and understory. We used the *kstest2* in MATLAB 2012b (Mathworks 2012) to conduct the two sample K-S. To compare whether the two distributions [field vs. VLM] were the same for the two sample K-S test we used trees taller than the height at which the VLM under-predicts the distribution (which was between 12 m and 16 m in the Valles Caldera and Pinaleño respectively). We set the least squares regression parameters with *a posteriori* values for both α and β.

3 Results

3.1 Empirical and Fitted Size Distributions

We normalized the inventory rank-size distribution based on total area measured from the Pinaleño [79 plots * 0.05 ha = 3.95 ha] and Valles Caldera [48 plots * 0.1 ha =

4.8 ha] dividing the total number of trees by the area, resulting in a frequency of trees per hectare (Figure 3). Regardless of cover percentage the shape of the distribution is very similar (Figure 3). In plots with a larger sample size, typically explained by a higher percentage ($<$ 80%) of cover, the rank-size distribution appeared to exhibit less variability between the linear bins than in the lower cover types.

Visually, neither forest inventory exhibit a log-log linear Pareto distribution [which would connote a positive scale parameter] (Figures 2, 3, SM2). The observed rank-size distribution has both lower and upper tails that are divergent (Figures 3, SM2). When we plot a tapered Pareto with a negative shape parameter (Figure 2) the distribution closely follows the entire observed rank-size distributions of both locations (Figure SM2).

Because the α (shape parameter) of the distribution was the same, regardless of cover percentage (Table 2) (Figure 3), and β is known, the unknown dependent parameters are θ and the normalization constant.

Both distributions by least squares were found to have similar negative shape parameters across the same range of height [2m $<$ x $<$ 35 m]; Valles Caldera $\alpha \cong$ -0.115 ± 0.44, $r^2 = 0.876$ and Pinaleño $\alpha \cong -0.123 \pm 0.403$, $r^2 = 0.876$ (Table 2). This suggests the distribution is declining hyper-exponentially ($\alpha < 0$) in the upper tail. A graph showing the variance between the observed and the modeled distribution are given in the Supplemental Materials (Figure SM3).

The taper parameter θ was estimated by least squares as $\theta \cong 9.5 \pm 4.8\, m$ (Valles Caldera) and $\theta \cong 9.7 \pm 3.9\, m$ (Pinaleño) (Table 2). The θ value represents the point at which the exponential decline begins and is expected to decrease for larger areas because the upper limit of the distribution cannot change. The two-tailed p-value of an unpaired t-test with different sample sizes and variance was p $<$ 0.001.

3.2 VLM vs. Observed rank-size distributions

The VLM precisely measured the height distribution $>$ 12 m in the Valles Caldera and $>$16 m in the Pinaleño (Figure 3,4, SM4a, SM4b). The VLM does not do well at identifying trees $<$ 12 m in either stand. The height at which the VLM lost accuracy co-

varied with canopy cover % (Figures SM4a and SM4b). As canopy cover increased the height at which the VLM was accurate changed from ~4 m in the most open plots to ~19 m in the densest (Figures SM4a and SM4b). The Pinaleño had on average denser canopy cover with 46/79 plots exceeding 60% cover, compared to 6/48 plots in the Valles Caldera (Table 2). In open ($<$ 40%) cover stands the VLM appeared to have segmented almost all individual trees with little error of omission (Figures SM4a, SM4b). Consequently, the reported confidence in the VLM inventory for areas with low percent canopy cover may be greater than for areas with high cover. In higher cover percentage plots the VLM did equally well for overstory but worse in the understory than the open condition stands. This result is in accordance with the prediction of MST which suggests trees below a certain height will have inter-connected canopies (Enquist et al. 2009).

In plots where canopy cover >50%, the observed height frequency distribution [1 m $<$ x $<$ 38 m] fit a tapered Pareto distribution [$r^2 = 0.88$], based on a two-sample Kolmogorov-Smirnov test [p-value $\geq$ 0.497, K-S statistic $\geq$ 0.184] (Figure 5, Table 3). We test whether the two distributions [field vs. VLM] in both locations were the same for the two sample K-S test based on the range of heights over which the VLM was accurate (Figure 6).

For the second part of the analysis we set the range of values identified by the least squares regression of the VLM segmentation with α and β parameter values bound *a posteriori* (Figure 7). The reported fits (Table 2) for the various cover percentage categories all had similar α estimates which hovered around the full inventory fit of $-0.123 < \alpha < -0.115$. For the VLM overstory regression fit we set α to the *a posteriori* value and kept $\beta = 1\ m$; Curve Fitting Tool determined the θ and normalization constant for each distribution. The resulting fits are reported in Figure 7.

We cannot reject the null hypothesis [that the distributions are the same] based on the two sample K-S tests for (1) the full observed distribution vs. Tapered Pareto, (2) the observed overstory vs. tapered Pareto, or (3) understory tapered Pareto vs. VLM tapered Pareto (Table 3, Figure 6). These results also suggest the Pinaleño and Valles Caldera

share similar distributions for their rank-size frequency, possibly making this behavior generalizable to other semi-arid forests (Table 3).

4 Discussion

In dense canopy semi-arid conifer forests the rank-size distribution of a primary size measures [e.g. maximum height] exhibited what we describe as a tapered Pareto distribution; the shape and scale parameters of this distribution are scale invariant, allowing us to predict the normalization constant [and thus an inventory number] when enough points of the overstory distribution are known. We know *a priori* that physical mechanisms define the scaling and shape parameters of the rank-size distribution in forests [which in these forests are well represented with field study data], and are shown to be statistically indistinguishable regardless of the overall canopy cover percentage (Figure 3). *A posteriori* aerial LiDAR local maxima segmentation are inaccurate at discriminating trees in the mid- and understory because of sampling limitations inherent to the technology. When we apply the tapered Pareto distribution to the overstory with known shape and scale parameters the modeled SFD of the overstory trees accurately predicts the frequency distribution of the mid- and understory trees.

In the semi-arid conifer forests examined in this study, plots $< 50\%$ canopy cover tended not to fit the power law distribution with a high coefficient of determination, possibly due to a lack of self-organization in the open conditions. In those stands, the local maxima segmentation fit the observed plots inventory with a low omission or commission error across all size classes, making regression fitting unnecessary (Figures SM4a and SM4b). In denser plots, which appear to exhibit self-organization [as predicted by MST (Enquist et al. 2009, West et al. 2009)], the application of a power law distribution became a useful tool for prediction of understory tree frequency.

We binned our data linearly [either with 1 m or 0.5 m widths] and found the tapered Pareto had a shape parameter whose value ranged $-0.22 < \alpha < -0.05$ (Table 2). Importantly, the α parameter may change significantly when a different binning

technique is used, along with the normalization constant (Newman 2005, White et al. 2008).

4.1 Tapered Pareto distribution

If we had fit only the overstory distribution by least-squares we would have wrongly found a positive α scaling exponent and over predicted the rank-size distribution of the understory. By bounding the shape and scale parameters, the least squares regression only determined the best fit of the θ parameter and the normalization constant C. The *a priori* assumption that the forest rank-size distribution exhibits scale invariance in its shape and scale parameters is critical to this step. One surprise finding that supports this idea is that regardless of overall canopy cover % the shape parameter of the rank-size distribution do not change significantly (Figure 2, Table 2).

Because the VLM diverged from the observed inventory at around $12\ m - 16\ m$ height there was a question of whether the sampling strategy in the two inventories [full inventory of only $1/10^{th}$ of the 0.1 ha plot in the Valles Caldera and 1/3 of the 0.05 ha plot in the Pinaleño] were sufficient at describing the overall population rank-size distribution before extrapolating the mid- and understory. The average height of an 18.5 cm DBH tree in both locations was 11 ± 4 m suggesting the sampling protocol for measuring large trees in the outer plot was sufficient for at least comparing the VLM maxima to the overstory inventory. The extrapolated frequency of understory relative to the overstory in the Valles Caldera has more uncertainty than the Pinaleño, yet both appear to be consistent in their rank-size distribution.

We did not fit the truncated Pareto (Burroughs and Tebbins 2001) in this study because we found the truncated Pareto does not diverge for the lower tail when the shape parameter is negative in the way the tapered Pareto does. Because we were interested in fitting the lower tail of a distribution to our data the upper-truncated Pareto was not useful in this exercise, regardless of how well it fit the upper-tail of the distribution in our data.

4.2 VLM vs. Observed SFDs

Our work builds on the findings of previous studies by Persson et al. (2002), Popescu and Wynne (2004), Chen et al. (2006), and Falkowski et al. (2006, 2008) which reported a decrease in the accuracy of the segmentation in the understory with increasing canopy cover. Notably, our study benefits from having denser aerial LiDAR data, a suggestion made in Falkowski et al. (2008).

The VLM segmentation was unable to segment trees in the observed mid- and understory for several reasons: (1) canopies are interconnected below a certain height making maxima determination impossible, and (2) some trees are located beneath other tree canopies, causing them to disappear in the 2D canopy height model. The first point is supported by the predictions of MST where the overlap of tree canopies is governed by their allometry where shorter tree canopies tend to be interconnected (Enquist and Niklas 2001, 2002, Niklas et al. 2003). Canonically, the total energy available for metabolic work is equal across space, so for any overstory tree with understory trees growing beneath it the total contribution of production done in the understory trees are likely to only be a reciprocal of the energy in the footprint of the large tree. The contribution of biomass from understory trees within each overstory tree's footprint to the total forest biomass is likely to be within the observed range of variability of biomass of overstory trees in general (i.e. an individual understory tree's total mass is less or the same as a single large branch of the overstory tree).

5 Conclusions and Applications

The techniques demonstrated here provide a robust way to determine the upper confidence boundary of a potential forest mid- and understory, and provide a consistent way for predicting the rank-size distribution of trees that are not observed in an aerial LiDAR local maxima forest inventory. These techniques may help forest ecologists interested in mapping the full size distribution of forests at landscape scales, and better quantify estimates of total forest biomass. Managers and scientists interested in generating virtual inventories of forests based on aerial LiDAR are urged to first parameterizing the rank-size distribution of their forest with local field observations. For

managers, aerial LiDAR derived inventories of forest stands could help in monitoring forest conditions and process without expending valuable resources on monitoring teams that can measure only tiny portions of complex landscapes with variable disturbance histories or local microclimates. For ecologists, the benefit of having landscape level inventories opens up new questions about the distribution of mass and energy flux utilization in forests. The scale of a complete landscape forest LiDAR inventory removes the need for statistical extrapolation from small plot areas, and provides the ability to understand structure and biomass distributions at both large spatial extents and fine spatial resolution.

Acknowledgements

References

ASPRS 2011. LAS specification version 1.4 – R12. 10 June 2012. 27 p. Bethesda, Maryland. www.asprs.org.

Balmat, J., and J. Kupfer. 2004. Assessment of timber resources and logging history of the Valles Caldera National Preserve. University of Arizona Technical Report VCT04011 for Valles Caldera Trust, Tucson, AZ.

Burroughs, S.M., Tebbens, S.F., 2001. Upper-truncated power laws in natural systems. Pure Appl. Geophys. 158, 741-757.

Chen, Q., Baldocchi, D., Gong, P., Kelly, M., 2006. Isolating individual trees in a savanna woodland using small footprint lidar data. Photogramm. Eng. Remote Sensing 72, 923-932.

Dralle, K., Rudemo, M., 1996. Stem number estimation by kernel smoothing of aerial photos. Canadian Journal of Forest Research 26, 1228-1236.

Deng, J., Zuo, W., Wang, Z., Fan, Z., Ji, M., Wang, G., Ran, J., Zhao, C., Liu, J., Niklas, K.J., 2012. Insights into plant size-density relationships from models and agricultural crops. Proceedings of the National Academy of Sciences 109, 8600-8605.

Enquist, B.J., Brown, J.H., West, G.B., 1998. Allometric scaling of plant energetics and population density. Nature 395, 163-165.

Enquist, B.J., Niklas, K.J., 2001. Invariant scaling relations across tree-dominated communities. Nature 410, 655-660.

Enquist, B.J., Niklas, K.J., 2002. Global allocation rules for patterns of biomass partitioning in seed plants. Science 295, 1517-1520.

Enquist, B.J., West, G.B., Brown, J.H., 2009. Extensions and evaluations of a general quantitative theory of forest structure and dynamics. Proceedings of the National Academy of Sciences 106, 7046-7051.

Enquist, B.J., West, G.B., Charnov, E.L., Brown, J.H., 1999. Allometric scaling of production and life-history variation in vascular plants. Nature 401, 907-911.

Falkowski, M.J., Smith, A.M., Gessler, P.E., Hudak, A.T., Vierling, L.A., Evans, J.S., 2008. The influence of conifer forest canopy cover on the accuracy of two individual tree measurement algorithms using lidar data. Canadian Journal of Remote Sensing 34, S338-S350.

Falkowski, M.J., Smith, A.M., Hudak, A.T., Gessler, P.E., Vierling, L.A., Crookston, N.L., 2006. Automated estimation of individual conifer tree height and crown diameter via two-dimensional spatial wavelet analysis of lidar data. Canadian Journal of Remote Sensing 32, 153-161.

Frazer, G., Magnussen, S., Wulder, M., Niemann, K., 2011. Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. Remote Sens. Environ. 115, 636-649.

Gatziolis, D., Fried, J.S., Monleon, V.S., 2010. Challenges to estimating tree height via LiDAR in closed-canopy forests: A parable from western Oregon. For. Sci. 56, 139-155.

Grissino-Mayer, H.D., Baisan, C.H. and Swetnam, T.W. 1995. Fire history in the Pinaleño Mountains of southeastern Arizona: Effects of human related disturbances. In Debano, L.F., Gottfried, G.J., Hamre, R.H., Edminster, C.B., Ffolliott, P.F. and Ortega-Rubio, A., editors, Biodiversity and management of the Madrean Archipelago: the Sky Islands of Southwestern United States and Northwestern Mexico, US Department of Agriculture Forest Service, General Technical Report RM-GTR-264, 399–407.

Hudak, A.T., Crookston, N.L., Evans, J.S., Hall, D.E., Falkowski, M.J., 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. Remote Sens. Environ. 112, 2232-2245.

Hyyppa, J., Kelle, O., Lehikoinen, M., Inkinen, M., 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. Geoscience and Remote Sensing, 39, 969-975.

Hyyppä, J., Mielonen, T., Hyyppä, H., Maltamo, M., Yu, X., Honkavaara, E., Kaartinen, H., 2005.Using individual tree crown approach for forest volume extraction with

aerial images and laser point clouds, In: Proceedings of ISPRS workshop laser scanning. Citeseer, 12-14.

Kempes, C.P., West, G.B., Crowell, K., Girvan, M., 2011. Predicting maximum tree heights and other traits from allometric scaling and resource limitations. PloS one 6, e20551.

Koch, G.W., Sillett, S.C., Jennings, G.M., Davis, S.D., 2004. The limits to tree height. Nature 428, 851-854.

Laes, D.; Reutebuch, S., McGaughey, B., Maus, P., Mellin, T., Wilcox, C., Anhold. J., Finco, M., Brewer, K., 2008. Practical LiDAR acquisition considerations for forestry applications. RSAC-0111-BRIEF1. Salt Lake City, UT: U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center. 7 p.

Laes, D., Mellin, T., Wilcox, C., Anhold, J., Maus, P., Falk, D.A., Koprowski, J., Drake, S., Dale, S., Fisk, H., Joria, P., Lynch, A.M., Alanen, M., 2009. Mapping vegetation structure in the Pinaleño Mountains using LiDAR. RSAC-0118-RPT1. Salt Lake City, UT: U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center. 22 p. 173-181

Lefsky, M.A., Cohen, W.B., Parker, G.G., Harding, D.J., 2002. Lidar Remote Sensing for Ecosystem Studies: Lidar, an emerging remote sensing technology that directly measures the three-dimensional distribution of plant canopies, can accurately estimate vegetation structural attributes and should be of particular interest to forest, landscape, and global ecologists. Bioscience 52, 19-30.

Li, W., Guo, Q., Jakubowski, M.K., Kelly, M., 2012. A new method for segmenting individual trees from the lidar point cloud. Photogramm. Eng. Remote Sensing 78, 75-84.

Massey Jr, F.J., 1951. The Kolmogorov-Smirnov test for goodness of fit. Journal of the American statistical Association 46, 68-78.

MATLAB version 8.0.0783 (R2012B) Natick, Massachusetts: The Mathworks Inc. 2012.

McGaughey, R., 2012. FUSION/LDV: Software for LIDAR Data Analysis and Visualization, Version 3.01. US Department of Agriculture, Forest Service,

Pacific Northwest Research Station, University of Washington.Available online at: (last accessed 24 August 2012).

McMahon, T., 1973. Size and shape in biology. Science 179, 1201-1204.

McMahon, T.A., Kronauer, R.E., 1976. Tree structures: deducing the principle of mechanical design. J. Theor. Biol. 59, 443-466.

Monnet, J., Mermin, É., Chanussot, J., Berger, F., 2010. Tree top detection using local maxima filtering: a parameter sensitivity analysis. Proceedings of Silvilaser 2010.

Muldavin, E., Tonne, P., 2003. A vegetation survey and preliminary ecological assessment of Valles Caldera National Preserve, New Mexico. Rep.Coop.Agree. 01CRAG0014.

Newman, M. E. 2005. Power laws, Pareto distributions and Zipf's law. Contemporary physics 46:323-351.

Niering, W.A., Lowe C.H., 1984. Vegetation of the Santa Catalina Mountains: community types and dynamics. Vegetatio 58, 3-28 (1984).

Niklas, K.J., Enquist, B.J., 2001. Invariant scaling relationships for interspecific plant biomass production rates and body size. Proceedings of the National Academy of Sciences 98, 2922-2927.

Niklas, K.J., Midgley, J.J., Rand, R.H., 2003. Tree size frequency distributions, plant density, age and community disturbance. Ecol. Lett.6, 405-411.

O'Connor, C., D.A. Falk, A.M. Lynch, C.P. Wilcox, T.W. Swetnam, T.L. Swetnam. 2010. Growth and demography of Pinaleño high elevation forests. RJVA 07-JV-11221615317 Performance Report. Tucson, AZ: University of Arizona, Laboratory of Tree-Ring Research; School of Natural Resources and the Environment. 21 p.

Pareto, V., 1896. Cours d'économie Politique, reprinted as a volume of Oeuvres Completes. Droz, Geneva1965.

Perona, P., Malik, J., 1990. Scale-space and edge detection using anisotropic diffusion. Pattern Analysis and Machine Intelligence, IEEE Transactions on12, 629-639.

Persson, A., Holmgren, J., Söderman, U., 2002. Detecting and measuring individual trees using an airborne laser scanner. Photogramm. Eng. Remote Sensing 68, 925-932.

Popescu, S.C., Wynne, R.H., 2004. Seeing the trees in the forest: using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height. Photogramm. Eng. Remote Sensing 70, 589-604.

Popescu, S.C., Wynne, R.H., Nelson, R.F., 2002. Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window size. Comput. Electron. Agric.37, 71-95.

Purves, D.W., Lichstein, J.W., Strigul, N., Pacala, S.W., 2008. Predicting and understanding forest dynamics using a simple tractable model. Proceedings of the National Academy of Sciences 105, 17018-17022.

Reineke, L.H., 1933. Perfecting a Stand-Density Index for Even-Aged Forests. US Government Printing Office.

Reutebuch, S.E., Andersen, H., McGaughey, R.J., 2005. Light detection and ranging (LIDAR): an emerging tool for multiple resource inventory. J. For. 103, 286-292.

Ryan, M.G., Yoder, B.J., 1997. Hydraulic limits to tree height and tree growth. Bioscience 47, 235-242.

Schoenberg, F.P., Patel, R.D., 2012. Comparison of Pareto and tapered Pareto distributions for environmental phenomena. The European Physical Journal Special Topics 205, 159-166.

Strigul, N., Pristinski, D., Purves, D., Dushoff, J., Pacala, S., 2008. Scaling from trees to forests: tractable macroscopic equations for forest dynamics. Ecol. Monogr. 78, 523-545.

Swetnam, T.L., Falk, D. A., Enquist, B. J., Lynch, A. M., Guertin, D. P., Yool, S.R., *in review*. The generality of Metabolic Scaling Theory in differing forest disturbance regimes. The American Naturalist.

Swetnam, T.L., Falk, D.A., *in review*. A Variable-area Local Maxima tool for segmenting individual trees from aerial LiDAR: Allometric scaling rules to reduce error in individual tree segmentation. Forest Ecology and Management.

Takahashi, T., Awaya, Y., Hirata, Y., Furuya, N., Sakai, T., Sakai, A., 2010. Stand volume estimation by combining low laser-sampling density LiDAR data with QuickBird panchromatic imagery in closed-canopy Japanese cedar (Cryptomeria japonica) plantations. Int. J. Remote Sens. 31, 1281-1301.

Trimble Navigation Limited. 2005. White Paper: H-Star Technology Explained. 9 p.

Ustin, S.L., Gamon, J.A., 2010. Remote sensing of plant functional types. New Phytol. 186, 795-816.

Van Leeuwen, M., Nieuwenhuis, M., 2010. Retrieval of forest structural parameters using LiDAR remote sensing. European Journal of Forest Research129, 749-770.

Wang, J., Tsang, W.W., Marsaglia, G., 2003. Evaluating Kolmogorov's distribution. Journal of Statistical Software 8.

West, G.B., Brown, J.H., Enquist, B.J., 1999. A general model for the structure and allometry of plant vascular systems. Nature 400, 664-667.

West, G.B., Enquist, B.J., Brown, J.H., 2009. A general quantitative theory of forest structure and dynamics. Proceedings of the National Academy of Sciences 106, 7040-7045.

White, E.P., Enquist, B.J., Green, J.L., 2008. On estimating the exponent of power-law frequency distributions. Ecology 89, 905-912.

Williams, A.P., Allen, C.D., Macalady, A.K., Griffin, D., Woodhouse, C.A., Meko, D.M., Swetnam, T.W., Rauscher, S.A., Seager, R., Grissino-Mayer, H.D., 2012. Temperature as a potent driver of regional forest drought stress and tree mortality. Nature Climate Change 3, 292-297.

Yao, W., Krzystek, P., Heurich, M., 2012. Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. Remote Sens. Environ.123, 368-380.

Yoda, K. 1963. Self-thinning in over-crowded pure stands under cultivated and natural conditions.(In-traspecific competition among higher plants. XI.). J Biol Osaka City Univ 14:107-129.

Zhao, K., Popescu, S., 2007. Hierarchical watershed segmentation of canopy height model for multi-scale forest inventory. Proceedings of the ISPRS working group Laser Scanning 436-442.

Zhao, K., Popescu, S., 2009. Lidar-based mapping of leaf area index and its use for validating GLOBCARBON satellite LAI product in a temperate forest of the southern USA. Remote Sens. Environ. 113, 1628-1645.

Zhao, K., Popescu, S., Nelson, R., 2009. Lidar remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. Remote Sens. Environ. 113, 182-196.

Zheng, G., Moskal, L.M., 2009. Retrieving leaf area index (LAI) using remote sensing: Theories, methods and sensors. Sensors 9, 2719-2745.

Figure Captions

Figure 1: The two aerial LiDAR study areas in Arizona and New Mexico. The Valles Caldera (VCNP) have 48 0.1 ha randomly located forest calibration plots; the Pinaleño Mountains (PM) have 78 0.05 ha plots located in a gridded transect.

Figure 2: Example of different tapered Pareto PDFs $f(x)$ with either positive-valued, zero, or negative-valued shape parameters α. When $\alpha > 0$ the distribution is log-log linear in the lower tail but diverges in the upper tail beyond θ; when $\alpha = 0$ the lower tail is flat, when $\alpha < 0$ the lower tail diverges toward an asymptote near β and is exponentially declining in the upper-tail beyond .

Figure 3: Empirical rank-size distribution normalized per unit area by cover percentage [estimated from the aerial LiDAR] in the two study areas. The Valles Caldera are on the left, Pinaleño on the right. Both distributions share a similar shape parameter for the tapered Pareto, regardless of cover percentage.

Figure 4: The observed tree height (black line) for Valles Caldera (left panel) and Pinaleño (right panel) with the VLM output for two runs with 1 and 3 iterations. The vertical dashed red lines indicate the height above which the VLM accurately captured tall tree abundance. The average cover percentage in the Valles Caldera was lower than in the Pinaleño, suggesting the accuracy at a lower height is an artifact of sampling and not forest structure.

Figure 5: Two-sample K-S test results for the observed height distribution versus the least squares regression from the tapered Pareto (values from Table 2) for PM and VCNP. Both distributions are not significantly different, meaning we cannot reject they are from the same distribution.

Figure 6: Two-sample K-S test results for the observed vs. VLM distribution for overstory trees >12 m in the Valles Caldera and >16 m in the Pinaleño. Both distributions are not significantly different, meaning we cannot reject they are from the same distribution.

Figure 7: Comparison of the observed tapered Pareto (black line) versus the VLM tapered Pareto (gray line) using a negative α from the observed data; the θ parameter and normalization constant are determined by least squares, the β parameter was set at 1 m. The 12 m cut-off height for the VLM is shown with a break line. The points are displayed with 0.5 m linear bin widths. Notably, the modeled distribution has a strong fit to the observed data in the shorter height classes not used to fit the distribution.

Figures

Figure 1

$$f(x) = (\alpha/x + 1/\theta)(\beta/x) \exp(\beta - x/\theta)$$
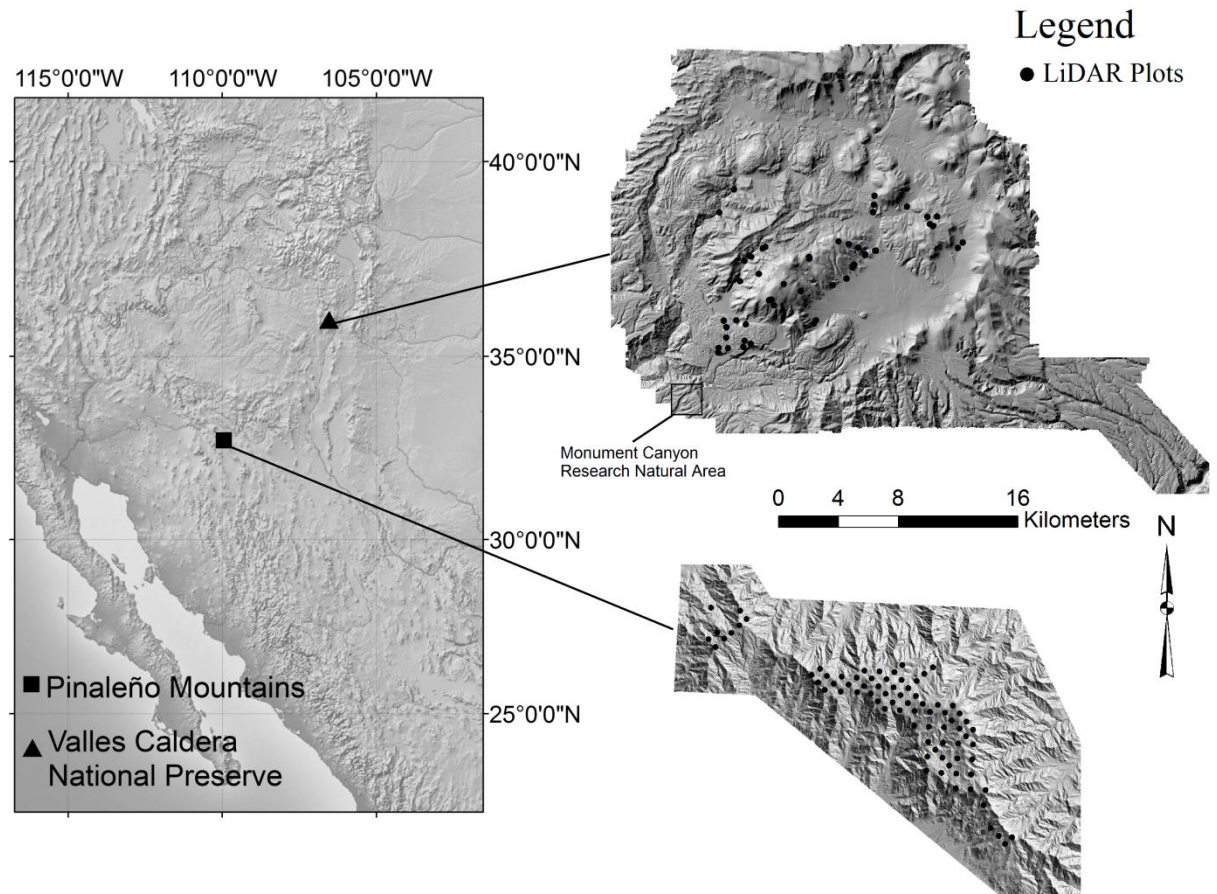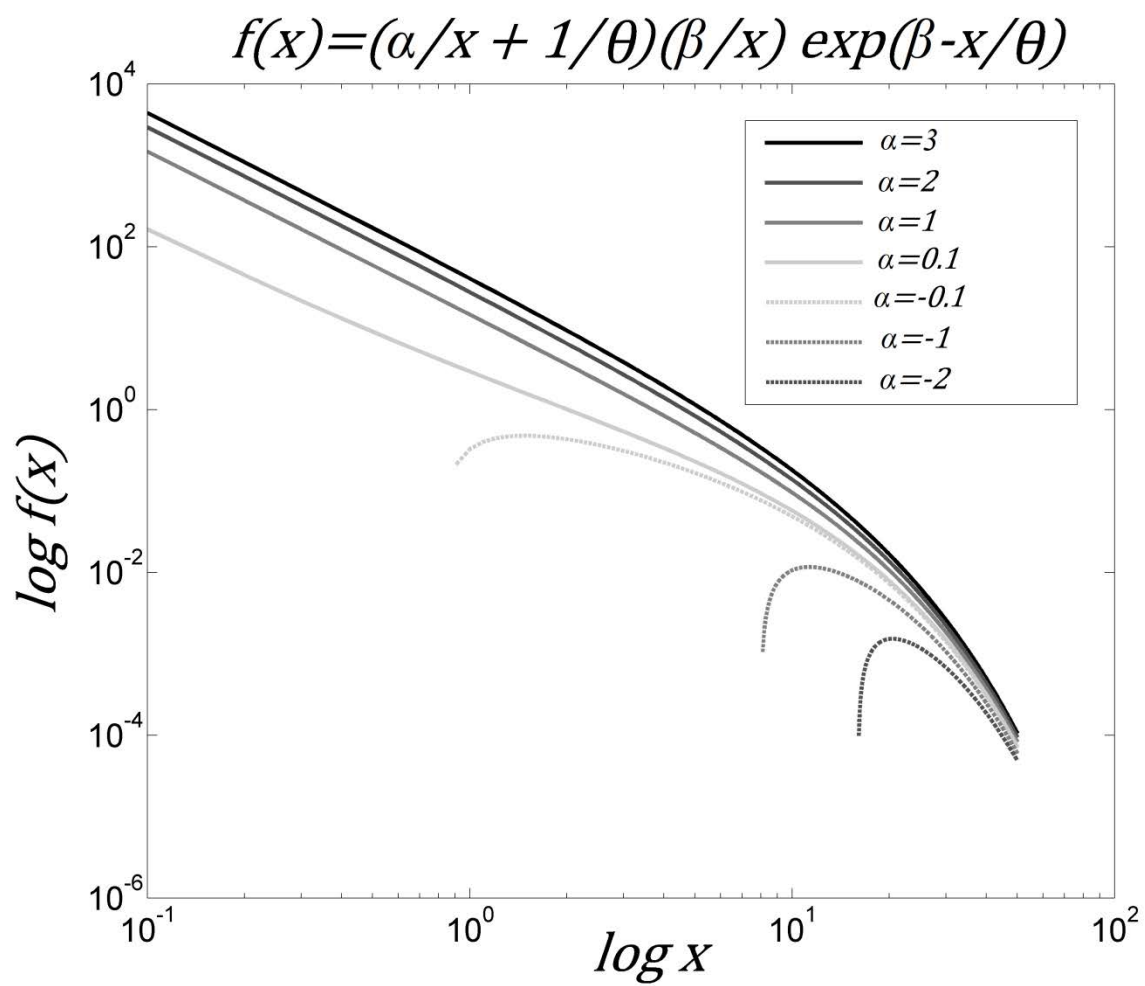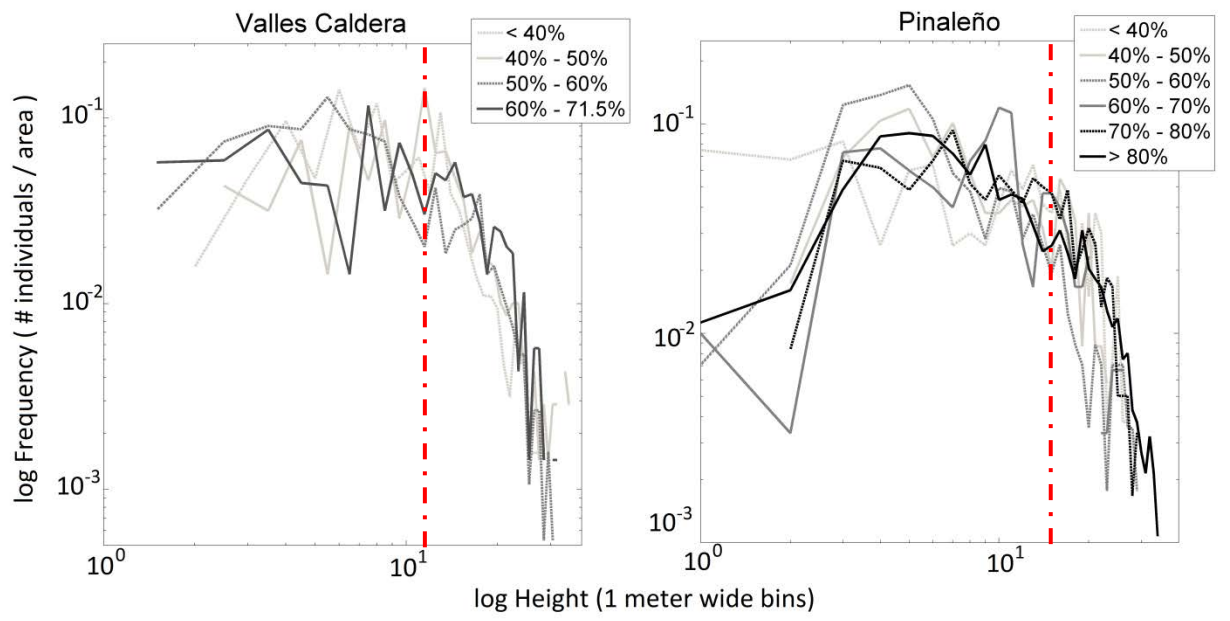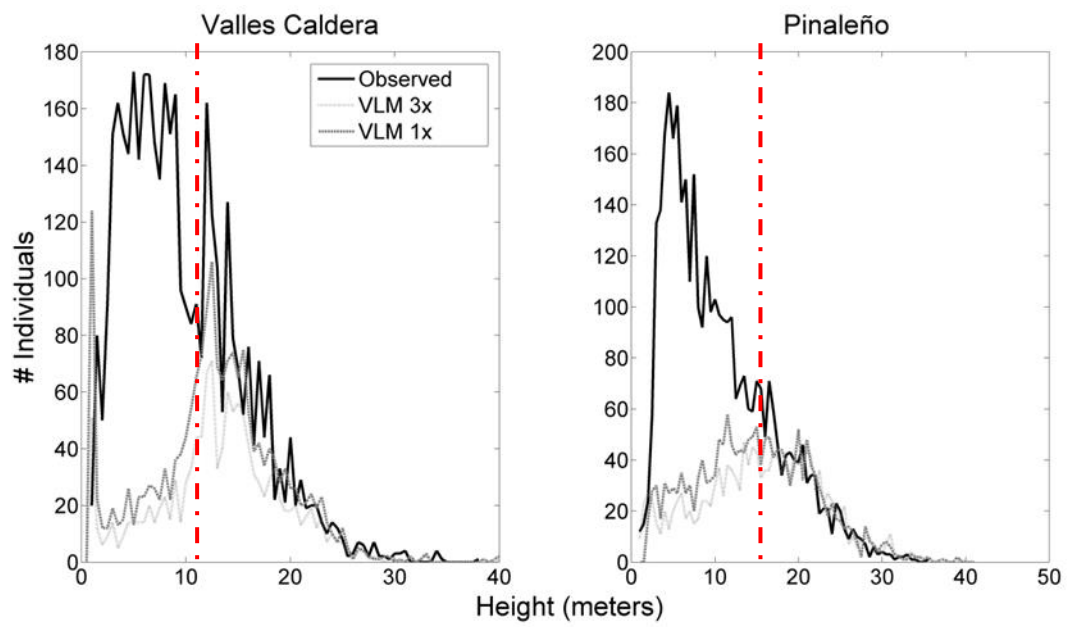
Figure 2
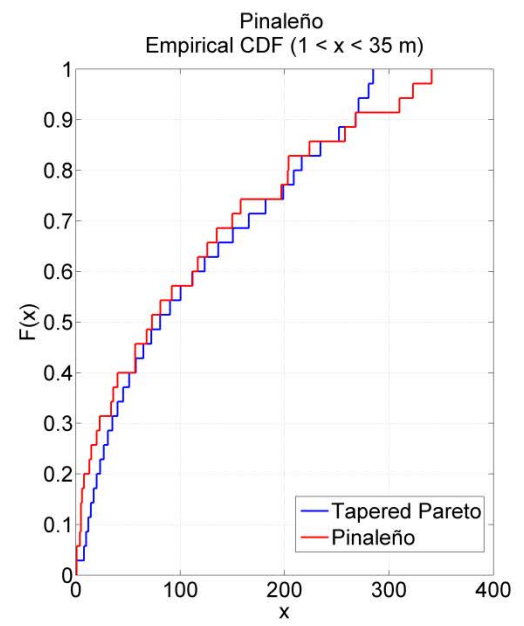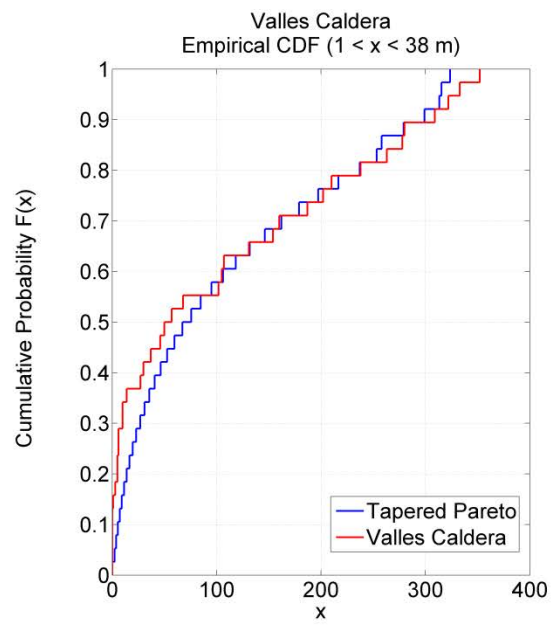
Figure 3
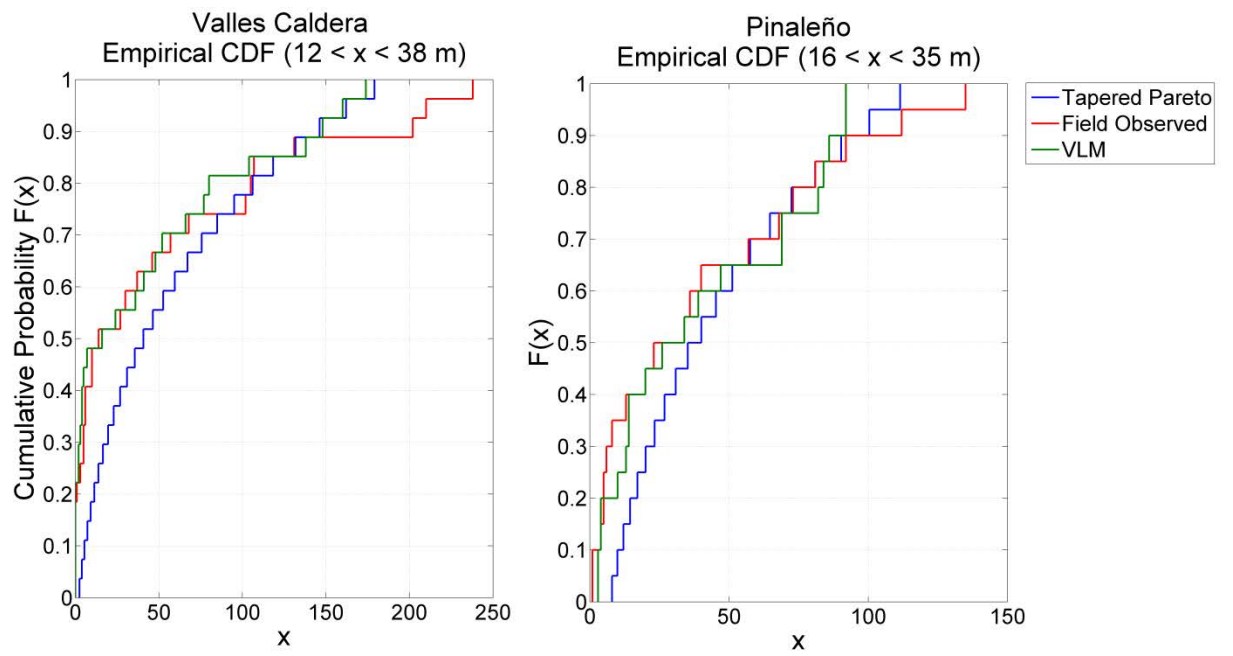
Figure 4
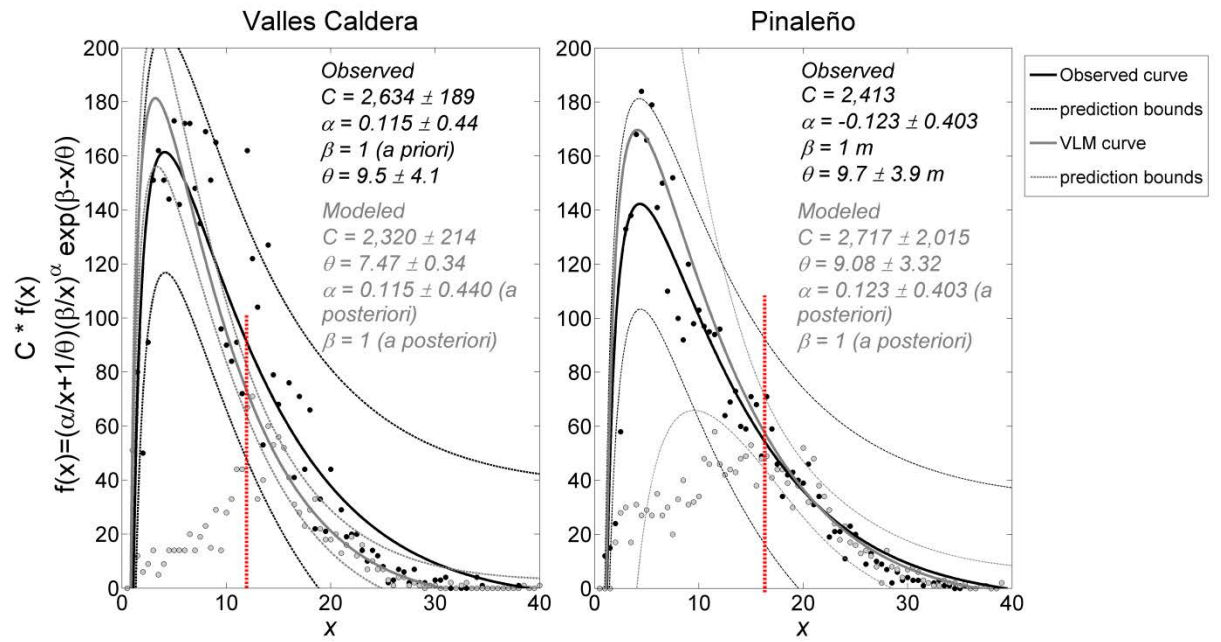
Figure 5

Figure 6

Figure 7

Tables

Table 1. Study area aerial LiDAR flight-parameters.

Scan Characteristic	Pinaleño	Valles Caldera National Preserve
Vendor/Provider	Watershed Sciences	NCALM
Acquisition Date	September 22-27, 2008	January 2010, July 2010
Scanner	Leica ALS50 Phase 2	Optech Gemini
Pulse Rate	70-90kHz	100kHZ
Scan Rate	52.2Hz	<100Hz
Pulse returns	1-4, + 8-bit Intensity (0-255)	1-4, + 8-bit Intensity (0-255)
Scan Angle	15 degrees	25 degrees
Divergence in milliradians (mrad)	0.22 mrad	0.25 mrad
Stated Accuracy (Vertical/Horizontal)	3.2cm/1.0m	7.0cm/1.0m
Flight above ground level	800-1,300 m	~1000 m
Flight line overlap	50% side lap	50% side lap
μ pulses per square meter (ppsm)	Leaf-on, 7.36 ppsm	Leaf-off (Snow): 8.86 ppsm Leaf-on: 5.91 ppsm
μ Bare ground spacing (ppsm)	0.98 ppsm	1.11 ppsm
Acquisition Area	85,518 Ac (34,608 Ha)	Leaf-off:72,648 acres (29,400 Ha) Leaf-on: 186,811 acres (75,600 Ha)
$\sum$ **Point returns**	2,892,925,979	Leaf-off: 2,541,885,987 Leaf-on: 7,754,915,628
Units	Meters	Meters
Projection, Datum	WGS84, NAD83	WGS84, NAD83

Table 2. Least squares regression of the tapered Pareto to tree height [1 m wide linear bins] for the observed SFD. The top portion of the table represents the entire distribution, as shown in Figure 3 and Figure SM2. NF means there was 'No Fit' found in the range of the data.

Full distribution	Plot n	Tree n	Range	r^2	β (m)	$\alpha \pm$SE	θ (m) $\pm$SE
Pinaleño	79	3,941	1.0 m < x < 35 m	0.88	1	-0.123 ±0.400	9.70 ± 3.90
Valles Caldera	48	3,909	1.0 m < x < 38 m	0.88	1	-0.105 ± 0.440	9.60 ± 4.1

Pinaleño	Plot n	Tree n	Range	r^2	β (m)	$\alpha \pm$SE	θ (m) $\pm$SE
<40%	13	266	1 m < x < 28 m	0.58	1	-0.001	17.03 ± 6.70
40 – 50 %	11	346	2.0 m < x < 30 m	0.72	1	-0.22 ± 2.51	7.78 ± 1.11
50 – 60 %	9	566	2.0 m < x < 32 m	0.77	1	-0.15 ± 0.05	7.00 ± 1.65
60 – 70 %	7	300	3.0 m < x < 31 m	0.51	1	-0.09± 0.06	11.76 ± 5.31
70 – 80 %	13	598	3.0 m < x < 32 m	0.82	1	-0.18 ± 1.21	10.47 ± 11.41
80 – 90 %	19	1,393	1.0 m < x < 35 m	0.86	1	-0.11 ± 0.66	9.33 ± 6.20
80 – 97 %	26	1,865	3.0 m < x < 30 m	0.86	1	-0.10 ± 0.63	10.10 ± 6.55

Valles Caldera	Plot n	Tree n	Range	r^2	β (m)	$\alpha \pm$SE	θ (m) $\pm$SE
<40%	11	636	1.0 m < x < 27 m	0.59	1	-0.13 ± 1.52	9.37 ± 14.27
40 – 50 %	14	697	1.0 m < x < 34 m	0.49	1	-0.09 ± 1.21	12.87 ± 6.31
50 – 60 %	17	1,881	1.0 m < x < 38 m	0.90	1	-0.12 ± 0.71	7.81 ± 5.35
60 – 70 %	6	695	1.0 m < x < 31 m	0.56	1	-0.05 ± 2.09	12.49 ± 2.78

Table 3: Two-sample Kolmogorov-Smirnov (K-S) test results for the observed inventory and the VLM vs. the best-fit tapered Pareto from the least squares. The empty cells are where comparisons cannot be made because they are not representative of one another (e.g. full Pinaleño vs. Valles understory).

	Tapered Pareto (Full)	
	p-value	K-S statistic
Pinaleño (Full)	0.640	0.171
Pinaleño Understory (2 m $< x <$ 16m)	0.912	0.188
Pinaleño LiDAR Overstory (15 m $< x <$ 35 m)	0.531	0.238
Valles Caldera (Full)	0.497	0.184
Valles Caldera Understory (2 m $< x <$ 12m)	0.991	0.167
Valles Caldera LiDAR Overstory (12 m $< x <$ 38 m)	0.153	0.296

Supplemental Figure Captions

Figure SM1: Observed vs. LiDAR height of trees from Pinaleño plots: $n = 399, r = 0.9837, MSE = 0.8097 \, m$.

Figure SM 2. Observed SFD for all plots in the Pinaleño (light gray) and Valles Caldera (dark gray), and a tapered Pareto distribution (black).

Figure SM3: Tthe observed frequency of individuals per hectare [with linear binning] versus the predicted tapered Pareto distribution with values given in Table 2.

Figure SM4a: Tree height distributions for the VCNP in 10% canopy cover classes.

Figure SM4b: Tree height distributions for the PM in 10% canopy cover classes.

Supplemental Figures

n = 399
r^2 = 0.9837
Mean Square Error = 0.8097
p = 0.0055
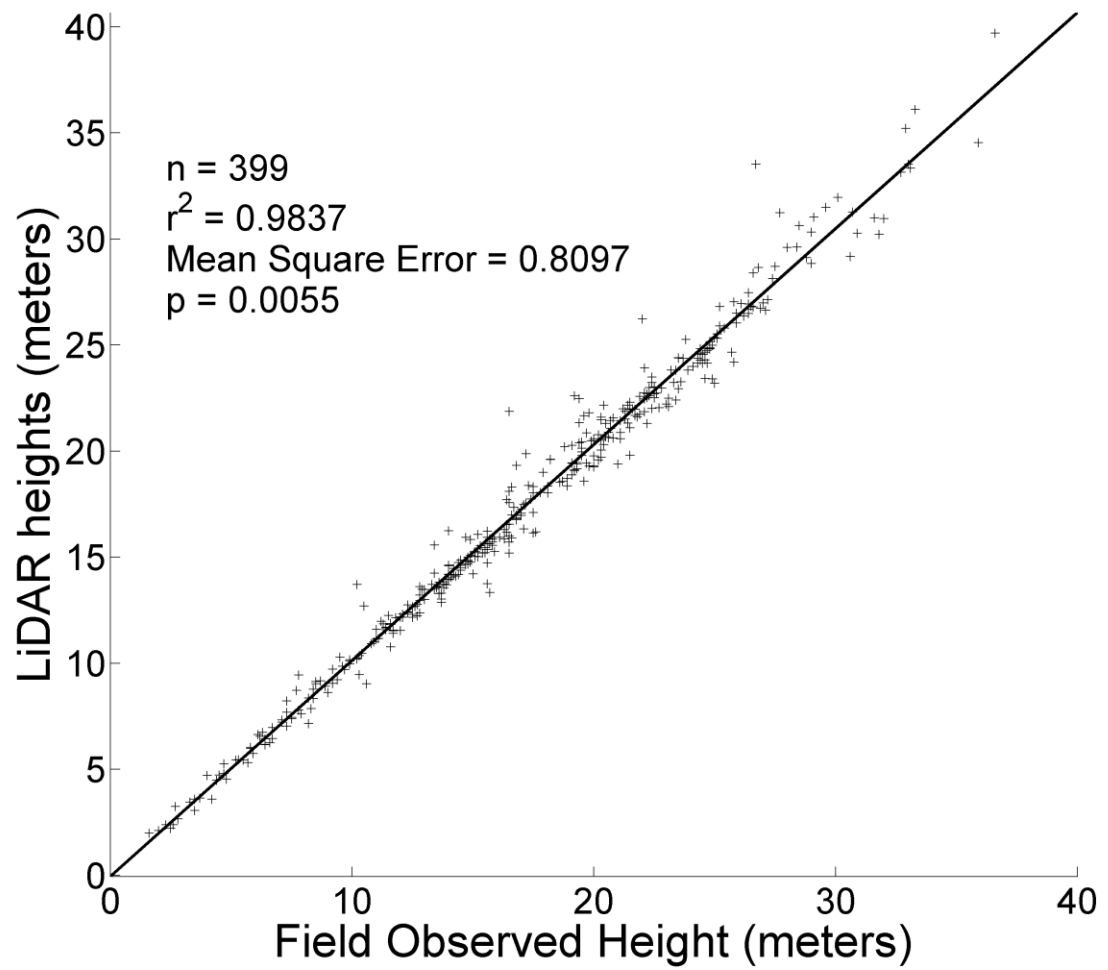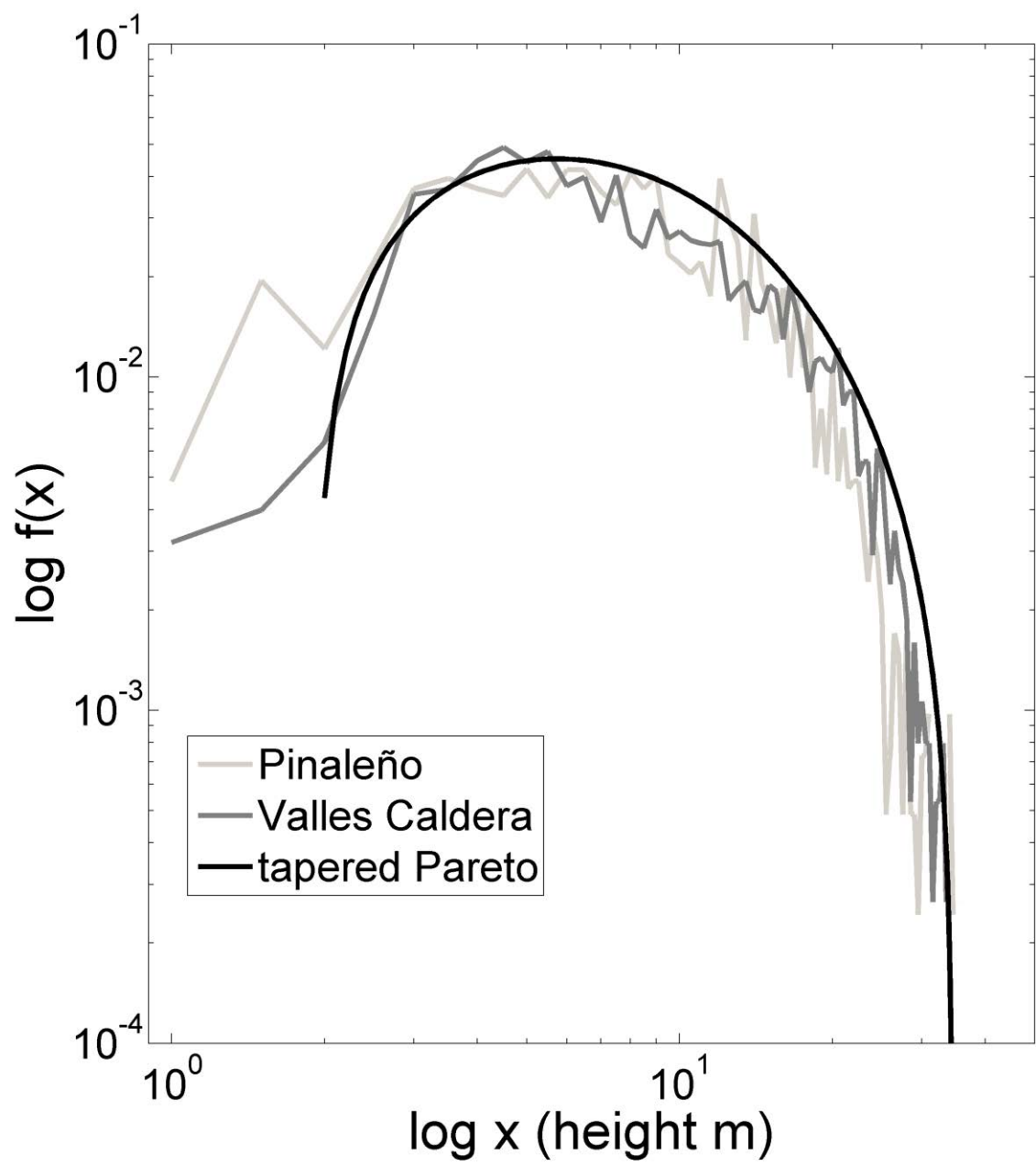
Figure SM1
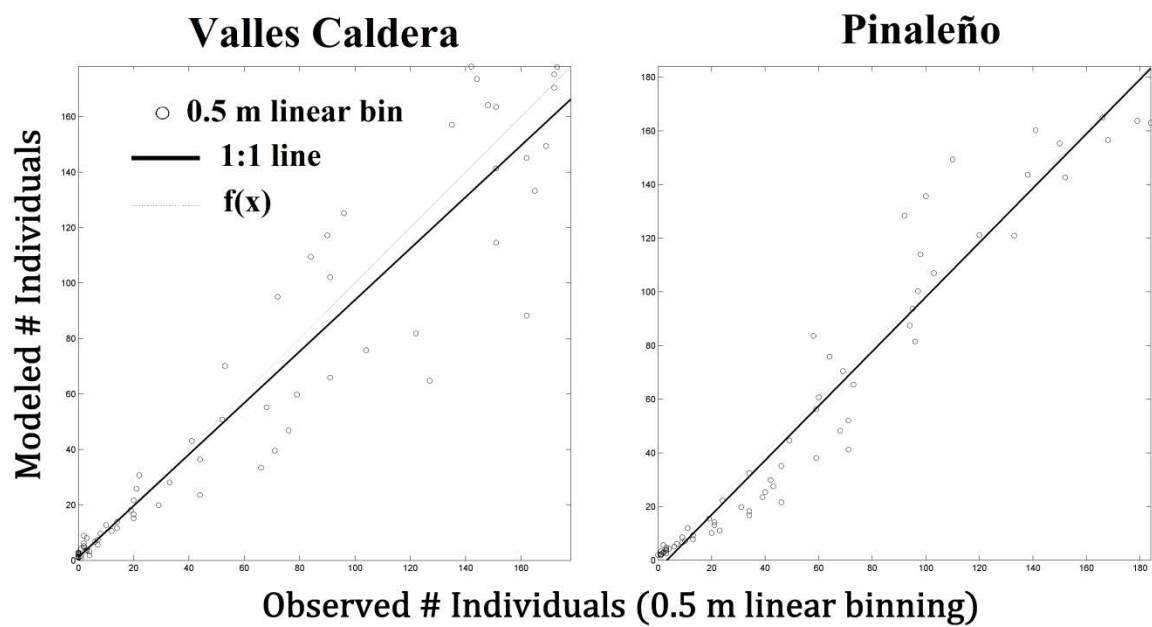
Figure SM2
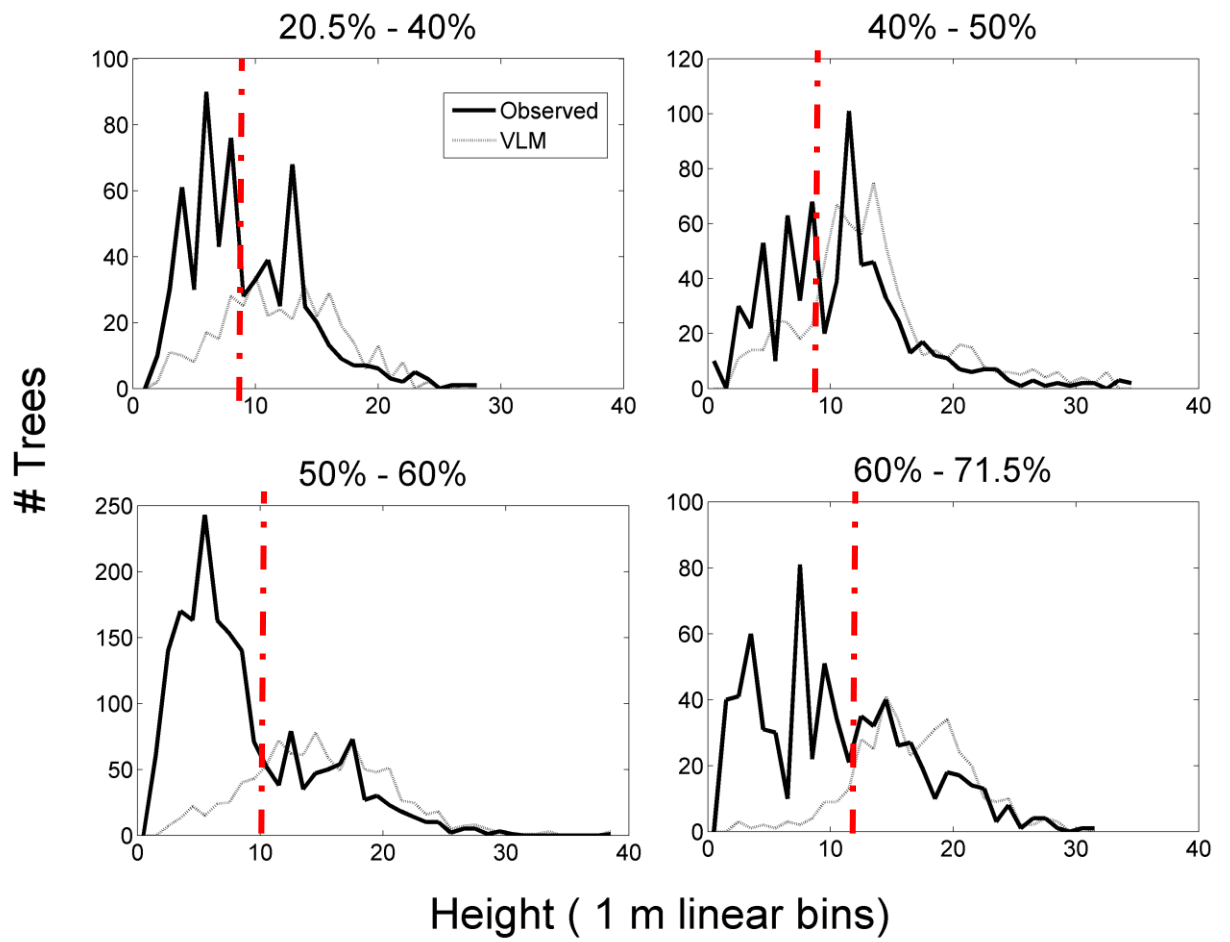
Figure SM3.
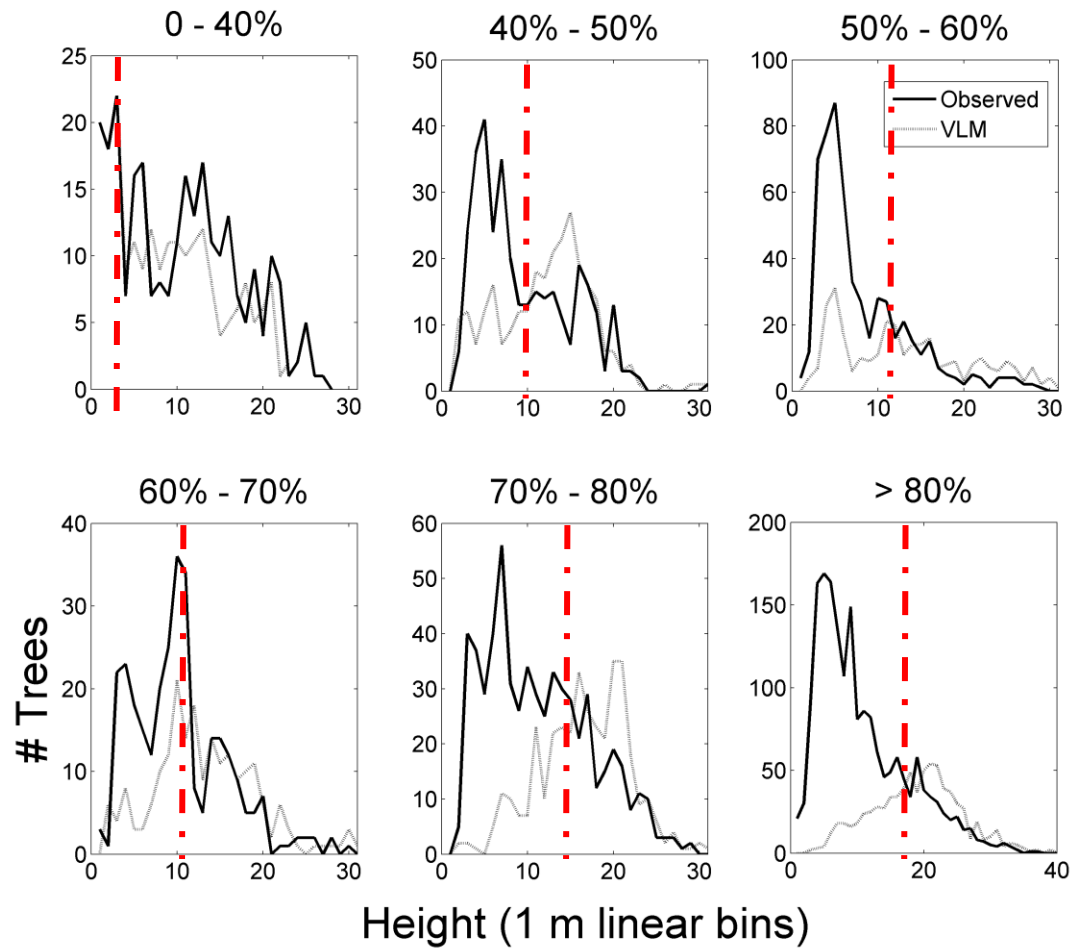
Figure SM4a.

Figure SM4b.

APPENDIX D: CHARACTERIZING FOREST STRUCTURE AND DISTURBANCE LEGACIES WITH AERIAL LIDAR USING GETIS-ORD G_I^* AND ANSELIN'S MORAN I LOCAL INDICATORS OF SPATIAL ASSOCIATION

Tyson L. Swetnam · Ann M. Lynch · Donald A. Falk · Stephen R. Yool · D. Phillip Guertin

Format of Submission: Landscape Ecology

Word count: 7,128

T. L. Swetnam (Corresponding author) · D. A. Falk · D. P. Guertin

School of Natural Resources and Environment, The University of Arizona

Biological Sciences East Building Room 325, 1311 E 4th St

Tucson, AZ 85721, USA

email: tswetnam@email.arizona.edu

phone: (520) 247-2293

A.M. Lynch

U.S. Forest Service, Rocky Mountain Research Station

1215 E. Lowell Street. Tucson AZ 85719.

S.Yool

School of Geography and Development, University of Arizona

P.O. Box 210076 Tucson, AZ 85721

Abstract

Standing forest structure reflects legacies of past disturbance, but only a limited number of characteristics can be quantified by conventional methods. Aerial LiDAR enables foresters, biometricians, and ecologists an unprecedented view of entire landscapes [e.g. ecosystems]. Here we use an individual tree inventory derived from multiple LiDAR flights over semi-arid coniferous forests in Arizona and New Mexico to identify structural legacies from historical disturbance. Parts of these landscapes were either preserved from logging and wildfires or were logged and burned, in some cases extensively, in the recent past through the present. We calculated two different local indicators of spatial association (LISA) to analyze a landscape-scale individual tree inventory: the Getis-Ord G_i^* and the Anselin Moran's I_i in a GIS. Our results show that landscape legacies of natural disturbances such as fires of different severities and dates, as well as human-caused disturbances such as logging and development, can be identified at the stand scale for young to old-growth stands. These LISA methods may be useful for landscape managers interested in defining stand boundaries, quantifying the level at which a landscape is modified from its historical range of variability, and possibly to locate historical disturbance legacies not included in modern records.

Keywords: spatial statistics, Anselin Moran's I, Getis-Ord G, disturbance, forest structure, LiDAR

Introduction

The need for landscape scale spatial analyses

Differentiating *in situ* whether vegetation communities have (1) been fragmented by human development or by natural processes, (2) developed in response to differences in landscape control features [e.g. soil depth, insolation, soil available water] or (3) are merely at an early- to mid- to late-developmental stage in time, provides ecologists and foresters with important clues toward understanding how patterns form in complex natural systems. Semi-arid forests across the western United States (US) are prone to episodic disturbance events which alter forest structure (Allen et al. 2002) and have been profoundly impacted by modern anthropogenic land use. Without a record of past events current forest structure cannot be placed into a developmental continuum that managers can understand. Identifying and tracking structural legacies of disturbance are one of the most critical components of managing forest ecosystems to be healthy, resilient, and resistant in the event of future disturbance and climate variability (Stephens et al. 2013).

Managing forests on restricted budgets, in the face of increased frequency and intensity of fire disturbance (Westerling et al. 2006), and increased climatic variability (Williams et al. 2012), require better information that spatially and quantitatively defines the effects of past disturbance and existing forest structure. Foremost is the need to first characterize and quantify existing forest structure across space: how does the forest vary, and how can we differentiate between natural and anthropogenic variation? Acquiring this new information at a relevant landscape scale is possible through the application of remote sensing (Ustin and Gamon 2010; van Leeuwen and Nieuwenhuis 2010).

Landscape-use and disturbance histories are not particularly well documented in western landscapes prior to the 1900s, and are not spatially geo-referenced, complicating disturbance reconstruction. Often only paper maps and reports survive with an unknown number of records having been lost or discarded; unfortunately some records do not separate planned management actions from actual outcomes. Historical records typically represent only the perimeter of a fire or other event with little additional quantitative information (Johnson and Wittwer 2008; Farris et al. 2010). In cases where no spatial

information about a forest was recorded it becomes necessary to use other lines of evidence such as: tree establishment dates (Margolis and Balmat 2009; Margolis et al. 2007; 2011), tree-ring evidence of insect outbreaks (Lynch 2012) and spatial fire history reconstruction (Farris et al. 2010; Swetnam et al. 2011; Falk et al. 2011). These proxy measures can help to delimit when and where a forest disturbance occurred, but they are spatially imprecise and temporally limited to the areas where tree-ring records are obtainable, time-consuming, and expensive to conduct.

The study areas we describe here in southeastern Arizona (Pinaleño Mountains) and north-central New Mexico (Jemez Mountains) are typical of forests across the southwestern US and northern Mexico. Historical fire history reconstructions for the Pinaleño (Grissino-Mayer et al. 1995) and the Jemez Mountains (Touchan et al. 1996) including the Monument Canyon Research Natural Area (Falk 2004) indicate that frequent low-severity fires dominated in the pine and dry mixed-conifer forest types, similar to ponderosa pine forests across the western US (Foster et al. 1998; Allen et al. 1998; 2002; Swetnam et al. 2009; Farris et al. 2010; O'Connor et al. *in prep*). There is evidence of less frequent high-severity wind- and topography-driven crown fires burning smaller areas in these forests on longer intervals at higher elevations (Margolis and Balmat 2009; Margolis et al. 2007; 2011). Other less frequent events (relative to surface fires) include semi- to multi-decadal drought (Swetnam and Betancourt 2010; Williams et al. 2012), and insect outbreaks (Swetnam and Lynch 1993; Negron et al. 2008; Lynch 2009; O'Connor et al. 2010).

Thinking globally, measuring locally

Legacies of past disturbance on contemporary forest structure are often readily visible both in the field and from above with aerial photography (Ustin and Gamon 2010). Conventional plot-based observations (Bechtold and Patterson 2005) continue to provide detailed plot-scale information to managers and scientists, but their inherent limitations for describing temporal and spatial patterns and processes restrain overall understanding of forest structure and dynamics which necessarily involves inevitable errors from spatial averaging over complex ecosystems (Chave et al. 2004; Tomppo et al.

177

2008). Specifically the uncertainty of linking landscape-scale processes to plot-scale measurements, what Levin (1992) called 'the problem of pattern and scale' makes applying plot scale data to landscapes inherently difficult. Recent innovations in a remote sensing technology (Ustin and Gamon 2010; van Leeuwen and Nieuwenhuis 2010) called 'LiDAR' for the portmanteau of 'Light' and 'RADAR'. Such data provide foresters and biometricians with individual tree scale precision at landscape scale numbering in the tens of millions (Andersen et al. 2005; Falkowski et al. 2009).

Here we examine whether individual tree height distributions are modified by either natural or anthropogenic disturbance across space can be distinguished in LiDAR-derived forest structure using two related LISA including Anselin local Moran's I (1995), and the Getis-Ord family of G-statistics (Getis and Franklin 1987; Getis and Ord 1992; 1996; Ord and Getis 1995; 2001; O'Sullivan and Unwin 2010).

In general almost any standard statistic can be decomposed from a global to a local statistic (Cressie 1993; O'Sullivan and Unwin 2010): global spatial autocorrelations are developed from weighted matrices and cover a region of interest but only support a single statistical value (Moran 1950); local spatial autocorrelations have a range of weights defined by the user at the scale of the data and are unique for each spatially defined location (point, cell, or polygon). Spatial autocorrelation at a global scale will have a significantly different statistic relative to autocorrelation measured at the local scale (O'Sullivan and Unwin 2010). A local statistic requires a user-defined distance or a number of near neighbors from which to draw a statistical distribution, where groups of individuals (e.g. trees) become spatially coherent, forming patterns of like-valued neighbors. Here we explore whether the LISA characterize and differentiate changes in forest structure across space not readily explained by changes in first order effects [i.e. area with the same elevation, aspect, soil type and depth, etc.].

Statistical tests of pattern analysis typically begin with a null hypothesis of complete spatial randomness, where subsequent rejection requires the existence of a pattern exhibiting either a dispersed or clustered distribution based on its z-score [a

positive or negative standard deviation about the mean] and its α-value of statistical significance, typically set at 0.05 or 0.01 [a 95% or 99% interval].

Study Areas

Pinaleño Mountains, Arizona

The Pinaleño Mountains are located southwest of Safford, Arizona at 32.7° N, 109.9° W, in the Coronado National Forest (N.F.) (Figure 1). The Pinaleño reach 3,267 m above mean sea level (amsl) atop Mount Graham. The mountains are characteristic of the Basin and Range province, and are a complex of steeply sided canyons with relatively gentle high elevation uplands above 2,700 m amsl (Figure 1). Lower elevation forests of the Pinaleño are consistent of Madrean Pine-Oak communities across the Sky Islands (1,830-2,440 m amsl) (Whittaker and Niering 1975, Niering and Lowe 1984). Several *Pinus* spp. coexist above 2,130 m amsl with Gambel oak (*Q. gambellii*), these include Ponderosa pine (*Pinus ponderosa*), and Southwestern White pine (*P. strobiformis*). More mesic species include Douglas-fir (*Pseudotsuga menziesii*) and white fir (*Abies concolor*). In the highest elevation forests (>2,740 m amsl) Engelmann spruce (*Picea engelmannii*), and corkbark fir (*A. lasiocarpa* var. *arizonica*) coexist (Whittaker and Niering 1975, Niering and Lowe 1984). Post-fire seral communities of quaking aspen (*Populus tremuloides*) are present throughout the upper elevations of the Pinaleño.

In 2008 the Pinaleño were flown with aerial LiDAR (Mitchell et al. 2012) to provide forest structure data for the Pinaleño Ecosystem Restoration Project. Thirty four thousand hectares of the upper elevation forest above 2,300 m were scanned with high density pulse returns to allow forest canopy structural modeling (Mitchell et al. 2012) (Table 1).

Logging was prevalent in the Pinaleño during the late 19th and early 20th century (Grissino-Mayer et al. 1995; Sanderson and Koprowski 2009).

Jemez Mountains, New Mexico

The Jemez Mountains are located west of Santa Fe, New Mexico at 35.8° N, 106.5° W (Figure 1). The most dominant feature of the Jemez Mountains is a collapsed volcanic caldera with a rim approximately 19 km wide and is managed by the USFS as

the Valles Caldera National Preserve. The Monument Canyon Research Natural Area (RNA) is located on the Santa Fe N.F. in the southwestern Jemez Mountains (Figure 1), at 35.8° N, 106.6° W. Monument Canyon RNA is unique, having been preserved from all logging and from fire since its establishment in the early 20th century (Falk 2004). It is one of a few stands of undisturbed 300+ year old natural ponderosa pine left in the western US (Falk 2004). Monument Canyon is bisected by Cat Mesa and San Juan Mesa, with the canyon draining to the southwest, elevation varies between 2,460 - 2,550 m amsl (Figure 3).

The Jemez share many of the same species as at the Pinaleño, with replacement of Madrean flora by species more typical of the southern Rocky Mountains. *P. ponderosa* is common between 2,100-2,400 m amsl, with some limber pine (*P. flexilis*) and Douglas-fir on mesic sites (Muldavin and Tonne 2003; Muldavin et al. 2006). Gambel oak is common in post-fire seral stands amongst ponderosa and quaking aspen. North aspects tend to be dominated by Douglas-fir and white fir, as well as by subalpine fir (*A. lasiocarpa* var. *lasiocarpa*) and Engelmann spruce.

Part of Monument Canyon recently underwent a mastication thinning treatment where trees <20 cm DBH were removed along the San Juan Mesa (southeastern portion visible in Figure 3). Sites with >30° slope were not thinned. Prior to thinning Monument Canyon was considered one of the worst examples of a 'dog-hair thicket' in ponderosa pine in the western US (Allen et al. 2002; Falk 2004).

In January and June of 2010 and again in June and July 2012, discrete return and waveform LiDAR were flown in 2010 over the Santa Fe N.F. and Valles Caldera National Preserve, by the National Center for Airborne Laser Mapping (NCALM) as part of the National Science Foundation's Critical Zone Observatory (CZO) Program, and again in 2012 by Watershed Sciences as part of the Southwest Jemez Collaborative Forest Landscape Restoration Program (Table 1). In 2010 data from a total are of 75,600 ha were collected, and in 2012 from 52,588 ha; a large portion of the 2012 collection overlaps the 2010 collection (Table 1, Figure 1).

Methods

Spatial disturbance histories

O'Connor et al. (2010) digitized historical aerial photographs drawn upon by past foresters to map the logging contracts awarded in the Pinaleño Mountains in the early to mid-20th century. Those images were ortho-rectified in ArcGIS 10.1 (ESRI 2012) to the logging history timber-sale polygons based on the historical marks and the visible stand conditions. The forest has recovered to the point that logged areas are no longer discernible in modern ortho-photography.

Modern fire perimeters and some insect activity polygons are available for both mountain ranges from both field observations and satellite imaging. The USFS GIS data service (USFS 2013) supports shape files and polygon layers available for each of the two N.F. (Coronado and Santa Fe) from 1970 - 2012.

Aerial LiDAR forest inventory

Aerial LiDAR canopy height models (Persson et al. 2002) were created in FUSION 3.30 (McGaughey 2012) and exported to MATLAB 2012b (The Mathworks 2012) where a variable area local maxima (VLM) algorithm located the tops of individual large trees (Swetnam and Falk in prep) (Figure 2) and subsequently exported into ArcGIS 10.1 (ESRI 2012) as a point layer where the local statistical analyses were performed.

Local Indicators of Spatial Association (LISA)

Statistics to evaluate the existence of clusters in spatial arrangements are considered to be indicators of spatial association. Global spatial autocorrelations include Moran's I (Moran 1950) calculated as:

$$I = \left[\frac{n}{\sum_{i=1}^{n}(y_i - \bar{y})^2}\right] * \left[\frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j} * (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}}\right] \quad \text{Equation 1}$$

where i and j are the different areal units of the study, y is the value of a unit at each location; $w_{i,j}$ is a spatial weight term from the matrix which affects the covariance element by the spatial closesnesss. Moran's I (Eq. 1) includes the covariance term in the numerator on the right side.

Global spatial autocorrelations such as Moran's I yield a single statistic for a study area or population, and implicitly assume no variation across space. In nature this condition is rarely satisfied, suggesting the need for statistics that vary over space (Legendre 1993). Moran's I is a summation of cross products for all sample points, and thus can be illustrated by Local Indicators of Spatial Association (LISA) where a Moran's I is calculated for each i,j pair's statistical associations (Anselin 1995, Zhang and Zhang 2007). For our analysis we assessed the autocorrelation of tree height for individual trees represented as either points or polygons, determined by the VLM in MATLAB 2012b (The Mathworks), in ArcGIS 10.1 'Outlier and Cluster Analysis' Tool (ESRI 2012). This local version of Moran's I (Anselin 1995) provides the z-score of the height distribution based on the provided tree data. In our example we used the standard deviation of the Hot-Spot Analysis Getis-Ord's G_i^* z-scores. Once each break point was chosen we used the Select Tool to create a set of polygons either greater than or less than the range of values specified. For example, z-scores significant at the 90% level are ±1.644 and at 99% are ±2.575.

Getis-Ord G_i and G_i^*

The Getis-Ord statistics (Getis and Ord 1992; Ord and Getis 1995) enable the identification of local concentrations of high to low z-scores in an attribute or 3x3 moving window, where each pixel is the center of a 3-dimensional kernel. A global G statistic is written as:

$$G(d) = \frac{\sum_i \sum_j w_{i,j}(d)x_i x_j}{\sum_{i=1}^n \sum_{j=1}^n x_i x_j} \; for \; all \; i \; \neq j. \qquad \text{Equation 2}$$

The local statistic described by a location i, and thus the local G_i value is:

$$G_i(d) = \frac{\sum_j w_{ij}(d)x_j}{\sum_{j=1}^n x_j} \; for \; all \; i \; \neq j \qquad \text{Equation 3}$$

where $w_{ik}(d)$ are spatial weights from a weight matrix and x_j is the value at location j. Because we are interested in local variations, we employ a local version of the statistic G_i, which is available in ESRI's ArcGIS Toolbox (ESRI 2012; Mitchell 2005) as the 'Hot-Spot Analysis' Tool:

$$G_i^* = \frac{\sum_k w_{i,j} x_j - \bar{X} \sum_{j=1}^n w_{i,j}}{S \sqrt{n \dfrac{[n \sum_{j=1}^n w_{i,j}^2 - (\sum_{j=1}^n w_{i,j})^2]}{n-1}}}$$

Equation 4

where x_j is the value for feature j, $w_{i,j}$ is the spatial weight between the features i and j, n is equal ot the total number of features where $\bar{X} = \frac{\sum_{j=1}^n x_j}{n}$ and the standard deviation $S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - (\bar{X})^2}$. G_i^* is a z-score where scores close to zero indicate no apparent clustering within the sample. Positive z-scores indicate clustering of high values, and negative z-scores indicate clustering of low values. We set Euclidean distance with an inverse distance spatial weight and allowed the tool to set the default neighborhood search threshold (Mitchell 2005).

Anselin Local Moran's I_i

Anselin (1995) used Moran's I (Eq. 1) with LISA to calculate a local I_i score. In our application, we use Anselin Moran's I_i in ArcGIS 10.1 Toolbox (ESRI 2012) 'Cluster and Outlier Analysis' Tool (Mitchell 2005) to describe the neighborhood features of a tree's relationship to its neighbors in the following five categories:

- Not significant: No relationship to the size of neighboring trees
- High – High: The point is a locally tall value, with locally tall neighbors
- High – Low: The point is a locally tall value, with locally short neighbors
- Low – High: The point is a locally short value, with locally tall neighbors
- Low – Low: The point is a locally short value, with locally short neighbors

When we view the z-score values of individual trees we begin to see areas where like-valued trees form self-organized neighborhoods based on a combination of: (1) their shared environmental context, (2) time since establishment [e.g. High-High and Low-Low], and (3) unique individuals [e.g. a High-Low value tree in a Low-Low field of saplings] in areas undergoing transition possibly because of recent disturbance or change.

The equation is weighted by the height of the neighboring trees. The calculation conducted in ArcGIS (ESRI 2012) uses a local Moran's I_i:

$$I_i = \frac{\overline{x_i - X}}{S_i^2} \sum_{j=1, j \neq i}^{n} w_{i,j}(x_j - \bar{X}) \qquad \text{Equation 5}$$

where x_i is an attribute of i, $\bar{X}$ is the mean of $w_{i,j}$ a spatial weight between i and j, where

$S_i^2 = \frac{\sum_{j=1, j \neq i}^{n}(x_j - \bar{X})}{n-1} - \bar{X}^2$ and n is the number of features. The z-score is computed as:

$$Z_{I_i} = \frac{I_{i - E[I_i]}}{\sqrt{V[I_i]}} \qquad \text{Equation 6}$$

where: $E[I_i] = -\frac{\sum_{j=1, j \neq i}^{n} w_{i,j}}{n-1}$ and $V[I_i] = E[I_i^2] - E[I_i]^2$ (Mitchell 2005). The spatial

relationship among features was set in ArcGIS 'Cluster and Outlier Analysis' (Mitchel 2005, ESRI 2012) to the inverse Euclidean distance method. The tool averaged the spatial separation between points to derive a neighborhood size for each tree. The neighbor maxima were weighted by the observed tree's local maximum height in meters.

Thiessen polygons for common stand characterization

Common characteristics of individual trees allow them to be aggregated into 'stands' of like valued individuals based on either: (1) available vegetation type model, (2) local topographic features, (3) soil type model, or (4) combinations of these parameters. Here we are interested in delineating the pattern of the existing forest structure, so we focus on the local statistic z-scores.

We generated Thiessen polygons (Brassel and Reif 1979) [also called Voronoi polygons (Boots 1986)] in ArcGIS 10.1 (ESRI 2012) about the individual points, which include the area of influence about a point relative to its neighbors. Thiessen polygons can be generated either before or after the Hot Spot Analysis (ESRI 2012), as the data are equivalently vector based [point or polygon].

We calculated the Getis-Ord G_i^* statistic and used this to determine the break points in stand architecture, and spatially connected neighborhoods of Thiessen polygons which represented the range of weighted tree neighborhoods in the sample dataset. We merged the individual Thiessen polygons into a single polygon layer, which became the boundary of our common stand based on structural characteristics.

In order to define common stand boundaries we re-classified the continuous data using pre-defined sets of break-lines [e.g. manual, equal interval, quantile, standard deviation, etc. in ArcGIS 10.1 (ESRI 2012)], here we used the standard deviation. Once each break point was chosen we used the Select Tool to create sets of polygons either greater than or less than the range of values specified. For example, z-scores significant at the 90% level are ± 1.644 and at 99% are ± 2.575. We chose to segment the Monument Canyon forest at the 90% level (Figure 5, panel 4).

Results

In both the Pinaleños and Jemez Mountains, the Getis-Ord G_i^* z-score for trees with short neighbors was significantly negative ($z < -2.58, p < 0.01$) and for trees with very tall neighbors significantly positive ($z > 2.58, p < 0.01$). Most stands share the same structural characteristics (i.e. all tall neighbors, or all short neighbors) and have z-scores that far exceeded ±2.58 ($p < 0.01$, 99% confidence interval). If the neighborhood was mixed, a random or non-significant z-score was typically reported ($-1.65 < z < 1.65$); in these stands we considered the structure to be heterogeneous (Figure 2, bottom left).

The Anselin Moran's I_i was superior to Getis-Ord G_i^* because it makes additional observations about each individual tree (Figure 2, bottom right). This had the benefit of helping to identify where, for example, old growth trees are still standing among young reproduction, or where young trees are filling gap space surrounded by old growth overstory trees. The Anselin Moran's I_i did a better job than Getis-Ord G_i^* at discriminating stands of only young trees versus stands where a few over-story trees may be present but are surrounded by young reproduction. The Getis-Ord G_i^* was still useful in clearly delineating changes in structure visually muddled by complex Anselin Moran's I_i index values.

Both of the LISA when applied across a variety of forested stands returned coherent patterns which were readily aligned with known spatial disturbance histories, i.e. the timber sale polygons, fence lines, historical fire perimeters, and with changes in topography (aspect reversals, drainages vs. mesa tops) (Figures 3 and 4).

Pinaleño Mountains: 1960's Timber Sales

 The forest structure of the logged and burned areas near Mt. Graham in the Pinaleño varies widely across the example area from old-growth mixed-conifer along the main riparian drainage (Big Creek), to completely burned stands along the northeast part of the figure (Figure 3). Getis-Ord G_i^* (Figure 3, bottom left) and Anselin Moran's I_i (bottom right) both reveal stands of trees along Big Creek at climax growth state (colored red); however the total proportion of the example being at climax is not continuous due to the logging and fire history.

 Within the areas delineated at timber sales identified as timber sales from the historical aerial photos (O'Connor et al. 2010), there are noticeable changes to stand structure in both the Getis-Ord G_i^* and Anselin Moran's I_i statistics that do not align with the hand drawn lines (Figure 3). The centrally located logging polygon (in yellow) from the 1964-1970 period (lower two panels) reveals in both Getis-Ord G_i^* (bottom left) and Anselin Moran's I_i (bottom right) that loggers harvested trees in ~4 ha patch cuts along the logging roads. Along the western edge, logging appears to have been much more intense, with many more trees removed. This is new information not previously detailed in any records. Where there is a change in aspect [from north to south] the logged patches appear to be recovering at different rates – this is clear in the two patch cuts in the southwest corner of the figure near the Swift Trail highway (Figure 3, bottom left, Table 2). All of the patch-cut areas are in close proximity to the blazed logging roads. These roads were not visible in the aerial photography but are clearly visible in the aerial LiDAR bare earth surface hill shade. The areas along the northeast side of the figure were burned by the 2004 Nuttall-Gibson Fire and the Anselin Moran's I_i (Figure 3, bottom right) reveals where a large stand of high-low and low-high trees exist – these are the standing snags with young reproduction coming up between the dead trees. The Getis-Ord G_i^*(Figure 3, bottom left) does not make this discrimination.

Jemez Mountains, Monument Canyon

 The largest trees (>40 m height) in the Monument Canyon are clustered within the drainages between mesa tops (Figure 4). These old-growth stands are continuous across

most of the research natural area and end abruptly along the southern border fence suggesting that old growth trees south of the fence were logged in the last 50-100 years. Trees along the top of the mesas are shorter than trees in the drainage bottoms, though they are the same age based on a gridded dendrochronological reconstruction of the entire Monument Canyon (Falk 2004). There are smaller patches of new recruitment both in the drainage bottoms and along the mesa tops across the research natural area. This structure is supported by the observation that Monument Canyon was historically more open and has recently in filled with young trees (Falk 2004).

In general, open stands of mature ponderosa in the canyon bottoms have highly positive z-scores for their Getis-Ord G_i^* and are associated with tall trees sharing tall neighbors for the Anselin Moran's I_i (Figure 4, Table 2). Along the slopes of the canyon are some recently established cohorts with negative z-scores and some mid-development clusters with mixed or non-significant Getis-Ord G_i^* z-scores.

Discussion

We found that both Anselin Moran's I_i and Getis-Ord G_i^* can be used to reveal change in forest structure from first order effects, second order effects, and disturbance legacies not readily apparent in other forms of remotely sensed data. Anselin Moran's I_i provides additional information over Getis-Ord G_i^* because it incorporates both the z-score of the neighbors (as Getis-Ord G_i^* does) as well as describing the condition of the point of reference (in this case, an individual tree). The application of LISA to ultra-high resolution, landscape-scale aerial LiDAR data presents unparalleled opportunities to understand forest structure and dynamics (Swetnam et al. 2011).

Disturbance identification

Cross validation with known disturbance histories revealed significant impacts not previously documented, mapped, or quantified explicitly. We intentionally selected areas with known disturbance history to evaluate the potential fine scale variability of aerial LiDAR data for describing forest structure. This research provides a significant advancement in the ability to (1) describe variability in forest structure, (2) map forest

disturbance legacies, (3) make quantitative measurements of past forest disturbance and (4) discover historic events lost or poorly recorded in the written record.

For both the Arizonan and New Mexican examples of forest structure the Anselin Moran's I_i and Getis-Ord G_i^* showed coherent differences that are readily apparent both by disturbance history (in the Pinaleño example), or topographic variability (in the Monument Canyon example). The areas in the Pinaleño logged with either clear cutting or selective patch removal are easily differentiated based on their Getis-Ord's G_i^* z-scores which are correlated with changes in stand structure (Figure 3). The presence of the old logging roads (visible in the aerial LiDAR bare earth hill shade model), and historical aerial photos support the observation that logging was the disturbance agent that modified the forest structure in those locations.

Identifying where structural change in vegetation is a response to a first order effect is most important for determining whether the forest is modified by disturbance. For example, in Monument Canyon there is a strong association of large trees with the more mesic canyon bottoms between mesa tops (Figure 4), whereas along the mesa slopes there are mostly insignificant associations [e.g. heterogeneous] or negative associations [e.g. recently established cohorts] of size neighborhoods. Based on the reconstructed evidence from tree rings (Falk 2004) frequent low severity fires historically spread across the entire Monument Canyon area until the late 19th century. It is plausible that fires were intensified by slope alignments as they spread out of the canyon bottoms, which likely led to continual tree mortality along the slopes preventing larger trees from establishing. The existence of cohorts of regeneration trees, i.e. dog-hair thickets, throughout MCN is another consequence of ponderosa pine simply taking advantage of available space and an absence of modern fires that historically would have modified stand density to create more open stands.

Potential Applications

There are many potential applications for using aerial LiDAR derived forest inventory structure data. Here we list a few examples of potential applications we consider most beneficial using LISA:

1) Help foresters and biometricians delimit common stand areas where trees exhibit similar structural characteristics.

2) Improve upon historical maps where only a perimeter for a disturbance or timber management plan was recorded.

3) Identify stands and quantify their level of departure from potential eco-physical maxima or historical range of variability (Morgan et al. 1994). This requires identifying spatially the potential net primary productivity. A thermodynamic model like the one described in Rasmussen (2012) has already shown promise for just such application (Pelletier et al. 2013).

4) Better quantify the impacts of recent wildfires and insect outbreaks at landscape scale.

5) Monitor and compare the rates of recovery in areas with known stand-replacing events.

As we have shown here, incorporating aerial LiDAR derived evidence from scales once impossible to collect afield [and executing them in a GIS all at one time] allows us ask new questions about patterns and processes that include concepts we could not have addressed before.

References

Allen CD, Betancourt JL, Swetnam TW (1998) Landscape changes in the southwestern United States: techniques, long-term data sets, and trends. Perspectives on the land use history of North America: a context for understanding our changing environment.US Geological Survey, Biological Science Report USGS/BRD/BSR-1998-0003:71-84

Allen CD, Savage M, Falk DA et al (2002) Ecological restoration of southwestern ponderosa pine ecosystems: a broad perspective. Ecol.Appl. 12:1418-1433

Andersen H, McGaughey RJ, Reutebuch SE (2005) Estimating forest canopy fuel parameters using LIDAR data. Remote Sens.Environ. 94:441-449

Anselin L (1995) Local indicators of spatial association—LISA. Geogr.Anal. 27:93-115

Baker WL, Veblen TT (1990) Spruce beetles and fires in the nineteenth-century subalpine forests of western Colorado, USA. Arct.Alp.Res.:65-80

Balmat J, Kupfer J (2004) Assessment of timber resources and logging history of the Valles Caldera National Preserve. University of Arizona Technical Report VCT04011 for Valles Caldera Trust, Tucson, AZ

Bechtold WA, Patterson PL. (2005) The enhanced forest inventory and analysis program: national sampling design and estimation procedures. USDA Forest Service, GTR-SRS-080, 85.

Boots BN (1986) Voronoi (Thiessen) Polygons. Geo books Norwich, UK

Brassel KE, Reif D (1979) A procedure to generate Thiessen polygons. Geogr.Anal. 11:289-303

Chave J, Condit R, Aguilar S et al (2004) Error propagation and scaling for tropical forest biomass estimates. Philosophical Transactions of the Royal Society of London.Series B: Biological Sciences 359:409-420

Cressie N (1993) Statistics for spatial data. Revised Ed. New York: Wiley.

ESRI (2012) ArcGIS Desktop: Release 10.1. Redlands, CA: Environmental Systems Research Institute.

Falk DA, Heyerdahl EK, Brown PM et al (2011) Multi-scale controls of historical forest-fire regimes: new insights from fire-scar networks. Frontiers in Ecology and the Environment 9:446-454

Falk DA (2004) Scaling rules for fire regimes. PhD Dissertation, Department of Ecology & Evolutionary Biology. University of Arizona, Tucson,AZ

Falkowski MJ, Evans JS, Martinuzzi S et al (2009) Characterizing forest succession with lidar data: An evaluation for the Inland Northwest, USA. Remote Sens.Environ. 113:946-956

Farris CA, Baisan CH, Falk DA et al (2010) Spatial and temporal corroboration of a fire-scar-based fire history in a frequently burned ponderosa pine forest. Ecol.Appl. 20:1598-1614

Foster DR, Knight DH, Franklin JF (1998) Landscape patterns and legacies resulting from large, infrequent forest disturbances. Ecosystems 1:497-510

Getis A, Franklin J (1987) Second-order neighborhood analysis of mapped point patterns. Ecology. 473-477

Getis A, Ord JK (1996) Local spatial statistics: an overview. Spatial analysis: modelling in a GIS environment 374

Getis A, Ord JK (1992) The analysis of spatial association by use of distance statistics. Geogr.Anal. 24:189-206

Grissino-Mayer HD, Baisan CH, Swetnam TW (1994) Fire history in the Pinaleño Mountains of southeastern Arizona: effects of human-related disturbances. USDA Forest Service RM-GTR-264, 399-407.

Johnson E, Wittwer D (2008) Aerial detection surveys in the United States. Australian Forestry 71:212-215

Levin SA (1992) The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. Ecology 73:1943-1967

Lynch A.M., Swetnam T.W. 1992. Old growth mixed-conifer and western spruce budworm in the southern Rocky Mountains. *In* Kaufmann, M.R.; Moir, W.H.; Bassett, T.R., eds. Proceedings of the workshop on old-growth forests in the

Southwest and Rocky Mountain Region, 9-13 March 1992, Portal AZ. USDA Forest Serv. Gen. Tech. Rep. RM-213:66-80.

Lynch AM (2004) Fate and characteristics of Picea damaged by Elatobium abietinum (Walker) (Homoptera: Aphididae) in the White Mountains of Arizona. West.N.Am.Nat. 64:7-17

Lynch AM (2009) Spruce Aphid, Elatobium abietinum (Walker) Life History and Damage to Engelmann Spruce in the Pinaleño Mountains, Arizona. In: Sanderson, H.R.; Koprowski, J.L., eds. The last refuge of the Mt. Graham red squirrel: Ecology of endangerment. Tucson, AZ: University of Arizona Press. p. 318-338.

Lynch AM (2012) What tree-ring reconstruction tells us about conifer defoliator outbreaks. Pp. 125-154 *in* Barbosa P., Letorneau D.K., and Agrawal A.A. eds. Insect Outbreaks – Revisited. Wiley-Blackwell, Chichester, West Sussex, UK.

Margolis EQ, Balmat J (2009) Fire history and fire–climate relationships along a fire regime gradient in the Santa Fe Municipal Watershed, NM, USA. For.Ecol.Manage. 258:2416-2430

Margolis EQ, Swetnam TW, Allen CD (2007) A stand-replacing fire history in upper montane forests of the southern Rocky Mountains. Canadian Journal of Forest Research 37:2227-2241

Margolis E, Swetnam T, Allen C (2011) Historical stand-replacing fire in upper montane forests of the Madrean Sky Islands and Mogollon Plateau, southwestern USA. Fire Ecol 7:88-107

McGaughey R (2012) FUSION/LDV: Software for LIDAR Data Analysis and Visualization, Version 3.01. US Department of Agriculture, Forest Service, Pacific Northwest Research Station, University of Washington.Available online at: http://forsys.cfr.washington.edu/fusion/fusionlatest.html (last accessed 24 August 2012)

Mathworks 2012. MATLAB and Statistics Toolbox Release 2012b, The MathWorks, Inc., Natick, Massachusetts, United States.

Mitchell A (2005) The ESRI guide to GIS analysis, Volume 2: Spatial Measurements and Statistics. Redlands CA.

Mitchell B, Walterman M, Mellin T et al (2012) Mapping vegetation structure in the Pinaleno Mountains using lidar-phase 3: Forest inventory modeling. RSAC-100007-RPT1. Salt Lake City, UT: U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center. 17 p.

Moran PA (1950) Notes on continuous stochastic phenomena. Biometrika 37:17-23

Morgan P, Aplet GH, Haufler JB et al (1994) Historical range of variability: a useful tool for evaluating ecosystem change. J.Sustainable For. 2:87-111

Muldavin E, Tonne P (2003) A vegetation survey and preliminary ecological assessment of Valles Caldera National Preserve, New Mexico. Rep.Coop.Agree. 01CRAG0014

Muldavin E, Tonne P, Jackson C et al (2006) A vegetation map of the Valles Caldera National Preserve, New Mexico. Final report for Cooperative Agreement (01CRAG0014).

Negron JF, Bentz BJ, Fettig CJ et al (2008) US Forest Service bark beetle research in the western United States: Looking toward the future. J.For. 106:325-331

O'Sullivan D, Unwin D (2010) Geographic information systems. John Wiley & Sons. 439 p.

O'Connor C, Falk DA, Lynch AM, Wilcox CP, Swetnam TW, Swetnam TL (2010) Growth and demography of Pinaleño high elevation forests. RJVA 07-JV-11221615317 Performance Report. Tucson, AZ: University of Arizona, Laboratory of Tree-Ring Research; School of Natural Resources and the Environment. 21 p.

O'Connor CD, Falk DA, Lynch AM, Swetnam TW, Wilcox CP (in prep) Site productivity mediates stability of species assemblages following fire exclusion. To be submitted: Ecology.

Ord JK, Getis A (2001) Testing for local spatial autocorrelation in the presence of global autocorrelation. J.Reg.Sci. 41:411-432

Ord JK, Getis A (1995) Local spatial autocorrelation statistics: distributional issues and an application. Geogr.Anal. 27:286-306

Pelletier JD, Barron-Gafford GA, Breshears DD et al (2013) Coevolution of nonlinear trends in vegetation, soils, and topography with elevation and slope aspect: A case study in the sky islands of southern Arizona. Journal of Geophysical Research: Earth Surface (2013): 1-18.

Persson A, Holmgren J, Söderman U (2002) Detecting and measuring individual trees using an airborne laser scanner. Photogramm.Eng.Remote Sensing 68:925-932

Rasmussen C (2012) Thermodynamic constraints on effective energy and mass transfer and catchment function. Hydrology and Earth Sys. Sci., 16(3), 725-739.

Sanderson HR, Koprowski JL. (2009) The Last Refuge of the Mt. Graham Red Squirrel: Ecology of Endangerment. University of Arizona Press.

Stephens SL, Agee JK, Fulé PZ, North MP, Romme WH, Swetnam TW, Turner MG (2013) Managing Forests and Fire in Changing Climates. 4 October 2013 Science 342. 41-42.

Swetnam TW, Baisan C, Grissino-Mayer H (2009) Tree-ring perspectives on fire regimes and forest dynamics in mixed conifer and spruce-fir forests on Mt. Graham. Sanderson HR, Koprowski JL (tech eds) The last refuge of the Mt.Graham red squirrel: ecology of endangerment, University of Arizona Press, Tucson

Swetnam TW, Betancourt JL. (2010) Mesoscale disturbance and ecological response to decadal climatic variability in the American Southwest. In: Tree Rings and Natural Hazards. Springer, pp 329-359

Swetnam TW, Lynch AM (1993) Multicentury, regional-scale patterns of western spruce budworm outbreaks. Ecol.Monogr.: 399-424

Swetnam T, Falk DA, Hessl AE et al. (2011) Reconstructing Landscape Pattern of Historical Fires and Fire Regimes. In: The Landscape Ecology of Fire. Springer, pp 165-192

Tomppo E, Olsson H, Ståhl G, Nilsson M, Hagner O, Katila M (2008) Combining national forest inventory field plots and remote sensing data for forest databases. Remote Sens. Environ. 112(5), 1982-1999.

Touchan R, Allen CD, Swetnam TW (1996) Fire history and climatic patterns in ponderosa pine and mixed-conifer forests of the Jemez Mountains, northern New Mexico. USDA Forest Service RM-GTR-286, 33-46.

USFS (2013) GIS Data for Coronado NF and Santa Fe NF (including the VCNP): http://www.fs.usda.gov/detail/r3/landmanagement/gis/?cid=stelprdb5202474

Ustin SL, Gamon JA (2010) Remote sensing of plant functional types. New Phytol. 186:795-816

Van Leeuwen M, Nieuwenhuis M (2010) Retrieval of forest structural parameters using LiDAR remote sensing. European Journal of Forest Research 129:749-770

Westerling AL, Hidalgo HG, Cayan DR et al (2006) Warming and earlier spring increase western US forest wildfire activity. Science 313:940-943

Whittaker RH, Niering WA (1975) Vegetation of the Santa Catalina Mountains, Arizona. V. Biomass, production, and diversity along the elevation gradient. Ecology: 771-790

Williams AP, Allen CD, Macalady AK et al (2012) Temperature as a potent driver of regional forest drought stress and tree mortality. Nature Climate Change 3, 292–297 (2013).

Zhang S, Zhang K (2007) Comparison between general Moran's index and getis-ord general G of spatial autocorrelation. Acta Scientiarum Naturalium Universitatis Sunyatseni 4:022

Tables

Table 1: Aerial LiDAR flight parameters for the two study areas as reported by the vendors and produced in USFS FUSION (McGaughey 2012) 'Catalog' Quality Control and Quality Assessment module. Density of the data is reported in pusles per square meter (ppsm).

	Pinaleño	Jemez	
Vendor (Area)	Watershed Sciences	NCALM (VCNP)	Watershed Sciences (SW Jemez)
Acquisition Dates leaf on | off	September 22-27, 2008; leaf-on	January 2010 leaf-off July 2010 leaf-on	June 18-19,22-27,29-30, July 1 2012; leaf-on
LiDAR Scanner	Leica ALS50 Phase 2	Optech Gemini	Leica ALS60
Pulse Rate	70-90kHz	100kHZ	105.9kHz
Scan Rate	52.2Hz	<100Hz	<100Hz
Returns per pulse Scan Angle	1-4, Intensity ±15°	1-4, Intensity ±25°	1-4, Intensity ±13°
Accuracy Vert. | Horiz.	3.2 cm | 1.0 m	7.0 cm | 1.0 m	2.8 cm | 1.0 m
Flight level	800-1,300 m	~1,000 m	650 - 1,100 m
μ pulse-return spacing	7.36 ppsm	January: 8.86 ppsm June: 5.91 ppsm	13.35 ppsm
μ Bare ground spacing	0.98 ppsm	1.11 ppsm	2.58 ppsm
Acquisition Area	34,608 Ha	January: 29,400 Ha June: 75,600 Ha	52,588 Ha
$\sum$ **Point returns**	2,892,925,979	January: 2,541,885,987 June: 7,754,915,628	7,020,498,000
Projection, Datum	UTM Zone 12, WGS84, NAD83	UTM Zone 13, WGS84, NAD83	UTM Zone 13, WGS84, NAD83
Units	Meters	"	"
Classification	Default 1, Ground 2, Withheld 11	"	"
Flight line overlap	+50 % side-lap	"	"

Table 2: Average $\bar{z}$-scores and the comparative types of stand conditions for both Getis-Ord G_i^* and Anselin Moran's I_i. In the third row the first letter is a target tree, the second is the type of neighborhood it is surrounded by, e.g. H-L is a tall tree with short neighbors, L-L is a short tree with short neighbors, NS is not significant (the neighborhood trees are heterogeneous in structure), etc.

Location / Structure	$\bar{z} - scores$		
	G_i^*	I_i	H = high/tall, L = low/short, NS = not significant
Pinaleño			
1950's cut	-6.2	3.5 | 0.25 | -2.5	L-L | NS | H-L
Old-growth adjacent patch cut	8.5	-3.3 | 1.2 | 8.2	L-H | NS | H-H
N-facing 1960's logged patch cut	-11.2	30.5	L-L
S-facing 1960's logged patch cut	-0.7	0.4	NS
High-severity fire	-14.2	-3.9 | -1.3 | 6.3	H-L | NS | L-H
Monument Canyon			
Even-aged young PIPO	-3.2	6.6	L-L
Even-aged mature PIPO	5.5	11.5	H-H
Primeval mature (+300 years)	4.3	-0.5 | 15.3	NS, H-H
Open-structure mature (120 stems/ac)	2.1	-5.2 | 0.0 | 6.6	L-H | NS | H-H
Closed-structure mid (800 stems/ac)	-1.5	1.5 | 2.8	NS | L-L

Figure captions

Figure 1: The locations of the Pinaleño and Jemez Mountains in the southwestern US (left panel) and their recently recorded disturbance histories within the aerial LiDAR coverage areas (gray hillshade). Logging history data are from USFS (2013) and Balmat and Kupfer 2004. Fire history perimeters are from USFS (2013).

Figure 2: The Monument Canyon (upper left) with VLM local maxima over a small (~1 ha) area (upper right). The reddest blobs are the tallest canopy trees; black boundary circles are based on the allometric relationship of canopy diameter to height for each tree, and blue is the surface. The results of the Getis-Ord G_i^* (bottom left) and Anselin Moran's I_i (bottom right) are shown for the same sampled window. Positive z-scores indicate assemblies of large neighborhoods for Getis-Ord G_i^* and large trees within large neighborhoods for Anselin Moran's I_i; negative z-scores indicate small neighborhood trees.

Figure 3: Modern logging and fire history in the Pinaleño Mountains. Timber sales are shown in yellow and orange, recent fire perimeters in red. The centrally located logging polygon from the 1964-1970 period (lower two panels) is visible in both Getis-Ord G_i^* (bottom left) and Anselin Moran's I_i (bottom right), where loggers harvested trees in ~4 ha patch cuts along the logging road.

Figure 4: Monument Canyon (top) with the Getis-Ord G_i^* (bottom left) and Anselin Moran's I_i (bottom right) LISA. The areas colored in red in both of the bottom panels represent stands of the largest trees [positive z-score or high-high valued, respectively]. Tall mature stands are predominantly located in the canyon bottoms between the two mesas bisecting the study area. There is evidence of recent reproduction along the tops of the mesas (blue for Getis-Ord G_i^* or green for Anselin Moran's I_i)

Figure 5: The four panels from top left to bottom right are (1) Mean canopy height model at 0.333 m^2, (2) Thiessen polygon weighted by height, (3) Hot Spot and Cluster Analysis Getis-Ord G_i^* for the Thiessen polygons, and (4) three break lines set at z = <-1.65, <1.65, and >1.65.

Figures

Legend
USA Federal Lands
- Forest Service
- Department of Defense
- Bureau of Land Management
- Fish and Wildlife Service
- National Park Service

Legend
- Logging
- Fire Perimeters
- Example Areas

Jemez Mountains

Monument Canyon RNA

Pinaleño Mountains

Mt Graham Logging + Fire

Figure 1.

Meters

High : 48   Low : 0

CR
O Meters

Hot Spot Analysis
GiZScore

- ● < -2.58 Std. Dev.
- ● -2.58 - -1.96 Std. Dev.
- ● -1.96 - -1.65 Std. Dev.
- ● -1.65 - 1.65 Std. Dev
- ● 1.65 - 1.96 Std. Dev.
- ● 1.96 - 2.58 Std. Dev.
- ● > 2.58 Std. Dev.

Anselin Moran's I

- ● Not Significant
- ● HH
- ● HL
- ● LH
- ● LL

0 10 20   40   60   80   100
Meters

Figure 2.

Legend
Logging
1964-1970
1946-1964
Fire Perimeters

Meters
High : 48   Low : 0

Hot Spot Analysis
GiZScore
< -2.58 Std. Dev.
-2.58 - -1.96 Std. Dev.
-1.96 - -1.65 Std. Dev.
-1.65 - 1.65 Std. Dev
1.65 - 1.96 Std. Dev.
1.96 - 2.58 Std. Dev.
> 2.58 Std. Dev.

Anselin Moran's I
Not Significant
HH
HL
LH
LL

0   0.25   0.5   1   1.5   2
Kilometers

Figure 3.

Meters

High : 48   Low : 0

CR
○ Meters

Hot Spot Analysis
GiZScore
- ● < -2.58 Std. Dev.
- ● -2.58 - -1.96 Std. Dev.
- ● -1.96 - -1.65 Std. Dev.
- ● -1.65 - 1.65 Std. Dev
- ● 1.65 - 1.96 Std. Dev.
- ● 1.96 - 2.58 Std. Dev.
- ● > 2.58 Std. Dev.

Anselin Moran's I
- ● Not Significant
- ● HH
- ● HL
- ● LH

0   0.5   1   1.5   2
Kilometers

Figure 4.

Figure 5.

APPENDIX E: BIOMASS AND CARBON MODELS FOR AERIAL LASER SWATH MAPPING INVENTORIES IN SOUTHWESTERN NORTH AMERICAN SEMI-ARID CONIFER FORESTS

Tyson L. Swetnam[1,7], Donald A. Falk[1,2] , Brian J. Enquist[3], Ann M. Lynch[2,4], and Stephen R. Yool[5]

[1] School of Natural Resources and Environment, The University of Arizona Tucson, AZ 85721, USA

[2]Laboratory of Tree-Ring Research, The University of Arizona Tucson, AZ 85721, USA

[3]Department of Ecology and Evolutionary Biology, The University of Arizona Tucson, AZ 85721, USA

[4]United States Forest Service, Rocky Mountain Research Station

[5]School of Geography and Development, The University of Arizona Tucson, AZ 85721, USA

[7]Corresponding Author: email: tswetnam@email.arizona.edu, phone: (520) 247-2293

Abstract

Models of above ground biomass (AGB) and above ground carbon (AGC) using individual tree height and aerial LiDAR mean canopy height (MCH) profiles as independent variables are created with inventory data from three forests in Arizona and New Mexico by species and mixed-species forest type [by dominant species plant functional types (PFT): Ponderosa pine, White-fir dominated, Mixed-Conifer, Spruce-fir, and seral Aspen forest]. The PFTs are taxonomically and climatically similar to other forests across the Southwestern USA, making these models potentially generalizable to a larger geographic area. The general model results of MCH of 500 m^2 and 1,000 m^2 circular plots estimated AGB to within ± 18% of the observed values at the 95% confidence interval. The accuracy of the MCH profile PFT models was higher ($r^2 = 0.770 - 0.840$) than the general model ($r^2 = 0.630$). Four of the five individual tree PFT models exhibited the same or higher accuracy ($r^2 = 0.685 - 0.760$) [except Aspen ($r^2 = 0.492$)] than the general model ($r^2 = 0.685$). Estimation of AGB at the individual tree scale was based on the Metabolic Scaling Theory (MST). Typically the independent variable for estimating AGB is bole radius: $AGB = \beta r^\alpha$ where β is a normalization constant and α is the scale parameter; here found by least squares regression to be $\alpha = 2.47 \pm 0.02$. This α was indistinguishable from published tropical tree allometry reported elsewhere to be $\alpha = 2.5$ and is a rejection of the MST prediction of $\alpha = 2.66\overline{6}$. However, AGB models generated by individual species found α varied between 2.22 ± 0.05 and 2.65 ± 0.04. Further, when health was considered scaling of individual bole radius to height $HT = \beta r^\alpha$, predicted by MST to be $\alpha = 0.66\overline{6}$, found the unhealthiest trees had $\alpha = 0.599 \pm 0.030$ and the healthiest $\alpha = 0.675 \pm 0.015$. Plot observed AGB totals are similar to other reports of forest biomass in the Southwestern USA, however some plots far exceeded what is predicted using conventional models, for example, in mature (+200 year old) mixed-conifer stands left undisturbed over the last century a maximum of $1,495\ Mg\ ha^{-1}$ was observed this far exceeds a model 99th percentile estimate ($\sim 640\ Mg\ ha^{-1}$).

Keywords: semi-arid, forest, carbon, biomass, allometry, LiDAR, metabolic scaling theory

Introduction

Remote sensing lets biometricians evaluate forest structure between inventory plots by correlating observations made in the field to numerical similarities across a range of spectral values. Spectral indices of vegetation (Kreigler et al. 1969, Huete 2002, Baccini et al. 2004) have had profound implications for understanding the carbon cycle, especially in tropical forests (Huete 2002); but have been limited in their ability to accurately characterize above ground biomass (AGB) and above ground carbon (AGC) locally due to scale limitations of the data [satellite data are between 30 meters (m) and 500 m spatial resolution]. To date, the greatest advancement in remote sensing of forest biomass has come from active sensor systems such as aerial LiDAR, and interferometric synthetic aperture radar (InSAR)(Anderson et al. 2003, Hyde et al. 2006). LiDAR and InSAR produce three dimensional observations which allow biometricians to measure trees and canopy structure from the individual to ecosystem scale (Leeuwen and Neuwenhuis 2009, Ustin and Gamon 2010).

Techniques for estimating the AGB and AGC with aerial LiDAR are available for tropical forests (Asner et al. 2011a, 2011b, 2012, Mascaro et al. 2011), and temperate forests (Reutebuch et al. 2003, 2005, Gatziolis and Anderson 2008, Erdody and Moskal 2010, Richardson and Moskal 2011, Falkowski et al. 2009, 2010; Hudak et al. 2008, 2012, Zhao et al. 2007, 2009). Most of those publications report AGB and AGC at a grid-scale, where a metric of the forest canopy height profile (Lefsky et al. 1999) is used to model the correlation of AGB measured in field plots. Importantly, as aerial LiDAR technology has advanced and specific computer software for analyzing the data written, the ability to discriminate and segment individual trees has also been demonstrated (Falkowski et al. 2006, 2008, Hudak et al. 2009, Yao et al. 2012). The next logical step for these individual tree-scale inventories is to apply a species specific model [rather than general model] to increase precision. Until very recently species identification has been difficult or impossible to achieve from the dearth of multi-spectral and hyper-spectral sensors capable of discriminating individual species at the same scale as the aerial LiDAR. Research into the fusion of hyper-spectral imagery and aerial LiDAR has shown

positive results at discrimination of individual tree species in other forest systems (Holmgren et al. 2008, Ke et al. 2008, Jones et al. 2010, Naidoo et al. 2012, Zhang and Qiu 2012, Bright et al. 2012, Lu et al. 2012). New platforms currently being developed, such as the National Ecological Observatory Network's Airborne Observation Platform (http://www.neoninc.org/science/aop) should allow for both individual tree and species segmentation at least at the overstory tree scale.

In the present study, aerial LiDAR data derivatives (Figure 1) are used to create models of AGB and AGC of ten common tree species and five common forest types [plant functional type (PFT) (Smith et al. 1993)] (Figures 2 and 3) in semi-arid conifer dominated forests from Arizona and New Mexico. These models are intended to inform the results of future aerial LiDAR inventory projects in the Southwestern USA. Field measures of AGB are derived from: (1) the primary size measures of individual trees observed in the field [e.g. tree height equivalent to bole length, and diameter at breast height (DBH)], and (2) the summation of the biomass distribution for all measured trees in each plot. A total of 22 AGB models are reported: two general models, ten common PFTs [five for mean canopy height profile data and five for individual tree heights], and ten species models for individual tree height (Figures 2 and 3).

The intent of this manuscript is to give the end user two options: (1) the ability to characterize Southwestern US forest AGB and AGC from aerial LiDAR using a grid-based method, and (2) the ability to characterize individual tree AGB and AGC from aerial LiDAR after applying an individual tree segmentation method (Hyyppä et al. 2005, Falkowski et al. 2006, 2008, Swetnam and Falk *in review*). The first technique has already been replicated for forests around the globe (Asner et al. 2011b, d'Oliveira et a. 2012, Hudak et al. 2012). We too use the same method, developing local-scale normalization constants and scaling parameters that correlate more closely to local forest allometry than the global general model (Asner et al. 2011b). The second technique we report on anticipates future methods in development for species-level identification in Southwestern USA forests using multi-spectral or hyper-spectral image fusion with aerial LiDAR. We provide baseline allometric measures of ten common Southwestern USA

species. These results may be useful to scientists working in the Southwestern USA who do not have the resources to also conduct their own rigorous field inventory. All the inventory plots, and in some cases the trees within the plots, have been geo-referenced with differential GPS or surveyed by back-azimuth to allow future repeat measurements by field technicians or aerial platform.

Methods

First, we describe where we conducted our study and what the common forest types and species are. Second, we show how Metabolic Scaling Theory (MST) (Brown et al. 2004, Enquist et al. 1999, Enquist and Niklas 2002, West et al. 1999, Savage et al. 2010, Bentley et al. 2013) and a general quantitative theory of forest structure (West et al. 2009, Enquist et al. 2009), predict tree-scale cubic volume using primary size measures; we also incorporate published estimates of wood density (Jenkins et al. 2004, Miles and Smith 2009) and carbon content (Lamlom and Savidge 2004) to determine the AGB and AGC. Third, we describe a grid-scale approach to estimating the AGB with mean canopy height (MCH) profiles (Lefsky et al. 1999) derived from aerial LiDAR metrics (McGaughey 2012). Last, we describe the model creation and validation process. The observed plot data with related aerial LiDAR height metric information are available in the Supplemental Materials.

Study area geology and climate

Two of the three study areas: the Pinaleño Mountains the Santa Catalina Mountains are located in the Coronado National Forest (N.F.) Arizona; the third area is the Jemez Mountains New Mexico and includes the Santa Fe N.F. and Valles Caldera National Preserve (Figure 4) (Table 1).

These study locations are currently monitored by the National Science Foundation's Santa Catalina Mountain – Jemez River Basin (SCM-JRB) Critical Zone Observatory (CZO) (www.criticalzone.org). The Jemez Mountains, New Mexico (Figure 4) are also one of the US Department of Agriculture's Collaborative Forest Landscape

Restoration Program sites; a ten year project aimed at restoring natural fire and improving the health of the Southwest Jemez portion of the Santa Fe N.F.

The Santa Catalina Mountains and Pinaleño Mountains, Arizona (Figure 4) have been studied intensively and extensively for over one hundred years, historical observations of the flora and fauna in the Santa Catalina and Pinaleño led the famous ecologists C. Hart Merriam and Forrest Shreve toward the theorization of 'gradient ecology' (Merriam and Stejneger1890, Shreve 1915) and later Robert Whittaker and William Niering to a theory linking temperature, precipitation, and solar insolation to net primary productivity, species distributions, and biodiversity analyses (Whittaker et al. 1974, Whittaker and Niering 1964, 1965, 1968a, 1968b, 1975, Niering and Lowe 1984) (Supplementary Materials). More recently, theories on the pedologic and geomorphic evolution of arid landscapes were developed in these same Sky Islands that account for the effect a forest has on mountain scale weathering (Rasmussen 2011, Pelletier et al. 2013).

Both the Santa Catalina and Pinaleño are characteristic of basin and range topography, exhibiting complexes of steeply sided canyons at lower elevations with relatively gentle high elevation uplands above 2,700 m above mean sea level (amsl). The Santa Catalina are located north of Tucson, Arizona, at 32.4° N, 110.7° W (Figure 4) (Table 1). The geology of their south slopes consists of a single granitic-gneiss; the north slopes are a complex of granite, diorite, schist, andesite, shale and slate, quartzite, and limestone (Du Bois 1959, Whittaker and Niering 1968). The Santa Catalina rise from a basin 900 m amsl to a maximum elevation of 2,791m amsl atop Mount Lemmon. The Pinaleño are located southwest of Safford, Arizona, at 32.7° N, 109.9° W (Figure 4). The geology of the Pinaleño consists of mylonitic gneiss, granite, and diorite on the north east side and un-deformed granite and granodiorite on the south side (Martin and Fletcher 1943). The Pinaleño are taller than the Santa Catalina, rising over 2,100 m from their base to reach 3,267 m amsl atop Mount Graham (Table 1). The Pinaleño exhibit a large gentle rolling upland above 2,700 m, covered by a mixed-conifer and spruce-fir forest.

Climatically the Santa Catalina and Pinaleño are nearly identical, being arid to semi-arid with two rainy seasons: a winter season (December-March) and summer monsoon (July-September). The average annual precipitation for the Santa Catalina and Pinaleño varies between $377 \frac{mm}{yr^{-1}}$ at the Sabino Canyon weather station (805m amsl) [data from the Western Regional Climate Center (WRCC], data averaged between 1981-2010); up to $850 \frac{mm}{yr^{-1}}$ at the Mount Lemmon station (2,374m). The average annual precipitation for the Pinaleño varies between $325 \frac{mm}{yr^{-1}}$ at the Fort Grant station (1,472m amsl) (WRCC) to $850 \frac{mm}{yr^{-1}}$ at the Columbine RAWS station (2902m amsl) (WRCC) (Table 1). Precipitation exhibits an approximate linear trend with elevation, a function of the dry-adiabatic lapse rates and orographic lift of mountain topography (Brown-Mitic et al. 2007, Whittaker and Niering 1975). The average atmospheric temperature lapse rate with elevation at the surface is $\sim 5.73 \pm 1.06° \frac{K}{km^{-1}}$ and is $\sim 7.625 \pm 1.51° \frac{K}{km^{-1}}$ for free air (Harlow et al. 2004). Temperatures are lowest in January-February (average min. at Sabino Canyon: $2.8°C$, Mount Lemmon: $-3.5°C$), the record low for Mount Lemmon is $-18.9°C$; the highest temperatures are in June-July (average high at Sabino Canyon: $39.8°C$, Mount Lemmon: $24°C$), the record high for Sabino Canyon is $46.1°C$.

The Valles Caldera National Preserve is located in the Jemez Mountains west of Santa Fe, New Mexico at 35.8° N, 106.5° W (Figure 4). The Preserve was created by the US Congress in the year 2000 and is managed for grazing, hunting, recreation, conservation, and research. The mountains are part of a collapsed volcanic caldera with a rim approximately 19km wide; within the caldera are resurgent domes over 200m high. Elevations range from 2,300m amsl in Redondo Meadow, to 3,431 m amsl atop Redondo Peak (Table 1). The Valles Caldera's geology consists mainly of Rhyolite, Ignimbrite, Tuff, and Pumice (Izett 1981).

The Valles Caldera is colder and dryer across the same range of elevations as the two Arizona sites, and is considered to be a temperate montane climate. The area receives

between 476 $\frac{mm}{yr^{-1}}$ at the Los Alamos station (2,234 m amsl)[WRCC 1981-2010) to 790 $\frac{mm}{yr^{-1}}$ annual precipitation at the Quemazon SnoTel site (2,895 m amsl)] (Table 1). Approximately two thirds of the precipitation falls as snow (October-April) and one third falls during the summer monsoon (July-September) (Liu et al. 2008, Molotch et al. 2009). Temperatures rarely exceed 27°C in summer, or fall below -15°C in winter; the average annual temperature is 9°C.

The Valles Caldera underwent widespread land use changes over the last eighty years. Extensive clear-cut logging took place between 1930 and 1980, and continued at a slower pace until the Preserve was established in 2000. Most of the Valles Caldera's extant forests are second growth or have been partially logged and contain some mid-development aged trees along with recent regrowth.

The Pinaleño and Santa Catalina both burned in landscape-scale high severity fires in the last decade: the Bullock Fire (2002) and Aspen Fire (2003) in the Santa Catalina, and the Clark Peak Fire (1996) and Nuttall-Gibson Fire (2004) in the Pinaleño burned large portions of the upper elevation mixed-conifer forests. In 2011the Los Conchas Fire burned the Jemez across portions of the Santa Fe National Forest, Valles Caldera National Preserve, Bandelier National Monument, and Department of Energy land in, in 2013 the Thompson Ridge Fire burned the western portion of the Valles Caldera.

Monitoring plots

Common observations made at all of the study locations included: tree species, condition, DBH at 4.5 feet or 1.37m above ground level, and maximum canopy height; additional observations made in the Valles Caldera and Santa Catalina included canopy base height, and canopy radius. DBH was measured with a metal forest diameter tape. Tree height was measured with a Laser Technologies Impulse 200 hypsometer (Pinaleño plots) and Nikon Forester 550 hypsometer (Santa Catalina and Valles Caldera). Plot center locations were determined with a sub-meter level GPS unit (Trimble GeoXH with

a Zephyr antenna and two-meter range pole) in NAD83 and NAVD88 coordinate systems. A differential correction or Trimble H-Star correction (Trimble Navigation Limited, 2005) increased plot location accuracy. GPS differential corrections were applied using H-Star (Trimble Navigation Limited, 2005) "H-Star Technology Explained". Horizontal root mean square error of $\sim 0.3\ m$ from plot center was acquired for each plot center. Geo-referencing of plot locations were completed either coincidental to vegetation sampling or at a later time. Basal areas are reported in square meters per hectare ($m^2 ha^{-1}$), and the AGB and AGC in metric tons per hectare ($Mg\ ha^{-1}$). Other AGB estimates of the common PFTs from historical and recent studies conducted in Arizona and New Mexico are reported in Table 4.

In the summers of 2008 and 2009 technicians measured 79 spatially arrayed 0.05 ha (500 m^2) radial ($r = 12.62\ m$) field plots in the Pinaleño for a dendrochronological study and aerial LiDAR project (O'Connor et al. 2010, Laes et al. 2008, Mitchell et al. 2012). The methods for establishing systematic gridded plot transects for dendrochronological studies around the western US are described in further detail in Brown et al. (2008a, 2008b) and Heyerdahl et al. (2011). Permanent staked, tree tagged, geo-referenced plots were located along a systematic Universal Transverse Mercator (UTM) 1km grid extending across the entire mountain range above 2,300 m ASL (Figure 4). For all project locations field crews navigated to the plots by handheld global positioning system (GPS) units. A total of 2,862 trees were measured. Diameter-to-height regressions of 2,173 healthy standing trees from all locations were made after stumps, logs, and damaged trees were removed from the analysis to account for physical damage that may have altered allometric scaling relationships.

The Santa Catalina plots were collected in the summers of 2010 and spring of 2012 along a 1km UTM grid, similar to the Pinaleño, however at the time of the grid creation the GPS datum for the plot locations were set to NAD1927 to correspond to existing USFS surface maps, rather than to NAD83 and later re-projected. Plot radii for the Santa Catalina and Valles Caldera were established at 0.1 ha (1000 m^2) radial plots

($r = 17.83$ m) with nested 0.01ha ($r = 5.62m$) inner plots. Field observations were identical to the Pinaleño project but included the measurement of canopy radius. Twenty four plots were collected across an elevation gradient from low elevation Sonoran desert, through the Madrean oak savannah-woodland, up to the mixed conifer forest. Of the 24 plots collected in the Santa Catalina only thirteen were in conifer forest and used in this analysis; a total of 226 trees were measured.

In summer 2010 technicians collected 42 0.1 ha radial plots, identical to those in the Santa Catalina, in the Valles Caldera. Locations were selected from a random distribution of points within 200 m of existing roads to enable safe access in the complex terrain (Figure 4), plots were permanent staked, tree tagged, and geo-referenced. On the Valles Caldera there are over one hundred kilometers of maintained dirt roads and several thousand kilometers of decommissioned logging roads. Of the 42 plots collected, 33 plots were geo-referenced with a high precision GPS (±0.3m), the other nine plot locations were estimated in the lab using the stem-map of observed trees with back azimuth estimates of plot center. In total 1,813 trees were measured of which 1,313 were alive and undamaged; 13,999 were counted (including seedlings and saplings) across the 42 Valles Caldera plots. All trees were tagged and numbered for future visitation and continued monitoring. After stumps, logs, and damaged trees were removed a total of 1,313 undamaged trees were applied to the diameter/height regression. As with the Pinaleño study site all standing dead and damaged trees are included for the aerial LiDAR validation.

We present all of the plot identification number, UTM locations, observations of basal area, live and dead AGB, CloudMetric MCH and max height, GridMetric MCH, percent canopy cover, PFT, and effective energy to mass transfer model value (Rasmussen et al. 2011) in the Supplemental Materials.

Plant Functional Types

The multi-species community forests in the study areas have been variously classified as: PFTs (Smith et al. 1993), existing vegetation types (EVT) (Comer et al.

2003, Brohman and Bryant 2005, NatureServe 2005, Muldavin et al. 2006, LANDFIRE 2010), potential natural vegetation types (PNVT) (Nature Conservancy 2006, 2007), plant associations (Stuever and Hayden 1997), mid-scale dominance types (Mellin et al. 2008), and 'Regional Gap analysis' (Lowry et al. 2007, Prior-Magee et al. 2007) (Table 2). Variously these can be generalized and cross-walked by a combination of the most common species present.

We used five general PFT for the biomass models:

1) **Ponderosa pine forest**: Sites are between 2,100 and 2,750 m amsl. Dominated by *Pinus ponderosa* (>80% by volume), other species typically make up less than 20% of total biomass and include: Southwestern white pine (variants in the literature are called *P. strobiformis, P. reflexa,* or *P. flexilis*, here we do not differentiate between them, though there are minor taxonomical differences), Gambel's oak (*Quercus gambelli*) uncommon in Santa Catalina and Pinaleño but common in Valles Caldera, New Mexico locust (*Robinia neomexicana*), and Silver-leaf oak (*Quercus hypoleucoides*) - present in Santa Catalina and Pinaleño absent in Valles Caldera. *P. ponderosa* makes up the greatest proportion of AGB at the mid-range of elevations, with *P. strobiformis,* and *P. menziesii* increasing in proportion (>30%) as elevation increases; along this PFT's lower elevation boundary additional species of Madrean pine and oak co-occur in the Santa Catalina and Pinaleño that include Chihuahua pine (*P. chihuahuana*), Mexican Pinyon (*Pinus cembroides*), *Q. hypoleucoides,* Arizona white oak (*Q. arizonica*), Alligator juniper (*Juniperus deppanea*), and point leaf manzanita (*Arctostaphylus pungens*) (Whittaker and Niering 1975). In the Valles Caldera the lowest elevation is at the mid-elevation range of the *P. ponderosa* forest PFT; also in the Valles Caldera large stands of *Q. gambelli* exist along southerly aspects, in what are believed to be seral post-fire stands (Muldavin and Tonne 2003).

2) **Mixed-conifer forest**: Sites are between 2,400-2,750 amsl. Douglas-fir (*P. menziesii*), white fir (*Abies concolor*), sub-alpine fir (*A. lasiocarpa*) dominate this

type though pine species may also be present. *P. menziesii* is dominant and makes up >70% of volume; *A. concolor*, *P. ponderosa, P. strobiformis, P. engelmannii, A. lasiocarpa, Q. gambelii,* and *P. tremuloides* may all co-occur in this PFT. These sites are generally more mesic than pine forest PFTs, and appear first along north aspects at lower elevation. At upper elevations the northerly aspects of the mixed-conifer PFT become shared by spruce and fir, while pine and oak are more common on the southerly aspects with *R. neomexicana* at lower elevation.

3) **White fir forest**: Sites are between 2,200-2,750 amsl. *A. concolor* dominates this PFT (>70%) in ravines and steeply sided drainages in the Santa Catalina and Pinaleño, and along streams in the Valles Caldera. *P. menziesii* and *P. flexilis* trees are a significant component (<30%), notably in the Santa Catalina and Pinaleño this PFT is where the largest *P. menziesii* trees were observed, suggesting this PFT is the most productive forest type in the study areas. *A. concolor* appears to have been limited historically due to low severity fires, of which it is a poor survivor and easily killed by a single low severity fire; in the modern period it has increased in number and area across the *P. ponderosa* forest type. Several deciduous tree species are also present in the Santa Catalina White-fir PFT: Big tooth maple (*Acer grandidentatum*), Douglas maple (*Acer glabrum*), and *R. neomexicana*, but make up a small proportion (<5%) of the total AGB. This PFT has increased its distribution in the modern era; possibly the due to the suppression of the naturally frequent fire regime.

4) **Spruce-fir forest**: The highest elevation PFT, only exhibited in the Pinaleño and Valles Caldera, occurs mainly above 2,900m amsl, and has the coldest climate. Engelmann spruce (*Picea engelmannii*) are dominant above 2800 amsl in the Pinaleño and Valles Caldera. The Santa Catalina have no endemic spruce; but do retain a small stand of corkbark fir (*A. lasiocarpa var. arizonica*). The most common species in this type are *P. engelmannii* and *A. lasiocarpa* in the Pinaleño and Valles Caldera, with some Colorado blue spruce (*Picea pungens*) in the Valles Caldera. In the Pinaleño, recent large wildfires and insect outbreaks have

severely damaged the extant spruce-fir forest PFT, with up to 80% of living trees killed by insects in unburned areas, and 100% mortality in burned areas (O'Connor 2013, Lynch 2009).

5) *Aspen forest*: *Populus tremuloides* occurs between 2,100 and 3350 m. *P. tremuloides* exists in early and late seral communities in all three locations. Aspen forest tends to occur on more mesic condition sites, but can be present at any elevation in the Valles Caldera. Aspen is shade intolerant and needs open conditions to grow and establish (Burns and Honkala 1990).

Species common to all three study areas occur in the Madrean Sky Island Archipelago, Mogollon Plateau, and Southern Rocky Mountains (Merriam and Stejneger 1890, Woolsey 1911, Shreve 1915, Whittaker and Niering 1975, Burns and Honkala 1990, Muldavin and Tonne 2003). The Santa Catalina and Pinaleño are both wholly within the Madrean Sky Island Archipelago; this floristic setting stretches south into Sonora and Chihuahua, Mexico. Species composition and frequency across each PFT for the two Arizona study areas are nearly identical. The Valles Caldera shares a similar species composition and frequency with the Arizona sites, but are more characteristic of the Southern Rockies. Notably, Madrean *Quercus* and *Pinus spp.* are absent in the Valles Caldera. PFT gradients are less pronounced in the Valles Caldera where elevations change more slowly with vertical relief relative to spatial extent. A unique feature of the Valles Caldera are its multiple inverted tree-lines along its 'valles', Spanish for 'valley', where night time air sinks and pools forming temperature inversions that freeze kill seedlings, preventing trees from establishing (Coop and Givnish 2007). Cold air drainages that exhibit inverted tree-line behavior are also present in the Pinaleño, though much less spatially extensive as those in the Valles Caldera, there are no examples of inverted tree lines in the Santa Catalina.

Across all three study areas the PFTs change along elevation gradients. Temperature gradients are mostly related to orographic lifting and atmospheric lapse rates: low elevations are hotter and drier, i.e. 'xeric', and high elevations are wetter and cooler, i.e.

'mesic'. Specific PFTs generally extend lower in elevation along north aspects and riparian drainages that tend to be more mesic than southerly aspects of the same elevation. Relatively warmer sites are dominated by species of *Pinus, Quercus,* and *Pseudotsuga*; while cooler wetter sites are dominated by *Abies* and *Picea* (Burns and Honkala 1990).

Existing vegetation type maps

The Santa Catalina and Pinaleño are part of a landscape-scale fire management assessment on the Coronado N.F., known as 'FireScape' (http://azfirescape.org) (Supplemental Materials). The FireScape vegetation type assessment includes 1) a geo-referenced map of Land Type Associations (LTAs) that are based on similar geology, soil, and landform types; and 2) ten types of Ecological Units (EU) based on plant associations. Because the geology, soil type, elevation, slope, and aspect of the mountains are highly variable across space the EU and associated LTAs can be split up into dozens of specific LTA and EU types or lumped [as we have done] into the five PFTs (Table 2). By convention forest cover percentages (%) are described by the proportions of grass, shrubs, and trees covering a site. Based on the FireScape EUs for the present study we considered open 'desert' any location without a significant grass component. 'Grassland' is typified by no trees and few shrubs. A 'savanna' is typified as having less than 10% cover by tree canopies and some short shrubs. 'Woodland' is typified as having more canopy cover than a savanna; mesquite or oak/juniper/pine systems in the Santa Catalina and Pinaleño are characterized as woodland, though their canopy cover may be greater than 80%. 'Forests' are typified as consisting of conifer trees including pine, fir, and spruce; a forest is generally not broken down into cover % though they can range from dense to open.

The Valles Caldera's existing vegetation type map was developed by Muldavin and Tonne (2003) and Muldavin et al. (2006) from 268 vegetation plot control points, aerial photography and Landsat satellite imagery. Muldavin et al. (2006) describe 20 map units distributed amongst forest, shrubland, grassland, and wetland ecosystems and dry-

and wet- mesic types (Supplemental Materials Appendix B), this classification scheme followed the NatureServe (2005) standard.

Tree Allometry

Trees transport water and nutrients from the soil to the leaves through capillary action of more negative atmospheric water potential versus the soil; this forces liquid water and solute down a gradient [e.g. vertically upward, against gravity] to the location where part of the plant's metabolic cycle [e.g. photosynthesis] takes place. The network of capillary 'pipes' bundle together to form a Lindenmeyer system (Przemyslaw et al. 1990) with fractal-like branching that is both constrained by the physical limitations of the hydraulic network (West et al. 1999, Enquist et al. 1999, Savage et al. 2010) and cavitation of the water column with increasing path length (Ryan and Yoder 1997, Koch et al. 2004, Kempes et al. 2011), and is area preserving in nature (Savage et al. 2010, Bentley et al. 2013). Allometric equations for estimating whole organism mass are based on primary size measures [e.g. height or bole length l, DBH where radius $r = \frac{DBH}{2}$, canopy diameter (CD), and whole tree CV]. We assume a tree's volume is equivalent to a cylinder based on the 'pipe-model' of Shinozaki et al. (1964), updated by Savage et al. (2010) who showed that trees support their phytomass through their network of capillary veins, and Bentley et al. (2013) who systematically measured every branching unit of several species of trees [including species from the Santa Catalina Mountains] and found that branching-levels are in-fact area preserving.

The CV of each individual tree is equivalent to the volume of a cylinder:

$$CV = \pi * r^2 * l \hspace{3cm} \text{Equation 1}$$

where r is bole radius $r = \frac{DBH}{2}$ and bole length l is the vertical measure of maximum height above ground level.

Jenkins et al. (2003, 2004) and Návar-Cháidez (2010) reported AGB using a logistic equation [equivalent to a power law]:

$$Exp(\beta + \alpha \ln x) \equiv \beta x^\alpha \cong AGB \hspace{3cm} \text{Equation 3a}$$

$$\ln\beta + \alpha \ln x \cong \ln AGB \qquad \text{Equation 3b}$$

where x is $> 2.5cm\ DBH$, Exp is an exponential function $e = 2.718282$, ln is the natural log, and β are scalar coefficients. Reported β and α for common species from Jenkins et al. (2004) and Návar-Cháidez (2010) are given in Supplemental Materials. We found the power law β and α by least square regression of our plot data in MATLAB 2012b Curve Fitting Tool (The MathWorks 2012).

Common forestry practices report the basal area (BA) of a tree which is equivalent to its cross sectional area:

$$BA = \frac{\pi}{4}D^2 \equiv \pi r^2 \qquad \text{Equation 4}$$

The total basal area of stand is calculated as the sum of all of individual trees basal area:

$$\sum_{i=m}^{n} BA_i = BA_m + BA_{m+1} + \cdots + BA_{n-1} + BA_n \qquad \text{Equation 5}$$

Similarly, the total AGB and ACD of stand is calculated as the sum of all of individual trees basal area:

$$\sum_{i=m}^{n} AGB_i = AGB_m + AGB_{m+1} + \cdots + AGB_{n-1} + AGB_n \qquad \text{Equation 6}$$

Importantly, wood density and carbon content also vary by species (Table 3) where the cell mass takes up a fraction of the total cubic volume and carbon content makes up a fraction of the cell mass. To estimate biomass M from the cubic volume the equation is:

$$M = CV * SG * MC = CV * WD \qquad \text{Equation 2}$$

where WD is wood density, SG is specific gravity, and MC is moisture content of the wood. Wood volume is made up of void [typically filled with aqueous solution] and the cell walls; here we considered the distribution of cellular wall tissue to be constant. Both wood density and specific gravity vary with moisture content, though we assumed the difference in the observed volume through the expansion and contraction of living and

dead woody cells to be invariant of the measures used. Hardwoods have a higher wood density than softwoods, with oak (*Quercus spp.*) and locust (*Robinia spp.*) between 0.72 and 0.61 (Table 3 & Supplementary Materials) (Jenkins et al. 2004). In the study areas wood density of the most common conifer species are reported to be between 0.35 (*Abies*) and 0.45 (*Pseudotsuga*)(Jenkins et al. 2004, Miles and Smith 2009)(Table 3). Because the present study did not measure live and dead wood moisture content or the shrinkage and expansion of DBH with moisture content we used a single normalization constant for each species specific gravity and wood density when estimating AGB from cubic volume. Only three of the ten species observed had published C density estimates (Lamlom and Savidge 2004); for species without published estimates we assumed the density available for the same genus. Based on the available information, variance was less amongst species of the same genus than amongst genera (Table 3), in general the average and variance in C density was 50% $\pm$ 3% of total biomass.

We estimated the AGB of individual trees with two different peer-reviewed techniques: (1) an allometric power-law model of primary size measures based on Metabolic Scaling Theory (MST) that follows the 'pipe model' (Shinkozaki et al. 1964, West et al. 2009, Enquist et al. 2009, Savage et al. 2010) and uses both observed DBH and height to estimate whole organism *CV* which is multiplied by wood density for each species or genera (Jenkins et al. 2004, Miles and Smith 2009), the AGC was estimated by multiplying AGB by carbon density (50 ±3% of total mass) (Lamlom and Savidge 2004) (Supplemental Materials); (2) we also report the published allometric models that use only DBH [equivalent to basal area] in the independent variable (Jenkins et al. 2004, Smith et al. 2006, 2007, Chave et al. 2005 , Miles and Smith 2009, and Návar-Cháidez 2010). The AGB of the individual trees was estimated by least squares in MATLAB 2012b Curve Fitting Tool (The Mathworks).We fit a power-law model [Eq. 3a] using *CV* or DBH as the independent variable for all trees (Table 3). We repeated this process for each PVT [based on the plot characterization] and by individual tree species.

221

Metabolic scaling

MST demonstrates plant's metabolism and physical scaling are related to thermodynamic energetic exchanges (Gilooly et al. 2001) and volumetric flow constraints of hierarchical branching networks (Enquist et al. 2009, West et al. 2009, Savage et al. 2010). The fractal relationship of folding surface areas and -fircation branching observed in plants and animals can be explained as an 'area-volume hybrid' (Sernetz et al. 1985). Thus vascular organisms exhibit a fractal three quarter power scaling for mass rather than a Euclidian surface-to-volume two thirds scaling. MST (Enquist et al. 1998, 2009, West et al. 2009) predicts tree mass M as the dependent variable and height h or radii r as independent variables:

$$M \propto r^{\frac{8}{3}} \propto h^4.$$
<div align="right">Equation 7</div>

Whereas when radii and height are written as the dependent variables, respectively, their relationship is: $r \propto h^{\frac{3}{2}}$ or $h \propto r^{\frac{2}{3}}$.

A commonly applied power law [Eq. 3] model for estimating biomass takes the form:

$$AGB = \beta x^\alpha$$
<div align="right">Equation 8</div>

where β is a normalization constant, α is the scaling parameter, and x is a primary size measure [typically height or radii] or with remote sensing the MCH profile (Zianis and Mencuccini 2004, Návar-Cháidez 2010, Asner et al. 2011b). The MST predicted scaling of individual mass by height is predicted to be $M = \beta_1 h^4$, and by radius: $M = \beta_2 r^{8/3}$. The observed scaling relationship reported by Chave et al. (2005) suggest the exponent is closer to $M = \beta_2 r^{2.5}$; meanwhile Jenkins et al. (2004) reported for North American conifers a value between $\alpha = 2.3323 - 2.4835$; and for Mexican pines Návar-Cháidez (2010) estimate $\alpha = 2.5498$. Those exponents are all significantly less than the theoretical $\alpha = 8/3$ suggested by MST (Enquist et al. 1998, 2009, West et al. 2009).

Allometric equations that incorporate a taper term typically do so when merchantable timber is the unit of interest. Because the observed variance between the predicted mass vs. radius or height were large the difference of a taper term would be

<div align="right">222</div>

negligible, therefore we have chosen not to include an extra taper parameter in our estimates of AGB and AGC.

Raster data

      The aerial LiDAR data were processed into raster (gridded) data at different spatial resolutions. The statistical distribution of point pulse returns in the height profile has been shown to be the most important predictor of tree biomass (Lefsky et al. 1999, 2002a, 2002b, Drake et al. 2003, Popescu et al. 2004, Asner et al. 2009, 2011b, Hudak et al. 2012). The MCH is calculated as:

$$MCH = \sum_{i=1}^{h} CHP(i) * i$$

<div align="right">Equation 9</div>

where CHP is canopy height profile at height *i* above ground to a maximum height *h* (Lefsky et al. 1999). The MCH for each field plot was calculated in USFS FUSION (McGaughey 2012) for both the CloudMetrics process and GridMetrics process. CloudMetrics was used to estimate the MCH of the differentially correct plots with a defined radii distance (circular polygon). The GridMetrics were run at 30 m (0.09 ha) pixel size. Because the plots are circular and the GridMetrics are square, there is some error associated with the two at the local scale.

      Gatziolis et al. (2010) conducted an exhaustive examination of the error between aerial LiDAR and field observations in closed canopy forests on steep slopes and reported that the absolute values of height or biomass had error $\pm 10\%$ the observed height. Hawbaker et al. (2009) report that low density LiDAR data in a Wisconsin conifer and deciduous forest had model accuracy varied between $65\% - 88\%$ by DBH, basal area, tree height, and biomass. Studies that use one primary size measure to define a second measure are likely to exhibit expanding uncertainties at larger size classes. These errors are likely present in our data as well, although we did not do a rigorous test to determine these errors. For 394 geo-referenced trees that were located in the plots a cross-examination of the aerial LiDAR estimated height and the field measured height suggested a correlation of $r = 0.984$ and error ± 0.81 m.

We replicated techniques presented in Asner et al. (2009, 2011b, 2012) and Mascaro et al (2011) for estimating AGB and AGC at a grid-scale. Those models use the MCH from aerial LiDAR as the independent variable and field measured AGB as the dependent variable. We conducted a least squares regression in MATLAB Curve Fitting Tool (The Mathworks 2012) of the observed AGB in our study plots versus the MCH cloud-metrics of the plot data and the sampled grid-metrics (McGaughey 2012) of the plot locations.

Asner et al. (2011b) reported models of AGC [called above ground carbon density (ACD) in Asner et al. (2011b)] using the same equation as Eq. 8 for the LiDAR MCH: $ACD_2 = \beta MCH^\alpha$ (Asner et al. 2011b); and $ACD_3 = \beta MCH^{\alpha_1} BA^{\alpha_2} WD^{\alpha_3}$ (Asner et al. 2011b) where basal area, and wood density are derived from known coefficients; notably both α_2, α_3 were very close to 1 suggesting they are nearly constants. Basal area was also modeled as a function of the MCH where: $BA = \beta * MCH$ (Asner et al. 2011b).

Note, the models presented in the tables and figures are in units of tons per hectare ($Mg\ ha^{-1}$); these estimates are from the plot data extrapolated from either 1/20th or 1/10th hectare. Also, the value of raster-pixels are in $Mg\ ha^{-1}$. If the user wishes to estimate actual biomass per pixel they must divide the area of the pixel by a hectare and multiply that proportion by the AGB or AGC $Mg\ ha^{-1}$ estimate.

Common Stand Exam Data

Forty-eight USFS common stand exam plots were also incorporated as part of the Santa Catalina analysis (Supplemental Materials Appendix A). The stand exam data make up a significant portion of the upper elevation forest areas in the Santa Catalinas. Forest types in the stand data are classified as: Ponderosa pine, Mixed-Conifer, White fir, and Sub-alpine fir. The stand exam data were collected in 2004 and 2005 and are available from the Coronado National Forest's data website. Because common stand exams do not record plot locations within specified perimeters (the area extent) we analyzed the statistics of the entire stand within the polygon boundary, rather than by individual plots, with ArcGIS 10.1 Zonal statistics (ESRI 2012).

Airborne Laser Swath Mapping

In April 2007 the Pima County Flood Control District acquired (by SANBORN Inc.) a discrete return aerial LiDAR collection over the Sabino watershed of the Santa Catalina in response to a large scale flood event that occurred in the summer of 2006. The Sabino drainage is the largest watershed in the Santa Catalina with its headwaters reaching the peak of the mountains including Mount Lemmon. These data were collected without LiDAR-based vegetation characterization in mind and are outside of USFS recommended vegetation characterization parameters for LiDAR (Gatziolis and Andreson 2008, Laes et al. 2008); the data are still robust enough however for estimating vegetation metrics, though with greater uncertainties than the other two datasets.

In September 2008 the Coronado N.F. acquired discrete return aerial LiDAR for the entire upper-elevation (>2300m amsl) of the Pinaleño (flown by Watershed Sciences, Inc., Portland, OR) (Laes et al. 2009). The LiDAR collection was parameterized with the US Forest Service LiDAR Vegetation standard protocol developed by Gatziolis and Anderson (2008) and Laes et al. (2008) in mind.

In January and June of 2010 discrete return and waveform aerial LiDAR were flown over the Santa Fe N.F. and Valles Caldera National Preserve, by the National Center for Airborne Laser Mapping (NCALM) as part of the Critical Zone Observatory (CZO) National Program. Over the Valles Caldera, Santa Fe N.F., and Bandelier National Monument a total of 75,600 ha were collected (Figure 2, Supplemental Materials Appendix B). The January flight made up a smaller subset of the June collection, resulting in a 'snow-on, leaf-off' condition and a 'snow-off, leaf-on' summer flight. All of the Valles Caldera vegetation survey plots were within the snow-on flight area. Because the recorded winter snow depth did not exceed 2 m except in the highest elevation locations of the Valles Caldera flight area both LAS layers can be used to measure forest crown structure >2 m above ground level and >3 m for the deepest snow areas; this extra dataset pseudo-doubles the sample density of the LiDAR. For this analysis we only use the discrete return data.

Point, polygon, and grid analysis of aerial LiDAR point cloud data

The validation of the AGB models were done using both cloud metrics (plot scale) and grid metrics (landscape scale) to (1) determine the normalization constant and dynamic exponent of the non-linear model by least squares in MATLAB (Mathworks 2012); and (2) determine the variance between plot based cloud metric and pixel based grid metrics. The minimum height above ground was set at 2m above ground level for the GridMetrics and CloudMetrics analyses to remove effects of understory and downed woody fuels. In Figure 1 we show the process tree of how the aerial LiDAR data were analyzed, and the derivative products created.

Cloud Metrics

From each geo-referenced plot center a circular buffer was produced with the ArcGIS Buffer function: 13 Santa Catalina and 48 Valles Caldera plots with a 17.86m radius (0.1ha), the 79 Pinaleño plots with a 12.62m radius (0.05ha). FUSION PolyClipData function was run for each of the circular plots: extracting the local pulse returns, these were next saved as individual LAS files. FUSION ClipData function was run on the individual files to normalize the z-heights of the LAS from elevation values to height above ground with Digital Elevation Models (DEM) derived from the LAS classified bare earth returns. With the plot data extracted and normalized the FUSION CloudMetrics function was run for each plot where the outputs were saved as CSV file formats and merged into .XLS spreadsheets for each study area.

The CloudMetric outputs only a single statistic. Species-level models were not developed for the grid scale because the grid scale aggregates all species that are present together into a single statistic. The dependent variable, AGB, was determined from the observed tree AGB at the plot level. AGB for individual trees was found by Eq. 3 and then summed for all trees in the plot. The independent variable, the aerial LiDAR derived MCH, was derived from CloudMetrics and Gridmetrics. GridMetrics were produced at 10x10 m for all project areas; a separate 25x25 m GridMetrics dataset was also available

for the Pinaleño (Laes et al. 2009). These data are available in ASCII file formats (Supplemental Materials Appendix A).

The CloudMetrics were done at every plot in all three study areas. The outputs are condensed by study location at the plot level and saved into CSV file formats.

Grid Metrics

FUSION GridMetrics process (McGaughey 2012) produces height above ground level statistics included: mean, max, min, standard deviation (SD), variance, covariance, skew, and kurtosis; and measures involving the vertical distribution that included the: % canopy cover, density, count, 1st, 5th, 10-90th, 90th minus 10th, and 95th minus 5th percentiles (Supplemental Materials Appendix B). These data were processed for the vertical profile points 2 m above ground level to avoid skewing values in open stands where most pulses are at surface level. The intensity values for point cloud data were also processed for the same statistics as the height values, though those results are not discussed further here.

The GM statistics were projected into each LiDAR dataset's datum and coordinate format (UTM or State Plane) and saved as ArcGIS ASC files; these were next merged with FUSION's 'MergeRaster' batch command, making all tiles into a single layer. These final outputs were imported into ArcGIS 10.1 (ESRI 2012). The plot locations were projected over the GridMetric layers and a buffer equal to the plot radius was created with the Buffer function in ArcGIS Toolbox. The grid pixel values were extracted for each plot and saved to a table using ArcGIS 'Zonal Statistics to Table' function. These values were output as .DBF files and opened in Microsoft Excel where the data for each plot was evaluated.

Results

This section is organized into three categories: (1) from field data observations of primary size measures [DBH and height] the derivation of above ground biomass, (2) the development of aerial LiDAR raster-based AGB models that use MCH profiles (Figure 2), and (3) the creation of individual tree AGB models for application with aerial LiDAR

derived tree height (Figure 3). Within each of the two model categories are subgroups: a general model, five PFT models, and for the individual tree models a third subgroup for ten species-specific models (Figures 2 and 3). The raster-based models cannot discriminate species because our study forests were not monocultures. In several cases estimated AGB totals are significantly different from other recent studies quantifying individual tree AGB (Brown 1997, Chave et al. 1995, Návar-Cháidez 2010), and modeled estimates of AGB from aerial LiDAR (Asner et al. 2011b, Mascaro et al. 2011).

Above Ground Biomass

The frequency of species sampled in each study area varied widely; only five species (*A. concolor, P. ponderosa, P. strobiformis, P. menzeisii, and Q. gambelli*) were sampled in all of the study areas, although three more are known to exist in all three areas; *A. lasioscarpa, P. tremuloides,* and *R. neomexicana* were not sampled in the Santa Catalina (Table 3). The assumed average wood density was between 0.404 and 0.489 and the average carbon density 50.55 – 51.09 % based on the distribution of species (Table 3). Estimated wood densities by species were found in Jenkins et al. (2004), Smith et al. (2006), Brown (2008), and Miles and Smith (2009); estimated carbon density were found in Lamlom and Savidge (2004).

From the observed study areas the average AGB at the plot level [Eq. 3] in the Valles Caldera and Santa Catalina were similar on average (219.5 ±140.3 $Mg\ ha^{-1}$ and 226.66 ±125.84 $Mg\ ha^{-1}$, respectively) while the Pinaleño had nearly double the AGB on average (384.4 ±284.5 $Mg\ ha^{-1}$)(Table 5). Average basal area was highest in the Pinaleño (49.5 ±26.5 $m^2\ ha^{-1}$) and lowest in the Valles Caldera (32.7 ±15.3 $m^2\ ha^{-1}$) (Table 5).

By PFT the White-fir dominated forest type had the highest AGB (485.9 ± 288.0 $Mg\ ha^{-1}$) and Ponderosa pine the lowest (173.6 ± 95.2 $Mg\ ha^{-1}$) (Table 6). Importantly, these values are averaged and encompass all growing conditions, cover percentages, and aged stands. At an individual plot location the highest observed AGB was a White-fir dominated stand in the Pinaleño which was measured to have 1,495

$Mg\ ha^{-1}$ of live biomass (Supplemental Materials Appendix A). In the Valles Caldera the greatest observed plot level AGB was 495 $Mg\ ha^{-1}$ and in the Santa Catalina 483 $Mg\ ha^{-1}$, both in Mixed-conifer plots. These observations are in good agreement with historical studies from the same areas (Whittaker and Niering 1975, Niering and Lowe 1984) and from the region (Woolsey 1911, Merkel 1954, Smith et al. 2006, Sesnie et al. 2009, Anderson-Teixeira et al. 2011) (Table 4). Woolsey (1911) reported the basal area in 'virgin stands' of *P. ponderosa* forest in central and northern Arizona and New Mexico to average between $13.8 - 45.53\ \frac{m^2}{ha}$ for 'black jack' [a discrimination made on the age of the trees] and 'yellow bark' trees, respectively with an 'average stand of the older yellow barked pine' said to average $23.4\ \frac{m^2}{ha}$ on the Prescott N.F., here we did not consider the physiognomy of *P. ponderosa*'s bark to be a significant feature though it helps in identifying the temporal development of a stand. Merkel (1954) reported basal area in a spruce-fir forest on the Kaibab N.F. in northern Arizona to average $55\frac{m^2}{ha}$. Sesnie et al. (2009) reported the basal area for mixed-conifer forests on the Kaibab between $29.2 - 34.1\frac{m^2}{ha}$. basal area estimates in the Santa Catalina by Whittaker and Niering (1975) and Niering and Lowe (1984) range between 31 - $118\frac{m^2}{ha}$ depending on PFT (Table 4). The USFS common stand exams from the Santa Catalina taken in 2005 reported forest types dominated by *A. concolor* as $\leq 55\frac{m^2}{ha}$, *P. mensezii* $\leq 54\frac{m^2}{ha}$, and *P. ponderosa* $\leq 34.4\frac{m^2}{ha}$.

The plot data collected from the Pinaleño are more numerous than in the Santa Catalina, and had greater observed basal area ($110\frac{m^2}{ha}$) in the mixed-conifer PFT, which corroborate with the Whittaker and Niering (1975) and Niering and Lowe (1984) estimates of the same PFTs in the same forests (Table 4). Smith et al. (2006) report general models for forests in the Southern Rockies and give 50th and 99th percentile estimates of AGC (Table 4). Anderson-Teixeira et al. (2011) measured biomass around

flux tower locations in New Mexico, two of which are in the Valles Caldera and report the total above ground carbon density (Table 4).

Tree Models

The AGB for individual trees was evaluated for both DBH and tree height and compared to other published models (Table 7). For the general model: radius had uncertainty with proportionate height ± 3.6 m (Table 8); and for height to radius the uncertainty was ± 11cm (Table 7). These uncertainties are consistent across PFTs (Table 9) and species (Table 10). The MST theoretical relationship of height and diameter [Eq. 7] was significantly different for four of the five PFTs [Pine forest being the exception] (Table 9) and all of the observed species (Table 10).

General

The general allometric model fit by least-squares regression for estimating basal area from height (n=3,702) was: $h^{1.07 \pm 0.03} \propto r$, and AGB from height: $h^{3.62 \pm 0.08} \propto M$ (Supplementary Materials Appendix C). These are both significantly less than the MST predictions of $h^{3/2} \propto r$ and $h^4 \propto M$. The observed differences in both the height to radius and height to AGB relationships, while significant, are still well behaved and consistent across SFDs, suggesting the models are accurate enough to be predictive, albeit with an increasing range of uncertainty for the largest height class.

Based on the observed departure from MST predicted scaling we wondered what effect tree health had on the scaling exponent. When we compare the scaling relationships of observed tree health based qualitatively on field observations of total leaf area, bole straightness, and the presence of disease, the scaling exponent for the healthiest trees was found to be $\alpha = 0.692 \pm 0.052$ and the least healthy trees $\alpha = 0.599 \pm 0.03$ (Table 8). Importantly, the average scaling was $\alpha = 0.619 \pm 0.016$; statistically similar to scaling exponents observed in the literature.

PFT

The scaling exponents for the PFT models estimating AGB from primary size measures varied widely (Table 9). We rejected the MST and Chave et al. (2005) models (as both are an over prediction) for the seral Aspen PFT because the observed $\alpha = 2.358 \pm 0.038$ is so low; that value however is similar to Jenkins et al. (2004): $\alpha = 2.3867$ for *Populus spp.*. For our other three PFT: the Mixed-Conifer (2.493±0.048), White Fir (2.448±0.033), and Spruce-fir (2.537±0.038) we rejected the MST prediction of $\alpha = 2.6667$, and failed to reject the general model from Chave et al. (2005).

Species

By species, the allometric scaling exponent of bole radii to height varied between $\alpha = 0.521 - 0.861$ (Supplemental Materials Appendix C) versus the general model: $\alpha = 0.619 \pm 0.016$ (Table 8). MST's theoretical scaling relationship of $r^{\frac{2}{3}} \propto h$ was statistically indistinguishable for 8 of the 10 most common species including all the conifers (Supplemental Materials Appendix C).

For a particular species [in the example in Figure 5 we use Ponderosa pine, which was one of the most common species sampled in the study and is the most common tree in Arizona and New Mexico], both the general Chave et al. (2005) and Jenkins et al. (2004) pine models were found to underestimate the AGB versus the observed data. We found by least-squares regression Eq. 8 had an $\alpha = 2.64 \pm 0.037$, which was closer to the Návar-Cháidez (2010) model of Mexican pines, and slightly less though not significantly so, than the MST prediction of $\alpha = 2.66\bar{6}$ (Figure 5).

Grid Models

General Models

We created a single general model using all of the study areas (n=133) (Figure 6), and general models for the Valles Caldera (n=48) and Pinaleño (n=72) but not for Santa Catalina (n=13) as there were only a few plots in the Santa Catalina from which to generate a regression model (Table 11). The related general model $GC = \beta MCH^{\alpha}$; where

AGC = AGB * % Carbon (Table 3) was found to be: $\beta = 0.721 \pm 0.618$ and $\alpha = 2.151 \pm 0.313$ and was not significantly different from the universal general model reported by Asner et al. (2011b): where $\beta = 0.844$ and $\alpha = 2$. In the Pinaleño the estimates for the AGC model were $\beta = 0.815 \pm 0.718$ and $\alpha = 2.156 \pm 0.317$, and for the Valles Caldera were $\beta = 1.583 \pm 1.984$ and $\alpha = 1.732 \pm 0.477$; both of which are within the 95% confidence interval of Asner et al.'s (2011b) reported values.

The difference between the MCH obtained from CloudMetrics and GridMetrics (McGaughey 2012) were significant at the plot scale. The correlation between the two metrics was in general a decrease in height along the 1:1 correlation line for the GridMetric height. This is partly due to the decrease in area of the 30 m GridMetric pixel [900 m^2] versus the 1,000 m^2 CloudMetric plots, and from edge effects that may have included splitting a plot into two or more pixels. For this reason we used the CloudMetric to estimate the correlation between observed AGB and the MCH of the aerial LiDAR.

PFT

The PFT-specific AGB models had similar or less uncertainty ($Mg \; ha^{-1}$) and higher r^2 than than did the general model: $r^2 = 0.635$, RMSE= 151 $Mg \; ha^{-1}$ (Table 11); for example, the Pine forest PFT: $r^2 = 0.83$, RMSE = 60.4 $Mg \; ha^{-1}$; and in White-fir PFT: $r^2 = 0.72$, RMSE = 156 $Mg \; ha^{-1}$ (Table 12).

Common Stand Exams

The because of the lesser number of plot data from the Santa Catalina we also incorporated a recent set of stand exams (2005-2006) to compare to the 2007 aerial LiDAR (Figure 7). The Stand Exams included estimates of standing cubic volume and basal area, along with GIS polygons of plot locations. With the ArcGIZ Zonal Statistics Toolbox we established the average of the 30 m pixel GridMetrics across each of the 48 stands (Appendix A). The MCH was found to have a greater coefficient of determination [$r^2 = 0.754$] than the max canopy height [$r^2 = 0.558$] for cubic volume (Figure 8).

Discussion

Uncertainty in height-diameter-biomass allometry

The uncertainty amongst biomass models that use primary size measures are potentially due to a host of factors. Plausible explanations include: (1) variation in response to tree health (Table 8); (2) a vertical height limitation imposed by hydraulic cavitation of the apical meristem (Ryan and Yoder 1997, Koch et al. 2004, Kempes et al. 2011); and (3) the bimodal distribution of young and old growth trees in stands undergoing gap-dynamic tree fall and replacement. Anecdotally, the scaling parameter α [Eq. 8] in the observed data decreases along with decreasing tree health (Table 8), suggesting the parameter value found by least-squares regression of the entire population [both healthy and dying trees] is why we reject the MST prediction. To our second point, the possibility of a vertical height limitation and associated divergence from its scaling relationship with bole radii and estimated biomass increases for trees on sites with limited resources and trees of surpassing ages. O'Connor (2013) observed that short old trees typically exist on xeric south aspects, on steep slopes, and along ridge-tops; conversely, the tallest observed trees were all in riparian drainages and in deep water saturated soils. To our third point, the increase in the uncertainty in height and radii relationships is amplified with age (Figure 9, O'Connor 2013); as a tree approaches its vertical height limit vertical growth slows to become asymptotic relative to the resistance of their meristem to water tension, yet the tree will continue to increase its bole radii for the rest of its lifetime. Observations of these same forests from aerial LiDAR (Swetnam and Falk, *in review*) reveal the tallest trees do not exceed 48 m height. Relative to more temperate forests [with similar species composition] the Arizona and New Mexico trees are less than ½ the height of the largest members of their species (Swetnam and Falk *in review*). If we assume the pipe-model (Savage et al. 2010) is correct, as was the case for at least one species of tree (*P. ponderosa*) [painstakingly measured and quantified for its branching volume (Bentley et al. 2013)], the average AGB of a mature semi-arid *P. ponderosa* tree is significantly greater than what the general model reported in Jenkins et

al. (2004) [for general *Pinus*], and Chave et al. (2005) [for tropical trees] predict using DBH as the independent variable (Figure 5).

A major contribution of the very large estimates in AGB at the plot scale is from the presence of very large trees. Lutz et al. (2012) reported on the disproportionate effect of large trees on over-all stand biomass noting that ~49% of biomass is stored in ~1.4% of the trees (by inventory count) in similar mixed-conifer forests. Additionally, larger cohorts of trees than were measured in the systematic and random plot locations are known to exist based on the aerial LiDAR canopy height model. For example, the Valles Caldera's 'History Grove', located near the former ranching headquarters and reserved from harvest, consists of mature trees >30 m in height with most >100 cm DBH, with some trees >40m height and 180 cm DBH. An AGB estimate for the History Grove was calculated using the pipe-model to exceed 1,000 $Mg\ ha^{-1}$ based on the count of the trees on a one hectare section of canopy height model from the aerial LiDAR. This estimate far exceeds the largest observed value for any PFT in the Valles Caldera. The nearest measured plot (#303) was noted to be less than ½ forested and had 450 $Mg\ ha^{-1}$ (Supplemental Materials Appendix A). Similarly, in the Pinaleño the largest *P. menziesii* specimens are known to exceed 40 m in height and 200 cm in diameter. Near the Clark Peak forest access road there are remnant stumps of *P. menziesii* that exceed 220 cm at stump height. These lines of evidence suggest that, historically, individuals and possibly stands with more biomass once existed in the Pinaleño.

Uncertainty in biomass models

The observed AGB and basal area in the study plots are in close agreement with other published values (Woolsey 1911, Merkel 1954, Whittaker and Nering 1975, Sesnie et al. 2009) (Tables 4 and 5). Notably, in the Pinaleño estimated AGB in several plots was greater than 1,100 $Mg\ ha^{-1}$, this exceeds the maximum estimates reported by Whittaker and Niering (1975) and Niering and Lowe (1984) for the Santa Catalina and Pinaleño.

Below ground biomass values contributes between 12-20% of the total biomass (Xiao et al. 2003, Santantonio et al. 1997, Cairns et al. 1997); for the Santa Catalina and Pinaleño values close to the 14-16% of total biomass in conifers and 37% for aspen were estimated by Whittaker and Niering (1975). On the other hand, Robinson (2004, 2007) suggests below ground biomass estimates in the literature may under predict total belowground biomass by up to 68% once fine roots and mycorrhizal fungi are accounted for. We considered the Whittaker and Niering estimates to be precise enough in the absence of new measurements. Of note, Woolsey (1911) showed *P. ponderosa* as having extensive lateral root systems occasionally exceeding 45 m, suggesting that in some locations the below ground biomass contribution may be greater than assumed in the literature. Below ground biomass is likely to vary widely across the study areas and five PFTs based on differences in soil type and soil depth. The law of large numbers suggests a median value might be appropriate for landscape scale estimates. A further examination of how allometry changes with soil type and depth is needed to better corroborate AGB and field observations of soil organic matter and soil C.

Why do our estimates of AGB and AGC vary so much more widely (Tables 5 & 6) than other recently published estimates of AGB and AGC (Asner et al. 2011b, Lu et al. 2012)? First, if we scale the area of our plot estimates, the inherent plot-level variation is more similar than first suspected. In Mascaro et al. (2011) the uncertainty of AGC for 0.04, 0.25, 1.0, and 6.25 ha plot areas ranged from 63.2, 20.6, 11.1, and 6.25 M$g\ ha^{-1}$, respectively; where predicted relative error for a 1,000 m^2 plot would be 39.7 M$g\ ha^{-1}$. Similarly, Hawbaker et al. (2009) reported an RMSE of 45.55 M$g\ ha^{-1}$ for coniferous forests and 39.32 $Mg\ ha^{-1}$ for deciduous forests. Our results are consistent with the concept reported in Mascaro et al. (2011) that relative error scales with plot size as $(plot\ area)^{-\frac{1}{2}}$. The Pinaleño use a smaller plot size (500 m^2) than the Valles Caldera (1,000 m^2), some of the inherent variation in our general model RMSE uncertainty can be attributed to these variable plot sizes. Second, the techniques used to estimate uncertainty vary from the other publications to the ones here, for example, Asner et al. (2011b) estimated AGB and AGC from the mean of MCH for 5x5m grid pixels at tenth hectare

plot scale. Further, tropical forest biomass estimates topped at 508 $g\ ha^{-1}$ (Asner et al. 2009); whereas our study forests, in particular the Pinaleño study area, had up to 1,495 $Mg\ ha^{-1}$. We suggest the presence of old-growth large diameter trees and that the semi-arid conifer trees appear to stop growing upward with age and tend to grow outward results in a disproportionate increase in the observed RMSE in our data. Also, the effect of disturbance low intensity disturbances (fire and herbivory) in the understory further increase the gap spacing between trees and increasing the variability.

First- and second-order effects on biomass

Accounting for the effect of first order stationarity, i.e. no variation over space, and second-order stationarity, i.e. no interaction between objects or occurrences, becomes critical when looking at large areas, specifically across stands where there are gradients in elevation, slope, aspect, and parent material. The observed heterogeneity of size frequencies within individual plots suggests that changes in first order effects over areas less than one tenth of a hectare can affect biomass estimates in these Southwestern forests. Landscape legacies also potentially lead to second order effects within stands that may not be accounted for at larger scale when only a single attribute like MCH is applied. The variation in MCH observed in the common stand exam observations is one example of where first or second order effects may be changing across a single stand. In the Santa Catalina CSE Stand #23 and in the Marshall Gulch watershed the MCH value varies from 0 to 30 meters along topographic differences [i.e. aspect reversals – a first order effect] (Figure 8). In other areas post-disturbance effects [i.e. low-to-high fire severity, a second order effect] radically change stand structure, as shown in the Marshall Gulch stand (Figure 8).

Conclusions

Two novel findings of this research are (1) when the condition of trees are taken into account the MST predicted scaling exponent is observed in the healthiest trees, however, MST predictions were found to fail when all trees regardless of condition were used in the least-squares regression. In those cases the observed estimates were

indistinguishable from the Jenkins et al. (2004) and Chave et al. (2005) models; and (2) AGB derived from height based observations (individual tree height or MCH) is shown to significantly under predict biomass in old-growth semi-arid conifer forests.

We suggest two mechanisms are at work which account for the greater than expected level of biomass with shorter than expected tree height and MCH profiles: (1) a truncation of maximum height based on the combination of species physiology and locally limiting factors [i.e. precipitation, temperature, and total plant available water in the soil profile], and (2) the particularly long lives [up to 450 years (Grissino-Mayer et al. 1995)] of trees in the observed old growth stands. A semi-arid conifer tree may reach a vertical height near its local eco-physiological limit early in its life cycle [50-150 years], after which and over the proceeding decades to centuries the tree continues to increase in cambial diameter and maximizing its canopy volume while maintaining the same vertical height relative to the amount of water it is able to obtain. Such a tree eventually is shorter than predicted using general models based on its diameter. Importantly, the age of the stand becomes a critical factor in determining whether the stand has an exceptional amount of AGB.

References

Allen, C.D., A.K. Macalady, H. Chenchouni, D. Bachelet, N. McDowell, et al. 2010. A global overview of drough and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management, 259*: 660-684.

Andersen, H.-E., Reutebuch, S.E., McGaughey, R.J. 2006. A rigorous assessment of tree height measurements obtained using airborne LIDAR and conventional field methods. *Canadian Journal of Remote Sensing, 32*(5): 355-366.

Andersen, H. E., McGaughey, R. J., Carson, W. W., Reutebuch, S. E., Mercer, B., & Allan, J. 2003. A comparison of forest canopy models derived from LIDAR and INSAR data in a Pacific Northwest conifer forest. *International Archives of Photogrammetry and Remote Sensing, 34*(3): 211-217.

Anderson-Teixeira, K. J., Delong, J. P., Fox, A. M., Brese, D. A., & Litvak, M. E. 2011. Differential responses of production and respiration to temperature and moisture drive the carbon balance across a climatic gradient in New Mexico. *Global Change Biology, 17*(1): 410-424.

Asner, G. P., Hughes, F.R., 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 GP, Hughes RF, Mascaro J, et al. 2011a. High-resolution carbon mapping on the million-hectare Island of Hawaii. *Frontiers in Ecology and the Environment, 9*(8): 434-439.

Asner G.P., J. Mascaro, H.C. Muller-Landau, G. Vieilledent, R. Vaudry, M. Rasamoelina, J. Hall, and M. van Breugel. 2011b. A universal airborne LiDAR approach for tropical forest carbon mapping. *Oecologia, 168*:1147-1160.

Asner G.P., J.K. Clark, J. Mascaro, R. Vaudry, K.D. Chadwick, G.Vieilledent, M. Rasamoelina, A. Balaji, T. Kennedy-Bowdoin, L. Matoug, et al. 2012. Human and environmental controls over above ground carbon storage in Madagascar. *Carbon Balance and Management, 7*(2). 13 p.

Baccini, A., M. A. Friedl, C. E. Woodcock, and R. Warbington 2004. Forest biomass estimation over regional scales using multisource data, *Geophys. Res. Lett., 31*, L10501.

Bentley, L.P., Stegen, J.C., Savage, V.M., Smith, D.D., Allmen, E.I., Sperry, J.S., Reich, P.B., Enquist, B.J., 2013. An empirical assessment of tree branching networks and implications for plant allometric scaling models. *Ecol. Lett. 16*(8):1069-78.

Bright, B. C., Hicke, J. A., & Hudak, A. T. 2012. Estimating aboveground carbon stocks of a forest affected by mountain pine beetle in Idaho using lidar and multispectral imagery. *Remote Sensing of Environment*, *124*: 270-281.

Brohman, R.; Bryant, L. eds. 2005. Existing Vegetation Classification and Mapping Technical Guide. Gen. Tech. Rep. WO–67. Washington, DC: U.S. Department of Agriculture Forest Service, Ecosystem Management Coordination Staff. 305 p.

Brown, S. Estimating biomass and biomass change of tropical forests. 1997. Food and Agricultural Organization of the United Nations Forestry Paper 134 (Rome, 1997).

Brown JH, Gillooly JF, Allen AP, Savage VM, West GB 2004. Toward a metabolic theory of ecology. *Ecology 85*(7): 1771-1789.

Brown, P.M., C.L. Wienk, and A.J. Symstad. 2008a. Fire and forest history at Mount Rushmore. *Ecological Applications 18*: 1984-1999.

Brown, P.M., E.K. Heyerdahl, S.G. Kitchen, and M.H. Weber. 2008b. Climate effects on historical fires (1630-1900) in Utah. *International Journal of Wildland Fire 17*: 28-39.

Brown-Mitic, C., W.J. Shuttleworth, R.C. Harlow, J. Petti, E. Burke, R. Bales. 2007. Seasonal water dynamics of a sky island subalpine forest in semi-arid southwestern United States. *Journal of Arid Environments 69*: 237-258.

Burns, R.M., H. Honkala, tech. coords. 1990. Silvics of North America: 1. Conifers; 2. Hardwoods. Agriculture Handbook 654. U.S. Department of Agriculture, Forest Service, Washington, DC. vol.2, 877 p.

Cairns, M.A., Brown, S., Helmer, E.H., and Baumgardner, G.A., 1997, Root biomass allocationin the world's upland forests, *Oecologia*, *111*:1–11.

Chave, J., Andalo, C., Brown, S., Cairns, M. A., Chambers, J. Q., Eamus, D., ... & Yamakura, T. 2005. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia*, *145*(1): 87-99.

Comer, P., D. Faber-Langendoen, R. Evans, S. Gawler, C. Josse, G. Kittel, S. Menard, M. Pyne, M. Reid, K. Schulz, K. Snow, and J. Teague. 2003. Ecological Systems of the United States: A Working Classification of U.S. Terrestrial Systems. NatureServe, Arlington, VA. 75 p.

Coop J.D., T.J. Givnish. 2007. Gradient analysis of reversed treelines and grasslands of the Valles Caldera, New Mexico. *Journal of Vegetation Science 18*: 43-53.

Drake JB, Knox RG, Dubayah RO, Clark DB, Condit R, Blair JB, Hofton M. 2003. Above-ground biomass estimation in closed canopy neotropical forests using lidar remote sensing: factors affecting the generality of relationships. *Glob Ecol Biogeography 12*:147–59.

d'Oliveira, M.V.N., S.E. Reutebuch, R.J. McGaughey, H-E. Anderson. 2012. Estimating forest biomass and identifying low-intensity logging areas using airborne scanning lidar in Antimary State Forest, Acre State, Western Brazilian Amazon. *Remote Sensing of Environment, 124*: 479-491.

Du Bois R.L., 1959. Geology of the Santa Catalina Mountains. Southern Arizona Guidebook II, L.A. Heindel, ed. *Arizona Geological Society* 106-116

Enquist, B. J., West, G. B., Charnov, E. L., & Brown, J. H. 1999. Allometric scaling of production and life-history variation in vascular plants. *Nature*, *401*(6756): 907-911.

Enquist, B. J., & Niklas, K. J. 2002. Global allocation rules for patterns of biomass partitioning in seed plants. *Science*, *295*(5559):1517-1520.

Enquist B.J., G.B. West, J.H. Brown. 2009. Extensions and evaluations of a general quantitative theory of forest structure and dynamics. *Proceedings of the National Academy of Sciences Early Edition vol. 106*(17) April 2009.

Erdody T. and L. M. Moskal, 2010. Fusion of LiDAR and Imagery for Estimating Forest Canopy Fuels, *Remote Sensing of Environment, 114*(4); 725-737.

ESRI 2012. ArcGIS Desktop: Release 10.1. Redlands, CA: Environmental Systems Research Institute.

Evans J.S., A.T. Hudak, R. Faux, A.M.S. Smith. 2009. Discrete Return Lidar in Natural Resources: Recommendations for Project Planning, Data Processing, and Deliverables. *Remote Sensing, 1*: 776-794

Falkowski MJ, PE Gessler, AMS Smith, AT Hudak, LA Vierling. 2006. Automatically measuring individual tree crown diameter and height from LiDAR data: A comprehensive evaluation of spatial wavelet analysis. Eos Transactions, American Geophysical Union, Fall Meeting, Dec. 2006.

Falkowski MJ, AMS Smith, PE Gessler, AT Hudak, LA Vierling, JS Evans. 2008. The influence of conifer forest canopy cover on the accuracy of two individual tree measurement algorithms using lidar data. *Canadian Journal of Remote Sensing, 34*, S2, S338-S350.

Falkowski M.J., J.S. Evans, S. Martinuzzi, P.E. Gessler, A.T. Hudak. 2009. Characterizing forest succession with lidar data: An evaluation for the Inland Northwest, USA. *Remote Sensing of Environment, 113*: 946–956.

Falkowski M.J., A.T. Hudak, N.L. Crookston, P.E. Gessler, E.H. Uebler, A.M.S. Smith. 2010. Landscape-scale parameterization of a tree-level forest growth model: a k-nearest neighbor imputation approach incorporating LiDAR data. *Canadian Journal Forestry Research, 40*: 184-199

Feldpausch, T. R., Banin, L., Phillips, O. L., Baker, T. R., Lewis, S. L., Quesada, C. A., ... & Bird, M. 2010. Height-diameter allometry of tropical forest trees. *Biogeosciences*, *8*(5): 1081-1106.

Ffolliot P.F., C.L. Stropki, D.G. Neary. 2008. Historical Wildfire Impacts on Ponderosa Pine Tree Overstories: An Arizona Case Study. Research Paper RMRS-RP-75. Rocky Mountain Research Station. U.S. Department of Agriculture, Forest Service, Washington, DC.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., ... & Snyder, P. K. 2005. Global consequences of land use. *Science, 309*(5734): 570-574.

Franklin O., J. Johansson, R.C. Dewar, U. Dieckmann, R.E. McMurtrie, Å. Brännström, R. Dybzinski. 2012. Modeling carbon allocation in trees: a search for principles. *Tree Physiology, 00:* 1-19

Friedlingstein P., P. Cox, R. Betts, L. Bopp, et al. 2006 Climate-Carbon Cycle Feedback Analysis: Results of the C^4MIP Model Intercomparison. *Journal of Climate, 19*: 3337-3353.

Gatziolis, D., H.E. Andersen. 2008. A guide to LIDAR data acquisition and processing for the forests of the Pacific Northwest. Gen. Tech. Rep. PNW-GTR-768. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 32 p

Gatziolis, D., Fried, J. S., & Monleon, V. S. 2010. Challenges to estimating tree height via LiDAR in closed-canopy forests: A parable from western Oregon. *Forest Science, 56*(2), 139-155.

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.* 13 p.

Gillooly J.F. J.H. Brown, G.B. West, V.M. Savage, E.L. Charnov. 2001. Effects of Size and Temperature on Metabolic Rate. *Science, 293*(5538): 2248-2251.

Goetz, S., & Dubayah, R. 2011. Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. *Carbon Management, 2*(3): 231–244.

Goodale, C.L., M.J. Apps, R.A. Birdsey, C.B. Field, L.S. Heath, R.A. Houghton, J.C. Jenkins, G.H. Kohlmaier, W. Kurz, S. Liu, G.-J. Nabuurs, S. Nilsson, S., and A.Z. Shvidenko. 2002. Forest carbon sinks in the northern hemisphere. *Ecological Applications* 12(3): 891-899.

Grissino-Mayer, H.D., Baisan, C.H. and Swetnam, T.W. 1995. Fire history in the Pinaleño Mountains of southeastern Arizona: Effects of human related disturbances. In Debano, L.F., Gottfried, G.J., Hamre, R.H., Edminster, C.B., Ffolliott, P.F. and Ortega-Rubio, A., editors, Biodiversity and management of the Madrean Archipelago: the Sky Islands of Southwestern United States and Northwestern Mexico, US Department of Agriculture Forest Service, General Technical Report RM-GTR-264, 399–407.

Harlow, R.C., E.J. Burke, R.L. Scott, W.J. Shuttleworth, C.M. Brown, J.R. Petti. 2004. Derivation of temperature lapse rates in semi-arid south-eastern Arizona. *Hydrology and Earth Systems Sciences, 8*(6): 1179-1185.

Hawbaker, T. J., Keuler, N. S., Lesak, A. A., Gobakken, T., Contrucci, K., & Radeloff, V. C. 2009. Improved estimates of forest vegetation structure and biomass with a LiDAR-optimized sampling design. *Journal of Geophysical Research, 114*.

Heyerdahl, E.K., P.M. Brown, S.G. Kitchen, M.H. Weber. 2011. Multi-century fire and forest histories across forest types in Utah and eastern Nevada. General Technical Report. GTR-RMRS-261 Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Holmgren, J., Persson, Å., & Söderman, U. 2008. Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images. *International Journal of Remote Sensing, 29*(5): 1537-1552.

Huete A., K. Didan, T. Miura, E.P. Rodriguez, X. Gao, L.G. Ferreira. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment 83*: 195-213.

Hudak A.T., N.L. Crookston, J.S. Evans, D.E. Hall, M.J. Falkowski 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Remote Sensing of Environment 112*: 2232-2245.

Hudak A.T., J.S. Evans, A.M.S. Smith 2009. LiDAR Utility for Natural Resource Managers. *Remote Sensing 1*: 934-951

Hudak, A. T., Strand, E. K., Vierling, L. A., Byrne, J. C., Eitel, J. U., Martinuzzi, S., & Falkowski, M. J. 2012. Quantifying aboveground forest carbon pools and fluxes from repeat LiDAR surveys. *Remote Sensing of Environment, 123*: 25-40.

Hyde P, R Dubayah, W Walker, JB Blair, M Hofton, C Hunsaker. 2006. Mapping forest structure for wildlife habitat analysis using multisensory (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. *Remote Sensing of Environment, 102*(1-2): 63-73.

Hyyppä, J., Kelle, O., Lehikoinen, M., Inkinen, M. 2001. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. IEEE *Transactions on Geoscience and Remote Sensing, 39*: 969–975.

Hyyppä, J., Mielonen, T., Hyyppä, H., Maltamo, M., Yu, X., Honkavaara, E., Kaartinen, H., 2005, Using individual tree crown approach for forest volume extraction with aerial images and laser point clouds. In Proceedings of the ISPRS Workshop on Laser Scanning, 12–14 September 2005, Enschede, The Netherlands, pp. 144–149

Izett, G.A., 1981. Volcanic ash beds: Recorders of Upper Cenozoic silicic pyroclastic volcanism in the western United States. *Journal of Geophysical Research, 86*(11), P. 10,200.

Jenkins, J.C.; Chojnacky, D.C.; Heath, L.S.; Birdsey, R.A. 2004. A comprehensive database of biomass regressions for North American tree species. Gen. Tech. Rep. NE-319. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northeastern Research Station. 45 p.

Ke, Y., Quackenbush, L. J., & Im, J. 2010. Synergistic use of QuickBird multispectral imagery and LIDAR data for object-based forest species classification. *Remote Sensing of Environment*, *114*(6), 1141-1154.

Kempes CP, West GB, Crowell K, Girvan M (2011) Predicting Maximum Tree Heights and Other Traits from Allometric Scaling and Resource Limitations. *PLoS ONE, 6*(6): e20551.

Koch G.W., S.C. Sillett, G.M. Jennings, S.D. Davis 2004. The limits of tree height. *Nature, 428*: 851-854.

Kraus, K., and N. Pfeifer. 1998. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing, 53*: 193-203.

Kriegler, F.J., Malila, W.A., Nalepka, R.F., and Richardson, W. 1969. 'Preprocessing transformations and their effects on multispectral recognition.' *Proceedings of the Sixth International Symposium on Remote Sensing of Environment,* p. 97-131.

Laes, D.; Reutebuch, S.; McGaughey, B.; Maus, P.; Mellin, T.; Wilcox, C.; Anhold. J.; Finco, M.; Brewer, K. 2008. Practical lidar acquisition considerations for forestry applications. RSAC-0111-BRIEF1. Salt Lake City, UT: U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center. 7 p.

Laes, D.; Mellin, T.; Wilcox, C.; Anhold, J.; Maus, P.; Falk, D.A.; Koprowski, J.; Drake, S.; Dale, S.; Fisk, H.; Joria, P.; Lynch, A.M.; Alanen, M. 2009. Mapping vegetation structure in the Pinaleño Mountains using lidar. RSAC-0118-RPT1. Salt Lake City, UT: U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center. 22 p. 173-181

Lamlom S.H., R.A. Savidge. 2003. A reassessment of carbon content in wood: variation within and between 41 North American species. *Biomass and Bioenergy, 25*: 381-388.

LANDFIRE: LANDFIRE 1.1.0 Vegetation Dynamics Models. 2010. [Homepage of the LANDFIRE Project, U.S. Department of Agriculture, Forest Service; U.S. Department of the Interior],
[Online]. Available: http://www.landfire.gov/index.php [2010, October 28].

Lefsky MA, Cohen WB, Acker SA, Parker GG, Spies TA, Harding D. 1999. Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. *Remote Sens Environ, 70*: 339–61.

Lefsky MA, Cohen WB, Harding DJ, Parker GG, Acker SA, Gower ST. 2002a. Lidar remote sensing of above-ground biomass in three biomes. *Glob Ecol Biogeogr, 11*:393–9.

Lefsky MA, Cohen WB, Parker GG, Harding DJ. 2002b. Lidar remote sensing for ecosystem studies. *Bioscience, 52*:19–30.

Lefsky MA, Harding DJ, Keller M, Cohen WB, Carabajal CC, Espirito-Santo FDB, Hunter MO, Oliveira R Jr. 2005. Estimates of forest canopy height and aboveground biomass using ICESat. *Geophys Res Lett, 32*:L22S02.

Liu, F., R. Parmenter, P.D. Brooks, M.H. Conklin, R.C. Bales. 2008. Seasonal and interannual variation of streamflow pathways and biogeochemical implications in semi-arid, forested catchments in Valles Caldera, New Mexico. *Ecohydrology, 1*: 239-252.

Lowry, R., T. Shrupp, S. Kirby, Waller, Schrader, Falzarano, Langs, Manis, Wallace, Schultz, Comer, Pohs, Rieth, Velasquez, Wolk, Kepner, Boykin, O'Brien, Bradford, Thompson, and Prior-Magee. 2007. Mapping moderate-scale land cover over very large geographic areas within a collaborative framework: A case study of the Southwest Regional Gap Analysis Project (SWReGAP). *Journal of Remote Sensing of Environment, 108*: 59-73.

Lu, D., Chen, Q., Wang, G., Moran, E., Batistella, M., Zhang, M., ... & Saah, D. 2012. Aboveground Forest Biomass Estimation with Landsat and LiDAR Data and Uncertainty Analysis of the Estimates. *International Journal of Forestry Research*, *2012*. 436537, 16 p.

Lutz, J. A., Larson, A. J., Swanson, M. E., & Freund, J. A. (2012). Ecological importance of large-diameter trees in a temperate mixed-conifer forest. *PloS one*, *7*(5), e36131.

Lynch AM (2009) Spruce Aphid, Elatobium abietinum (Walker) Life History and Damage to Engelmann Spruce in the Pinaleño Mountains, Arizona. In: Sanderson, H.R.; Koprowski, J.L., eds. The last refuge of the Mt. Graham red squirrel: Ecology of endangerment. Tucson, AZ: University of Arizona Press. p. 318-338.

Malhi, Y., Wood, D., Baker, T. R., Wright, J., Phillips, O. L., Cochrane, T., ... & Vinceti, B. 2006. The regional variation of aboveground live biomass in old-growth Amazonian forests. *Global Change Biology*, *12*(7): 1107-1138.

Martin W.P., J.E. Fletcher 1943. Vertical Zonation of great soil groups on Mt. Graham, Arizona, as correlated with climate, vegetation, and profile characteristics. Technical Bulleton No. 99. January 1, 1943. University of Arizona College of Agriculture.

Mascaro J., D. Matteo, G.P. Asner, H.C. Muller-Landau 2011. Evaluating uncertainty in mapping forest carbon with airborne LiDAR. Remote Sensing of Environment. 115 (2011) pp. 3770–3774.

MATLAB R2012a. 2012. The MathWorks, Inc. Version 7.14.0.739.

Mitchell, B., Walterman, M., Mellin, T., Wilcox, C., Lynch, A. M., Anhold, J., ... & Fisk, H. 2012. Mapping vegetation structure in the Pinaleno Mountains using lidar-phase 3: Forest inventory modeling. RSAC-100007-RPT1. Salt Lake City, UT: U.S. Department of Agriculture, Forest Service, Remote Sensing Applications Center. 17 p.

McGaughey, R. (2012) FUSION / LDV: Software for LiDAR Data Analysis and Visualization. USDA Forest Service, Pacific Northwest Research Station.

Mellin, T., Triepke, F. J., & Joria, P. 2008. Mapping existing vegetation at the mid-scale level in the Forest Service Southwestern Region. In *Proceedings of the twelfth biennial USDA Forest Service remote sensing applications conference. Salt Lake City, Utah* pp. 15-19.

Merkel J. 1954. An Analysis of the Spruce-fir community on the Kaibab Plateau, *Arizona. Ecology, 35*(3) : 316-322.

Merriam C.H., L.H. Stejneger. 1890. Results of a biological survey of the San Francisco Mountain region and desert of the Little Colorado, Arizona. Washington DC, Government Printing Office. 136 p.

Miles, P.D., W.B. Smith. 2009. Specific Gravity and Other Properties of Wood and Bark for 156 Tree Species Found in North America. 2009. Northern Research Station. Research Note NRS-38. p.39

Molotch, N.P., P.D. Brooks, S.P. Burns, M. Litvak, R.K. Monson, J.R. McConnell, K. Musselman. 2009. Ecohydrological controls on snowmelt partitioning in mixed-conifer sub-alpine forests. *Ecohydrology, 2*: 129-142.

Muldavin. E. and P. Tonne. 2003. A Vegetation Survey and Preliminary Ecological Assessment of Valles Caldera National Preserve, New Mexico. *Natural Heritage New Mexico final report to Valles Caldera National Preserve, New Mexico*. 73 p.

Muldavin E., P. Neville, C. Jackson, T. Neville. 2006. A vegetation map of the Valles Caldera National Preserve, New Mexico. Final report submitted in April 4, 2006 in partial fulfillment of National Prak Service Award No. 1443-CA-1248-01-001 and Valles Caldera Trust Contract No. VCT-TO 0401.

Naidoo, L., Cho, M. A., Mathieu, R., & Asner, G. 2012. Classification of savanna tree species, in the Greater Kruger National Park region, by integrating hyperspectral and LiDAR data in a Random Forest data mining environment. *ISPRS Journal of Photogrammetry and Remote Sensing*, *69*: 167-179.

Nature Conservancy. 2006. Southwest Forest Assessment Project. Available at: http://azconservation.org/projects/southwest_forest_assessment

Nature Conservancy. 2007. Historical Range of Variation for Potential Natural Vegetation Types of the Southwest. Available at: http://azconservation.org/downloads/historical_range_of_variation_for_potential_natural_vegetation_types

NatureServe. 2005. International Ecological Classification Standard: Terrestrial Ecological Classifications. NatureServe Central Databases. Arlington, VA. U.S.A. Data current as of 19 October 2005.

Návar-Cháidez, J. D. J. 2010. Biomass allometry for tree species of northwestern Mexico. *Tropical and Subtropical Agroecosystems*, *12*(3): 507-519.

Niering, W.A., C.H. Lowe. 1984. Vegetation of the Santa Catalina Mountains: community types and dynamics. *Vegetatio, 58*: 3-28.

O'Connor, C., D.A. Falk, A.M. Lynch, C.P. Wilcox, T.W. Swetnam, T.L. Swetnam. 2010. Growth and demography of Pinaleño high elevation forests. RJVA 07-JV-

11221615317 Performance Report. Tucson, AZ: University of Arizona, Laboratory of Tree-Ring Research; School of Natural Resources and the Environment. 21 p.

O'Connor, C., 2013. Spatial and temporal dynamics of disturbance interactions along an ecological gradient of the Pinaleno Mountains, Arizona, USA. Ph.D. Dissertation. School of Natural Resources and Environment. University of Arizona, Tucson AZ.

Pelletier J., G. Barron-Gafford , D. Breshears , P. Brooks , J. Chorover , M. Durcik , C.Harman , T. Huxman , K. Lohse , R. Lybrand , T. Meixner , J. McIntosh , S. Papuga , C.Rasmussen , M. Schaap , T. Swetnam , P. Troch. 2013. Coevolution of nonlinear trends in vegetation, soils, and topography with elevation and slope aspect: A case study in the sky islands of southern Arizona. *Journal of Geophysical Research - Earth Surface* (2013): 1-18.

Popescu SC, Wynne RH, Scrivani JA. 2004. Fusion of small footprint lidar and multispectral data to estimate plot-level volume and biomass in deciduous and pine forests in Virginia, USA. *Forest Sci 50*:551–65.

Stuever, M., J. Hayden 1997 Plant Associations of Arizona and New Mexico, ed. 3, 7/1997. Vol 1: Forests, Vol 2: Woodlands (an update of the USDA Forest Serv SW Region Habitat Typing Guides. 9/1996, rev 7/1997. Contract R3-95-27).

Prior-Magee, J.S., K.G. Boykin, D.F. Bradford, W.G. Kepner, J.H. Lowry, D.L. Schrupp, K.A. Thomas, and B.C. Thompson, Editors. 2007. Southwest Regional Gap Analysis Project Final Report. U.S. Geological Survey, Gap Analysis Program, Moscow, ID. 422 pp.

Prusinkiewicz, P., Lindenmayer, A., Hanan, J. S., Fracchia, F. D., Fowler, D. R., de Boer, M. J., & Mercer, L. (1990). *The algorithmic beauty of plants* (Vol. 2, No. 6). New York: Springer-Verlag.

Rasmussen C., Troch P.A., Chorover J., Brooks P., Pelletier J., and Huxman T. 2011. An open system framework for integrating critical zone structure and function. *Biogeochemistry*, *102*(1-3): 15-29.

Reutebuch, S.E., R.J. McGaughey, H.-E. Andersen, and W.W. Carson. 2003. Accuracy of a high-resolution LIDAR terrain model under a conifer forest canopy. *Canadian Journal of Remote Sensing 29*(5): 527-535.

Reutebuch, S.E., H.-E. Andersen, and R.J. McGaughey. 2005. Light detection and ranging (LIDAR): An emerging tool for multiple resource inventory. *Journal of Forestry 103*(6): 286-292.

Richardson J. and L. M. Moskal, 2011. Strengths and Limitations of Assessing Forest Density and Spatial Configuration with Aerial LiDAR, *Remote Sensing of Environment, 114*(4): 725-737.

Robinson, D. 2004. Scaling the depths: below-ground allocation in plants, forests and biomes. *Functional Ecology*, *18*(2): 290-295.

Robinson, D. 2007. Implications of a large global root biomass for carbon sink estimates and for soil carbon dynamics. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1626): 2753-2759.

Ryan, M.G. and B.J. Yoder 1997. Hydraulic limits to tree height and growth. *BioScience 47*: 235–242.

Santantonio, D., Hermann, R.K., and Overton, W.S., 1997, Root biomass studies in forest ecosystems, *Pedobiologia 17*:1–31.

Savage, V. M., Bentley, L. P., Enquist, B. J., Sperry, J. S., Smith, D. D., Reich, P. B., & von Allmen, E. I. (2010). Hydraulic trade-offs and space filling enable better predictions of vascular structure and function in plants.*Proceedings of the National Academy of Sciences*, *107*(52): 22722-22727.

Schimel, D., J. Melillo, H. Tian, A.D. McGuire, D. Kicklighter, T. Kittel, N. Rosenbloom, S. Running, P. Thornton, D. Ojima, W. Parton, R. Kelly, M. Sykes, R. Neilson, and B. Rizzo. 2000. Contribution of increasing CO_2 and climate to carbon storage by ecosystems in the United States. *Science 287*: 2004-2006.

Sesnie S.E., B.G. Dickson, J.M. Rundall, T.D. Sisk 2009. Preliminary stratification and characterization of the mixed conifer forest type on the Kaibab National Forest, North Kaibab Ranger District. Final Report. Northern Arizona University

Sherril KR, Lefsky MA, Ryan MG 2008. Forest structure estimation and pattern exploration from discrete-return lidar in subalpine forests of the central Rockies. *Can J Res 38*:2081–2096.

Shinozaki K, Yoda K, Hozumi K, Kira T 1964. A quantitative analysis of plant form– the pipe model theory. I. Basic analysis. Japan Journal of Ecology 14:97–105.

Shreve F. 1915 The vegetation of a desert mountain range as conditioned by climatic factors. Carnegie Institution of Washington.

Sibly, R.M., J.H. Brown, A. Kodric-Brown, Eds. 2012. Metabolic Ecology: A Scaling Approach. John Wiley & Sons. 2012. 375 pages.

Smith, T. M., Shugart, H. H., Woodward, F. I., & Burton, P. J. 1993. Plant functional types. In *Vegetation Dynamics & Global Change* (pp. 272-292). Springer US.

Smith J.E., L.S. Heath, K.E. Skog, R.A. Birdsey. 2006. *Methods for Calculating Forest Ecosystem and Harvested Carbon With Standard Estimates for Forest Types of the United States.* GTR-NE-343. Newtown Square, PA; USDA Forest Service: www.nrs.fs.fed.us/pubs/8192.

Smith J.E., L.S. Heath, M.C. Nichols. 2007. *U.S. Forest Carbon Calculation Tool: Forestland Carbon Stock and Net Annual Change* [CD-ROM]. GTR-NRS-13. Newtown Square, PA: USDA Forest Service: www.nrs.fs.fed.us/pubs/2394.

Sundquist, E. T., Ackerman, K. V., Bliss, N. B., Kellndorfer, J. M., Reeves, M. C., & Rollins, M. G. 2009. Rapid assessment of US forest and soil organic carbon storage and forest biomass carbon sequestration capacity. *US Geological Survey Open-File Report*, *1283*: 15.

Swetnam, T.L., Falk, D.A., in review. Application of Metabolic Scaling Theory to reduce error in local maxima tree segmentation from aerial LiDAR. Submitted: *Forest Ecology and Management.*

Thomas, C. D. et al. 2004. Extinction risk from climate change. Nature 427: 145-148.

Trenberth K.E. 1998. Atmospheric moisture residence times and cycling: Implications for rainfall rates and climate change. *Climatic Change, 39*: 667-694.

Trimble Navigation Limited. 2005. White Paper: H-Star Technology Explained. 9 p.

Turner, D.P., M. Guzy, M.A. Lefsky, W.D. Ritts, S. Van Tuyl, B.E. Law. 2004
Monitoring Forest Carbon Sequestration with Remote Sensing and Carbon Cycle
Modeling. *Environmental Management 33*(4): 457-466.

Ustin, S., J.A. Gamon 2010. Remote sensing of plant functional types. *New Phytologist.*
186(4): 795-816.

West, G. B., Brown, J. H., & Enquist, B. J. 1999. The fourth dimension of life: fractal
geometry and allometric scaling of organisms. *Science*, *284*(5420): 1677-1679.

West G.B., B.J. Enquist, J.H. Brown (2009) A general quantitative theory of forest
structure and dynamics. *Proc Natl Acad Sci 106*(17): 7040–7045

Western Regional Climate Center. 2013. Data available from: http://www.wrcc.dri.edu

Whittaker, R. H., F. H. Bormann, G. E. Likens and T. G. Siccama. 1974. The Hubbard
Brook Ecosystem Study: forest biomass and production. *Ecol. Monogr. 44*(2):
233-254.

Whittaker RH, WA Niering 1964. Vegetation of the Santa Catalina Mountains, Arizona:
I. Ecological Classification and Distribution of Species. *Journal of the Arizona*
Academy of Science, 3(1): 9-34.

Whittaker RH, WA Niering 1965. Vegetation of the Santa Catalina Mountains, Arizona:
II. A Gradient Analysis of the South Slope. *Ecology 46*(4): 429-452.

Whittaker RH, WA Niering 1968a. Vegetation of the Santa Catalina Mountains, Arizona:
III. Species distribution and floristic relations of the North Slope. *Journal of the*
Arizona Academy of Science, 5(1): 3-21.

Whittaker RH, WA Niering 1968b. Vegetation of the Santa Catalina Mountains, Arizona:
IV. Limestone and Acid Soils. *Ecology 56*(2): 523-544.

Whittaker RH, WA Niering 1975. Vegetation of the Santa Catalina Mountains, Arizona
V. Biomass, Production and Diversity along the Elevation Gradient. *Ecology*
56(4):771-790.

Williams, A.P., C.D. Allen, A.K. Macalady, D. Griffin, C.A. Woodhouse, D.M. Meko, et
al. 2012. Temperature as a potent driver of regional forest drought stress and tree
mortality. *Nature Climate Change 3*: 292-297.

Woolsey T.S. 1911. Western Yellow Pine in Arizona and New Mexico. US Department of Agriculture Forest Service Bulletin 101. November 24, 1911.

Woodbury P.B., Smith J.E., Heath L.S. (2007). Carbon sequestration in the US forest sector from 1990 to 2010. *Forest Ecology and Mana*gement*, 241*(1), 14-27.

Wooldridge, S. A. XXXX. Mass extinctions past and present: a unifying hypothesis, *Biogeosciences Discuss., 5*: 2401-2423,

Xiao C.W., J.C. Yuste, I.A. Janssens, P. Roskams, L. Nachtergale, A. Carrara, B.Y. Sanchez, R. Cuelemans. 2003. Above- and belowground biomass and net primary productivity in a 73-year-old Scots pine forest. *Tree Physiology 23*: 505-516.

Zhang, C., & Qiu, F. 2012. Mapping individual tree species in an urban forest using airborne lidar data and hyperspectral imagery. *Photogrammetric Engineering & Remote Sensing, 78*(10): 1079-1087.

Zhao, G., G. Shao, K. M. Reynolds, M.C. Wimberly, T. Warner, J.W. Moser, K. Rennolls, S. Magnussen, M. Kohl, H.-E. Andersen, G.A. Mendoza, L. Dai, A. Huth, L. Zhang, J. Brey, Y. Sun, R. Ye, B.A. Martin, and F. Li. 2005. Digital Forestry: A White Paper. *Journal of Forestry 103*(1):47-50.

Zhao, K., S. Popescu, R. Nelson. 2009. LiDAR remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. *Remote Sensing of Enviornment 113* : 182-196.

Zianis D., M. Mencuccini. 2004. On simplifying allometric analysis of forest biomass. *Forest Ecology and Management 187* : 311-332.

Figure Captions

Figure 1: Process tree of the aerial LiDAR data turned into data derivatives in USFS FUSION and MATLAB 2012B. The green boxes are software programs, the orange colored boxes represent files types, the blue boxes are the executable processes or scripts used in the software program to create the derivative file types, and the red boxes are the scalars.

Figure 2: Process tree of the Grid level biomass models. The Mean Canopy Height is used as a dependent variable, the plot measured above ground biomass (and carbon) are the independent variable. Each PFT model is reported for basal area, above ground biomass, and above ground carbon.

Figure 3: Process tree of the individual tree level biomass models. The individual tree heights are either broken down by forest type average or by species.

Figure 4: LiDAR coverage of the three study areas hill shaded with LiDAR plot locations and Common Stand Exam boundaries shown.

Figure 5: *Pinus* AGB model from Jenkins et al. (2004) (dashed pink line); Návar-Cháidez (2010) model for Mexican pines (dashed blue line); tropical tree model from Chave et al. (2005) where $D^{2.5}$, (dashed brown line); West et al.'s (2009) MST theoretical $r^{\frac{8}{3}}$, (dashed red line); and our observed model for *P. ponderosa* with 95% confidence intervals in solid black and thin dashed black lines, respectively. The observed trees, *P. ponderosa*, are open circles.

Figure 6: The above ground biomass estimated from 132 plots in the three study areas: Santa Catalina (n=12), Valles Caldera (n=48), and Pinaleno (n=72). The black line is the fit of the general model to the data found by least squares regression.

Figure 7: Example of MCH for the common stands in the Santa Catalina. Colors represent height above ground. The Marshall Gulch watershed is part of the CZO where the two smaller sub-basins (Granit and Schist) are of different geologies and are heavily

instrumented for studying multiple earth processes.

Figure 8: Data from the 2004-2005 Santa Catalina Common Stand Exams versus the 2007 Sabino aerial LiDAR mean and max canopy height metric derived at 30m resolution. All data have been converted from feet to meters.

Figure 9: The range in age of trees with the measured height and diameter varies widely in the Pinaleño, n=368, data provided by Christopher O'Connor (O'Connor 2013).
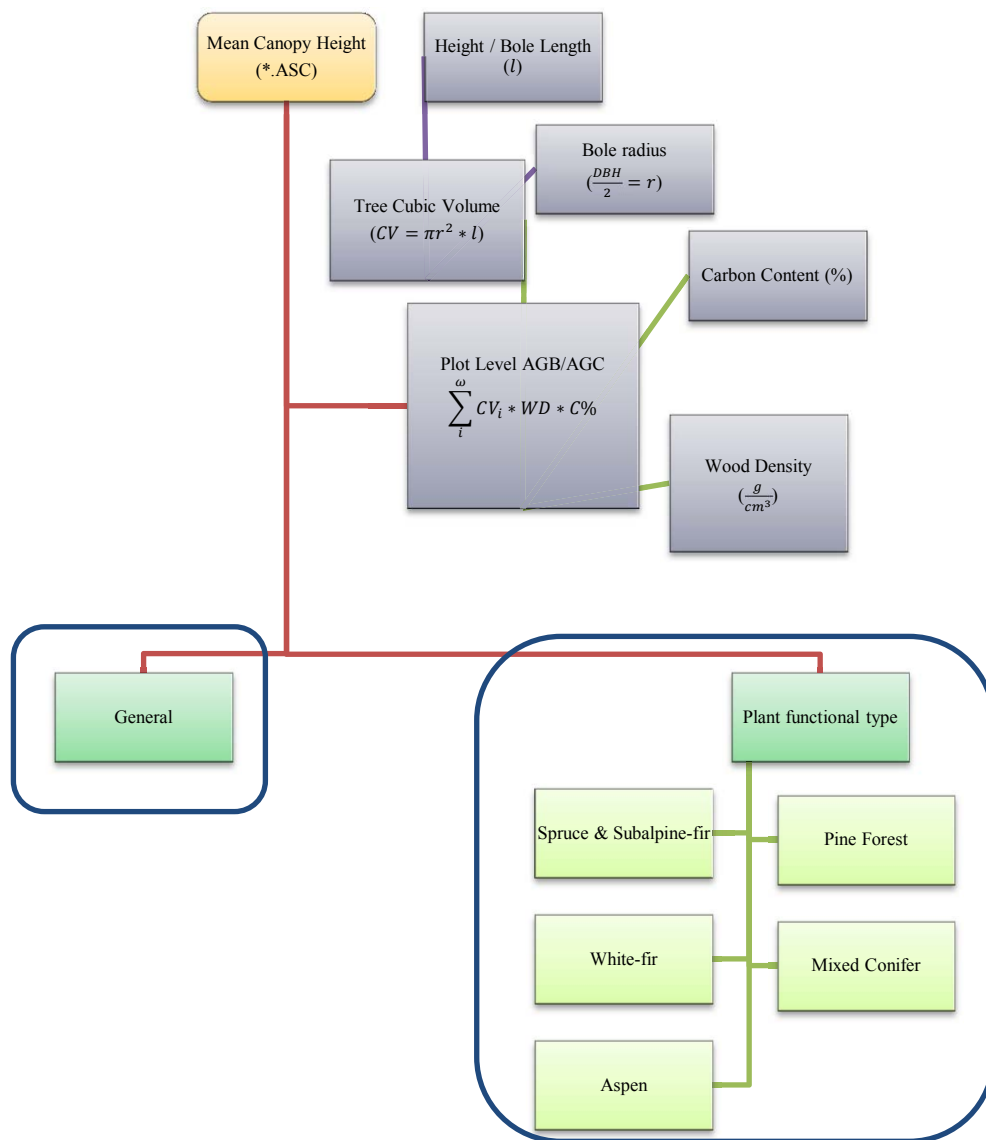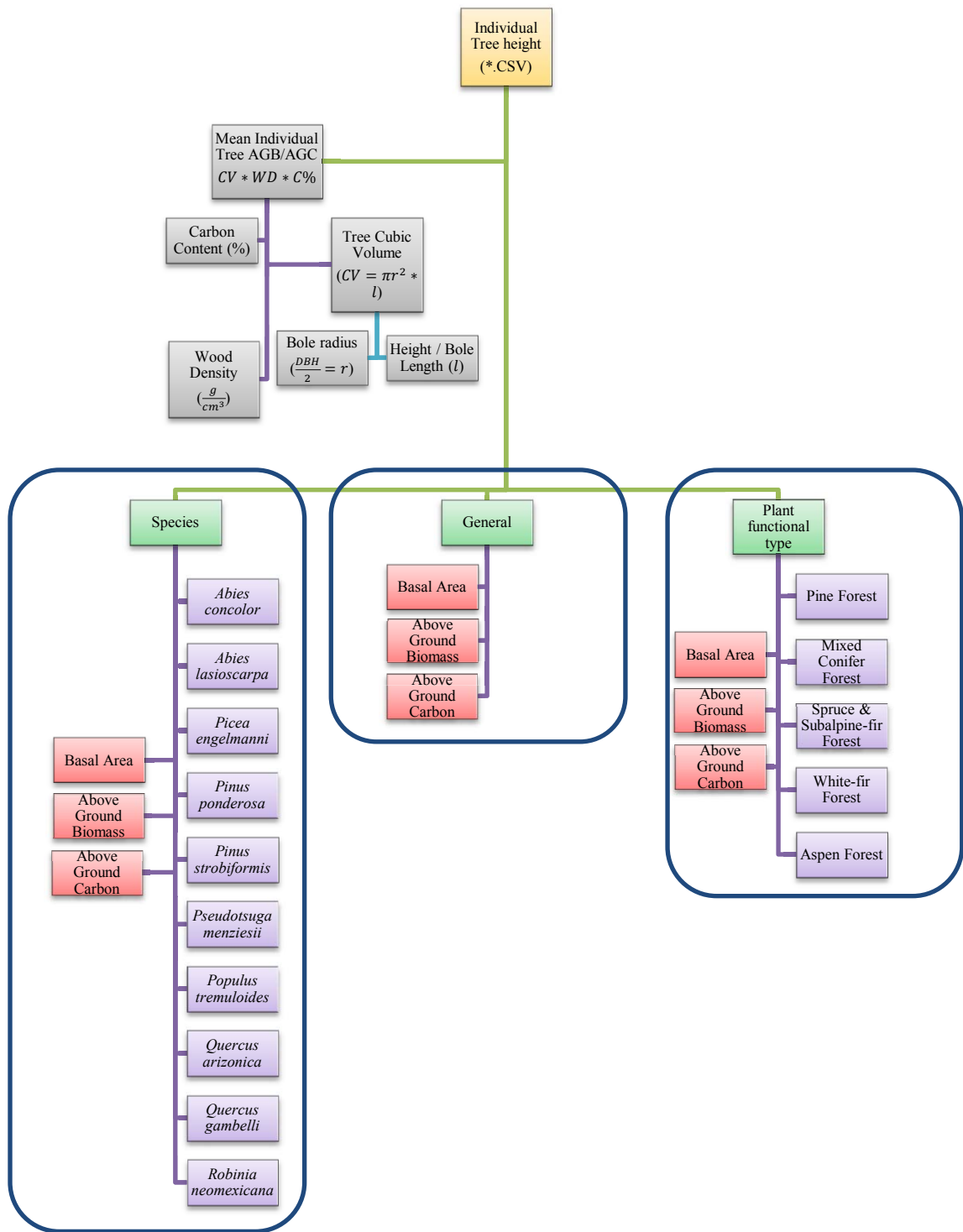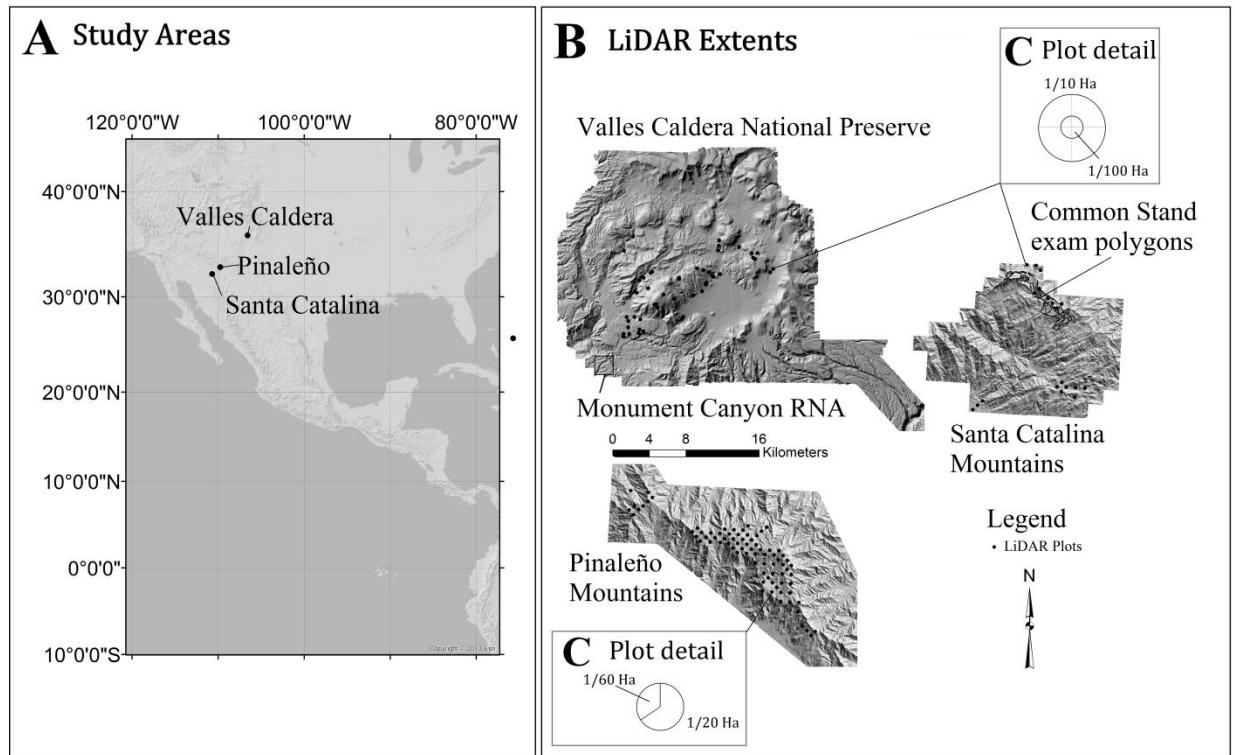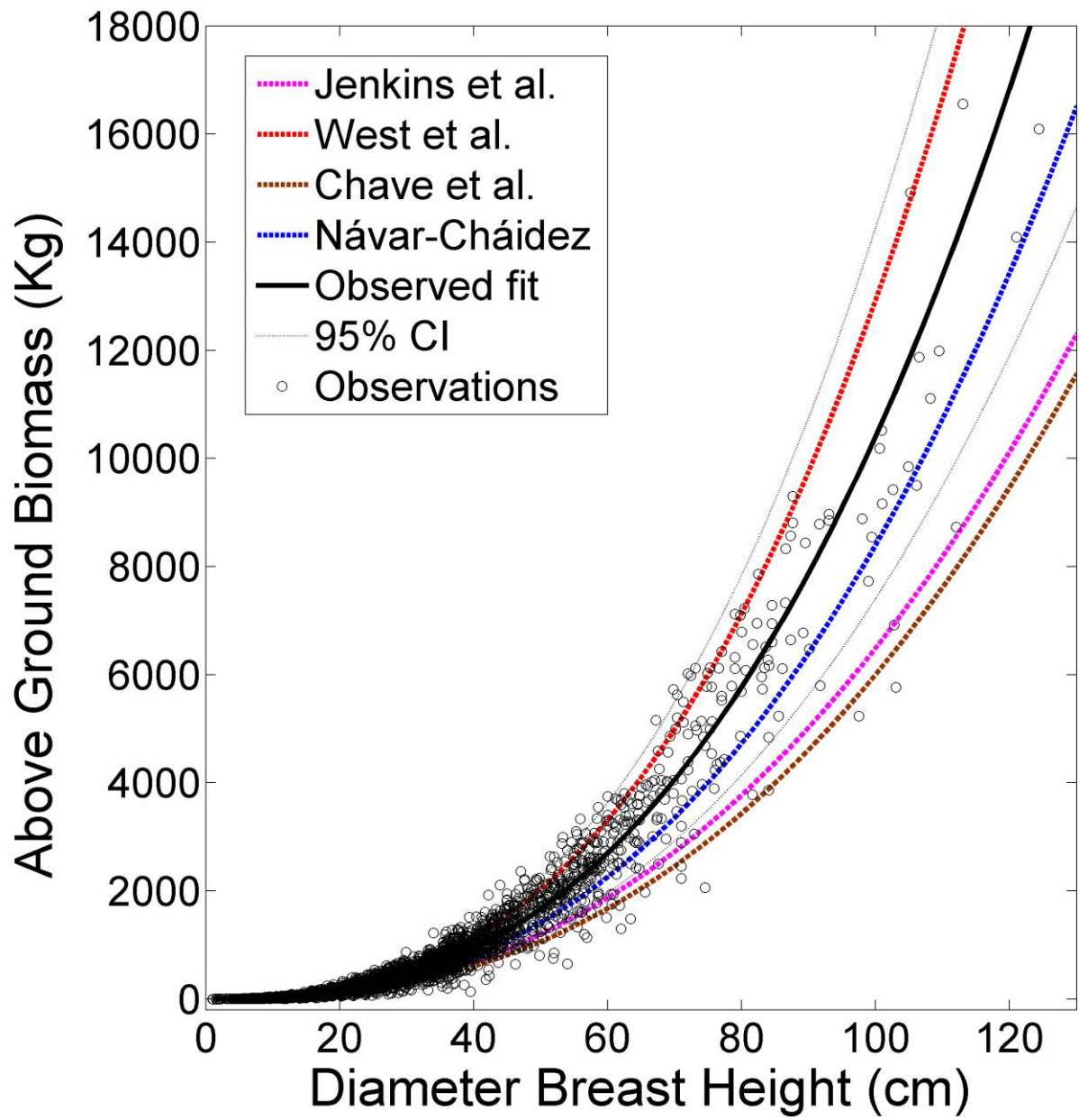
Figures

Figure 1.

Figure 2.

Figure 3.
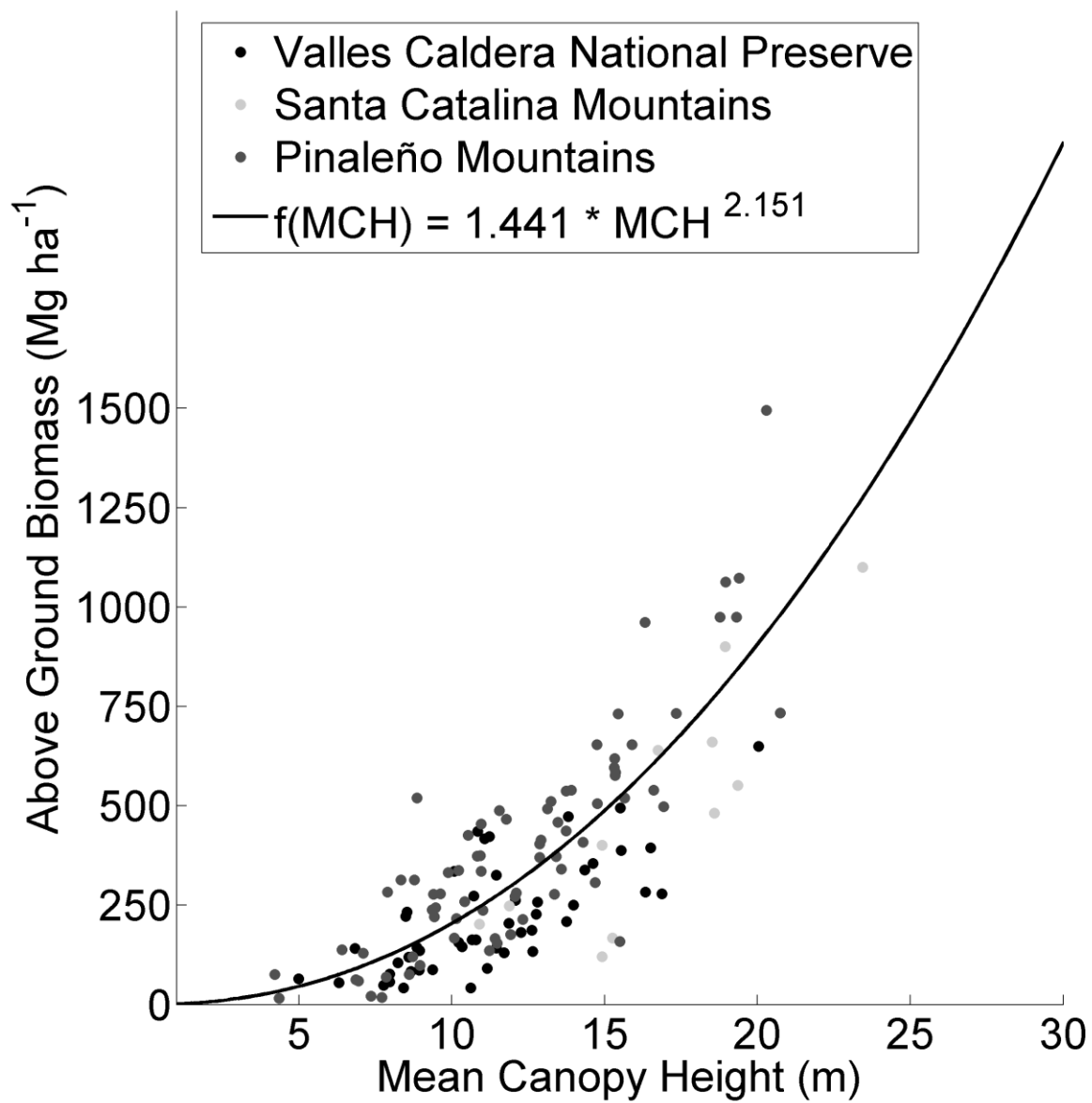
Figure 4.
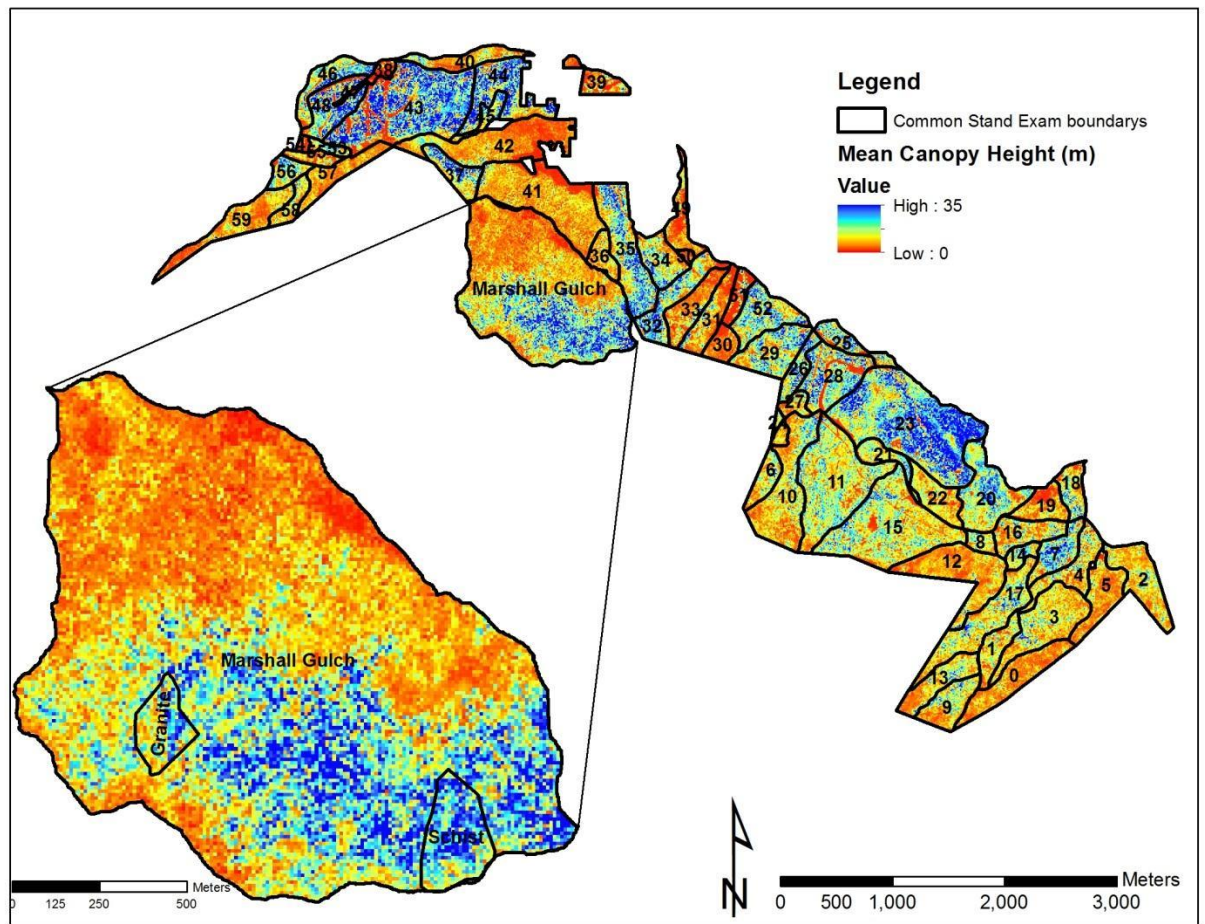
Figure 5.
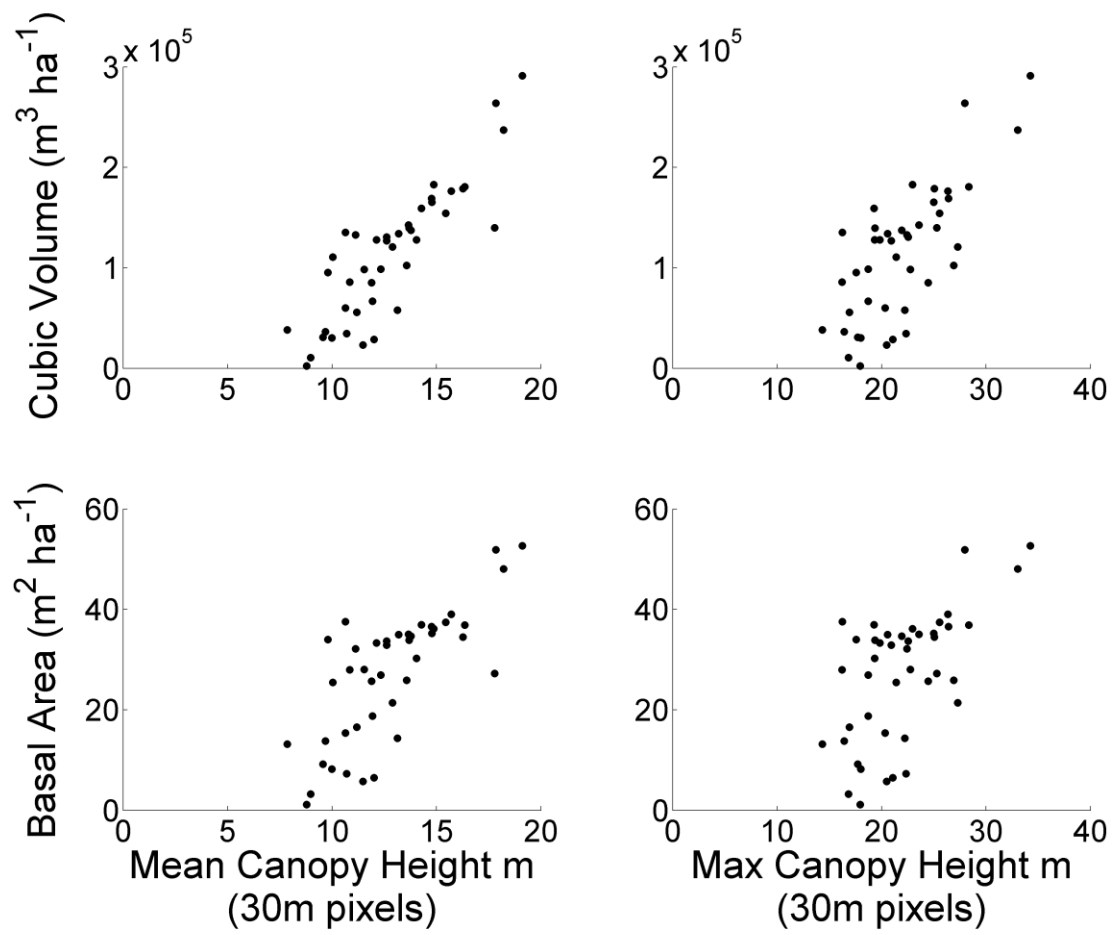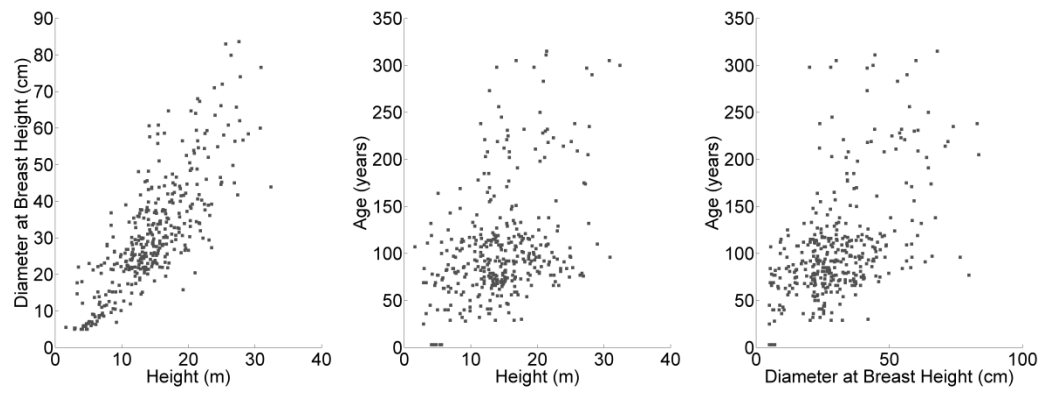
Figure 6.

Figure 7.

Figure 8.

Figure 9.

Tables

Table 1: Location, climate, geology, and five possible PFT-settings (Ponderosa Pine = PP, Mixed-Conifer = MC, White-fir dominated = WF, Spruce and Fir = SF, and Aspen disclimax = AD) of the three study areas.

Site	Latitude, Longitude	Elevation m amsl	Mean annual temp. (low-high C°)	Mean annual precipitation $\frac{mm}{yr^{-1}}$	Soils	Eco-region	Plant Functional Types
Santa Catalina Mountains	32.4° N, 110.7° W	900 – 2,700	-3.5°– 24°	377 – 850	granite, diorite, schist, andesite, shale and slate, quartzite, and limestone	Madrean Sky Island	PP, WF, MC
Pinaleño Mountains	32.7° N, 109.9° W	1,472 – 3,267	-6° – 24°	322 – 850	mylonitic gneiss, granite, and diorite	Madrean Sky Island	PP, WF, MC, SF, AD
Jemez Mountains	35.8° N, 106.5° W	2,234 – 3,431	-15° – 27°	476 – 790	Rhyolite, Ignimbrite, Tuff, and Pumice	Southern Rockies	PP, WF, MC, SF, AD

Table 2: Vegetation Classification schemes common for all three study areas with the suggested biomass model for each classification scheme: PFT (Smith et al. 1993), EVT (Comer et al. 2003, Brohman and Bryant 2005, Muldavin et al. 2006, LANDFIRE 2010), mid-scale dominance (Mellin et al. 2008), plant associations (Stuever and Hayden 1997), PNVT (Nature Conservancy 2006, 2007), and ReGap (Lowry et al. 2007, Prior-Magee et al. 2007) .

Biomass PFT Model	Landfire Existing Vegetation Type (EVT)	USFS Mid-Scale Dominance Type	USFS Plant Habitat Type (Potential Associations)	USFS Potential Natural Vegetation Type (PNVT)	Southwest Regional ReGAP Ecological System	Whittaker and Niering (1975), Niering and Lowe (1984)
Spruce-fir	Inner-Mountain Basins Aspen-Mixed Conifer forest and woodland	Upper evergreen forest tree mix	Engelmann Spruce, Corkbark Fir Series	Spruce-fir forest	Rocky Mountain Subalpine Mesic Spruce-fir forest and woodland	Subalpine fir forest
Mixed-Conifer	Southern Rocky Mtn. Dry-Mesic Montane Mixed Conifer forest and woodland, Rocky Mtn. Montane Riparian	Upper evergreen forest tree mix	Douglas-fir and Limber Pine Series	Mixed conifer – frequent fire, Mixed conifer with aspen	Madrean Upper Montane Conifer-Oak Forest and Woodland, Rocky Mtn. Aspen forest and woodland	North-slope montane fir forest, Drier montane fir forest
Ponderosa pine	Madrean Lower Montane Pine-Oak forest and woodland, Madrean Pinyon-Juniper woodland	Ponderosa pine, Oak/Juniper/Pinyon mix, Upper Pine-Oak	Ponderosa pine Series, Apache Pine and Chihuahua Pine Series	Madrean Pine-Oak Woodland, Ponderosa pine-Evergreen Oak, Ponderosa pine forest	Madrean Lower Montane Pine-Oak Forest and Woodland, Madrean Pinyon-Juniper woodland, Madrean Encinal, Mogollon Chaparral	High-elevation pine forest, Low-elevation pine forest, Pine-oak forest, Pine-oak woodland
White fir		Upper evergreen forest tree mix	White Fir Series			Mesic ravine fir forest
Aspen				Mixed conifer with aspen	Rocky Mountain Aspen forest and woodland	Successional aspen

Table 3: Average Wood Density (µWD) and average carbon percentage (µC%) based on the distribution of species observed. The species are *Abies concolor* – ABCO, *Abies lasciocarpa* – ABLA, *Picea pungens* – PIPU, *Picea engelmannii* – PIEN, *Pinus ponderosa* – PIPO, *Pinus strobiformis* – PIST, *Populus tremuloides* – POTR, *Pseudotsuga menziesii* – PSME, *Quercus spp.* – QU**, *Robinia neomexicana* – RONE. Wood density estimates were taken from Jenkins et al. (2004), Smith et al. (2006), Brown (2008) ,and Miles and Smith (2009); carbon values are from Lamlom and Savidge (2004) (Supplementary Materials Appendix B).

Species	ABCO	ABLA	PIPU	PIEN	PIPO	PIST	POTR	PSME	QU**	RONE		
Wood density	0.35	0.35	0.38	0.38	0.44	0.44	0.37	0.45	0.64	0.66		
Est. C % content	50.08	50.08	50.39	50.39	52.47	52.47	47.09	50.5	49.63	49.63		
Study Location					Observed Frequency (%)						**µWD**	**µC%**
Santa Catalina	8.64	--	--	--	39.09	4.94	--	19.75	27.57	--	0.489	51.09
Pinaleño	15.88	8.61	--	10.3	8.83	13.81	12.79	23.74	4.28	1.75	0.418	50.34
Valles Caldera	13.33	4.19	2.02	23.68	29.36	0.95	12.45	12.58	1.42	--	0.404	50.55

Table 4: Reported AGB and AGC data from other studies in Arizona and New Mexico. The Woolsey data are from the Mogollon Rim of Arizona and New Mexico. The Merkel (1954) and Sesnie et al. 2009 are from the Kaibab Plateau. The Whittaker and Niering (1975) and Niering and Lowe (1984) evaluated the same study plots in the Santa Catalina and Pinaleño. The estimates from Smith et al. 2006 (which were reported in T C/ha) were doubled to reflect AGB instead of AGC and are a representation of national level estimates. The Anderson-Teixera et al. (2011) data are from across New Mexico.

Forest/ Woodland Types	Woolsey 1911	Merkel 1954	Sesnie et al. 2009	Niering and Lowe 1984	Whittaker and Niering 1975	Smith e al. 2006 AGB $Mg\ ha^{-1}$	Smith e al. 2006 AGC $Mg\ ha^{-1}$	Anderson-Teixeira et al. 2011 AGB $Mg\ ha^{-1}$	Anderson-Teixeira et al. 2011 AGC $Mg\ ha^{-1}$
	Basal Area estimates $m^2\ ha^{-1}$				(AGB /TB) $Mg\ ha^{-1}$ %				
							50th to 99th % Living + Dead		
Pinus/Quercus (hypoleucoides)	--	--	--	30.5	34.9	163 / 200 81.5%	--	--	59.6 +66.2	29.7 +33.1 (Pinon-Juniper)
Pinus ponderosa / strobiformis	13.8 − 45.5	--	--	34.9	39.4	161 / 190 84.2%	106 +14 282 +22	53 +7 141 +11	152.8 +53.4	76.4 +26.7
Pseudotsuga menz.	--	--	--	72.2	70.5	438 / 520 84.2%	230 +40 540 +80	115 +20 270 +40	--	--
Pseudotsuga / Abies con.	--	~55	29.2 − 34.1	98.5	118.5	790 / 920 85.9%			--	--
Abies concolor	--	--	--	66.1	58.6	361 / 420 85.95%	376 +64 542 +98	188 +32 271 +49	266 +262	133 + 131
Abies lasiocarpa / Picea engelmannii	--	--	--	92.2	57.8	357 / 420 85.0%				
Populus tremuloides	--	--	--	N/A	31.6	126 / 200 63.0%	178 +34 404 +64	89 +17 202 +32	--	--

Table 5: Observed mean and standard error of the observed AGB, AGC, and basal area by study area where N is the number of plots. Average wood density and carbon content were estimated by the proportion of individual species observed in each study area (Supplemental Materials Appendix B).

Site	N	Size (radius)	AGB $\pm$ SE ($Mg\ ha^{-1}$)	AGC $\pm$ SE ($Mg\ ha^{-1}$)	Basal area $\mu \pm SE$ ($m^2\ ha^{-1}$)	μ Wood Density ($g\ cm^{-3}$)	μ C%
Santa Catalina	13	$1000\ m^2$ (17.86 m)	226.66 ± 125.84	113.33 ± 62.92	41.2 ± 20.3	0.489	51.09
Pinaleño	72	$500\ m^2$ (12.62 m)	384.38 ± 284.52	193.29 ± 143.83	49.5 ± 26.3	0.404	50.55
Valles Caldera	48	$1000\ m^2$ (17.86 m)	219.05 ± 140.26	109.54 ± 70.13	32.7 ± 15.3	0.418	50.34

Table 6: The observed AGB, AGC, and basal areas by PFT where N is the number of plots. Average wood density and carbon content were estimated by the proportion of individual species observed in each PFT (Supplemental Materials Appendix B).

PFT	N	μ AGB $\pm$ SE ($Mg\ ha^{-1}$)	μ AGC $\pm$ SE ($Mg\ ha^{-1}$)	μ Basal area $\pm SE$ ($m^2\ ha^{-1}$)	μ Wood Density ($g\ cm^{-3}$)	μ C%
Ponderosa pine	21	173.6 ± 95.2	91.1 ± 50.0	27.15 ± 13.93	0.44	52.47
Mixed-Conifer	36	275.6 ± 158.7	137.8 ± 79.4	37.94 ± 17.35	0.45	50.0
White-fir	41	453.9 ± 265.0	242.9 ± 144.0	55.32 ± 27.76	0.40	50.0
Spruce-fir	32	304.6 ± 237.5	152.9 ± 119.2	39.94 ± 21.96	0.36	50.2
Aspen	10	388.4 ± 272.8	186.4 ± 130.9	51.11 ± 25.97	0.35	48.0

Table 7: Individual tree DBH to AGB, by study area; other published models for AGB listed below, the '--' represent values not reported in the other publications.

MODEL:		$AGB = B_2 DBH^{\alpha_2}$				$AGB = B_3 HT^{\alpha_3}$			
Location/Study	N	$\beta_2 \pm SE$	$\alpha_2 \pm SE$	$RMSE\ kg$	R^2	$\beta_3 \pm SE$	$\alpha_3 \pm SE$	$RMSE\ kg$	R^2
Santa Catalina	226	0.074 ± 0.018	2.599 ± 0.056	335.3	0.977	0.007 ± 0.007	3.940 ± 0.282	927.6	0.822
Pinaleño	2,173	0.154 ± 0.016	2.383 ± 0.024	294.2	0.950	0.028 ± 0.012	3.584 ± 0.136	809.3	0.625
Valles Caldera	1,313	0.063 ± 0.007	2.615 ± 0.028	241.2	0.953	0.017 ± 0.006	3.709 ± 0.112	581.1	0.729
General	3,514	0.111 ± 0.009	2.470 ± 0.019	307.5	0.9468	0.023 ± 0.006	3.622 ± 0.081	748.1	0.685
Jenkins et al. 2004	--	0.0793	2.435	--	--	--	--	--	--
Chave et al. 2005	--	--	2.500	--	--	--	--	--	--
Brown 1997	--	0.124	2.530	--	--	--	--	--	--
Návar-Cháidez 2010	--	0.0597	2.574± 0.026	--	0.86	--	--	--	--
MST Theoretical	--	--	2.666	--	--	--	4	--	--

Table 8: Qualitative tree health categories with least-squares regression of DBH versus height $HT = \beta D^{\alpha}$; these models do not differentiate between species or location. The healthier the tree was considered to be [e.g. having a large decadent canopy volume with a straight standing bole] the greater the α and the closer in value to the predicted MST $\alpha = 2/3$.

DBH to Height	n	$\beta \pm SE$	$\alpha \pm SE$	RMSE m	R^2
Excellent	293	1.366 ± 0.269	0.696 ± 0.052	3.52	0.711
Good	857	1.626 ± 0.164	0.661 ± 0.027	3.46	0.702
Fair	871	2.07 ± 0.189	0.590 ± 0.026	3.14	0.733
Poor	1,109	1.83 ± 0.208	0.599 ± 0.030	4.00	0.635
All	3,130	1.826 ± 0.108	0.619 ± 0.016	3.65	0.689
MST	--	--	0.6667	--	--

Table 9: PFT-level models for individual tree AGB based on either the observed DBH or height (HT).

MODEL:		$AGB = B_2 DBH^{\alpha_2}$				$AGB = B_3 HT^{\alpha_3}$			
PFT	n	$\beta_2 \pm SE$	$\alpha_2 \pm SE$	RMSE kg	R^2	$\beta_3 \pm SE$	$\alpha_3 \pm SE$	RMSE kg	R^2
Pine Forest	511	0.055 ± 0.009	2.640 ± 0.037	186.3	0.970	0.010 ± 0.006	3.935 ± 0.190	526.7	0.760
Mixed-Conifer	760	0.096 ± 0.020	2.493 ± 0.048	363.5	0.921	0.027 ± 0.013	3.605 ± 0.143	668.2	0.717
White-fir	1,177	0.118 ± 0.017	2.448 ± 0.033	345.6	0.956	0.039 ± 0.017	3.455 ± 0.129	779.8	0.755
Spruce & Fir	1,001	0.072 ± 0.012	2.537 ± 0.038	195.6	0.952	0.036 ± 0.019	3.443 ± 0.164	504.5	0.685
Aspen	378	0.157 ± 0.025	2.358 ± 0.038	160.8	0.968	0.006 ± 0.009	3.947 ± 0.479	643.5	0.492

Table 10: Species models for individual tree aboveground biomass. AGB models by species based on observed DBH.

MODEL:		$AGB = B_2DBH^{\alpha_2}$				$AGB = B_3HT^{\alpha_3}$			
Species	*n*	$\beta_2 \pm SE$	$\alpha_2 \pm SE$	*RMSE kg*	R^2	$\beta_3 \pm SE$	$\alpha_3 \pm SE$	*RMSE kg*	R^2
Abies concolor	565	0.144 ± 0.032	2.358 ± 0.052	215.2	0.944	0.054 ± 0.033	3.357 ± 0.171	456.9	0.747
Abies lasciocarpa	247	0.064 ± 0.018	2.571 ± 0.076	75.8	0.942	0.252 ± 0.171	2.597 ± 0.223	147.8	0.735
Picea engelmanni	546	0.155 ± 0.039	2.334 ± 0.063	1740.0	0.923	0.195 ± 0.094	2.837 ± 0.156	293.1	0.734
Pinus Ponderosa	654	0.054 ± 0.008	2.651 ± 0.035	251.6	0.965	0.036 ± 0.020	3.503 ± 0.167	721.5	0.708
Pinus strobiformis	343	0.091 ± 0.028	2.510 ± 0.073	282.2	0.935	0.450 ± 0.336	2.696 ± 0.234	665.0	0.640
Pseudotsuga menziesii	730	0.181 ± 0.038	2.371 ± 0.047	517.0	0.949	0.046 ± 0.026	3.477 ± 0.167	1,123.0	0.752
Populus tremulodies	404	0.258 ± 0.047	2.215 ± 0.048	98.3	0.939	0.357 ± 0.341	2.457 ± 0.313	266.5	0.524
Quercus arizonica	55	0.107 ± 0.161	2.323 ± 0.410	122.3	0.751	11.210 ± 6.581	1.714 ± 0.248	139.1	0.677
Quercus gambelli	65	0.057 ± 0.039	2.563 ± 0.168	93.6	0.965	0.117 ± 0.392	3.629 ± 1.359	383.8	0.401
Robinia neomexicana	36	0.044 ± 0.029	2.763 ± 0.210	11.0	0.983	0.240 ± 0.284	3.059 ± 0.516	25.22	0.907

Table 11: Study area and general AGB model $AGB = \beta_1 MCH^{\alpha_1}$, where the independent variable is the FUSION CloudMetric estimated mean canopy height (MCH), β_1 is the normalization constant, α_1, is the scaling parameter or dynamic exponent. The area of the plots in the Pinaleño was 500 m^2 and in the Valles Caldera 1,000 m^2; the General model also uses 14 1,000 m^2 plots from the Santa Catalina.

Location	N	$\beta_1 \pm SE$	$\alpha_1 \pm SE$	r^2	RMSE Units
Valles Caldera	48	3.165 ± 3.970	1.732 ± 0.477	0.545	95.6 $Mg\ ha^{-1}$
Pinaleño	72	1.613 ± 1.420	2.156 ± 0.317	0.749	140.5 $Mg\ ha^{-1}$
General	133	1.441 ± 1.236	2.151 ± 0.313	0.635	151.1 $Mg\ ha^{-1}$

Table 12: PFT level AGB models from all three study areas by plot (N) with least-squares regression of the MCH: $AGB = \beta_1 MCH^{\alpha_1}$ of each plot.

PFT	N	$\beta_1 \pm SE$	$\alpha_1 \pm SE$	r^2	RMSE units
Pine	18	0.949 ± 1.057	2.142 ± 0.491	0.828	60.4 $Mg\ ha^{-1}$
Mixed Conifer	27	2.362 ± 4.712	1.901 ± 0.769	0.536	106.2 $Mg\ ha^{-1}$
White fir	31	1.461 ± 2.200	2.196 ± 0.537	0.722	155.9 $Mg\ ha^{-1}$
Spruce & Fir	26	1.430 ± 1.785	2.239 ± 0.478	0.815	92.87 $Mg\ ha^{-1}$
Aspen	9	0.456 ±1 .481	2.578 ± 1.151	0.837	120.7 $Mg\ ha^{-1}$