AN ABSTRACT OF THE DISSERTATION OF

Krishna P. Poudel for the degree of Doctor of Philosophy in Forest Resources presented on May 28, 2015.

Title: Strategies for Sampling and Estimation of Aboveground Tree Biomass

Abstract approved:

Temesgen Hailemariam

The issue of global climate change and an increasing interest in the reduction of fossil fuel carbon dioxide emissions by using forest biomass for energy production has increased the importance of quantifying forest biomass in recent years. The official U.S. forest carbon reporting is based on the forest biomass estimates obtained from the equations, sample tree measurements, and forest area estimates of the U.S. Forest Service, Forest Inventory and Analysis (FIA). These biomass estimates differ from the estimates obtained from regional and other commonly used biomass equations and the difference is more evident in the component biomass estimates.

In this dissertation, I assessed the efficiency of different sampling strategies to estimate crown biomass using data collected destructively from sampled trees. In terms of bias and root mean squared errors (RMSE), the stratified random sampling with probability proportional to branch basal diameter was better than other methods when 3 or 6 branches per tree are sampled but a systematic sampling with ratio estimation technique produced the smallest RMSE when 9 or 12 branches per tree are sampled.

Total and component aboveground biomass estimates were obtained using the existing approaches and locally fitted equations based on the data collected in this study. The use of existing equations resulted in biased component biomass estimates along with higher RMSE. The locally fitted system of component biomass equations with seemingly unrelated regression (SUR) provided better estimates than existing equations. The need to use other explanatory variables in addition to the diameter at breast height (DBH) to estimate component biomass was justified by decrease in RMSE. Beta, Dirichlet, and multinomial loglinear regressions that predict proportion of biomass in each component were unbiased and produced lower RMSEs compared to the SUR methods for most of the species-component combinations.

Three different methods for adjusting regional volume and component biomass equations were applied. All the adjustment methods were able to improve the performance of regional equations. Based on the leave one out cross validation, the RMSEs in cubic volume including top and stump (CVTS) and component biomass estimation were similar for the adjustments from a correction factor based on ordinary least square (OLS) regression through origin and an inverse approach. The adjustment based on OLS with intercept did not perform as well as the other two adjustment methods. When only one tree is available for calibration of regional models, we found it useful to use the tree with maximum DBH to calibrate regional CVTS and bark biomass equations and the dominant tree to calibrate bole, foliage, and branch biomass rather than to use randomly selected one tree.

Strategies for Sampling and Estimation of Aboveground Tree Biomass

by
Krishna P. Poudel

A DISSERTATION

Submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented May 28, 2015
Commencement June 2015

Doctor of Philosophy dissertation of Krishna P. Poudel presented on May 28, 2015.

APPROVED:

Major Professor, representing Forest Resources

Head of the Department of Forest Engineering, Resources, and Management

Dean of the Graduate School

I understand that my dissertation will become a part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Krishna P. Poudel, Author

ACKNOWLEDGEMENTS

I am indebted to my major professor Dr. Temesgen Hailemariam for his constant support and guidance throughout my PhD. Completion of this dissertation would not have been possible without his continued help and encouragement. My special gratitude goes to other members of my advisory committee Drs. Lisa Madsen, Andy Gray, Glen Murphy, and Cynthia Ocamb. I would also like to extend my sincere thanks to my MS Statistics committee members Drs. Paul Murtaugh and Alix Gitelman of the Department of Statistics.

Many thanks to my current and previous Forest Measurement and Biometrics lab mates including Jacob Strunk, Mike Shettles, Lacey Jeroue, Joonghoon Shin, Elijah Allensworth, Steve Huff, Michael Hoe, and Will Pollock. This journey would not have been as pleasant without those small chats, coffee breaks, and small walks around the campus.

I would also like to thank Patrick Duyck, Jeremiah Moore, and Claire Cvitanovich for their help in data collection. Special thanks to Jeff Wimer and his student logging crew of summer 2012 for their help in felling, bucking, and weighing trees.

I am also thankful to my fellow Nepalese Students at Oregon State University. Special thanks to Subin Sahu, Deepak Rayamajhi, Satish Serchan, Birat Sapkota, Gehendra Bhattarai, Mukesh Bhattarai, Dinesh Timilsina, G.S. Khatiwada, Suman Pant, and Roshan Adhikari.

I am most grateful to my wife Anusha Rijal and my family and friends in Nepal for their love, support, and inspiration throughout this journey.

CONTRIBUTION OF AUTHORS

Dr. Temesgen Hailemariam contributed at all phases of the project including the critical revision of manuscripts, coding and data analysis, and edited the manuscripts. Dr. Andrew Gray and Dr. Lisa Madsen provided valuable statistical and analytical advice and critical revision of the manuscripts. Dr. Glen Murphy and Dr. Cynthia Ocamb provided extensive critique on the dissertation.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF TABLES

Strategies for Sampling and Estimation of Aboveground Tree Biomass

Krishna P. Poudel

1. Introduction

1.1. Background

The potential contribution of forest ecosystems in mitigating climate change can be realized through sustainable forest management. Although there is a consensus that climate change is affecting forests around the world, less is known of forest dynamics under the changing climate (Liang et al. 2011). Accordingly, forest management will face new and evolving challenges in balancing interests between wood production and ecosystem services (Temesgen et al. 2015). Amount of biomass available in a forest is a critical factor that determines its role in climate change mitigation (Brown et al. 1995). Additionally, information on forest biomass is essential to calculate the carbon flux due to deforestation and carbon sink that results from reforestation (Houghton 2005).

Biomass, in general, includes both above and below ground living and dead mass of trees, shrubs, vines, and roots (Lu 2006) and is expressed as oven-dry weight (Brown 1997). The components of aboveground tree biomass are main stem, stem bark, bare branch, and leaves or needles while the belowground tree biomass components are coarse roots and fine roots (Fonseca et al. 2012). The biomass in these different components can serve different purposes. For example, the biomass in stem wood is important in timber sale while the biomass in crown is useful in fuel load assessment, formulating fire management strategies and in developing wildfire models (He et al. 2013). The biomass in small branches and leaves or needles is essential in assessing the available feedstock for bioenergy plants.

Forest management decisions rely on the quality, quantity, and distribution of forest resources. This information is obtained from systematic collection of the data that characterize stands, a process known as forest inventory. Due to the extensive nature of forest resources, measurement of the entire resources is neither practical nor economically feasible. Sampling is the process of selecting a subset of a

population to observe so as to estimate the characteristics of the whole population (Thompson 2002). The use of sampling strategies in forestry has a long history. Budget limitations and lack of resources are two major reasons that we use sampling instead of total enumeration. Additionally, measurement of some variables requires the destruction of itself, an excellent example of which is the measurement of biomass.

Because the direct estimation of crown biomass is not only tedious and laborious but also economically not feasible, different sampling approaches have been used to quantify this biomass component. These approaches range from the methods that do not require any additional measurements such as simple random sampling, systematic sampling, and stratified random sampling to more complex designs that require detailed auxiliary information such as probability proportional to size sampling and randomized branch sampling (Temesgen et al. 2011).

Simple random sampling is the basic building block and is generally used as a reference to evaluate the performance of other sampling designs. The stratified sampling provides samples that are more representative of the target population by grouping the population into homogenous strata. Snowdon (1986) found improved accuracy from stratification based on crown position in estimating crown biomass compared to the simple random sampling. Temesgen (2003) found stratified random sampling to result in the lowest mean squared error in estimating total leaf area in a tree. Temesgen et al. (2011), on the other hand, found the systematic sampling with ratio estimation as the most efficient method to estimate of individual tree foliage biomass.

The randomized branch sampling (RBS), originally introduced by Jessen (1955) to estimate fruit counts on individual orchard tree, has also been used by many researchers to quantify biomass in tree crown (e.g. Valentine et al. 1984; Williams 1989; Gregoire et al. 1995; Schlecht and Affleck 2014). One advantage of the RBS is that the samples are sequentially collected along the path starting from the

lowest branch thus counting the total number of branches beforehand is not necessary. However, as the RBS was designed for trees with decurrent crown structures, the optimal placement of RBS decision nodes is less apparent in the conifers (Schlecht and Affleck 2014). They found the performance of RBS to be intermediate between the probability proportional to size sampling and the simple random sampling in estimating green crown mass in Douglas-fir and western larch. Chapter 2 of this dissertation evaluates different sampling strategies for estimating crown biomass i.e. total dry weight of foliage and branch with bark.

The U.S. forest carbon inventories until 2009 were based on the tree biomass estimates obtained from Jenkins et al. (2003) biomass equations along with the U.S. Forest Service, Forest Inventory and Analysis (FIA) sample tree measurements and forest area estimates (Heath et al. 2008). Recent U.S. official carbon inventories are based on tree biomass estimates obtained from the component ratio method described by Woodall et al. (2011) and understory vegetation biomass obtained from Jenkins et al. (2003) equations (U.S. EPA, 2015).

The Jenkins et al. (2003) equations were developed by modified meta-analysis of the compiled diameter-based equations for total aboveground biomass. Merchantable stem wood, stem bark, and foliage biomass in their method is estimated as the proportion of total aboveground biomass while the biomass in stump and branches is calculated by subtraction. The Jenkins et al. (2003) equation for total aboveground biomass is a single entry equation that only uses DBH as predictor variable. With their equation form, biomass continues to increase as diameter increase (Heath et al. 2008) and does not account for the variation in stem form (Zhou and Hemstrom 2009).

The need for accurate and consistent methods of estimating total and component aboveground biomass is undoubted. In 2009, the FIA updated its biomass estimation protocols by switching to the component ratio method (CRM) to estimate biomass of medium and large trees. The CRM was proposed

for consistent national biomass estimation based on FIA volume estimates (Zhou and Hemstrom 2009) however because different FIA units use different volume equations, the biomass for trees of same species with same diameter differs by regions (Heath et al. 2008). The CRM calculates biomass in merchantable bole by converting sound wood volume to bole biomass using wood specific gravity factors compiled by Miles and Smith (2009). Similarly, the biomass in bole bark is based on the percent bark and bark specific gravities given in Miles and Smith (2009). Biomass in stump wood and bark is based on volume equations in Raile (1982) and the compiled set of wood and bark specific gravities. The biomass of tops and limbs is calculated by subtraction and using a CRM adjustment factor. Total aboveground biomass is obtained by summing these component masses.

The Pacific Northwest unit of the FIA uses its specific set of equations to calculate total and component biomass. The stem wood biomass is calculated from the cubic volume estimates and wood density factors. Each tree species is associated with a set of local volume and component biomass equations (Zhou and Hemstrom 2009). The regional models, however, may not be unbiased at the local scale if there is spatial variation in the tree form due to one or more unknown predictors and this regional bias could be reduced or removed if the models are localized to each sub-region or subarea (Räty and Kangas 2008).

When the aboveground biomass is estimated, the accuracy of the component models is different than the accuracy of the models that estimate total aboveground biomass. The biomass in each of the components can be estimated as the proportion of the total aboveground biomass. Proportions are bounded between zero and one thus can be modeled using regressions that assume distributions other than normal. Two commonly used distributions to model percentage or proportions are the beta and Dirichlet distributions. The beta distribution has been used in forestry to characterize the understory vegetation cover of the riparian area by Eskelson et al. (2011) and to model percent

canopy cover in a conifer forest by Korhonen et al. (2007). Chapter 3 of this dissertation compares the performances of different approaches to estimate total and component aboveground biomass.

Since biomass estimation requires destructive sampling, the use of allometric equations is essential. Additionally, the assessment of the performance of these methods using independent dataset is tedious in itself. On the other hand, the amount of data and the number of equations needed to precisely estimate biomass is unknown. If a small sample of biomass data is available for the area for which the biomass estimates are desired, the regional models can be adjusted or calibrated.

In an effort to develop models that are suitable for prediction of biomass at large scale, de-Miguel et al. (2014) proposed mixed-effects meta-models that can be calibrated for local conditions by sampling at least one tree for component biomass. Previously, adjustment factors based on ordinary least squares regression have been used to adjust regional height-diameter models (e.g. Temesgen et al. 2008; Garber et al. 2009). However, these methods have not been used in adjusting biomass equations. Chapter 4 of this dissertation uses three different methods for adjusting regional volume and component biomass equations for Douglas-fir and lodgepole pine.

1.2. Organization of Dissertation

The overall objective of this dissertation is to explore the methods for sampling and estimating total and component aboveground biomass. The specific objectives are to (1) evaluate the sampling strategies to estimate crown biomass, (2) develop methods for estimating aboveground biomass and its components for five Pacific Northwest tree species, and (3) calibrate the volume and component biomass equations for Douglas-fir and lodgepole pine forests in Western Oregon. These objectives are addressed in Chapters 2, 3, and 4, respectively. Figures and tables are appended at the end of each Chapter.

1.3. References

Brown, I.F., L.A. Martinelli, W.W. Thomas, M.Z. Moreira, C.A. Cid Ferreira, and R.A. Victoria. 1995. Uncertainty in the biomass of Amazonian forests: An example from Rondônia, Brazil. *For. Ecol. Manage.* 75(1): 175-189.

Brown, S. 1997. Estimating biomass and biomass change of tropical forests: a primer. Rome: FAO Forestry Paper 134. Food and Agriculture Org.

de-Miguel S., T. Pukkala, N. Assaf, and Z. Shater. 2014. Intra-specific difference in allometric equations for aboveground biomass of eastern Mediterranean Pinus brutia. *Ann. For. Sci.* 71(1): 101-112.

Eskelson, B.N., L. Madsen, J.C. Hagar, and H. Temesgen. 2011. Estimating Riparian understory vegetation cover with Beta regression and copula models. *For. Sci.* 57(3): 212-221.

Fonseca, W., F.E. Alice, and J.M. Rey-Benayas. 2012. Carbon accumulation in aboveground and belowground biomass and soil of different age native forest plantations in the humid tropical lowlands of Costa Rica. *New Forests.* 43(2): 197-211.

Garber, S.M., H. Temesgen, V.J. Monleon, and D.W. Hann. 2009. Effects of height imputation strategies on stand volume estimation. *Can. J. For. Res.* 39(3): 681-690.

Gregoire, T.G., H.T. Valentine, and G.M. Furnival. 1995. Sampling methods to estimate foliage and other characteristics of individual trees. *Ecology*. 76(4): 1181-1194.

He, Q., E. Chen, R. An, and Y. Li. 2013. Above-ground biomass and biomass components estimation using LiDAR data in a coniferous forest. *Forests*. 4(4): 984-1002.

Heath, L.S., M. Hansen, J.E. Smith, P.D. Miles, and B.W. Smith. 2008. Investigation into calculating tree biomass and carbon in the FIADB using a biomass expansion factor approach. *McWilliams, WH, Moisen, GG & Czaplewski, RL (eds.)*.

Houghton, R.A. 2005. Aboveground forest biomass and the global carbon balance. *Glob. Change Biol.* 11(6): 945-958.

Jenkins, J.C., D.C. Chojnacky, L.S. Heath, and R.A. Birdsey. 2003. National-scale biomass estimators for United States tree species. *For. Sci.* 49(1): 12-35.

Jessen, R.J. 1955. Determining the fruit count on a tree by randomized branch sampling. *Biometrics.* 11(1): 99-109.

Korhonen, L., K.T. Korhonen, P. Stenberg, M. Maltamo, and M. Rautiainen. 2007. Local models for forest canopy cover with beta regression. *Silva Fennica.* 41(4): 671.

Liang, J.J., M. Zhou, D.L. Verbyla, L.J. Zhang, A.L. Springsteen, and T. Malone. 2011. Mapping forest dynamics under climate change: A matrix model. *For. Ecol. Manage.* 262(12): 2250-2262.

Lu, D. 2006. The potential and challenge of remote sensing-based biomass estimation. *Int. J. Remote Sens.* 27(7): 1297-1328.

Miles, P.D., and W.B. Smith. 2009. Specific gravity and other properties of wood and bark for 156 tree species found in North America. *Res. Note. NRS-38*. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, p 35.

Raile, G.K. 1982. Estimating stump volume. *Res. Pap. NC-224.* St Paul, MN: U.S. Department of Agriculture, Forest Service, p 4.

Schlecht, R., and D.L. Affleck. 2014. Branch aggregation and crown allometry condition the precision of randomized branch sampling estimators of conifer crown mass. *Can. J. For. Res.* 44(5): 499-508.

Snowdon, P. 1986. Sampling strategies and methods of estimating the biomass of crown components in individual trees of Pinus radiata D. Don. *Aust. For. Res.* 16(1): 63-72.

Temesgen, H. 2003. Evaluation of sampling alternatives to quantify tree leaf area. *Can. J. For. Res.* 33(1): 82-95.

Temesgen, H., V.J. Monleon, and D.W. Hann. 2008. Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. *Can. J. For. Res.* 38: 553-565.

Temesgen, H., D. Affleck, K. Poudel, A. Gray, and J. Sessions. 2015. A review of the challenges and opportunities in estimating above ground forest biomass using tree-level models. *Scand. J. Forest Res.* 30(4): 326-335.

Temesgen, H., V. Monleon, A. Weiskittel, and D. Wilson. 2011. Sampling strategies for efficient estimation of tree foliage biomass. *For. Sci.* 57(2): 153-163.

Thompson, S.K. 2002. Sampling. 2nd edn. John Wiley & Sons, Inc.: New York.

U.S. Environmental Protection Agency. 2015. Inventory of U.S. greenhouse gas emissions and sinks: 1990–2013. EPA 430-R-15-004., U.S. Environmental Protection Act, Washington, DC: U.S.

Valentine, H.T., L.M. Tritton, and G.M. Furnival. 1984. Subsampling trees for biomass, volume, or mineral content. *For. Sci.* 30(3): 673-681.

Williams, R.A. 1989. Use of randomized branch and importance sampling to estimate loblolly pine biomass. *South. J. Appl. For.* 13(4): 181-184.

Woodall, C., L.S. Heath, G.M. Domke, and M.C. Nichols. 2011. Methods and equations for estimating aboveground volume, biomass, and carbon for trees in the US forest inventory, 2010. Gen. Tech. Rep. NRS-GTR-88. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, p 30.

Zhou, X., and M.A. Hemstrom. 2009. Estimating aboveground tree biomass on forest land in the Pacific Northwest: a comparison of approaches. Res. Pap. PNW-RP-584. Portland, OR: U.S. Department of Agriculture, Forest Service. Pacific Northwest Research Station, p. 18.

2. Manuscript 1

Evaluation of Sampling Strategies to Estimate Crown Biomass

Krishna P. Poudel, Hailemariam Temesgen, and Andrew N. Gray

Springer-Verlag, GmbH

Forest Ecosystems

Tiergartenstr. 17

69121 Heidelberg, Germany

2:1

Abstract

Depending on tree and site characteristics crown biomass accounts for a significant portion of the total aboveground biomass in the tree. Crown biomass estimation is useful for different purposes including evaluating the economic feasibility of crown utilization for energy production or forest products, fuel load assessments and fire management strategies, and wildfire modeling. However, crown biomass is difficult to predict because of the variability within and among species and sites. Thus the allometric equations used for predicting crown biomass should be based on data collected with precise and unbiased sampling strategies. Using data collected from 20 destructively sampled trees, we evaluated 11 different sampling strategies using six evaluation statistics: bias, relative bias, root mean square error (RMSE), relative RMSE, amount of biomass sampled, and relative biomass sampled. We also evaluated the performance of the selected sampling strategies when different numbers of branches (3, 6, 9, and 12) are selected from each tree. Tree specific log linear model with branch diameter and branch length as covariates was used to obtain individual branch biomass.

Compared to all other methods stratified sampling with probability proportional to size estimation technique produced better results when 3 or 6 branches per tree were sampled. However, the systematic sampling with ratio estimation technique was the best when at least nine branches per tree were sampled. Under the stratified sampling strategy, selecting unequal number of branches per stratum produced approximately similar results to simple random sampling, but it further decreased RMSE when information on branch diameter is used in the design and estimation phases. Use of auxiliary information in design or estimation phase reduces the RMSE produced by a sampling strategy. However, this is attained by having to sample larger amount of biomass. Based on our finding we would recommend sampling 9 branches per tree to be reasonably efficient and limit the amount of fieldwork.

2.1. Introduction

The global issue of climate change and an increasing interest in the reduction of fossil fuel carbon dioxide emissions by using forest biomass for energy production has increased the importance of forest biomass quantification in recent years. Different national and international reports have presented the amount of carbon sequestered by forest ecosystems. For example, the Intergovernmental Panel on Climate Change reports that forests contain about 80% of aboveground and 40% of belowground carbon stock (IPCC 2001). Additionally, it is reported that the amount of carbon stored in dry wood is approximately 50% by weight (Brown 1986; Paladinic et al. 2009; Sedjo and Sohngen 2012).

Biomass, in general, includes both above and below ground living and dead mass of trees, shrubs, vines, and roots. However, most of the researches on biomass estimation have focused on aboveground biomass because of the difficulty in collecting belowground data (Lu 2006). The amount of biomass in a forest is influenced by various site factors such as stand density and site productivity; soil characteristics such as texture and moisture content; and tree characteristics such as species and age. On the other hand, distribution of crown biomass affects the carbon cycle, soil nutrient allocation, fuel accumulation, and wildlife habitat environments in terrestrial ecosystems and it governs the potential of carbon emission due to deforestation (Lu 2005). The major components of aboveground tree biomass are merchantable stem biomass (bole including bark and wood), stump biomass, foliage biomass, and branches/top biomass (Zhou and Hemstrom 2009).

The common biomass estimation approach selects some trees, which are representative of the populations of interest, for destructive sampling and weighs their components. Regression models are then fit to relate some easily measurable attributes, such as diameter at breast height and total tree height, with tree (or component) biomass. The amount of biomass distributed in different components is dependent on species and their geographic location (Pooreter et al. 2012), management practices

(Tumwebaze et al. 2013) and tree size and stand density (Jenkins et al. 2003). Ritchie et al. (2013) found that for the given DBH and crown ratio, thinned stands had more foliage biomass but slightly less branch biomass than unthinned stands. Similarly, the contribution of component biomass to the total aboveground biomass varies by tree size (de Miguel et al. 2014). Henry et al. (2011) found differences in biomass due to floristic composition, tree species and growth strategies for the tree species within a given climatic zone. Thus, the component biomass estimations, for example branch or crown biomass, bole biomass, and bark biomass, are important to account for the variability within the tree. The common understanding among researchers and practitioners is that an accurate carbon stock estimate requires improved and consistent methods for tree and component biomass estimation (Hansen 2002; Zhou and Miles 2009).

Crown biomass is the oven dry weight of the entire crown, including the leading shoot above the last-formed whorl, excluding the main bole (Hepp and Brister 1982). The components of crown biomass are wood, bark, and foliage weights. Crown biomass accounts for a significant portion of total tree biomass but the amount and its distribution varies by tree and site characteristics. Using the data from two Alaskan *Picea mariana* ecosystems, Barney et al. (1978) reported that foliage comprised 17% to 37% of the total tree mass for the lowland stand and 17% to 50% of the total tree mass in the upland stand. Total bole mass ranged from 11% to 58% in lowland stands and 21% to 61% in the upland stands. In a study to determine the patterns of biomass allocation in dominant and suppressed loblolly pine, Naidu et al. (1998) found that the dominant trees allocated 24.5% of biomass to the crown (13.2% in branch and 11.3% in needle) and the suppressed trees allocated 12.3% (6.7% in branch and 5.6% in needle). Kuyah et al. (2013) found that crown biomass formed up to 26% (22% in branch and 4% in needle) of aboveground biomass in farmed eucalyptus species. In assessing the importance of crown dimensions to improve tropical tree biomass estimate, Goodman et al. (2014) found the trees in their study to have nearly half of the total aboveground tree biomass in branches (44 ± 2%).

Estimates of crown biomass for each stand condition is necessary to understand nutrient depletion and for evaluating the economic feasibility of crown utilization for energy production or forest products (Hepp and Brister 1982). Furthermore, estimates of crown biomass aid in fuel load assessments and fire management strategies (He et al. 2013) because it is one of the important input variables in most wildfire models (Saatchi et al. 2007). Much of the focus in estimating crown biomass has been in the form of regression models and in the selection of predictor variables rather than in the methods of sample selection. In addition, comparisons of sampling strategies have been carried out mainly for foliar biomass sampling rather than the total crown (branch wood, bark, and foliage) biomass. Thus, the evaluation of different sampling designs and sample size in estimating crown biomass is an important aspect of aboveground biomass estimation.

Common sampling strategies used in aboveground biomass estimation include simple random sampling, systematic sampling, stratified random sampling, and randomized branch sampling. The suitability of a technique is determined by the availability of funds, required accuracy, structure and composition of vegetation, and desired specificity of estimation (Catchpole and Wheeler 1992). Additionally, the amount of time a particular technique takes to implement in the field is also important. The simple random sampling is generally used as the basis to evaluate the performance of other sampling designs (e.g. Snowdon 1986; Temesgen 2003).

Gregoire et al. (1995) have proposed a number of sampling procedures (randomized branch sampling, importance sampling, control-variate sampling, two-stage and three-stage sampling) that can be used to estimate foliage and other characteristics of individual trees. The randomized branch sampling (RBS) was originally introduced by Jessen (1955) to determine the fruit count on orchard trees. Valentine and Hilton (1977) used this method to obtain estimates of leaf counts, foliar area, and foliar mass of mature *Quercus species*. Good et al. (2001) have employed RBS with importance sampling for estimating tree component biomass. Since the sample is accumulated sequentially along the path, RBS

does not require locating and counting the total number of branches beforehand. However, Chirici et al. (2014) posed some doubts on the effectiveness of RBS in sampling big trees or trees with irregular forms. According to Valentine and Hilton (1977), the accuracy of RBS is largely dependent on the probability assignment and the time required to take RBS samples depends on the size of the trees and experience of those taking the samples.

Swank and Schreuder (1974) compared stratified two-phase sampling, two-phase sampling with a regression estimator, and two-phase sampling with a ratio-of-means estimator. They found the stratified two-phase sampling as the most precise and appropriate method for estimating surface area and biomass for a young eastern white pine forest. Temesgen (2003) found that stratified random sampling produced the lowest mean squared error value in comparing five sampling designs to quantify tree leaf area. Stratification in branch biomass sampling can be done in many different ways. Snowdon (1986) showed improved accuracy of estimates by stratification based on crown position compared to those obtained by simple random sampling, especially at low sampling intensities. Their findings suggest that stratification by whorl was slightly but not significantly inferior to stratification based on crown position or branch diameter. Another approach used in selecting branches for estimating crown biomass is to divide the bole into sections and pile up the branches from each section into different size class and randomly select a number of branches proportional to the total number of branches in each size class (e.g. Devine et al. 2013, Harrison et al. 2009). In an evaluation of ten different sampling strategies, Temesgen et al. (2011) found that systematic sampling with ratio estimation as the most efficient estimate of individual tree foliage biomass. de-Miguel et al. (2014) developed generalized , calibratable, mixed-effects meta-models for large-scale biomass prediction. One of their objectives was to investigate and demonstrate how the biomass prediction differed when calibration trees were selected using different sampling strategies. They found that a stratified sampling was better compared to the simple random sampling. Thus there is no strong rationale to support one method as being superior to another.

Crown biomass is difficult to predict because of the variability within and among species and various sites. A good allometric equation for predicting aboveground biomass should be based on data collected with an appropriate (precise and unbiased) sampling method. In this context, the objective of this study was to evaluate different sampling strategies to estimate crown biomass. We also evaluated how the performance of different methods was affected when different number of branches (3, 6, 9, and 12) per tree was sampled in estimating crown biomass.

2.2. Methods

2.2.1. Study Area

This study was conducted in the McDonald-Dunn Forest, an approximately 4550 ha property, managed by the Oregon State University in the western edge of the Willamette Valley in Oregon and on the eastern foothills of the Coast Range (123°15’ W, 44°35’ N, 120 m elevation). The forest consists predominantly of the Douglas-fir (*Pseudotsuga menziesii* (Mirbel) Franco) and a small portion of Grand fir (*Abies grandis* (Dougl. ex D. Don) Lindl.) and has a wide range of overstory age-class distribution with majority of the stands less than 80 years old and some stands that are 80 to 120 years old. The forest receives approximately 110 cm of annual rainfall and average annual temperature ranges from 6° to 17° C.

2.2.2. Data

Twenty sample trees (11 Douglas-fir and 9 Grand fir) were subjectively selected from stands of different ages for destructive sampling avoiding the trees with obvious defects and trees close to stand edges. The field work was carried out between the first week of July and third week of September 2012. Trees that were forked below breast height and with damaged tops were not included in sampling. Tree level attributes including total height, height to the base of first live branch, crown width, and main stem diameter at 0.15 m, 0.76 m, 1.37 m, and 2.4 m above ground, and every 1.22 m afterwards were

recorded. The branches were divided into four diameter classes (1.3 cm class = 0 – 2.5 cm, 3.8 cm class = 2.6 – 5.1 cm, 6.4 cm class = 5.2 – 7.6 cm, 8.9 cm class = 7.9 – 10.2 cm). For all first order branches, height to- and diameter- at branch base were measured.

For the first and every third branch, when proceeding from the base, in each diameter class, length and weight of both live and dead branches were recorded. From those selected branches, four branches per diameter class were weighed with and without foliage. The needles were removed in the field to obtain the green weight of foliage and branch wood with bark. Two of these four branches were taken to the lab, keeping branch and foliage in separate paper bags, for drying. The branches were chipped in to small pieces to expedite the drying process and placed in a kiln for drying at 105° C. The oven dry weight was recorded by tracking the weight lost by each sample until no further weight was lost. Table 2.1 presents the tree and branch level summary of the felled-tree data used in this study.

2.2.3. Individual Branch Biomass

Kershaw and Maguire (1995) developed a tree specific log linear model (equation 1) using branch diameter (BD) and depth into the crown (DINC: the distance from tip to the base of the branch) as covariates to estimate branch foliage biomass. Temesgen et al. (2011) successfully used this model in comparing sampling strategies for tree foliage biomass estimation.

$$\ln(y_{ij}) = \beta_{0i} + \beta_{1i}\ln(BD_{ij}) + \beta_{2i}\ln(DINC_{ij}) + \varepsilon_{ij} \quad (1)$$

This model was modified by replacing DINC with branch length (Equation 2). The modified model provided the best fit (Adj-R^2 = 0.93), therefore was used to predict individual branch biomass within each tree.

$$\ln(y_{ij}) = \beta_{0i} + \beta_{1i}\ln(BD_{ij}) + \beta_{2i}\ln(BL_{ij}) + \varepsilon_{ij} \quad (2)$$

Where, y_{ij}, BD_{ij} and BL_{ij} are oven dry weight (kg) of branch (wood, bark, and foliage combined), branch diameter (cm), and branch length (m) of the j^{th} branch on i^{th} tree respectively; $\beta_{ij}{}'s$

are regression parameters to be estimated; ln(·) is the natural logarithm; and $\varepsilon_{ij}{}'s$ are the random error. The full model included other variables such as height to the base of the branch, crown width and crown length, but were dropped because they were not statistically significant (p-value > 0.05). Lengths for the 2/3rd branches (not measured in the field) were obtained by fitting the following log linear model (Adj-R^2 = 0.74):

$$\ln(BL_{ij}) = \beta_{0i} + \beta_{1i}\ln(BD_{ij}) + \beta_{2i}\ln(RBD_{ij}) + \varepsilon_{ij} \quad (3)$$

Where, BD_{ij}, BL_{ij} , and ε_{ij} are same as defined in equation (2) and RBD_{ij} is the relative branch depth (relative position of the subject branch from the crown base) of *j*th branch in *i*th tree and is computed as follows (Ishii and Wilson 2001):

$$RBD = \frac{total\ tree\ height - height\ to\ the\ base\ of\ subject\ branch}{total\ tree\ height - height\ to\ the\ base\ of\ lowest\ live\ branch}$$

The RBD is 1.0 for the first live branch. The logarithmic regressions are reported to result in a negative bias when data are back transformed to arithmetic scale. The commonly used remedy to this is to multiply the back transformed results by a correction factor $\left[exp\left(\frac{MSE}{2}\right)\right]$, where *MSE* is the mean squared error obtained by the least-squares regression. However, there are conflicting remarks about the correction factor itself. For example, Beauchamp and Olson (1973) and Flewelling and Pienaar (1981) suggested that this correction factor was still biased because the sample variance is consistent but biased for finite sample size. We did to not use the correction factor in our study. The trend in the relationship between crown biomass and branch diameter and length was similar but the variability in biomass increased with increasing branch length (Figure 2.1). All statistical procedures were performed using statistical software R (R Core Team 2014).

2.2.4. Methods for Crown Biomass Sampling

We evaluated 11 sampling methods to select branches for estimating crown biomass. The 11 sampling strategies belonged to three main categories: simple random sampling, systematic sampling, and stratified sampling. Methods 1 and 2 are based on simple random sampling (SRS) strategy. In each of these methods, each branch was chosen randomly such that each individual branch has equal probability of selection at any stage of selection. The difference in these methods is in the estimation of total tree biomass: method 1 uses SRS estimator while method 2 (SRS-RAT) uses the ratio estimator with squared branch diameter as auxiliary information. Method 1 is also the basis for comparing the performance of other methods.

Method 3, probability proportional to size (PPS), uses branch size as auxiliary information in sample selection. Total crown biomass in this method was calculated using Horvitz-Thompson estimator (Horvitz and Thompson 1952). Methods 4 (SYS) and 5 (SYS-RAT) are systematic sampling with similar design phase but different estimation phase. Method 4 uses the SRS estimator while method 5 uses the ratio estimator. The fractional interval systematic sample selection procedure was used in the systematic selection of the branches because it ensures the equal probability of selection for all the branches (Temesgen et al. 2011). The interval was determined based on the total number of branches in each tree. In fractional interval systematic sample selection, first a random starting point between 1 and total number of branches was randomly chosen, the interval then is added obtaining exactly n (sample size) branches. Then the numbers are divided by the sample size and rounded to the nearest whole number to get the selected samples.

Methods 6-11 belonged to different stratified sampling strategies. The stratified sampling method divides the population into subpopulations of size n_h, where n_h is the number of elements in stratum *h*. The total crown length was divided into three sections having equal number of branches as three strata. In methods 6 (STR) and 7 (STR-RAT), *n*/3 branches were randomly selected with equal

probability, where n is the sample size. Again, the difference between these two methods lies in the estimation of total crown biomass. STR method uses the SRS estimation technique while STR-RAT method uses the ratio estimation technique to obtain the total crown biomass. Method 8 (STR-PPS), stratified sampling with PPS, selected branches in each stratum with probability proportional to the square of branch diameter. Total crown biomass in this method was obtained by summing the stratum total crown biomass calculated using Horvitz-Thompson unequal probability estimator (Horvitz and Thompson 1952).

Methods 9-11 (stratified, unequal) are based on the idea that the distribution of crown biomass in different strata depends on the relative position of the branches in the tree. Ishii and McDowell (2001) found that mean branch volume increased from upper- to lower-crown. For a given density, biomass (oven dry weight) is the function of volume. Therefore, the stratified sampling method was modified to incorporate the variability of biomass distribution within a tree. Trees were first divided into three sections having equal number of branches. Then 4, 3, and 2 branches from the lower, middle, and upper section of the trees were selected respectively. This corresponds that the number of branches selected in each section is proportional to the observed biomass in that section of the tree. Because stratification based on crown length resulted in the biased estimation of crown biomass, the balanced stratification method was applied. The total number of branches selected in each tree (nine) was determined based on the amount of biomass sampled. Total crown biomass in each stratum was computed using the SRS estimation technique in method 9 (Un-STR), PPS in method 10 (Un-PPS), and ratio estimation in method 11 (Un-STRRAT). Total crown biomass in each tree was computed by summing the crown biomass in each stratum. The unequal branch selection strategy was also evaluated using similar evaluation statistics used for the other 8 methods.

Performances of first eight methods were evaluated by selecting four different sample sizes (3, 6, 9, and 12) in each tree. These sample sizes were chosen for the ease of distributing samples into three

different strata in stratified sampling with equal number of branches per stratum. Methods 9-11 were based on selecting nine branches in each tree. Table 2.2 summarizes the inclusion probability, selection probability, and the estimator of the total crown biomass in each of the sampling strategies evaluated in this study.

2.2.5. Evaluation of Sampling Strategies

We evaluated the performance of 11 sampling strategies to estimate crown biomass using the following six statistics estimated from 5000 iterations. These measures were successfully used to evaluate the performance of sampling strategies to estimate foliage biomass in Temesgen et al. (2011).

1. **Bias:** For each tree the bias (kg) was calculated as the mean difference between observed and predicted total crown biomass for that tree as follows:

$$B_i = \frac{1}{5000}\sum_{s=1}^{5000}(\tau_{is} - \hat{\tau}_{is})$$

Where, τ_{is} and $\hat{\tau}_{is}$ are the observed and predicted total crown biomasses for i^{th} tree in s^{th} iteration, respectively.

2. **Relative Bias:** Relative bias percentage is the ratio of bias to the total observed crown biomass for that tree and computed as follows:

$$RB_i = \frac{1}{5000}\sum_{s=1}^{5000}\frac{(\tau_{is} - \hat{\tau}_{is})}{\tau_{is}}$$

Where, all the variables are same as defined previously.

3. **Root Mean Square Error (RMSE):**

$$RMSE_i = \sqrt{\frac{1}{5000}\sum_{s=1}^{5000}(\tau_{is} - \hat{\tau}_{is})^2}$$

4. **Relative RMSE:**

$$R - RMSE_i = \sqrt{\frac{1}{5000}\sum_{s=1}^{5000}\left(\frac{\tau_{is} - \hat{\tau}_{is}}{\tau_{is}}\right)^2}$$

5. **Biomass Sampled (BS):** Amount of cost for crown biomass estimation is directly proportional to the amount of crown biomass sampled. Therefore the amount of crown biomass sampled was also used as a criterion for the evaluation of sampling strategies. The amount of crown biomass sampled (sampling intensity) is calculated as follows:

$$BS_i = \frac{1}{5000}\sum_{s=1}^{5000}\sum_{j\in S} y_{ijs}$$

6. **Relative Biomass Sampled (RBS %):** indicates the proportion of crown biomass sampled with respect to the total crown biomass measured and is calculated as follows:

$$RBS_{ij} = \frac{1}{5000}\sum_{s=1}^{5000}\sum_{j\in S}\frac{y_{ijs}}{\tau_{ijs}}$$

2.3. Results and Discussion

Except for the ratio estimators, the estimators of population totals were unbiased, with biases close to zero for all sample sizes (Tables 2.3 and 2.4). The squared bias for these methods ranged from zero to 0.435 kg. Ratio estimators resulted in greater bias than the other methods. The absolute bias of the ratio estimators decreased with increasing sample size as expected.

As expected, the RMSE (and relative RMSE) decreased with increasing sample size (Tables 2.5 and 2.6) for all sampling strategies. Based on the RMSE values obtained from 5000 simulations, the stratified sampling with PPS estimation was the superior method compared to all other methods when sample size is 3 or 6 branches per tree. However, while using PPS, stratification of the crown into sections did not reduce the RMSE and relative RMSE significantly. On the other hand, when at least nine

branches per tree were sampled, the SYS-RAT was the best and the SRS-RAT was the second best method. Number of branches required to achieve desired precision is another important aspect of estimating crown biomass. On average, the RMSE decreased by 34.3% when the sample size increased from three branches per tree to six branches per tree. The RMSE further decreased by 22.1% and 15.4% when the sample size increased from 6 to 9 and 9 to 12 respectively.

The amount of biomass sampled determines the cost that would be incurred in estimating crown biomass. Biomass sampled and relative biomass sampled in different sampling strategies are presented in Tables 2.7 and 2.8. The Strategy-Cost-Accuracy graph (Figure 2.2) shows the efficiency trade-off across the strategies compared in the study. The SRS and SYS method resulted in the lowest amount of biomass sampled. On average, the amount of biomass sampled using the PPS method was 1.6, 1.5, 1.4, and 1.4 times higher than the amount of biomass sampled in stratified random sampling when 3, 6, 9, and 12 branches per tree were selected respectively.

On average, selecting 12 instead of 9 branches per tree increased the amount of biomass sampled by 29.2%. Therefore, nine branches in each tree were selected in evaluating the performance of unequal stratified sampling strategy. Results from unequal branch selection are presented in Table2.9. This strategy reduced the relative RMSE by 0.6%, 4.5% and 3.5% compared to selecting 9 branches using stratified random sampling, stratified sampling with ratio estimation, and stratified sampling with PPS respectively. This reduction in relative RMSE is obtained by sampling just a little more amount of biomass (1.03 times on average).

Use of allometric equations is inevitable in aboveground biomass estimation because weighing the trees and their components for direct biomass determination is destructive and prohibitively expensive. Choice of biomass sampling strategy determines the quality of data available for fitting such equations. Use of auxiliary information in design and/or estimation phase (PPS and ratio estimation)

produced better results in terms of RMSE compared to the methods that do not make use of such information in this study. Previous researches (e.g. Temesgen et al. 2011) have also shown the benefits from using auxiliary information in the design and/or estimation concerning tree biomass.

The model used in estimating branch biomass which is later used as a dependent variable in the test population, was a logarithmic model (equation 2). There is an inherent negative bias in this method because the dependent variable is transformed prior to estimation (Snowdon 1991). The ratio estimation strategies, SRS-RAT, SYS-RAT, and STR-RAT in this study, were negatively biased. However, in terms of RMSE, these strategies were clearly superior methods compared to the SRS approach. As noted in Temesgen et al. (2011), however, the efficiency of sampling strategies with ratio estimation may be affected by the amount of work and difficulty in implementing these techniques in the field.

The amount of biomass sampled determines the cost that would be incurred in estimating crown biomass. Choice of a sampling strategy determines the amount of biomass and relative biomass sampled. This ultimately determines the amount of time and cost required for a biomass estimation project. The SRS and SYS method resulted in the lowest amount of biomass sampled. Our results in terms of RMSE values reported and the amount of biomass sampled by each strategy are consistent with the findings of Temesgen et al. (2011) in estimating foliar biomass of Douglas-fir and Ponderosa pine.

2.4. Conclusions

Crown biomass estimation is a complex process that requires intensive manual field work involving destructive sampling. The amount of fieldwork required and the accuracy of biomass estimation is dependent on the sampling strategy used. Furthermore, the accuracy of the estimation can be improved by adopting appropriate techniques in both the design and estimation phases, beginning with the selection of sample plots and sample trees through model development. In this study, we

evaluated 11 different sampling strategies that belonged to three main categories: simple random sampling, systematic sampling and stratified sampling. The SRS, PPS, and ratio estimation techniques were used to obtain the total crown biomass in each tree.

Based on the RMSE values obtained from 5000 simulations, the stratified sampling with PPS estimation produced better results as compared to all other methods when 3 or 6 branches per tree were sampled. However, the SYS-RAT was the best and the SRS-RAT was the second best method when at least nine branches per tree were sampled. It should also be noted that the lower RMSE values in the PPS estimation techniques are obtained with an increased amount of biomass sampled in each tree. On the other hand, if the auxiliary information on branch size is not used, the systematic sampling provided better results than the SRS or STR method when at least 6 branches per trees were selected. Thus the selection of a specific sampling strategy is dependent on the availability of the time and cost for the given biomass sampling project. Based on our finding we would recommend sampling 9 branches per tree to obtain reasonable efficiency and amount of work involved in the field.

The logic for selecting unequal numbers of branches per stratum within a tree is justified by the fact that the biomass distribution within a tree is not uniform. Selecting equal branches per stratum produced approximately similar results to unequal sampling when the SRS estimation technique was used. However, making use of auxiliary information on branch size in the design and estimation phases further decreased the relative RMSE. Once again, the decreased RMSE by use of auxiliary information is attained by having to sample slightly higher amount of biomass. Findings of this study should prove beneficial for the stakeholders working in the field of aboveground biomass and carbon estimation. Additional work using the data from different species and location should be done to further validate the findings in this study.

2.5. Acknowledgements

We thank Professors Lisa Madison and Glen Murphy (both at Oregon State University) for their insights and comments on an earlier draft, and the Forest Inventory Analysis Unit for funding the data collection and analysis phases of this project.

2.6. References

Barney, R.J., K. VanCleve, and R. Schlentner. 1978. Biomass distribution and crown characteristics in two Alaskan Picea mariana ecosystems. *Can. J. For. Res.* 8(1): 36-41.

Beaucham, J.J., and J.S. Olson. 1973. Corrections for bias in regression estimates after logarithmic transformation. *Ecology*. 54(6): 1403-1407.

Brown, S. 1997. Estimating biomass and biomass change of tropical forests: a primer. Rome: FAO Forestry Paper 134. Food and Agriculture Org.

Catchpole, W., and C. Wheeler. 1992. Estimating plant biomass: a review of techniques. *Aust. J. Ecol.* 17(2): 121-131.

Chirici, G., N. Puletti, R. Salvati, F. Arbi, C. Zolli, and P. Corona. 2014. Is randomized branch sampling suitable to assess wood volume of temperate broadleaved old-growth forests? *For. Ecol. Manage.* 312: 225-230.

de-Miguel, S., L. Mehtätalo, and A. Durkaya. 2014. Developing generalized, calibratable, mixed-effects meta-models for large-scale biomass prediction. *Can. J. For. Res.* 44(6): 648-656.

de-Miguel, S., T. Pukkala, N. Assaf, and Z. Shater. 2014. Intra-specific difference in allometric equations for aboveground biomass of eastern Mediterranean Pinus brutia. *Ann. For. Sci.* 71(1): 101-112.

Devine, W.D., P.W. Footen, R.B. Harrison, T.A. Terry, C.A. Harrington, S.M. Holub, and P.J. Gould. 2013. Estimating tree biomass, carbon, and nitrogen in two vegetation control treatments in an 11-year-old Douglas-fir plantation on a highly productive site. *Res. Pap. PNW-RP-591*. Portland, OR: U.S. Department of Agriculture, Forest Service, p 29.

Flewelling, J.W., and L.V. Pienaar. 1981. Multiplicative regression with lognormal errors. *For. Sci.* 27(2): 281-289.

Good, N.M., M. Paterson, C. Brack, and K. Mengersen. 2001. Estimating tree component biomass using variable probability sampling methods. *J. Agric. Biol. Environ. Stat.* 6(2): 258-267.

Goodman, R.C., O.L. Phillips, and T.R. Baker. 2014. The importance of crown dimensions to improve tropical tree biomass estimates. *Ecol. Appl.* 24(4): 680-698.

Gregoire, T.G., H.T. Valentine, and G.M. Furnival. 1995. Sampling methods to estimate foliage and other characteristics of individual trees. *Ecology*. 76(4): 1181-1194.

Hansen, M. 2002. Volume and biomass estimation in FIA: national consistency vs. regional accuracy. P. 109 –120 in Proc. of the third annual forest inventory and analysis symposium, McRoberts, R.E., G.A. Reams, P.C. Van Deusen, and J.W. Moser (eds.). *US For. Serv. Gen. Tech. Rep. NC-230*. p 216.

Harrison, R.B., T.A. Terry, C.W. Licata, B.L. Flaming, R. Meade, I.A. Guerrini, B.D. Strahm, D. Xue, M.R. Lolley, A.R. Sidell, G.L. Wagoner, D. Briggs, and E.C. Turnblom. 2009. Biomass and stand characteristics of a highly productive mixed Douglas-Fir and Western Hemlock plantation in Coastal Washington. *West. J. Appl. For.* 24(4): 180-186.

Henry, M., N. Picard, C. Trotta, R.J. Manlay, R. Valentini, M. Bernoux, and L. Saint-André. 2011. Estimating tree biomass of sub-Saharan African forests: a review of available allometric equations. *Silva Fennica.* 45(3B): 477-569.

He, Q., E. Chen, R. An, and Y. Li. 2013. Above-ground biomass and biomass components estimation Using LiDAR data in a coniferous forest. *Forests.* 4(4): 984-1002.

Hepp, T.E., and G.H. Brister. 1982. Estimating crown biomass in loblolly pine plantations in the Carolina Flatwoods. *For. Sci.* 28(1): 115-127.

Horvitz, D.G., and D.J. Thompson. 1952. A generalization of sampling without replacement from a finite universe. *J. Am. Stat. Assoc.* 47: 663-685.

Ishii, H., and N. McDowell. 2001. Age-related development of crown structure in coastal Douglas-fir trees. *For. Ecol. Manage.* 169: 257-270.

Ishii, H., and M.E. Wilson. 2001. Crown structure of old-growth Douglas-fir in the western Cascade Range, Washington. *Can. J. For. Res.* 31(7): 1250-1261.

Jessen, R. J. 1955. Determining the fruit count on a tree by randomized branch sampling. *Biometrics.* 11(1):99-109.

Jenkins, C.J., D.C. Chojnacky, L.S. Heath, and R.A. Birdsey. 2003. National-scale biomass estimators for United States tree species. *For. Sci.* 49(1): 12-35.

Kershaw Jr, J.A., and D.A. Maguire. 1995. Crown structure in western hemlock, Douglas-fir, and grand fir in western Washington: trends in branch-level mass and leaf area. *Can. J. For. Res.* 25(12): 1897-1912.

Kuyaha, S., J. Dietz, C. Muthuri, M.V. Noordwijk, and H. Neufeldt. 2013. Allometry and partitioning of above- and below-ground biomass in farmed eucalyptus species dominant in Western Kenyan agricultural landscapes. *Biomass and Bioenergy.* 55: 276-284.

Lu, D. 2005. Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon. *Int. J. Remote Sens.* 26(12): 509-2525.

Lu, D. 2006. The potential and challenge of remote sensing-based biomass estimation. *Int. J. Remote Sens.* 27(7): 1297-1328.

Naidu, S.L., E.H. DeLucia, and R.B. Thomas. 1998. Contrasting patterns of biomass allocation in dominant and suppressed loblolly pine. *Can. J. For. Res.* 28(8): 1116-1124.

Paladinic, E., D. Vuletic, I. Martinic, H. Marjanovic, K. Indir, M. Benko, and V. Novotny. 2009. Forest biomass and sequestered carbon estimation according to main tree components on the forest stand scale. *Periodicum Biologorum.* 111(4): 459-466.

Pooreter, H., K.J. Niklas, P.B. Reich, J. Oleksyn, P. Poot, and L. Mommer. 2012. Biomass allocation to leaves, stems and roots: meta-analyses of interspecific variation and environmental control. *New Phytologist*. 193: 30-50.

R Core Team. 2014. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Ritchie, M.W., J. Zhang, and T.A. Hamilton. 2013. Aboveground tree biomass for Pinus ponderosa in northeastern California. *Forests.* 4(1): 179-196.

Saatchi, S., K. Halligan, D. G. Despain, and, R.L. Crabtree. 2007. Estimation of forest fuel load from radar remote sensing. *IEEE Trans. Geosci. Remote Sens.* 45(6): 1726-1740.

Sedjo, R., and B. Sohngen. 2012. Carbon sequestration in forests and soils. *Annu. Rev. Resour. Econ.* 4: 127-144.

Snowdon, P. 1986. Sampling strategies and methods of estimating the biomass of crown components in individual trees of *Pinus radiata* D. Don. *Aust. For. Res.* 16(1): 63-72.

Snowdon, P. 1991. A ratio estimator for bias correction in logarithmic regressions. *Can. J. For. Res.* 21(5): 720-724.

Swank, W. T., and H. T. Schreuder. 1974. Comparison of three methods of estimating surface area and biomass for a forest of young eastern white pine. *For. Sci.* 20(1): 91-100.

Temesgen, H. 2003. Evaluation of sampling alternatives to quantify tree leaf area. *Can. J. For. Res.* 33(1): 82-95.

Temesgen, H., V. Monleon, A. Weiskittel, and D. Wilson. 2011. Sampling strategies for efficient estimation of tree foliage biomass. *For. Sci.* 57(2): 153-163.

Tumwebaze, S. B., E. Bevilacqua, R. Briggs, and T. Volk. 2013. Allometric biomass equations for tree species used in agroforestry systems in Uganda. *Agroforest Syst.* 87: 781-795.

Valentine, H. T., and S. J. Hilton. 1977. Sampling oak foliage by the randomized-branch method. *Can. J. For. Res.* 7: 295-298.

Zhou, X., and M.A. Hemstrom. 2009. Estimating aboveground tree biomass on forest land in the Pacific Northwest: a comparison of approaches. *Res. Pap. PNW-RP-584*. Portland, OR: U.S. Department of Agriculture, Forest Service. Pacific Northwest Research Station, p. 18.

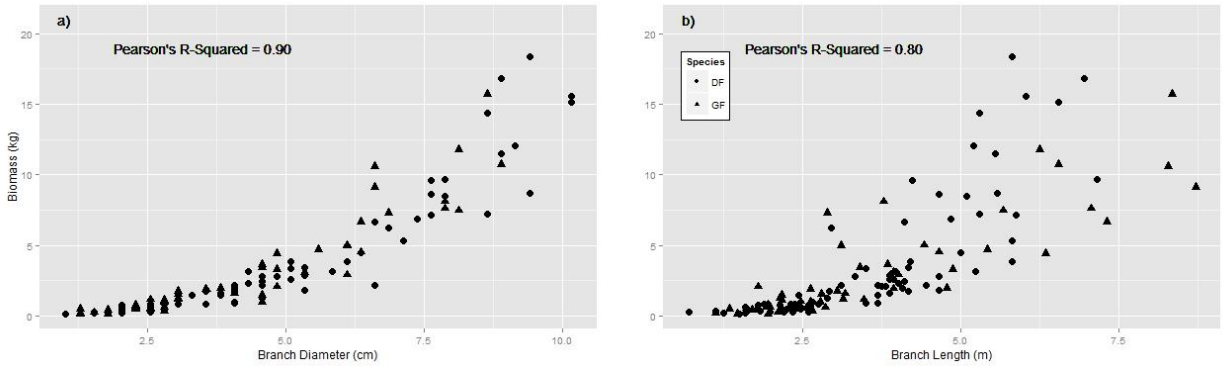

Figure 2.1. Scatterplot of dry biomass (kg) against branch diameter (a) and branch length (b) by species (DF=Douglas-fir, GF=Grand fir).

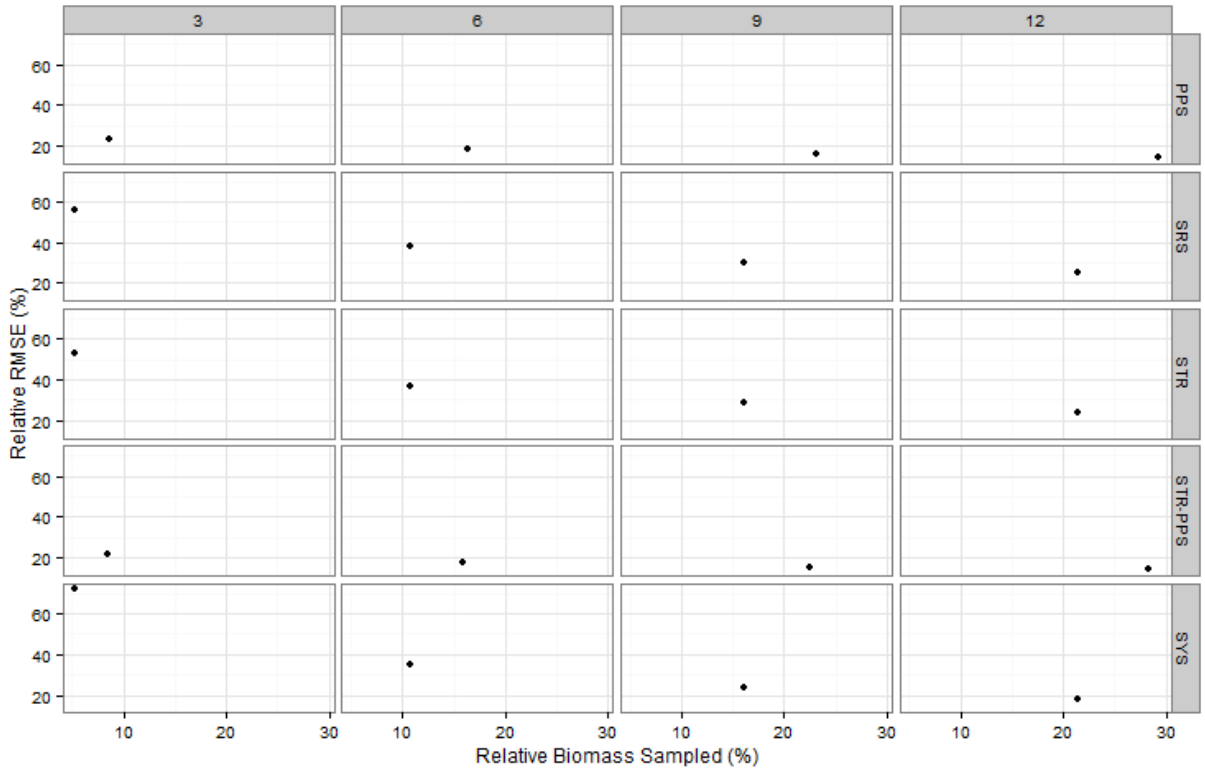

Figure 2.2. Relative RMSE (%) produced Vs. relative biomass sampled (percent of total crown mass) in different sampling strategies and sample sizes.

Table 2.1. Summary of felled-tree and branch-level attributes used in this study.

Variable	n	Mean	SD	Minimum	Maximum
Tree data					
DBH (cm)		62.8	27	23.9	114
Height (m)		34.6	10.3	18.3	46.5
Crown base height (m)		10.2	6.3	0.8	23.7
Crown length (m)		24.4	6.2	12.2	34.4
Crown width (m)		8.6	2.4	5.2	13.4
Branch data					
Diameter (cm)	3464	3.7	2.1	0.8	10.4
Length (m)	1178	2.8	1.6	0.1	10.2
Total green weight (kg)	1102	4.14	5.61	0.05	46.99
Green weight of branch wood (kg)	326	3.98	5.61	0.05	38.01
Green foliage weight (kg)	264	1.8	1.84	0.05	9.71
Total dry weight (kg)	128	3.44	4.08	0.12	18.4
Dry weight of branch wood (kg)	128	2.68	3.49	0.07	15.37
Dry foliage weight (kg)	128	0.76	0.79	0.04	4.58

Table 2.2. Summary of methods used for crown biomass estimation in this study.

Methods	Equations for Total Crown Biomass	Selection Probability	Inclusion Probability
Simple Random Sampling			
SRS	$\hat{\tau}_i = NB_i\left(\frac{1}{n}\sum_{j=1}^{n} y_{ij}\right)$	$\left(\frac{1}{NB_i}\right)^n$	$\frac{n}{NB_i}$
SRS-RAT	$\hat{\tau}_i = \frac{\sum_{j=1}^{n} y_{ij}}{\sum_{j=1}^{n} BD_{ij}^2}\sum_{j=1}^{NB_i} BD_{ij}^2$	$\left(\frac{1}{NB_i}\right)^n$	$\frac{n}{NB_i}$
PPS	$\hat{\tau}_i = \sum_{j\in S}\frac{y_{ij}}{\pi_{ij}^{PPS}}$	$\pi_{ij} = \frac{BD_{ij}^2}{\sum_{j=1}^{NB} BD_{ij}^2}$	$\pi_{ij}^{PPS} = 1-\left(1-\pi_{ij}\right)^n$
Systematic Sampling			
SYS	$\hat{\tau}_i = NB_i\left(\frac{1}{n}\sum_{j=1}^{n} y_{ij}\right)$	$\frac{n}{NB_i}$	$\frac{n}{NB_i}$
SYS-RAT	$\hat{\tau}_i = \frac{\sum_{j=1}^{n} y_{ij}}{\sum_{j=1}^{n} BD_{ij}^2}\sum_{j=1}^{NB_i} BD_{ij}^2$	$\frac{n}{NB_i}$	$\frac{n}{NB_i}$
Stratified Sampling			
STR	$\hat{\tau}_i = \sum_{h=1}^{H}\sum_{j=1}^{n_h}\frac{N_{ih}}{n_h} y_{ijh}$	$\frac{BD_{ij}^2}{\sum_{j=1}^{NB} BD_{ij}^2}$	$\frac{n_h}{N_{ih}}$
STR-RAT	$\hat{\tau}_i = \frac{\sum_{h=1}^{H}\sum_{j=1}^{n_h}\frac{N_{ih}}{n_h} y_{ijh}}{\sum_{h=1}^{H}\sum_{j=1}^{n_h}\frac{N_{ih}}{n_h} BD_{ijh}^2}$	$\frac{BD_{ij}^2}{\sum_{j=1}^{NB} BD_{ij}^2}$	$\frac{n_h}{N_{ih}}$
STR-PPS	$\hat{\tau}_{i(STR-PPS)} = \sum_{h=1}^{H}\sum_{j\in S_h}\frac{y_{ijh}}{\pi_{ijh}^{(STR-PPS)}}$	$\frac{BD_{ijh}^2}{\sum_{j=1}^{NB_{ih}} BD_{ijh}^2}$	$\pi_{ij}^{STR-PPS} = 1-\left(1-\pi_{ij}\right)^n$

Notation: $\hat{\tau}_i$is the estimated total crown biomass for i^{th} tree; y_{ij} is the oven dry weight of j^{th} branch on i^{th} tree; NB_i is the number of branches on i^{th} tree; n is number of branches sampled; N_{ih} is the number of branches in h^{th} stratum on i^{th} tree; and n_h= number of branches sampled in h^{th} stratum.

Table 2.3. Average bias (kg) produced by different sampling methods and sample sizes based on 5,000 simulations.

Sample Size	Sampling Strategies							
	SRS	**SRS-RAT**	**PPS**	**SYS**	**SYS-RAT**	**STR**	**STR-RAT**	**STR-PPS**
3	0.237	-4.091	-0.158	0.433	-9.320	-0.659	-3.554	-0.104
6	0.215	-1.711	0.010	0.139	-1.922	0.399	-1.387	-0.101
9	-0.081	-0.937	-0.068	0.133	0.166	0.191	-0.857	-0.094
12	-0.030	-0.894	0.158	0.043	-0.078	-0.242	-0.776	-0.073

Table 2.4. Relative bias (percent) produced by different sampling methods and sample sizes based on 5,000 simulations.

Sample Size	Sampling Strategies							
	SRS	**SRS-RAT**	**PPS**	**SYS**	**SYS-RAT**	**STR**	**STR-RAT**	**STR-PPS**
3	0.101	-2.437	-0.169	0.107	-6.086	-0.321	-2.107	-0.034
6	-0.001	-1.115	0.024	0.061	-1.061	0.197	-0.900	-0.012
9	0.063	-0.600	-0.027	0.038	0.033	0.148	-0.543	-0.069
12	-0.011	-0.523	0.100	0.037	0.032	-0.035	-0.439	-0.040

Table 2.5. Average RMSE produced by different sampling methods and sample sizes based on 5,000 simulations.

Sample Size	Sampling Strategies							
	SRS	**SRS-RAT**	**PPS**	**SYS**	**SYS-RAT**	**STR**	**STR-RAT**	**STR-PPS**
3	88.57	42.56	38.04	113.17	48.54	85.05	42.08	34.95
6	60.49	30.01	29.23	57.94	29.98	58.54	29.56	27.99
9	48.60	23.92	25.19	37.88	22.12	46.48	23.55	24.46
12	40.77	20.18	22.78	30.77	17.91	38.89	19.75	22.33

Table 2.6. Relative RMSE percent produced by different sampling methods and sample sizes based on 5,000 simulations.

Sample Size	Sampling Strategies							
	SRS	**SRS-RAT**	**PPS**	**SYS**	**SYS-RAT**	**STR**	**STR-RAT**	**STR-PPS**
3	55.70	25.72	23.46	72.02	28.80	53.41	25.61	21.57
6	38.07	17.93	18.23	35.23	17.53	36.63	17.78	17.59
9	30.38	14.19	15.83	23.58	13.29	29.03	14.07	15.55
12	25.42	11.92	14.40	18.37	11.11	24.27	11.76	14.22

Table 2.7. Amount of biomass sampled (kg) by different sampling strategies and sample sizes.

Sample Size	Sampling Strategies				
	SRS	**PPS**	**SYS**	**STR**	**STR-PPS**
3	7.74	12.58	7.73	7.68	12.40
6	15.49	23.79	15.46	15.50	23.29
9	23.17	33.77	23.18	23.19	33.05
12	30.91	42.83	30.92	30.86	41.79

Table 2.8. Relative amount of biomass sampled (%) by different sampling strategies and sample sizes.

Sample Size	Sampling Strategies				
	SRS	**PPS**	**SYS**	**STR**	**STR-PPS**
3	5.36	8.67	5.35	5.33	8.45
6	10.71	16.37	10.71	10.72	15.87
9	16.08	23.16	16.06	16.07	22.45
12	21.42	29.27	21.43	21.40	28.33

Table 2.9. Evaluation statistics produced when selecting 4, 3, and 2 branches from lower, middle, and upper stratum.

Method	Bias	Relative Bias	RMSE	Relative RMSE	Biomass Sampled	Relative Biomass Sampled
Un-STR	0.100	0.061	46.16	28.87	24.25	16.87
Un-STRRAT	-0.731	-0.491	22.58	13.44		
Un-PPS	0.140	0.063	23.67	15.00	33.85	23.04

3. Manuscript 2

Methods for Estimating Aboveground Biomass and its Components for Five Pacific Northwest Tree Species

Krishna P. Poudel and Hailemariam Temesgen

Canadian Journal of Forest Research

Canadian Science Publishing (NRC Research Press)

65 Auriga Drive, Suite 203

Ottawa, ON K2E 7W6, Canada

(Manuscript Submitted March 2015)

Abstract

Estimating aboveground biomass and its components requires sound statistical formulation and evaluation. Using data collected from 90 destructively sampled trees in different parts of Oregon, we evaluated the performance of three groups of methods to estimate total aboveground biomass and/or its components based on the bias and root mean squared error (RMSE) they produced. The first group of methods used analytical approach to estimate total and component biomass using existing equations, and produced biased estimates for our dataset. The second group used a system of equations fitted with seemingly unrelated regression (SUR), and were superior to the group I methods in terms of bias and RMSE. The third group of methods predicts the proportions of biomass in each component using beta, Dirichlet, and multinomial loglinear regression (MLR). The predicted proportions are then applied to the total aboveground biomass to obtain amount of biomass in each component. The MLR approach produced smaller RMSE compared to the SUR approaches except for grand fir branch biomass while the beta and Dirichlet regressions provided smaller RMSE compared to the SUR approaches for 85 percent of the species-component combinations.

3.1. Introduction

Forests play dual role in the global carbon cycle as an important sink by removing carbon dioxide through photosynthesis, converting that photosynthate to forest biomass, and as a carbon source by releasing carbon dioxide through respiration, wildfires, and decomposition. If forests are properly managed and timber is used for long-term products, such as buildings, forest management could result in a net reduction of atmospheric carbon while burning of wood for residential and commercial uses increases carbon emission rates. Thus there is a great concern about the uncertainty over whether forests will be a sink or a source of carbon in the future. Forest growth rates, harvest activities, natural disturbances such as wildfire, and loss of forest cover due to landuse changes are key factors that alter the carbon stock and absorption ability of a forest. It is essential to have knowledge of carbon stocks and fluxes to understand the current state and future course of carbon cycle in response to changing land uses and climatic conditions (Hollinger 2008). The amount of biomass as living vegetation or dead wood and debris is important factor that relates forestry function in regulating atmospheric carbon (Brown 2002).

Total carbon stocks in forest ecosystems can be divided in to two main parts as aboveground pools and fluxes and belowground pools and fluxes (Hoover 2008). Aboveground biomass constitutes the major portion of carbon pools in forest ecosystems (Vashum and Jayakumar 2012). Xiao et al. (2003) found 0.14 ratio of belowground biomass to aboveground biomass in 73-year-old Scots pine (*Pinus sylvestris* L.) forest. Czapowskyj et al. (1985) found 80 percent biomass in aboveground components and 20 percent in the root and stump of black Spruce (*Picea mariana* B.S.P. (Mill.)) in Maine. Aboveground biomass estimation has received significant attention in recent years because of the fact that the change in aboveground biomass is associated with the components of climate change (Lu et al. 2002). It depicts the forest's potential to sequester and store carbon. Therefore, aboveground biomass estimation is central in quantifying and monitoring of the amount of carbon stored in forest ecosystems.

The amount of biomass contained in a tree is influenced by various site factors such as stand density, site productivity; soil characteristics such as texture and moisture content; and tree characteristics such as species and age. Generally, aboveground biomass is estimated through the use of allometric equations that relate easily measurable attributes such as diameter at breast height (DBH), height, etc. to total tree biomass or component biomass. It is critical to note, however, that different regions and species link these attributes to biomass through different functions such as logarithmic, linear, or quadratic form (Zhou and Hemstrom 2009). Thus, site and species specific models might be better than regional models in predicting aboveground biomass at the local level.

Aboveground biomass is commonly divided into three major components: bole (main stem), stem bark, and crown (branch wood and foliage). The component biomass models are useful to account for the variability within the tree. Moreover, the use of component biomass for various purposes such as bioenergy production requires the estimate of biomass in each component to determine the associated cost of transportation and processing biomass at a bioenergy plant.

The amount of biomass allocated to different tree components varies by species and their geographic locations (Luo et al. 2013). In an effort to quantify aboveground biomass and net primary production of pine, oak and mixed pine-oak forests, Nunes et al. (2013) found that in pine, the greater distribution of biomass was on the stem whereas in oak and mixed stands the foliage was the component with the greater distribution of biomass. Repola (2006) found that wood density in pine decreased from the butt to the top, and the gradient in wood density was steep at the butt but decreased in the upper part of the stem. Even though the vertical dependence was similar in birch (a hardwood species), the density gradient was much smaller. Because mass is a function of volume and density, it also affects the amount of biomass contained in trees. In addition, the differences in management practices influence the allocation of biomass to different components (Tumwebaze et al. 2013). Therefore locally derived species specific biomass equations may yield biomass estimates that are

considerably different from the estimates obtained by using more generalized regional biomass equations. Weighing trees in the field would be the most accurate method for estimating tree biomass. However, it is time consuming and very expensive thus the use of allometric equations is essential. The type and amount of data and the number of equations required to accurately quantify biomass is unknown. Cautionary measures should be taken in developing and evaluating the methods for estimating aboveground biomass and its components.

When component models are fitted, the strength of relationship exhibited by bark, branch, and foliage models is nowhere near that for the stem wood (Boudewyn et al. 2007). On the other hand it is desired that the prediction from component regression equations add to the prediction from total tree regression model (Parresol 2001). The relationship between component mass and easily measurable tree attributes differ considerably in conifers and hardwoods. Component biomass modeling in hardwood requires more innovative approach due to decurrent form (Westfall et al. 2012).

With substantial increase in the demand of forest biomass information in recent years, considerable efforts have been made to estimate aboveground biomass and its components. Stem wood biomass accounts for the major portion of aboveground biomass. In comparing different approaches of aboveground biomass estimation, Zhou and Hemstrom (2009) found the proportion of softwood merchantable biomass ranging from 72 to 83 percent of the total aboveground biomass with little variation among species. Using LiDAR data from a *Picea crassifolia* (a coniferous tree) stand, He et al. (2013) found that, on average 72 percent of the total above-ground biomass of a tree is contained within the stem. The branches, foliage, and fruits in their study accounted 11, 13, and 4 percent of total aboveground biomass respectively. However, Kuznetsova et al. (2011) found only 31 and 27 percent of aboveground biomass in the main stem of 8-year old Scots pine and lodgepole pine respectively in the oil shale post-mining landscapes in Estonia. They found 32 and 28 percent in shoots and 37 and 45 percent in the needles of those 8-year old scots pine and lodgepole pine respectively.

These differences in proportion of biomass in different components among species warrant for species specific component models. More than 90 percent of aboveground live tree biomass in Oregon is contributed by softwood species (Zhou and Hemstrom 2009). Douglas-fir by itself contributes about 51 percent of total aboveground biomass of trees larger than 12.5 cm (5 inch) DBH and western hemlock and ponderosa pine contribute about 8 percent each (Zhou and Hemstrom 2009). Additionally, Douglas-fir, western hemlock, grand fir, red alder, and lodgepole pine are among the top eight tree species in Oregon. In this study, we have developed and evaluated different approaches for estimating aboveground biomass and its components for these five tree species using the data from destructively sampled trees.

Approaches in biomass estimation depends on scale of analysis, need for detail, user group interests and purpose of estimation (Zhou and Hemstrom 2009). Although there is a need for consistent methods of biomass estimation, there is no strong rationale to justify one method of estimation as being superior to another. In this study, we compared different methods that belonged to three major groups. The first group of methods (Group I methods) uses an analytical approach to estimate total and component biomass using existing equations. Group I methods are the component ratio method of the USDA Forest Service's Forest Inventory and Analysis (FIA-CRM), the regional approach for the Pacific Northwest (FIA-PNW), and an approach developed by Jenkins et al. (2003) (Jenkins). The second group (Group II methods) is a regression based approach that used a system of equations fitted with seemingly unrelated regression. The dependent variables in the system of equations were component and total biomass and the independent variables were DBH, DBH and total height, or DBH and crown length. These two groups of methods give the amount of biomass contained in each component and also the total aboveground biomass. The third group of methods (Group III methods) predicts the proportions of biomass in each component using beta, multinomial loglinear, and Dirichlet regression. Predicted

proportions are then applied to the observed total aboveground biomass to obtain amount of biomass in each component.

3.2. Materials and Methods

3.2.1. Data and Study Area

A detailed biomass data collection was carried out by destructively sampling 90 trees in different forests within the state of Oregon (Figure 3.1). Efforts were made to select trees to give an approximately equal representation across a range of size class while avoiding the trees with severe defects and close to the stand edges. Trees that were forked below breast height and with damaged tops were also not included in sampling. The 90 trees belonged to five different species: Douglas fir (*Pseudotsuga menziesii* (Mirbel) Franco), Grand fir (*Abies grandis* (Dougl. ex D. Don) Lindl.), Lodgepole Pine (*Pinus contorta)*, Western Hemlock (*Tsuga heterophylla)*, and Red Alder (*Alnus rubra).* The field work was carried out between the first week of July and third week of September 2012 and 2013. Tree level attributes DBH, total height, crown base height (height to the base of the first live branch), crown width, and main stem diameter at 0.15 m, 0.76 m, 1.37 m, and 2.4 m above ground, and every 1.22 m afterwards were recorded. Diameter measurements were made on a total of 11783 and 472 branches from conifers and hardwood species respectively. The average DBH ranged from 24.6 cm to 54.9 cm and average height ranged from 17 m to 33 m (Table 3.1).

The crown of the sample tree was divided in to three equal length strata. For all first order branches, height to- and diameter- at branch base were measured and for the first and every third branch in each stratum, branch length and green weight of both live and dead branches were recorded. Four, three, and two branches respectively from lower, middle, and upper stratum were randomly selected for weighing with and without foliage. The needles were removed in the field to obtain separate green weight of branch wood and foliage. These branches were then brought to the lab,

keeping branch wood and foliage in separate paper bags, for drying. Three to five centimeter thick disks were removed from the top of the stump and every 5.18 meter. Green weight of disks with and without bark as well as four measurements 90° apart along the disk edge for thickness was recorded in the field. Five to ten cm long bark sample was removed from each disk. The width, thickness, and length of the bark samples were also recorded in the field.

The sample branches were chipped in to small pieces to expedite the drying process and placed in a kiln for drying at 105° Celsius. Oven dry weight was recorded by tracking the weight lost by each sample (disk, branch wood, needle, and bark) until no further weight was lost.

Total tree biomass computation involved additional steps. Volume of each 5.18 m sections was converted into biomass by multiplying it by the average density of the disks taken from two ends. Total bole biomass was obtained by summing the section masses. Individual branch wood and foliage biomass was obtained by fitting species specific log linear model of the following form:

$$\ln(y_{ij}) = \beta_{0i} + \beta_{1i} \ln(BD_{ij}) + \varepsilon_{ij} \quad (1)$$

Where, y_{ij}, and BD_{ij} are oven dry weight (kg) of branch wood or foliage and branch diameter (cm) at base of the j^{th} branch on i^{th} tree respectively; $\beta_{ij}{}'s$ are regression parameters to be estimated from the data; ln(·) is the natural logarithm; and $\varepsilon_{ij}{}'s$ are the random error. The logarithmic regressions are reported to result in a negative bias when data are back transformed and the commonly used remedy to this is to multiply the back transformed results by a correction factor $[exp\left(\frac{MSE}{2}\right)]$, where MSE is the mean squared error obtained by the least-squares regression. However, there are conflicting remarks about the correction factor itself (e.g. Beaucham and Olson 1973, Flewelling and Pienaar 1981) and the effects of adjustment might be negligible (e.g. Harrison et al. 2009). Therefore, we did not use the correction factor in this study. Total branch wood and foliage biomass in each tree was obtained by summing these fitted values. The distribution of aboveground biomass in different components differed

among species with the majority of the aboveground biomass (73 to 82 percent on average) being present in the main bole (Figure 3.2).

3.2.2. Methods for Estimating Aboveground Biomass

There are many different methods available for the calculation of aboveground biomass and its components. The Forest Inventory and Analysis (FIA) of the U.S. Department of Agriculture (USDA) Forest Service uses the component ratio method (CRM) described in Heath et al. (2008) to produce national-level biomass and carbon estimates. The FIA-CRM approach is a multi-step process that involves converting sound volume of wood in the merchantable bole using a compiled set of wood specific gravities. The biomass in bole bark is calculated using a compiled set of percent bark and bark specific gravities. The set of wood and bark specific gravities and percent bark are compiled in Miles and Smith (2009). The biomass of tops and limbs is calculated as a proportion of the bole biomass based on component proportions from Jenkins et al. (2003). Biomass in stump wood and bark is based on volume equations in Raile (1982) and the compiled set of wood and bark specific gravities. Total aboveground biomass is obtained by summing these component masses.

The FIA program of the USDA Forest Service has four regional units: Northern, Southern, Interior West, and Pacific Northwest (FIA-PNW). The FIA-PNW collects and maintains the data on plots in coastal Alaska, California, Hawaii, Oregon, Washington, and U.S.-affiliated Pacific Islands. To calculate aboveground biomass, the FIA-PNW uses its specific set of equations. Tree stem biomass is calculated from the cubic volume and wood density factor. The specific equations used by FIA-PNW for volume and aboveground biomass in the Pacific Northwest can be found in Zhou and Hemstrom (2010).

The Jenkins et al. (2003) equations were derived by fitting regressions on pseudo-data generated from previously published equations. Jenkins et al. (2003) equations for total aboveground biomass are single entry equations that only use DBH as predictor variable. Biomass of the components

is predicted as the proportions of total aboveground biomass using an exponential function of DBH, differently for hardwood and softwood species groups.

Aboveground biomass equations are used to convert forest inventory data to biomass estimates. These equations combine biomass data obtained from destructive sampling with the dendrometric information through regression. Regression models for total and component aboveground biomass have been fitted using different fitting approaches with linear and nonlinear functions. The sets of component biomass equations can be fitted independently or as a system of equations. When a system of equations is fitted simultaneously, the residuals are correlated because the component biomasses come from the same tree. Therefore the SUR approach that allows the inclusion of dependencies among the error terms of the component biomass equations is commonly used to estimate component and total aboveground biomass (e.g. Parresol 1999; Lambert et al. 2005; Ritchie et al. 2013). The SUR models can be constrained such that the prediction of component equations sum to the prediction of total tree regression.

The single entry DBH-based allometric models for component biomass are much more complex than for total aboveground biomass (Jenkins et al. 2003). Zhang et al. (2004) found diameter at the base of the live crown as a better predictor of crown foliage biomass than DBH. For a mixed-species Atlantic forest stands, a transformed nonlinear biomass equation that used squared DBH, total tree height, and wood density as predictor variable was the most accurate model for aboveground biomass (Junior et al. 2014). Therefore we extended the single entry DBH-based SUR models to include other predictor variables. Both single-and multiple-entry SUR models were constrained such that the prediction of component equations sum to the prediction of total tree regression. The models for component and total biomass in the simple and extended SUR method were in the following form:

$$Bole = exp(a11 + a12 * X1 + a13 * X2 + a14 * X3) \quad (2)$$

$$Bark = exp(a21 + a22 * X1 + a23 * X2 + a24 * X3) \quad (3)$$

$$Branch = exp(a31 + a32 * X1 + a33 * X2 + a34 * X3) \quad (4)$$

$$Foliage = exp(a41 + a42 * X1 + a43 * X2 + a44 * X3) \quad (5)$$

$$Total = \sum_{i,j} exp(aij + aij * X1 + aij * X2 + aij * X3) \quad (6)$$

Where $aij, (i = 1, 2, 3, 4, \text{and } j = 1, 2, 3, 4)$ are parameters to be estimated from the data and X1, X2, and X3 are natural logarithms of DBH, total tree height and crown length respectively. Note that for the simple SUR, only X1 was used as explanatory variable. SAS procedure PROC MODEL (SAS Institute Inc 2013) was used to fit both simple and extended SUR models. We made necessary computational adjustments to match the component definitions in all methods and applied these methods to estimate component and total aboveground biomass of all five species.

3.2.3. Estimating Component Proportions

The component biomass can be estimated as the proportions of total aboveground biomass. A proportion is bounded between 0 and 1, and therefore the effect of explanatory variables tends to be nonlinear and the variance tends to decrease when the mean get closer to one of the boundaries. In this study we evaluated three different methods for estimating proportions namely the beta regression, Dirichlet, and the multinomial loglinear regression.

3.2.3.1. Beta Regression

The Beta regression model was introduced by Ferrari and Cribari-Neto (2004) and is useful when the variable of interest is continuous, restricted to the interval (0, 1) such as percentages, proportions and fractions or rates, and related to other variables through a regression structure. Since then it has -been used in many fields including medicine (Hubben et al. 2008), economics (De Paola et al. 2010), education (Smithson and Verkuilen 2006), and forestry (Korhonen et al. 2007; Eskelson et al. 2011).

Korhonen et al. (2007) used this technique for modeling percent canopy cover in a conifer dominated study area in central Finland. Eskelson et al. (2011) used beta regression to estimate riparian percent shrub cover. This method produced smaller mean squared prediction error and absolute bias compared to the ordinary least squares and generalized least squares regression models used in their study. The usual beta distribution is of the following form:

$$f(y;\ \alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1},\ \text{for } 0 < \mathrm{y} < 1 \tag{7}$$

Where, $\alpha, \beta > 0$ are two shape parameters and $\Gamma(\cdot)$ is the gamma function.

In order to obtain a regression structure for the mean of the response and a precision (dispersion) parameter, Ferrari and Cribari-Neto (2004) used a different parameterization of beta distribution. With mean and precision parameters defined as $\mu = \frac{\alpha}{(\alpha+\beta)}$ and $\phi = (\alpha+\beta)$ respectively, the beta density function has the following form under new parameterization:

$$f(y;\ \mu,\phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1},\ \text{for } 0 < \mathrm{y} < 1 \tag{8}$$

Where, $0 < \mu < 1 \text{ and } \phi > 0$.

With this parameterization the beta regression model can be written as:

$$g(\mu_i) = \mathbf{x}_i^{\mathrm{T}}\boldsymbol{\beta} = \boldsymbol{\eta}_i \tag{9}$$

Where, $g(\cdot)$ is strictly increasing and double differentiable link function that maps (0, 1) in to the real line $\mathbb{R}$, $\mathbf{x}_i = (x_{i1}, \dots\dots\dots, x_{ik})^{\mathrm{T}}$ is a vector of k explanatory variables, $\boldsymbol{\beta} = (\beta_1, \dots\dots\dots, \beta_k)^{\mathrm{T}}$ is a vector of unknown k unknown regression parameters ($k < n$), and $\boldsymbol{\eta}_i$ is a linear predictor (i.e. $\boldsymbol{\eta}_i = \beta_1 x_{i1} + \cdots + \beta_k x_{ik}$, usually $x_{i1} = 1$ for all i so that the model has an intercept) (Cribari-Neto and Zeileis 2010).

We used various combinations of explanatory variables to predict proportion of aboveground biomass present in different components. The logit link function $g(\mu) = \log\left(\frac{\mu}{1-\mu}\right)$ was used, thus the predicted proportions are obtained as $\mu_i = \frac{exp(\eta_i)}{1+\exp(\eta_i)}$. The beta regression was performed in R 3.1.2 (R Core Team 2014) with function `betareg` in package `betareg` (Cribari-Neto and Zeileis 2010).

3.2.3.2. Dirichlet Regression

The Dirichlet distribution is a multivariate generalization of the beta distribution and takes the following form:

$$f(\boldsymbol{y};\boldsymbol{\alpha}) = \frac{1}{\mathrm{B}(\boldsymbol{\alpha})}\prod_{c=1}^{C} y_c^{(\alpha_c-1)} \tag{10}$$

Where, α_c are the shape parameters for each variable, $\alpha_c > 0$, for all c, $y_c \in (0,1)$, $\sum_{c=1}^{C} y_c = 1$ for all c, and C is the number of variables. $\mathrm{B}(\alpha_c) = \frac{\prod_{c=1}^{C}\Gamma(\alpha_c)}{\Gamma(\sum_{c=1}^{C}\alpha_c)}$ is the multinomial beta function and $\Gamma(\cdot)$ is the gamma function. If $C = 2$, then the Dirichlet distribution reduces to the beta distribution. Maier (2014) used the generalization of Ferrari and Cribari-Neto (2004) and re-parameterized the Dirichlet distribution with mean and precision parameters $\mu_c = \frac{\alpha_c}{\phi}$ and $\phi = \alpha_0 = \sum_{c=1}^{C}\alpha_c$ respectively. Then the Dirichlet density has the following form:

$$f(\boldsymbol{y};\boldsymbol{\mu},\phi) = \frac{1}{\mathrm{B}(\boldsymbol{\mu}\phi)}\prod_{c=1}^{C} y_c^{(\mu_c\phi-1)} \tag{11}$$

Where, $0 < \mu < 1$ and $\phi > 0$.

The Dirichlet regression is useful for modeling data that represent the components as percentage of total. With the usual parameterization, the regression model can be formulated as:

$$g(\alpha_c) = \eta_c = \boldsymbol{X}^{[c]}\boldsymbol{\beta}^{[c]} \tag{12}$$

Where $g(\cdot)$ is the link-function which is log$(\cdot)$ for the model with usual parameterization (Maier 2014), the superscript $[c]$ represent predicted proportion of component c. The predicted values are obtained as $\mu_c = exp(\eta_c)$. The Dirichlet regression was performed in R 3.1.2 (R Core Team 2014) with function `DirichReg` in package `DirichletReg` (Maier 2014a).

3.2.3.3. Multinomial Loglinear Regression

In this method, four components (bole, bole bark, branch, and foliage) were set to four nominal values. The models to predict proportions of total tree biomass found in bole wood, stem bark, branch and foliage were fit simultaneously using a multinomial logit model. The models for component proportions were:

$$p_{Bole} = \frac{1}{1+e^{(a1+a2\times X_1+a_3\times X_2)}+e^{(b1+b2\times X_1+b_3\times X_2)}+e^{(c1+c2\times X_1+c_3\times X_2)}} \quad (13)$$

$$p_{Bark} = \frac{e^{(a1+a2\times X_1+a_3\times X_2)}}{1+e^{(a1+a2\times X_1+a_3\times X_2)}+e^{(b1+b2\times X_1+b_3\times X_2)}+e^{(c1+c2\times X_1+c_3\times X_2)}} \quad (14)$$

$$p_{Foliage} = \frac{e^{(b1+b2\times X_1+b_3\times X_2)}}{1+e^{(a1+a2\times X_1+a_3\times X_2)}+e^{(b1+b2\times X_1+b_3\times X_2)}+e^{(c1+c2\times X_1+c_3\times X_2)}} \quad (15)$$

$$p_{Branch} = \frac{e^{(c1+c2\times X_1+c_3\times X_2)}}{1+e^{(a1+a2\times X_1+a_3\times X_2)}+e^{(b1+b2\times X_1+b_3\times X_2)}+e^{(c1+c2\times X_1+c_3\times X_2)}} \quad (16)$$

Where, p_{Bole}, p_{Bark}, $p_{Foliage}$ and p_{Branch} are proportions of total aboveground biomass in bole, bark, foliage, and branch respectively; X_1= DBH; X_2= total tree height; and ai, bi, ci $(i = 1, 2, 3)$ are model parameters. The multinomial logit fit provides the "probability" of observing these components and can be considered as the proportion of biomass in each component and estimated by model parameters (Boudewyn et al. 2007). The multinomial loglinear regression (MLR) was performed in R 3.1.2 (R Core Team 2014) with function `multinom` in package `nnet` (Venables and Ripley 2002). The biomass present in each component was used as frequency weight and component bole was used as reference group.

3.3. Results and Discussion

The methods used to estimate aboveground biomass and/or its components in this study were applied to all five species. Performance of all the methods was evaluated based on the bias and RMSE produced by each method. The FIA-CRM, FIA-PNW, and the Jenkins methods were biased and produced the highest values for root mean squared error. The average bias and RMSE produced by these methods are given in Table 3.2. These methods produced similar estimates for total aboveground biomass except for Douglas fir. The Jenkins method for Douglas fir produced total aboveground biomass that was respectively 18.4 and 23.7 percent higher than the estimates provided by the FIA-PNW and FIA-CRM methods. Despite their similar predictions for total aboveground biomass, these methods showed discrepancies in component biomass estimates. However, none of these methods were consistent in over- or under-predicting the component masses (Figure 3.3).

It is important to note that the component biomass estimates obtained from these methods were similar for lodgepole pine and red alder. Both lodgepole pine and red alder were the trees with smaller diameter (average diameter 24.6 and 27.1 cm for lodgepole pine and red alder respectively) in our study. Indeed, these methods were more sensitive to tree size compared to other methods. For example, the RMSE percent for total aboveground biomass using Jenkins equations for Douglas fir dropped from 57.7 percent to 11.1 percent when this approach was applied to the trees that are less than 94 cm in DBH. With the FIA-CRM, the RMSE percent for total aboveground biomass of Douglas fir decreased from 10.2 to 7.1 percent for trees that are less than 94 cm in DBH. Similarly, the RMSE percent for total aboveground biomass of Douglas fir decreased from 16.3 to 8.5 percent in using FIA-PNW approach for trees that are less than 94 cm in DBH.

The parameter estimates and their approximate standard errors of the simple and extended SUR models for Douglas fir are presented in Table 3.3. The average bias and RMSE produced by simple

and extended SUR approaches are presented in Table 3.4. Both of these methods consistently provided lower RMSE compared to the FIA-CRM, FIA-PNW, and Jenkins methods. Including additional explanatory variables than just DBH in the SUR model resulted in the decrease in RMSE percent from 10.7 to 8.3 for Douglas fir, 4.7 to 4.3 for grand fir, 22.8 to 20.5 for lodgepole pine, 10.7 to 1.9 for Western hemlock, and 14.0 to 8.0 for red alder total aboveground biomass respectively. The RMSE for bole biomass estimation was reduced by 2.3, 0.2, 6.9, 10.1, and 2.0 percent for Douglas fir, grand fir, lodgepole pine, western hemlock, and red alder respectively by using the extended SUR approach instead of the simple SUR. It is logical because one would, for example, expect differences, at least, in the bole biomass for a same DBH tree with different height which would not be accounted for by DBH only models.

However, it should be noted that even though the RMSE for total aboveground biomass is decreased by using the extended SUR approach, the RMSE for some component biomass increased (Figure 3.4). This could have been avoided by not constraining the extended SUR models i.e. fitting independent component models rather than fitting a system of equations which in turn would have affected the additivity of the component models.

The beta, Dirichlet, and multinomial loglinear regressions provided the predicted proportions of each component biomass. The predicted proportions were then applied to observed total aboveground biomass to obtain predicted biomass estimates in different components ($\hat{B}_c$) i.e. $\hat{B}_c = \hat{p} * AGB$, where $\hat{p}$ and AGB are predicted proportions and observed total aboveground biomass (kg) respectively. The parameter estimates for these models (including parameters estimates for SUR models for other species) are available on request. The bias and RMSE produced by these methods are given in Table 3.5. These methods unbiasedly predicted component proportions for all species. Additionally, these methods consistently produced smaller values for bias and RMSE compared to the FIA-CRM, FIA-PNW, and Jenkins methods but there were some exceptions when these methods were compared against the simple and extended SUR methods. However, there was no clear winner within this group of methods.

The beta regression produced smaller RMSE compared to the simple SUR models except for grand fir foliage and bark biomass and Douglas fir branch biomass while it produced smaller RMSEs than the extended SUR models except for grand fir foliage, bark, and branch biomass. It is unclear whether the poor performance of beta regression in grand fir component proportion estimation is due to smaller sample size (n=9) because it performed better than both SUR methods for bole mass and better than the simple SUR for branch biomass for this species. In case of other species-component combinations, beta regression produced up to 24.6 and 17.7 percent lower RMSE for conifers and up to 46.8 and 40.9 percent lower RMSE for red alder (hard wood) compared to the simple and extended SUR methods respectively.

Similarly, the Dirichlet regression also produced smaller RMSE compared to the simple and extended SUR methods with some exceptions. It specifically performed poorly for red alder producing up to 32.1 and 22.3 percent higher RMSE compared to the simple and extended SUR methods respectively. In the case of conifers, it produced smaller RMSE compared to SUR except for Douglas fir branch biomass while it performed better than extended SUR except for western hemlock bark and grand fir branch biomass estimation. One advantage of using Dirichlet regression over beta regression is that the Dirichlet regression allows simultaneous fitting of the component proportions and therefore the predicted proportions sum to 1.

The MLR consistently produced smaller RMSE compared to the SUR method for all species and all components. It also produced smaller RMSE compared to the extended SUR for all species and components except for grand fir branch biomass for which it produced 2.7 percent higher RMSE compared to the extended SUR method. Once again, one of the reasons for this could have been smaller sample size (n=9) for grand fir. In a simulation study, Peduzzi et al. (1996) showed that with less than 10 events per predictive variables, the logistic regression model produced biased coefficients in both positive and negative directions. However, this method provided better estimates (up to 4.4 and 6.8

percent smaller RMSE compared to simple and extended SUR approaches respectively) for other components even in the grand fir.

3.4. Summary and Conclusion

The methods to estimate aboveground biomass and/or its components used in this study provided differing estimates for total aboveground biomass and its components. Both simple and extended SUR approaches performed better than the FIA-CRM, the FIA-PNW, and the Jenkins approaches for our dataset. The FIA-CRM, FIA-PNW, and Jenkins methods provided results comparable to both SUR approaches for smaller trees (less than 94 cm DBH) but they were highly biased for bigger trees. We would like to note that the larger trees (greater than 94 cm DBH) in our sample were all Douglas-fir and the extension of this result to other species might need further validation. The methods for estimating proportions were clearly superior to the FIA-CRM, the FIA-PNW, and the Jenkins methods in terms of bias and RMSE. These methods were also superior to both simple and extended SUR approaches with some exceptions. However, none of these methods was clearly superior to the other.

The Jenkins method for our Douglas fir trees produced total aboveground biomass that was 18.4 and 23.7 percent higher than the estimates provided by the FIA-PNW and FIA-CRM methods. Zhou and Hemstrom (2009) reported similar differences (17 and 20 percent respectively) for the major softwood species in Oregon. The FIA-CRM, FIA-PNW, and Jenkins methods produced up to 3.7, 2.7, and 3.4 times higher bark biomass estimates. The bole biomass estimates were within 10 percent for Douglas fir and grand fir and within 20 percent for lodgepole pine, red alder, and western hemlock. These methods were very inconsistent for branch and foliage biomass, over-estimating for some species and under-estimating for others. The FIA-CRM and Jenkins methods were developed for larger scale biomass estimation, therefore the estimates obtained from these methods are inconsistent at the smaller scale, perhaps due to generalized nature of the equations.

In fitting the SUR models, the use of explanatory variables other than DBH was justified by the reduction in RMSE produced by the extended SUR approach in estimating total aboveground biomass. This also reduced RMSE for bole biomass. However, for some species, it did not improve prediction for other components. This could be because the added variable provided more information for the bole and total biomass but not for other components. We also found that, for Douglas fir, the addition of crown length rather than total height to the DBH only model was useful in estimating foliage and branch biomass. By using crown length instead of total height, the adjusted R-squared of the foliage and branch models increased from 0.44 and 0.55 to 0.81 and 0.89 respectively.

The knowledge of biomass distributions in different tree components is essential to determine which portion of the tree can provide what amount of biomass for different purposes. The proportion of component biomass can be predicted using the beta, Dirichlet, and multinomial loglinear regressions. These methods were not only unbiased and but also produced very small RMSE values compared to the FIA-CRM, FIA-PNW, and Jenkins methods. They also provided better results compared to both SUR approaches for most of the species and components. This was expected because the measurement of proportion take values on the open interval (0, 1) and the influence of explanatory variables on continuous responses bounded between 0 and 1 can be investigated with the beta regression (Ferrari and Cribari-Neto (2004)) and the Dirichlet regression is multivariate extension of the beta. Additionally, the simultaneous fitting in the Dirichlet regression ensures that the predicted component proportions sum to 1.

Even though the methods or models that are capable of predicting forest biomass at the large-scale are desired, the use of such models without local calibration could lead to serious bias. With increasing interest in the utilization of component biomass for bioenergy, the importance of component models is also increasing. The differences in the scale of development and application of model could result in higher bias for the component biomass estimation. The findings of the study should provide

information on the efficiency of selected methods in quantifying component and total aboveground biomass. Application of the methods to predict component proportions for other species and locations and with the larger dataset would further validate their accuracy.

3.5. References

Beaucham, J.J., and J.S. Olson. 1973. Corrections for bias in regression estimates after logarithmic transformation. *Ecology*. 54(6): 1403-1407.

Boudewyn, P., X. Song, S. Magnussen, and M. Gillis. 2007. Model-based, volume-to-biomass conversion for forested and vegetated land in Canada. Canadian Forest Service, Victoria, Canada. Information Report BC-X-411.

Brown, S. 2002. Measuring carbon in forests: current status and future challenges. *Environmental pollution*. 116(3): 363-372.

Cribari-Neto, F., and A. Zeileis. 2010. Beta Regression in R. *J. Stat. Softw.* 34(2): 1-24.

Czapowskyj, M.M., D.J. Robison, R.D. Briggs, and E.H. White. 1985. Component biomass equations for Black Spruce in Maine. *Res. Pap. NE-564.* Broomal, PA: U.S. Department of Agriculture, Forest Service, p. 7.

De Paola, M., V. Scoppa, and R. Lombardo. 2010. Can gender quotas break down negative stereotypes? Evidence from changes in electoral rules. *Journal of Public Economics*. 94(5): 344-353.

Eskelson, B.N., L. Madsen, J.C. Hagar, and H. Temesgen. 2011. Estimating Riparian understory vegetation cover with Beta regression and copula models. *For. Sci.* 57(3): 212-221.

Ferrari, S., and F. Cribari-Neto. 2004. Beta regression for modelling rates and proportions. *J. App. Stat.* 31(7): 799-815.

Flewelling, J.W., and L.V. Pienaar. 1981. Multiplicative regressions with lognormal errors. *For. Sci.* 27(2): 281-289.

Harrison, R.B., T.A. Terry, C.W. Licata, B.L. Flaming, R. Meade, I.A. Guerrini, B.D. Strahm, D. Xue, M.R. Lolley, A.R. Sidell, G.L. Wagoner, D. Briggs, and E.C. Turnblom. 2009. Biomass and stand characteristics of a highly productive mixed Douglas-fir and Western Hemlock plantation in Coastal Washington. *West. J. Appl. For.* 24(4): 180-186.

He, Q.S., E. Chen, R. An, and Y. Li. 2013. Above-ground biomass and biomass components estimation using LiDAR data in a coniferous forest. *Forests*. 4(4): 984-1002.

Heath, L.S., M. Hansen, J.E. Smith, P.D. Miles, and B.W. Smith. 2008. Investigation into calculating tree biomass and carbon in the FIADB using a biomass expansion factor approach. *McWilliams, WH, Moisen, GG & Czaplewski, RL (eds.)*.

Hollinger, D.Y. 2008. Defining a Landscape-Scale Monitoring Tier for the North American Carbon Program. In *Field Measurements for Forest Carbon Monitoring.* C. Hoover (ed.), Springer Netherlands, p 3-16.

Hoover, C.M. 2008. Field measurements for forest carbon monitoring: A landscape-scale approach. Springer: New York, NY, USA, 2008.

Hubben, G.A.A., D. Bishai, P. Pechlivanoglou, A.M. Cattelan, R. Grisetti, C. Facchin, F. Compostella, J. Bos, M. Postma, and A. Tramarin. 2008. The societal burden of HIV/AIDS in Northern Italy: An analysis of costs and quality of life. *AIDS care*. 20(4): 449-455.

Jenkins, J.C., D.C. Chojnacky, L.S. Heath, and R.A. Birdsey. 2003. National-scale biomass estimators for United States tree species. *For. Sci.* 49(1): 12-35.

Junior, L.R.N., V.L. Engel, J.A. Parrotta, A.C.G.d. Melo, and D.S. Ré. 2014. Allometric equations for estimating tree biomass in restored mixed-species Atlantic Forest stands. *Biota Neotropica*. 14(2): 1-9.

Korhonen, L., K.T. Korhonen, P. Stenberg, M. Maltamo, and M. Rautiainen. 2007. Local models for forest canopy cover with beta regression. *Silva Fennica*. 41(4): 671.

Kuznetsova, T., M. Tilk, H. Parn, A. Lukjanova, and M. Mandre. 2011. Growth, aboveground biomass, and nutrient concentration of young Scots pine and lodgepole pine in oil shale post-mining landscapes in Estonia. *Environ. Monit. Assess.* 183(1-4): 341-350.

Lambert, M.C., C.H. Ung, and F. Raulier. 2005. Canadian national tree aboveground biomass equations. *Can. J. For. Res.* 35(8): 1996-2018.

Lu, D., P. Mausel, E. Brondizio, and E. Moran. 2002. Above-ground biomass estimation of successional and mature forests using TM images in the Amazon Basin. In *Advances in Spatial Data Handling*, Springer, p 183-196.

Luo, Y., X. Wang, X. Zhang, Y. Ren, and H. Poorter. 2013. Variation in biomass expansion factors for China's forests in relation to forest type, climate, and stand development. *Ann. For. Sci.* 70(6): 589-599.

Maier, M.J. 2014. DirichletReg: Dirichlet Regression for Compositional Data in R. Research Report Series / Department of Statistics and Mathematics, 125. WU Vienna University of Economics and Business, Vienna.

Maier, M.J. 2014a. DirichletReg: Dirichlet Regression in R. R package version 0.5-2.

Miles, P.D., and W.B. Smith. 2009. Specific gravity and other properties of wood and bark for 156 tree species found in North America. *Res. Note. NRS-38*. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, p 35.

Nunes, L., D. Lopes, F.C. Rego, and S.T. Gower. 2013. Aboveground biomass and net primary production of pine, oak and mixed pine-oak forests on the Vila Real district, Portugal. *For. Ecol. Manage.* 305: 38-47.

Parresol, B.R. 1999. Assessing tree and stand biomass: A review with examples and critical comparisons. *For. Sci.* 45(4): 573-593.

Parresol, B.R. 2001. Additivity of nonlinear biomass equations. *Can. J. For. Res.-Rev. Can. Rech. For.* 31(5): 865-878.

Peduzzi, P., J. Concato, E. Kemper, T.R. Holford, and A.R. Feinstein. 1996. A simulation study of the number of events per variable in logistic regression analysis. *Journal of clinical epidemiology*. 49(12): 1373-1379.

R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Raile, G.K. 1982. Estimating stump volume. *Res. Pap. NC-224.* St Paul, MN: U.S. Department of Agriculture, Forest Service, p 4.

Repola, J. 2006. Models for vertical wood density of Scots pine, Norway spruce and birch stems, and their application to determine average wood density. *Silva Fennica*. 40(4): 673-685.

Ritchie, M.W., J. Zhang, and T.A. Hamilton. 2013. Aboveground Tree Biomass for Pinus ponderosa in Northeastern California. *Forests*. 4(1): 179-196.

SAS Institute Inc. 2013. Base SAS® 9.4 Procedures Guide: Statistical Procedures, Second Edition. Cary, NC: SAS Institute Inc.

Smithson, M., and J. Verkuilen. 2006. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological methods*. 11(1): 54.

Tumwebaze, S.B., E. Bevilacqua, R. Briggs, and T. Volk. 2013. Allometric biomass equations for tree species used in agroforestry systems in Uganda. *Agroforest Systems*. 87(4): 781-795.

Vashum, K.T., and S. Jayakumar. 2012. Methods to Estimate Above-Ground Biomass and Carbon Stock in Natural Forests-A Review. *J. Ecosyst. Ecogr.* 2(4): 1-7.

Venables, W.N., and B.D. Ripley. 2002. Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0.

Westfall, J.A., D.W. MacFarlane, and A.R. Weiskittel. 2012. Biomass measurement and modeling challenges for hardwood species in the northern region. *In Moving from status to trends: Forest Inventory and Analysis (FIA) symposium 2012. December 4-6, 2012, Baltimore, MD*.

Xiao, C.W., J.C. Yuste, I.A. Janssens, P. Roskams, L. Nachtergale, A. Carrara, B.Y. Sanchez, and R. Ceulemans. 2003. Above- and belowground biomass and net primary production in a 73-year-old Scots pine forest. *Tree Physiology*. 23(8): 505-516.

Zhang, Y.J., B.E. Borders, R.E. Will, and H.D. Posadas. 2004. A model for foliage and branch biomass prediction for intensively managed fast growing loblolly pine. *For. Sci.* 50(1): 65-80.

Zhou, X., and M.A. Hemstrom. 2009. Estimating aboveground tree biomass on forest land in the Pacific Northwest: a comparison of approaches. *Res. Pap. PNW-RP-584*. Portland, OR: U.S. Department of Agriculture, Forest Service. Pacific Northwest Research Station, p. 18.

Zhou, X., and M.A. Hemstrom. 2010. Timber volume and aboveground live tree biomass estimations for landscape analyses in the Pacific Northwest. *Gen. Tech. Rep. PNW-GTR-819*. Portland, OR: U.S. Department of Agriculture, Forest Service, p. 31.

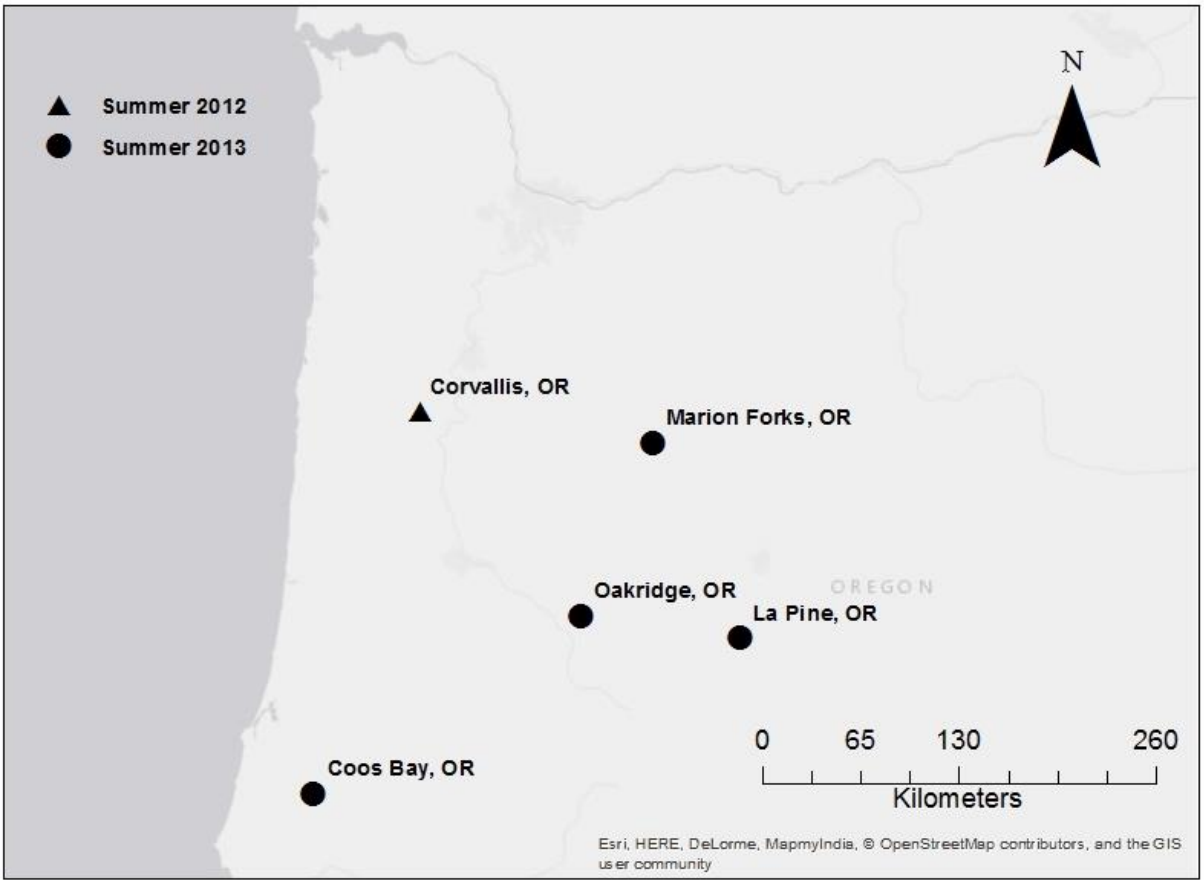

Figure 3.1. Locations of study sites. Data collection was carried out in the summer of 2012 and 2013.

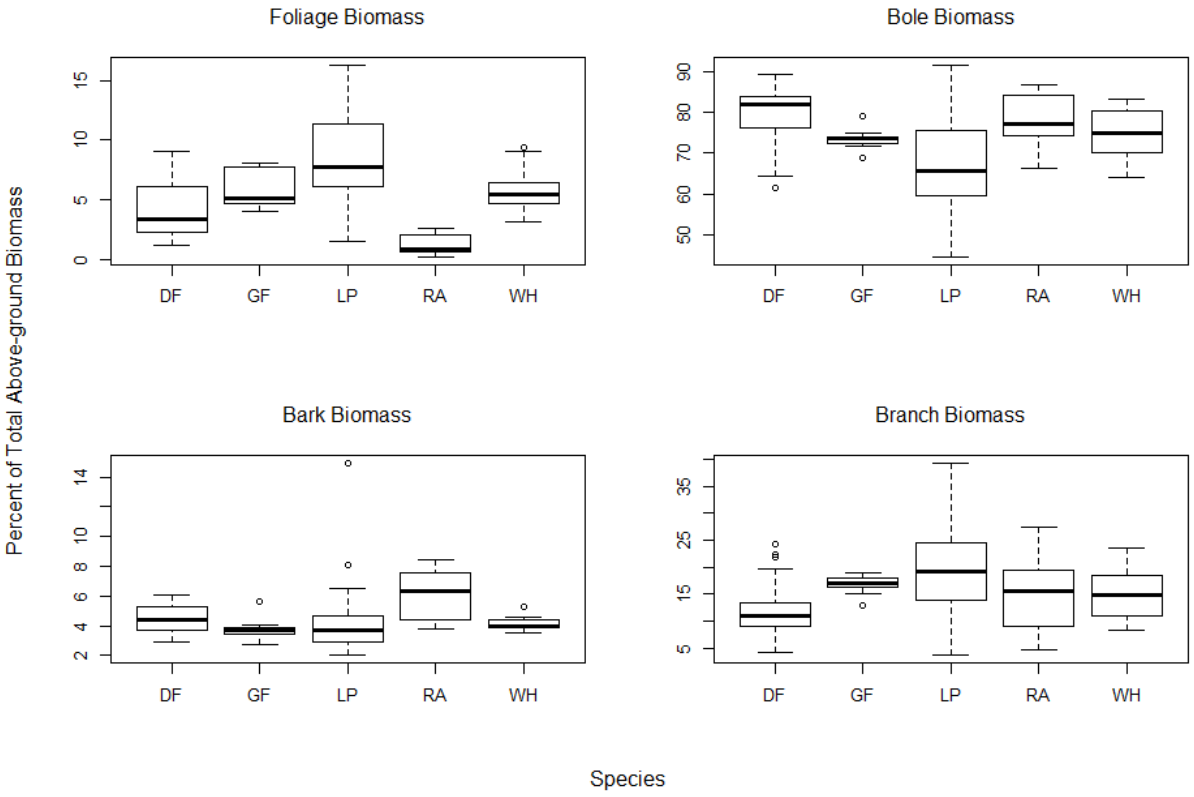

Figure 3.2. Aboveground biomass distribution in different tree components (percentage of total aboveground biomass) in different species (Species are DF = Douglas-fir; GF = Grand Fir; LP = Lodgepole Pine; RA = Red Alder; WH = Western Hemlock).

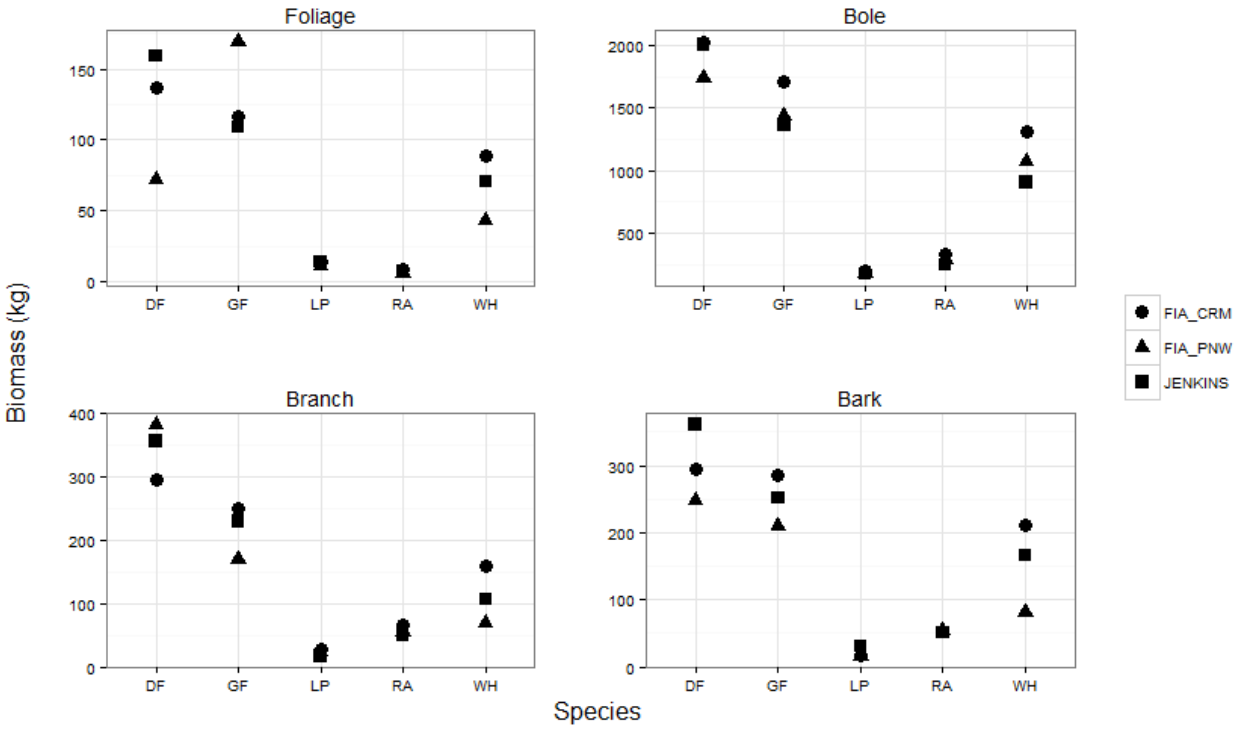

Figure 3.3. Average component biomass estimates produced by the FIA-CRM, FIA-PNW, and Jenkins methods in different species (Species are DF = Douglas-fir; GF = Grand Fir; LP = Lodgepole Pine; RA = Red Alder; WH = Western Hemlock).

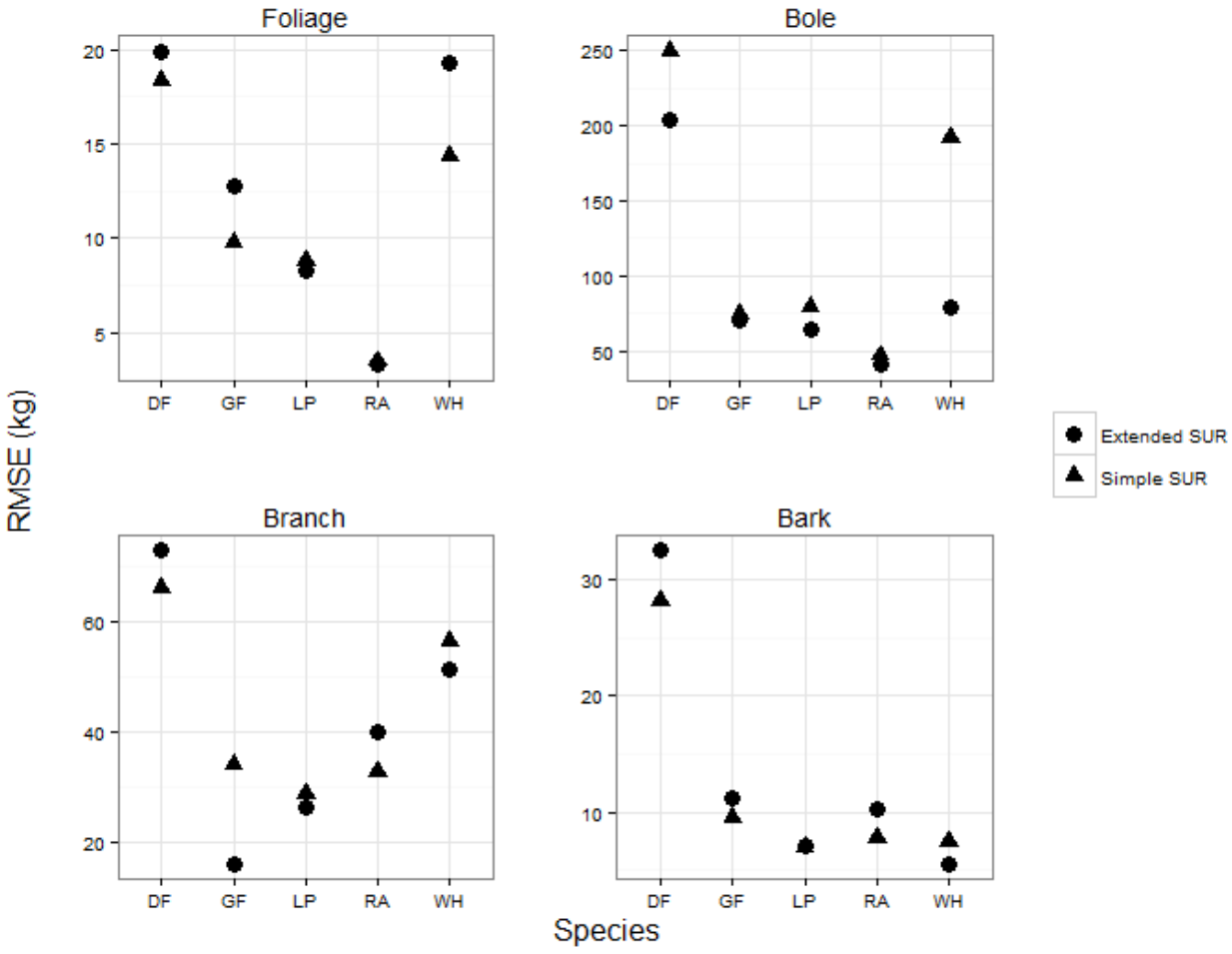

Figure 3.4. RMSE produced by simple and extended SUR approaches in estimating component biomass for different species (Species are DF = Douglas Fir; GF = Grand-fir; LP = Lodgepole Pine; RA = Red Alder; WH = Western Hemlock).

Table 3.1. Summary statistics for tree data sampled (Species are DF = Douglas-fir; GF = Grand Fir; LP = Lodgepole Pine; RA = Red Alder; WH = Western Hemlock).

Species	Number of trees	DBH (cm)			Total Height (m)		
		Min	Max	Mean	Min	Max	Mean
DF	22	19.3	114.0	54.9	16.6	48.8	33.0
GF	9	23.9	84.3	53.1	18.3	44.4	32.7
LP	33	13.5	42.9	24.6	9.2	31.9	17.0
RA	14	16.3	51.8	27.1	12.8	31.6	23.7
WH	12	18.0	69.9	44.7	14.4	39.8	31.3

Table 3.2. Average bias and RMSE for component and total aboveground biomass produced by the FIA-CRM, FIA-PNW, and Jenkins approaches (Species are DF = Douglas-fir; GF = Grand Fir; LP = Lodgepole Pine; RA = Red Alder; WH = Western Hemlock).

Method	Species	Bias (kg)					RMSE (kg)				
		Foliage	Bark	Branch	Bole	Total	Foliage	Bark	Branch	Bole	Total
FIA-CRM	DF	-79.6	-189.6	-73.3	-109.7	-32.5	126.0	281.1	136.8	289.8	234.5
	GF	-14.4	-208.7	105.4	-161.2	107.3	24.9	270.6	166.8	255.6	281.0
	LP	3.2	-4.8	19.4	25.1	72.4	12.7	12.5	37.9	70.5	101.8
	WH	-11.2	-148.8	77.7	-184.0	17.6	20.9	184.1	154.3	240.2	88.3
	RA	-4.8	-29.4	3.5	-27.7	-2.5	6.4	41.2	79.6	53.8	87.7
FIA-PNW	DF	-14.9	-143.2	-161.2	182.8	-136.6	37.7	232.5	280.3	311.7	376.1
	GF	-67.0	-132.4	185.4	114.6	100.7	88.4	175.1	263.6	229.8	265.7
	LP	6.6	-3.9	22.7	38.8	64.2	12.0	9.4	40.6	78.7	99.7
	WH	34.5	-19.6	166.4	45.0	226.4	49.4	26.9	237.3	81.8	305.7
	RA	-1.8	-29.6	13.8	19.6	2.1	2.5	41.9	71.0	35.8	71.5
Jenkins	DF	-102.1	-255.5	-134.8	-93.8	-586.2	176.0	418.7	246.5	597.6	1327.7
	GF	-7.0	-174.7	125.9	180.0	124.2	16.3	228.5	188.1	249.4	203.9
	LP	3.5	-18.7	29.7	49.1	63.6	9.9	26.7	44.6	102.5	104.2
	WH	7.0	-103.8	128.9	216.8	248.9	13.6	136.2	185.0	305.8	323.9
	RA	-3.6	-26.8	17.5	61.0	48.2	5.0	38.8	81.3	80.3	105.6

Table 3.3. Parameter estimates, their approximate standard errors, and p-values of the simple and extended SUR models for Douglas-fir.

Method	Parameter	Estimate	Approximate Standard Error	Approximate Pr > |t|
Simple SUR	a11	-0.34906	0.0961	0.0018
	a12	1.925357	0.0210	<.0001
	a21	-3.24695	0.7009	0.0002
	a22	1.919085	0.1536	<.0001
	a31	-2.60374	0.8215	0.0050
	a32	1.942273	0.1799	<.0001
	a41	-3.24152	0.8960	0.0018
	a42	1.75849	0.1973	<.0001
Extended SUR	a11	-1.99252	0.0618	<.0001
	a12	1.597243	0.0115	<.0001
	a13	0.826614	0.00861	<.0001
	a21	-4.45886	0.8061	<.0001
	a22	2.168904	0.1756	<.0001
	a31	-3.85644	0.4539	<.0001
	a32	1.795146	0.0931	<.0001
	a33	0.611343	0.0459	<.0001
	a41	-4.51727	0.5643	<.0001
	a42	1.496697	0.1467	<.0001
	a43	0.778996	0.1709	0.0002

Table 3.4. Average bias and RMSE for component and total aboveground biomass produced by the simple and extended SUR approaches (Species are DF = Douglas Fir; GF = Grand-fir; LP = Lodgepole Pine; RA = Red Alder; WH = Western Hemlock).

Method	Species	Bias (kg)					RMSE (kg)				
		Foliage	Bark	Branch	Bole	Total	Foliage	Bark	Branch	Bole	Total
Simple SUR	DF	4.0	0.4	0.3	-44.1	-39.3	18.4	28.2	66.2	249.2	245.8
	GF	-1.2	1.6	-6.6	1.8	-4.4	9.8	9.5	34.2	74.3	98.4
	LP	0.7	0.1	2.3	-2.7	0.5	8.8	7.1	28.7	79.7	67.9
	WH	-5.8	-0.8	-20.1	12.1	-14.8	14.4	7.5	56.5	192.4	159.7
	RA	-2.8	-6.1	24.4	-16.3	-0.5	3.5	7.8	32.8	47.1	56.2
Extended SUR	DF	-0.7	12.7	-16.6	-16.3	-20.9	19.9	32.5	72.9	204.3	190.9
	GF	-4.9	-5.1	6.6	2.1	-1.4	12.8	11.3	16.1	70.8	90.4
	LP	2.5	0.2	6.8	-10.9	-1.2	8.3	7.1	26.2	64.4	60.9
	WH	0.4	-1.4	-0.8	-0.7	-2.7	19.3	5.6	51.2	79.4	27.8
	RA	0.5	-2.3	5.8	-3.3	1.0	3.3	10.3	39.9	40.9	32.3

Table 3.5. Average bias and RMSE of component biomass produced by the beta, multinomial loglinear, and Dirichlet regression approaches (Species are DF = Douglas-fir; GF = Grand Fir; LP = Lodgepole Pine; RA = Red Alder; WH = Western Hemlock). Predicted component biomass was obtained by applying predicted proportions to the observed total aboveground biomass.

Method	Species	Bias (kg)				RMSE (kg)			
		Foliage	Bark	Branch	Bole	Foliage	Bark	Branch	Bole
Beta	DF	-0.218	-2.619	-1.221	2.994	15.3	20.0	72.5	79.8
	GF	1.109	-0.392	-0.152	-0.320	13.7	14.3	24.3	36.0
	LP	-0.122	0.105	-0.111	-0.766	5.9	6.6	22.2	25.5
	WH	0.811	-0.386	1.330	-1.567	13.6	5.4	32.8	43.2
	RA	0.350	-0.258	3.801	-3.230	1.9	4.2	22.4	20.8
MLR	DF	0.001	-0.005	-0.004	-0.083	17.2	23.2	63.7	74.3
	GF	-0.002	0.000	0.001	0.001	9.4	6.1	25.7	35.4
	LP	-0.007	-0.006	-0.025	-0.114	7.0	6.5	21.6	27.9
	WH	0.009	0.008	0.042	0.108	11.4	5.1	29.8	40.8
	RA	-0.001	-0.017	-0.023	-0.173	1.1	2.6	13.6	12.9
Dirichlet	DF	-1.948	-2.971	6.807	-1.979	17.5	24.3	70.8	83.0
	GF	-0.832	0.177	-0.312	0.968	9.4	6.0	24.7	33.5
	LP	-1.515	-1.356	0.729	1.990	7.5	7.0	21.7	28.6
	WH	1.473	-2.130	5.325	-4.501	13.3	7.1	34.9	45.7
	RA	-0.859	-5.289	15.286	-9.353	1.6	15.8	51.6	35.6

4. Manuscript 3

Calibration of Volume and Component Biomass Equations for Douglas-fir and Lodgepole Pine Forests in Western Oregon

Krishna P. Poudel and Hailemariam Temesgen

Abstract

Volume and biomass equations developed for national or regional scales are commonly used at the local scale. However, these models may not be unbiased at the local scale if there is spatial variation in the tree form due to one or more unknown predictors and this regional bias could be reduced or removed if the models are localized to each sub-region or subarea. The performance of regional volume and component biomass equations was evaluated using data collected destructively from sampled Douglas-fir and lodgepole pine trees. Upper stem diameters and volume estimates were also obtained by fitting a variable-exponent taper function. The volume and component biomass equations were calibrated using three different adjustment methods (a) use of a correction factor based on ordinary least square regression through origin (OLS-RTO method), (b) use of a correction factor based on OLS with intercept (OLS-WI method), and (c) an inverse approach. The regional volume equations performed fairly well and produced similar results as the locally fitted volume equations of the same form. The taper based volume estimates for Douglas-fir were comparable (root mean squared error (RMSE) within 3%) to the volume estimates obtained from regional volume equations but for lodgepole pine taper based volume estimates were highly biased. The regional predicted component biomass estimates were highly biased. All the adjustment methods were able to improve the performance of regional equations. Based on the leave one out cross validation, the RMSEs in cubic volume including top and stump (CVTS) and component biomass estimation was similar for the adjustments from a correction factor based on OLS-RTO and an inverse approach. The OLS-WI method did not perform as well as other two adjustment methods. When only one tree is available for calibration of regional models, we found it useful to use the tree with maximum DBH to calibrate regional CVTS and bark biomass equations and the dominant tree to calibrate bole, foliage, and branch biomass. Selecting trees of such characteristics to calibrate the regional models was more useful than randomly selecting one tree. The decreasing trend in RMSE by using randomly selected m trees slowed down significantly after five trees.

4.1. Introduction

The estimation of total and merchantable stem volume is crucial for land managers to ascertain continuous production. Two approaches are used to obtain these estimates (1) use of a direct volume equation and (2) integration of stem profile equations or taper functions. The volume equations are typically species specific regression models that range from a single entry simple linear regression models that relate tree volume with diameter at breast height (DBH) to multiple entry nonlinear models that use DBH, height, and crown ratio etc. The taper functions provide additional advantages in the sense that they can provide diameter inside or outside bark at certain height on the stem, estimates of total and merchantable stem volume, and volume of an individual log (Kozak 1988) or a height at given diameter (Li et al. 2012).

Several published volume and profile equations are available and are being used at different scales of forest management. Generally, these volume and profile equations are developed for either a national scale or a regional scale applications but commonly used to predict tree volume at local scales. The Forest Inventory and Analysis program in the Pacific Northwest Research Station (FIA-PNW) uses the direct volume equations for tree volume estimation. Particularly, for Douglas-fir and lodgepole pine cubic volume estimation, the FIA-PNW uses equations and tarif system based on Brackett (1973). The differences in stand conditions affect the tree form and thus tree volume (Bluhm et al. 2007). The evaluation of existing volume and taper equations itself is a rare opportunity (Omule et al. 1987) because it requires destructively sampled data.

Biomass of a tree or its component is defined as its oven-dry weight. Quantifying biomass in a forest ecosystem has received significant attention in recent years because the carbon content in biomass is approximately 50%. The amount of biomass present in a forest determines the potential of that forest to sequester and store carbon dioxide. Various climate change agreements and action plans

at scales ranging from local to international level such as the United Nations Framework Convention on Climate Change and in particular the Kyoto Protocol recognize the importance of forest carbon sink and the need to monitor, preserve and enhance terrestrial carbon stocks (Zianis et al. 2005). Additionally, a number of voluntary and regulated carbon markets have provisions for credits for carbon sequestration from forest management projects that meet certain criteria and verification.

In addition to total aboveground biomass estimate, the knowledge of biomass present in different components is useful for different purposes. Stem wood is important in timber sale; knowledge of crown biomass is useful in fuel load assessment, formulating fire management strategies, and in developing wildfire models while the biomass in small branches and leaves /foliage determines the possibility of installing bioenergy plants in that area. The FIA-PNW calculates the biomass in main stem, whether merchantable or total, from the cubic volume estimates and previously compiled wood density factors. The biomass in other components is calculated from published models that are derived from local tree studies as functions of DBH and total height (Zhou and Hemstrom 2009).

Three major methods used in aboveground biomass estimation in the United States are: 1) the component ratio method (CRM) of the FIA, 2) the equations developed by Jenkins et al. (2003) based on meta-analysis of the previously published models, and 3) the models developed for regional applications. In comparing these three methods, Zhou and Hemstrom (2009) found similar estimates for total aboveground biomass but substantially different estimates for the merchantable biomass for softwood species in Oregon with Jenkins method producing estimates 17 percent higher than the regional equations. We, ourselves, found that the Jenkins method for our Douglas fir trees produced total aboveground biomass that was 18.4 and 23.7 percent higher than the estimates provided by the FIA-PNW and FIA-CRM methods respectively (see Chapter 3).

Accurate estimates of forest aboveground biomass are needed to reduce uncertainties in global and regional terrestrial carbon fluxes (Pflugmacher et al. 2008). Direct measurement of biomass on the ground is time consuming and expensive (Houghton 2005) and the component biomass estimates obtained as percentages of total or stem biomass from the published information might be very simplistic (Hansen 2002). Similar to volume and taper equations, the evaluation of biomass models, in itself, is time consuming and expensive because it requires destructively sampled data.

A stem profile based system with an allometric model for biomass components would have more biological basis than the empirical models now being used (Hansen 2002). Stating the problem of compatibility of forest biomass estimates among different scales, Zeng et al. (2011) used the dummy variable model approach to construct compatible single-tree biomass equations at different scales. They found these models with local parameters to perform better than a population-average model. The regional models may not be unbiased at the local scale if there is spatial variation in the tree form due to one or more unknown predictors and this regional bias could be reduced or removed if the models are localized to each sub-region or subarea (Räty and Kangas 2008).

Model calibration is needed when the predictive validity of model is in question or when the data are inadequate to estimate model inputs. It is the process of systematic adjustment of model parameters such that the adjusted model predicts the observed outputs more accurately (Taylor et al. 2012). Two options are available when sample volume and biomass data is available at the local scale: one is to fit new volume or biomass equation and the other is to use the available data to calibrate the existing regional models, however, the sample size required for fitting new models is relatively larger than that required for calibration of existing equations (Garber et al. 2009).

Calibration itself is sometimes criticized because it requires destructive sampling if the original model formation is based on the destructive sampling (Ketterings et al. 2001). Mixed effects models

have been widely used in forestry applications so as to describe the mean effect by a certain model form and to account the variability due to the stand or plot effects through random effects (e.g. de-Miguel et al. 2014). The calibration process in this approach uses the best linear unbiased predictor of the random effects parameter. Temesgen et al. (2008) used this approach to calibrate nonlinear height-diameter equations while Garber et al. (2009) used this approach in evaluating the effects of height imputation strategies on stand volume estimation. Another commonly used method for calibrating regional models is based on an ordinary least square (OLS) correction factor. A multiplicative correction factor based on an OLS regression is obtained based on observed and regional predicted variable of interests (e.g. Hanus et al. 1999; Hann 2006; Temesgen et al. 2008; and Garber et al. 2009). Temesgen et al. (2008) obtained results from OLS correction factors that were comparable with the results from mixed effects model when four or more trees were subsampled for height while Garber et al. (2009) obtained 95% decrease in root mean squared error (RMSE) when the regional models were calibrated with a subsample of three tree heights.

The objectives of this study are to evaluate the performance of (1) existing volume and biomass equations in terms of their prediction errors; (2) a locally fitted taper equation in estimating inside bark cubic volume and upper stem diameters; (3) evaluate the performance of three different calibration methods (a) use of a correction factor based on OLS regression through origin (OLS-RTO method), (b) use of a correction factor based on OLS with intercept (OLS-WI method), and (c) an inverse approach. The abbreviation RTO for regression through origin was adopted from (Eisenhauer, 2003).

4.2. Materials and Methods

4.2.1. Data

Only the Douglas-fir and lodgepole pine tree data that was also used in Chapter 2 was used in this study. In summary, the tree level attributes DBH, total height, crown base height (height to the base

of the first live branch), crown width, and main stem diameter at 0.15 m, 0.76 m, 1.37 m, and 2.4 m above ground, and every 1.22 m afterwards were recorded. The average diameter for the Douglas-fir trees was 54.9 cm and for lodgepole pine was 24.6 cm and average height was 33 m and 17 m respectively. The details of methods for obtaining total and component biomass is described in Chapter 3. The distribution of aboveground biomass in different components differed among species with the majority of the aboveground biomass (73 to 82 percent on average) being present in the main bole (Figure 3.2).

The "observed" inside bark cubic volume including top and stump (CVTS) in the felled trees was calculated as follows. After the subject tree is felled, it was sectioned into 5.18 m long sections. The inside-bark cubic volume in these sections was calculated using the Smalian's formula below:

$$V = \frac{A_1 + A_2}{2} \times L \quad (1)$$

Where, V is the volume of 5.18 m section in cubic meter, $A_i's$ $(i = 1, 2)$ are the cross-sectional areas at the small and large end of the sections in m^2 with $A_i = \frac{\pi D_i^2}{4}$ where D_i = diameter inside bark (m) at the i^{th} end of the section, and L is the length of the section in meters (5.18 m).

Inside bark volume of stump was calculated as cylinder while the volume of top section was calculated as cone.

$$V_s = \frac{\pi D_s^2 H_s}{4} \quad (2)$$

$$V_t = \frac{\pi D_t^2 H_t}{12} \quad (3)$$

Where, V_s and V_t are inside bark volume of stump and top, respectively, in cubic meter, D_s and D_t are the inside bark diameter (m) at the stump top and bottom of the top section, respectively, and H_s

and H_t are the stump height and length of top section in meters. Total inside-bark CVTS was calculated by summing volume of all sections, stump, and top.

4.2.2. Methods

Tables 4.1 and 4.2 summarize the different sets of volume and biomass equations used by FIA-PNW to obtain CVTS and component biomasses respectively. The FIA-PNW estimates lack biomass for foliage (Melson et al. 2011) therefore we used the foliage biomass equation from Standish et al. (1985) to compute FIA-PNW equivalent foliage biomass estimates. The FIA-PNW volume and biomass equations were also refitted with our dataset to compare the prediction error differences in using published coefficients against the refitted models of the same form.

4.2.2.1. Taper Equation

The Kozak (2004) Model 02 taper equation that has been successfully used by many researchers (e.g. Li et al. 2012, Rojo et al. 2005) was used as the base model for fitting upper stem diameters. Because the diameter inside bark (DIB) was only measured at 5.18 m intervals and diameters outside bark were measured at every 1.22 m, the inside-bark diameter at every 1.22 m was first obtained by fitting a linear model for which the inside bark diameter was calculated by adding e_{ij} drawn from normal distribution with mean zero and variance equal to the residual standard error of the linear model.

$$\widehat{DIB} = \hat{\beta}_0 + \hat{\beta}_1 \times DOB + \hat{\beta}_2 \times RH + e_{ij} \quad (13)$$

Where, $\hat{\beta}_i's$ $(i = 1, 2, 3)$ are regression coefficients estimated from the data and $RH = \frac{h}{HT}$ is the relative height above ground. The adjusted R^2 value for this model was 0.996 and 0.993 for Douglas-fir and lodgepole pine, respectively (n = 673 and 523, residual standard error 1.32 and 0.73). A mixed

effects model with individual tree random effect did not perform (higher Bayesian Information Criterion (BIC)) as well as the fixed effects model.

The variable-exponent taper model of Kozak (2004) Model 02 is of the following form:

$$\hat{d}_i = a_0 D^{a_1} H^{a_2} X_i^{b_1 Z_i^4 + b_2\left[1/e^{D/H}\right] + b_3 X_i^{0.1} + b_4[1/D\,] + b_5 H^{Q_i} + b_6 X_i} \quad (14)$$

Where, $X_i = \left[1.0 - (h_i - H)^{\frac{1}{3}}\right] / \left[1.0 - p^{\frac{1}{3}}\right]$;

$Q_i = \left[1.0 - (h_i - H)^{\frac{1}{3}}\right]$;

$\mathrm{p} = \frac{1.3}{\mathrm{H}}$; and

$Z_i = \frac{h_i}{H}$ with h_i = height on bole, H = total tree height, *D* = DBH, and $\hat{d}_i$ = predicted diameter inside bark at h_i.

Different fitting approaches namely a nonlinear fixed effects model (NFEM), nonlinear mixed effects model (NMEM), nonlinear mixed effects model with first and second order continuous autoregressive error structures (CAR1 and CAR2) were tested. An error model of the following form was fitted.

$$e^2 = a + b * DBH + c * HT + d * TOTHT \quad (15)$$

Where, e is the residual obtained from fitting Kozak 2004 Model 02 with NFEM, *HT* = height on bole, *TOTHT* = total tree height, and a, b, c, and d are regression parameters. The reciprocal of the predicted values of e^2 were used as weights in model fitting. The final model was selected based on the BIC. After the upper stem diameter is obtained from the fitted taper model, the CVTS calculation was done in the similar fashion as in calculating measured volume using Smalian's formula.

4.2.2.2. Calibration of Regional Models

Model calibration is needed when the predictive validity of model is in question or when the data are inadequate to estimate model inputs. We, first, graphically compared the regression lines produced by regional volume equations with the locally refitted volumes and bole biomass equations. The empirical cumulative distribution function (ECDF) of the component biomass estimates obtained from the FIA-PNW equations and the ECDFs of the measured volume and biomass estimates in different aboveground components were also graphically compared. The ECDF is the cumulative distribution function associated with the empirical measure of the sample observations. For independent and identically distributed random variables $x_1, \ldots, x_n$, the ECDF is defined as:

$$\hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\{X_i \leq x\} \quad (16)$$

Where, $\mathbf{1}\{A\}$ is the indicator of an event A.

Additionally, a nonlinear regression model of the following form was fitted to determine if the calibration of the regional models is necessary or not.

$$y = a * y_R^b \quad (17)$$

Where, y and y_R are the measured and regional estimates of the inside bark volume and component biomass estimates and a and b are regression parameters. A joint hypothesis H_0: $(a, b) = (1, 1)$ for both species was tested using a Wald test at 5% level of significance in SAS procedure PROC MODEL (SAS Institute Inc 2013).

4.2.2.3. OLS-RTO Method

If a sample of volume and biomass estimates for *n* new trees is available, the regional models can be calibrated using a correction factor based on an ordinary least squares regression through origin. The technique is described in Draper and Smith (1998, p. 225) and has been implemented by Temesgen

et al. (2008) to calibrate nonlinear height-diameter models and by Garber et al. (2009) to estimate stand volume. With *n* sample trees from a new stand, the OLS correction factor through origin can be calculated as:

$$k = \frac{\sum_{i=1}^{n} \hat{X}_i X_i}{\sum_{i=1}^{n} \hat{X}_i^2} \quad (17)$$

Where, X_i and $\hat{X}_i$ are observed and regional predicted values of volume or component biomasses for the i^{th} tree respectively. Then the corrected (calibrated) value of volume or component biomass for new tree is be calculated as $k \times \hat{X}_i$. The OLS correction factor can be obtained based on the measurement of a single tree. In that case, the correction factor is just a ratio of measured and regional predicted volume or biomass of that tree. This correction factor was calculated with all trees; leave one (tree) out method, randomly selected *m* trees, with dominant tree, and a tree with maximum DBH.

4.2.2.4. OLS-WI Method

Let X be the measured volume or biomass and Y be the estimate of X obtained from a regional model, then the relation between X and Y is given by a simple linear regression model of the following form:

$$y = \alpha + \beta x + \varepsilon \quad (18)$$

Where, α and β are the parameters of linear regression and ε is the random error and for i observations, the model in (18) can be written as:

$$y_i = \alpha + \beta x_i + \varepsilon_i \quad (19)$$

Using the notation by Krutchkoff (1967), the least squares estimators b and a of β and α with $u_i = x_i - \bar{x}$, and $v_i = y_i - \bar{y}$ respectively are:

$$b = \frac{\sum_{i=1}^{N} u_i v_i}{\sum_{i=1}^{N} u_i^2} \quad (20)$$

and $$a = \bar{y} - b\bar{x} \quad (21)$$

Then the calibration equation is $x = \frac{y-a}{b}$ (Krutchkoff 1967; Kutner et al. 2004, p. 168)

Thus, if now one obtains Y as the regional predicted value of volume or biomass, the unknown value (X) of local volume or biomass estimate is obtained as:

$$\hat{X} = \frac{Y-a}{b} \quad (22)$$

4.2.2.5. Inverse Approach

The inverse approach of model calibration has been described by (Krutchkoff 1967). The process is as follows. With, $\gamma = -\frac{\alpha}{\beta}$, $\delta = \frac{1}{\beta}$, and $\varepsilon' = -\frac{\varepsilon}{\beta}$, the model in (18) can be rewritten as:

$$x = \gamma + \delta y + \varepsilon' \quad (23)$$

And for i observations, the model can be written as

$$x_i = \gamma + \delta y_i + \varepsilon_i' \quad (24)$$

Once again, using the notation by Krutchkoff (1967), the least squares estimators d and c of δ and γ respectively are:

$$d = \frac{\sum_{i=1}^{N} u_i v_i}{\sum_{i=1}^{N} v_i^2} \quad (25)$$

and $$c = \bar{x} - d\bar{y} \quad (26)$$

The calibration equation then is

$$x = c + dy \quad (27)$$

Thus, if now one obtains Y as the regional predicted value of volume or biomass, the unknown value of local volume or biomass estimate can be obtained as (Krutchkoff 1967):

$$\hat{X} = c + dY \quad (28)$$

The results obtained from these three methods were calculated using all data and validated with the leave one (tree) out method. Performance of all the methods was evaluated based on the bias, bias percent, RMSE, and RMSE percent (collectively called "evaluation statistics") produced by each method. In statistics, bias is defined as the difference between the true value of an unknown parameter and the expected value of its estimator. Bias in this study is defined as the mean difference between the measured/observed and model predicted value of the variable of interest.

4.3. Results and Discussion

The volume estimates obtained from FIA-PNW and refitted local equations are similar for Douglas-fir trees that are smaller than approximately 80 cm DBH. For the trees larger than 80 cm DBH, the FIA-PNW equation slightly underestimated the cubic volume (Figure 4.1). On the other hand, for lodgepole pine, the FIA-PNW equation consistently over-estimated the CVTS even for the trees that are 25 cm DBH (Figure 4.1).

The summary of the evaluation statistics in inside bark diameter prediction obtained from the taper equation fit with different fitting approaches are given in Table 4.3. The mixed effects model fitted with second order autoregressive correlation structures, the model with smallest BIC compared to all models was considered the final taper model. Therefore, volume calculation is also based on the diameter inside bark predicted from this model. This fitting technique was also found useful by Rojo et al. (2005) in comparing 31 taper functions for maritime pine (*Pinus pinaster Ait*.) in Galicia, Northwestern Spain. The predicted relative diameter (predicted DIB/DBH) is plotted against the relative height (height / total height) in Figure 4.2.

The locally refitted volume equations provided better, in terms of the evaluation statistics used, estimates of inside-bark CVTS for both Douglas fir and lodgepole pine (Table 4.4) compared to the

regional volume equation and taper based volume estimates. The taper based volume estimates might have been affected by the fact that inside bark diameter itself was first predicted from the outside bark diameter.

The FIA-PNW bole biomass estimates are product of the inside-bark CVTS and the compiled sets of species specific wood density values. Thus the curves of inside-bark bole biomass estimates against DBH are similar to inside bark CVTS curves but scaled by the wood density i.e. they have same shape but different scale (Figure 4.3).

The ECDFs of component biomasses obtained from regional models and measured component biomasses differed for both species and the differences are more pronounced in components other than bole (Figures 4.4 and 4.5). This implies that the probability of observing a tree with certain amount of component biomass based on FIA-PNW equations is different than the probability of what actually was observed in this study.

If one rejects a joint null hypothesis H0: (a, b) = (1, 1) for the model in equation (17), this implies that the regional model needs to be calibrated. In case of Douglas-fir, a Wald test of this null hypothesis was highly significant for bark and foliage biomass, marginally significant for branch biomass, and statistically non-significant for CVTS and bole biomass while for the lodgepole pine, same test was statistically significant for CVTS, bole biomass, and bark biomass and non-significant for foliage and branch biomass (Table 4.5). Note, however, that none of the coefficients were exactly 1 (Table 4.5).

The evaluation statistics in volume estimates obtained from regional and locally refitted volume and component biomass equations are given in Table 4.6. The FIA-PNW regional Douglas-fir bark biomass equation does a very poor job with over-estimating it by 221.52 percent (Table 4.6). This difference could possibly be attributed to the fact that the FIA-PNW bark biomass equation for Douglas-

fir was based on the data with maximum tree DBH of 66 cm. Hence, a careful consideration might be necessary in applying the regional models to estimate bark biomass in Douglas-fir.

4.3.1. OLS-RTO Method

The evaluation statistics obtained by using a correction factor based on OLS-RTO method in calibrating regional volume and component biomass equations using all data are given in Table 4.7 and similar statistics obtained from leave one (tree) out cross validation method are given in Table 4.8.

The values of OLS-RTO correction factor ranged from .02770 (for Douglas-fir bark biomass) to 1.7849 (for lodgepole pine branch biomass). When all available data is used, the OLS-RTO correction reduced the RMSE by 2.4, 3.3, 335.3, 29.1 and 8.1 percent for Douglas-fir CVTS, bole, bark, foliage, and branch biomass and 5.0, 5.2, 16.2, 10.8, and 16.3 percent for lodgepole pine CVTS, bole, bark, foliage, and branch biomass respectively. While the results from the leave one out cross validation showed that, this correction factor increased RMSE by 1.3 percent for Douglas-fir CVTS and reduced RMSE in bole, bark, foliage, and branch biomass by 4.4, 325.4, 24.1, and 2.5 percent. For the lodgepole pine, the RMSE was reduced by 3.3, 2.1, 7.4, 6.2, and 10.1 percent for CVTS, bole, bark, foliage, and branch biomass respectively. The performance of adjustment based on OLS-RTO correction factor calculated using randomly selected m trees in a Douglas-fir and lodgepole pine stand is shown in Figures 4.6 and 4.7 respectively.

The RMSE decreases significantly by using two trees compared to just one and the rate of decrease in RMSE slows down after five trees for both species. It should be noted that, measuring biomass in a large number of trees to calibrate the regional models is impractical (Temesgen et al. 2008) because one would rather fit a local model than to calibrate regional model with the available data set. Therefore, we also looked at the performance of this adjustment based on dominant tree and the tree with maximum DBH (Tables 4.9 and 4.10). For our lodgepole pine tree, however, the tree with highest

DBH was the same as dominant tree while using the second largest tree in terms of DBH, the RMSE for branch biomass was more than 100 percent. Therefore the results based on second largest trees are not reported.

The performance of the adjusted volume equation did not improve when only one (dominant tree or a tree with maximum DBH) is used to calculate the correction factor. With the correction factor based on dominant tree, the RMSE for Douglas-fir bole, bark, and branch biomass was reduced by 5.8, 288.6, and 17.8 percent but for branch biomass it was increased by 1.4 percent. However, for the lodgepole pine, the correction factor reduced RMSE in branch biomass estimate by 8.4 percent but it did not improve the estimates of any other components. Using the tree with maximum DBH for Douglas-fir further decreased RMSE in CVTS and bark biomass (3.4 and 39.6 percent respectively) but produced slightly higher RMSE in other components (0.6, 3.3, and 4.7 percent for bole, foliage, and branch biomass respectively). The reason behind this could be that there is more volume and bark biomass in the lower sections of the trees and is more affected by the diameter of the tree rather than the height while on the other hand biomass in branch and foliage is more dependent on the height of the tree.

4.3.2. OLS-WI Method

The adjustment based on a correction factor based on OLS with intercept provided unbiased estimates of CVTS and component biomass when all the data was used in calibrating the regional equation. The absolute bias obtained from the leave one out cross validation was less than 2 percent for CVTS and all component biomass. The RMSE and RMSE percent obtained with this adjustment based on all data are shown in Table 4.11. The results from leave one out cross validation is presented in Table 4.12.

This adjustment, using all data, reduced RMSE in all Douglas-fir CVTS and component biomass estimation and lodgepole pine CVTS and bole biomass estimates but it did not improve other

component biomass equations for lodgepole pine. The leave one out cross validation results were similar for trend in RMSEs and showed small negative bias.

4.3.3. Inverse Approach

The inverse approach of calibration of regional equations provided unbiased estimates of CVTS and component biomass when all the data was used in calibrating the regional equation. The absolute bias obtained from the leave one out cross validation was less than 2 percent for CVTS and all component biomass from this approach as well (same in OLS-WI). The RMSE and RMSE percent obtained with this adjustment based on all data are shown in Table 4.13. The results from leave one out cross validation is presented in Table 4.14.

It is expected that the calibrated models perform better than the unadjusted regional models. However, from practical perspective, it is desired that the calibrated models perform as well as the locally fitted models. Figures 4.8 and 4.9 compare the performance of calibrated models against the locally refitted models. The OLS-RTO adjustment performed as well as the locally refitted models for CVTS and component biomass models except for Douglas-fir foliage biomass when all the data set is used to calculate the correction factor. The results from leave one out cross validations for this adjustment method were intermediate between the FIA-PNW estimates and the local predictions. The results from using dominant tree or tree with maximum DBH were intermediate between the FIA-PNW estimates and the local models.

Even though having an intercept in the calibration provides more flexibility in adjustment, the OLS-WI adjustments, when all data is used, performed as well as local models for bole and foliage biomass in Douglas-fir and CVTS and bole biomass in lodgepole pine but its performance was intermediate between FIA-PNW estimates and local model for other variables. The RMSE percent from leave one out validation was higher compared to the local models for CVTS and all component biomass

equations for both species. The inverse approach with all data performed as well as the local models except for Douglas-fir CVTS and bark biomass. It produced higher RMSE compared to the local models based on the leave-one out cross validation results.

When all data is used to calibrate the regional model, the OLS-RTO method produced smallest RMSE for Douglas-fir CVTS and bark biomass, the OLS-WI method produced smallest RMSE for foliage biomass and the inverse approach produced smallest RMSEs for bole and branch biomass. For lodgepole pine, the inverse approach produced smallest RMSEs for CVTS and all component biomass. However, from the leave one out validation, the RMSEs from OLS-RTO and inverse approach were very similar for both species CVTS and component biomass.

4.4. Summary and Conclusion

We evaluated the performance of regional volume and component biomass equations in terms of bias and RMSE using the data from destructively sampled Douglas-fir and lodgepole pine trees. A taper equation adopted from Kozak 2004 was fitted and its performance in estimating inside bark CVTS was also assessed. Three different adjustment methods based on ordinary least squares regression were applied to the regional volume and component biomass equations. The regional volume equations performed fairly well and produced similar results as the locally fitted volume equations of the same form. The taper based volume estimates for Douglas fir were comparable (within 3% RMSE) to the volume estimates obtained from regional equations but for lodgepole pine taper based volume estimates were highly biased and produced RMSE that that was three times higher than the RMSE from the regional equation. Part of the error might have come from the fact that the inside bark diameters itself were first predicted using a model.

The regional component biomass estimates were highly biased producing up to 360 percent higher estimates for bark biomass in Douglas fir. We would like to make cautionary note on bark

biomass estimates because our calculation of "observed" bark biomass is based on a 3 – 5 cm bark sample obtained from the disks taken at every 5.18 m intervals. The density of these bark samples is based on their volume calculated as rectangles. This could have been erroneous assumption since the bark sample lacks regular shape. To check this, we calculated total bark volume by subtracting inside bark CVTS from outside bark CVTS and used the compiled sets of density factors (same as the ones used by the FIA-PNW) for Douglas-fir and lodgepole pine. The total bark biomass estimates obtained from this approach was similar to what we had obtained before. Better method to estimate bark volume would be to use a water displacement method but that is both time consuming and expensive compared to the method we used.

All the adjustment methods used in this study were able to improve the performance of regional equations. Based on the leave one out cross validation, the RMSEs in CVTS and component biomass estimation was similar for the adjustments from a correction factor based on OLS-RTO and an inverse approach. When all data is used to calibrate the regional model, the OLS-RTO method produced smaller RMSE for Douglas-fir CVTS and bark biomass, the OLS-WI method produced smaller RMSE for foliage biomass and the inverse approach produced smaller RMSEs for bole and branch biomass. For lodgepole pine, the inverse approach produced smaller RMSEs for CVTS and all component biomass.

The OLS-RTO correction factor can be calculated using only one tree with desired characteristics such as a dominant tree in the stand or the tree with maximum DBH and therefore might be considered better than the OLS-WI and inverse approach. Choice between the dominant tree and a tree with maximum DBH should be governed by which component model is desired to calibrate. We found it useful to use the tree with maximum DBH to calibrate regional CVTS and bark biomass equations and the dominant tree to calibrate bole, foliage, and branch biomass. Selecting trees of such characteristics to calibrate the regional models was useful than randomly selecting one tree. The decreasing trend in RMSE by using randomly selected m trees slowed down significantly after five trees.

4.5. References

Bluhm, A. A., S.M. Garber, and D.E. Hibbs. 2007. Taper equation and volume tables for plantation-grown red alder. *Gen. Tech. Rep. PNW-GTR-735*. Portland, OR: U.S. Department of Agriculture, Forest Service, p. 74.

Brackett, M. 1973. Notes on TARIF tree-volume computation. Resources Management Report No. 24. Washington Department of Natural Resources, Olympia, WA, p 26.

de-Miguel, S., L. Mehtätalo, and A. Durkaya. 2014. Developing generalized, calibratable, mixed-effects meta-models for large-scale biomass prediction. *Can. J. For. Res.* 44(6): 648-656.

Draper, N.R., and H. Smith. 1998. Applied regression analysis. 3rd ed. John Wiley & Sons, New York.

Eisenhauer, J. G. 2003. Regression through the origin. *Teaching Statistics*. 25(3): 76-80.

Garber, S.M., H. Temesgen, V.J. Monleon, and D.W. Hann. 2009. Effects of height imputation strategies on stand volume estimation. *Can. J. For. Res.* 39(3): 681-690.

Hann, D.W. 2006. ORGANON user's manual eidtion 8.2 [computer-manual]. Oregon State University, Department of Forest Resources, Corvallis, Ore., USA.

Hansen, M. 2002. Volume and biomass estimation in FIA: national consistency vs. regional accuracy. P. 109 –120 in Proc. of the Third annual forest inventory and analysis symposium, McRoberts, R.E., G.A. Reams, P.C. Van Deusen, and J.W. Moser (eds.). *US For. Serv. Gen. Tech. Rep. NC-230*. p 216.

Hanus, M.L., D.W. Hann, and D.D. Marshall. 1999. Predicting height for undamaged and damaged trees in southwest Oregon. Oregon State Univ. For. Res. Lab. Res. Cont. 27.

Houghton, R.A. 2005. Aboveground forest biomass and the global carbon balance. *Glob. Change Biol.* 11(6): 945-958.

Ketterings, Q.M., R. Coe, M. van Noordwijk, Y. Ambagau, and C.A. Palm. 2001. Reducing uncertainty in the use of allometric biomass equations for predicting above-ground tree biomass in mixed secondary forests. *For. Ecol. Manage.* 146(1): 199-209.

Kozak, A. 2004. My last words on taper equations. *For. Chron.* 80(4): 507-514.

Krutchkoff, R. G. 1967. Classical and inverse regression methods of calibration. *Technometrics*. 9(3): 425-439.

Kutner, M.H, C. J. Nachtsheim, and J. Neter. 2004. Applied linear regression models. 4th ed. McGraw-Hill Irwin, Boston.

Li, R., A. Weiskittel, A. R. Dick, J. A. Kershaw Jr., and R. S. Seymour. 2012. Regional stem taper equations for eleven conifer species in the Acadian Region of North America: development and assessment. *North. J. Appl. For.* 29(1): 5-14.

Melson, S.L., M.E. Harmon, J.S. Fried, and J.B. Domingo. 2011. Estimates of live-tree carbon stores in the Pacific Northwest are sensitive to model selection. *Carbon Balance Manag.* 6(2): 1-16.

Omule, S.A.Y., V.E. Fletcherand, and K.R. Polsson. 1987. Total and merchantable volume equations for small coastal Douglas-fir. Economic & Regional Develop. Agree., B.C. Min. For. and Lands, Res. Br., FRDA Rep. 010, Victoria, B.C.

Pflugmacher, D., W. Cohen, R. Kennedy, and M. Lefsky. 2008. Regional Applicability of Forest Height and Aboveground Biomass Models for the Geoscience Laser Altimeter System. *For. Sci.* 54(6): 647-657.

Räty, M., and A. Kangas. 2008. Localizing general models with classification and regression trees. *Scand. J. Forest Res.* 23(5): 419-430.

Rojo, A., X. Perales, F. Sanchez-Rodriguez, J.G. Alvarez-Gonzalez, and K. von Gadow. 2005. Stem taper functions for maritime pine (*Pinus pinaster* Ait.) in Galicia (Northwestern Spain). *Eur. J. For. Res.* 124(3): 177-186.

SAS Institute Inc. 2013. Base SAS® 9.4 Procedures Guide: Statistical Procedures, Second Edition. Cary, NC: SAS Institute Inc.

Standish, J.T., G.H. Manning, and J.P. Demaerschalk. 1985. Development of biomass equations for British Columbia tree species. Can. For. Serv., Info. Rep. BC-X-264, Pacific Forest Resource Center, Victoria, BC, Canada. p 47.

Taylor, D.C.A., V. Pawar, D.T. Kruzikas, K.E. Gilmore, M. Sanon, and M.C. Weinstein. 2012. Incorporating calibrated model parameters into sensitivity analyses deterministic and probabilistic approaches. *Pharmacoeconomics.* 30 (2): 119-126.

Temesgen, H., V.J. Monleon, and D.W. Hann. 2008. Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. *Can. J. For. Res.* 38: 553-565.

Zeng, W.S., H.R. Zhang, and S.Z. Tang. 2011. Using the dummy variable model approach to construct compatible single-tree biomass equations at different scales - a case study for Masson pine (Pinus massoniana) in southern China. *Can. J. For. Res.* 41(7): 1547-1554.

Zianis, D., P. Muukkonen , R. Mäkipää, and M. Mencuccini. 2005. Biomass and stem volume equations for tree species in Europe. *Silva Fennica Monographs.* 4. p 63.

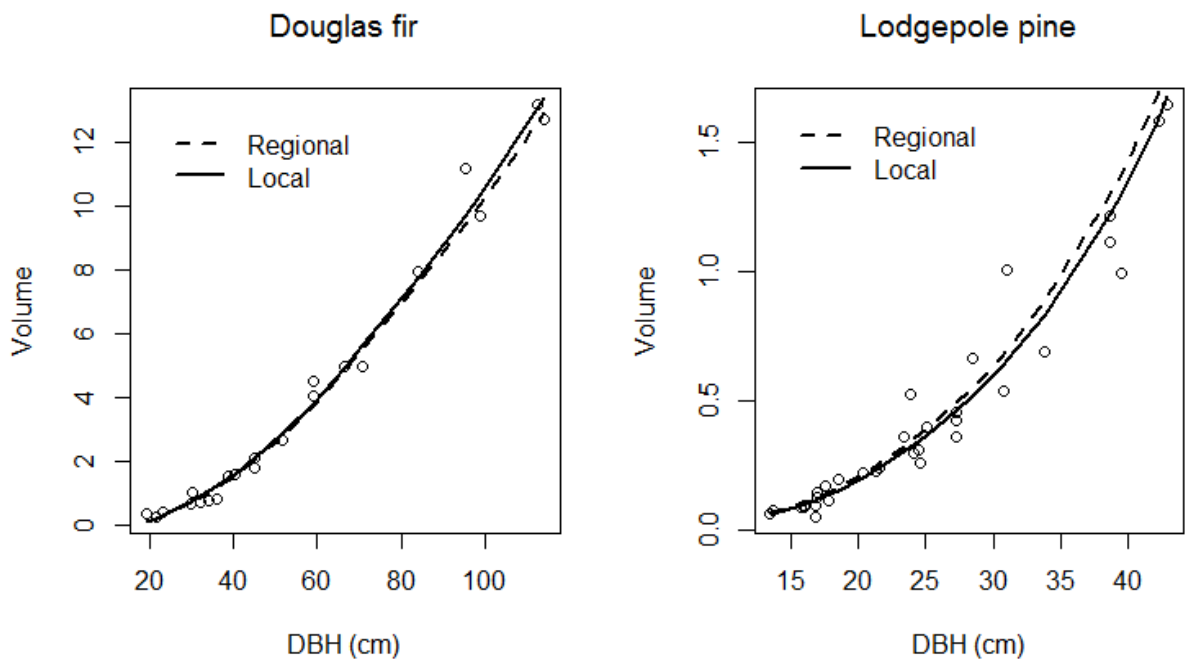

Figure 4.1. Graphical comparison of CVTS estimates obtained by using regional model coefficients and locally refitted volume equations of the same form.

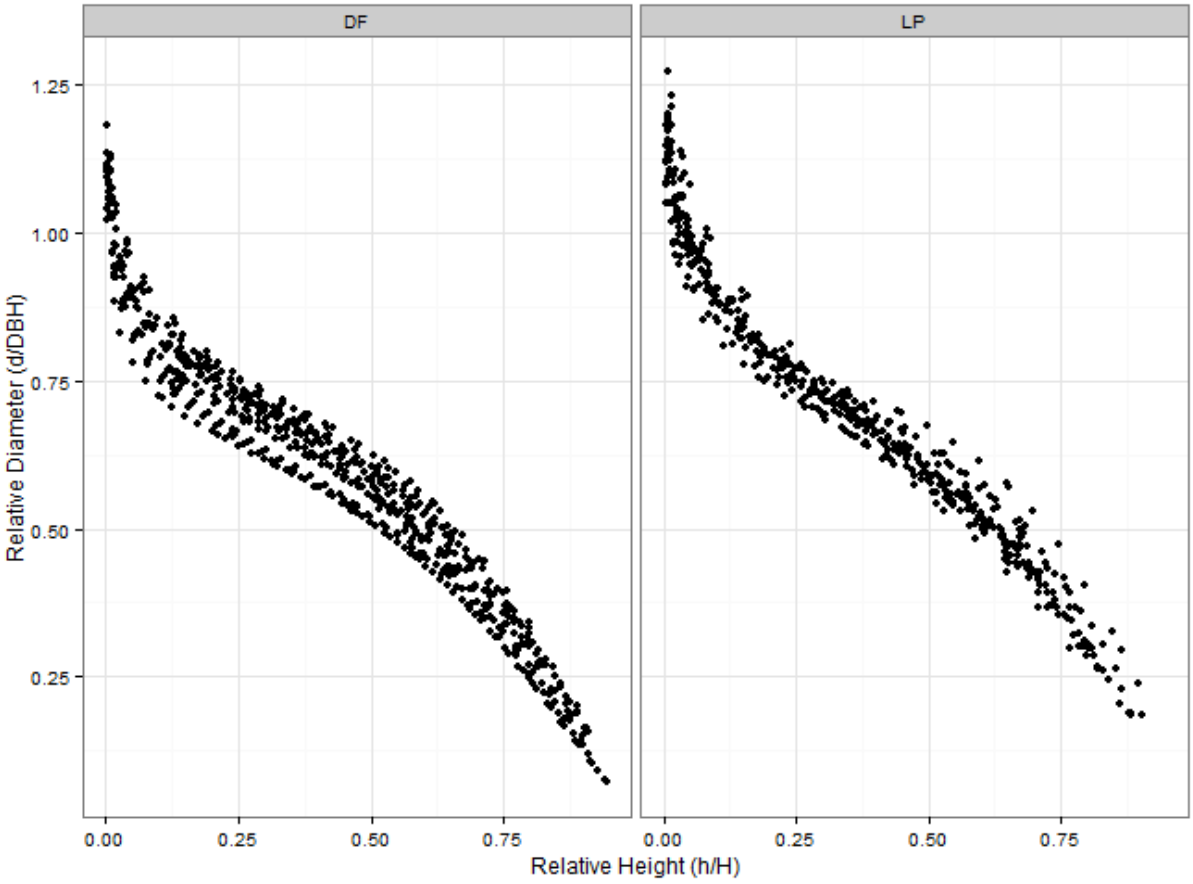

Figure 4.2. Predicted profiles for Douglas-fir (DF) and lodgepole pine (LP) plotted against relative height (height/total height).

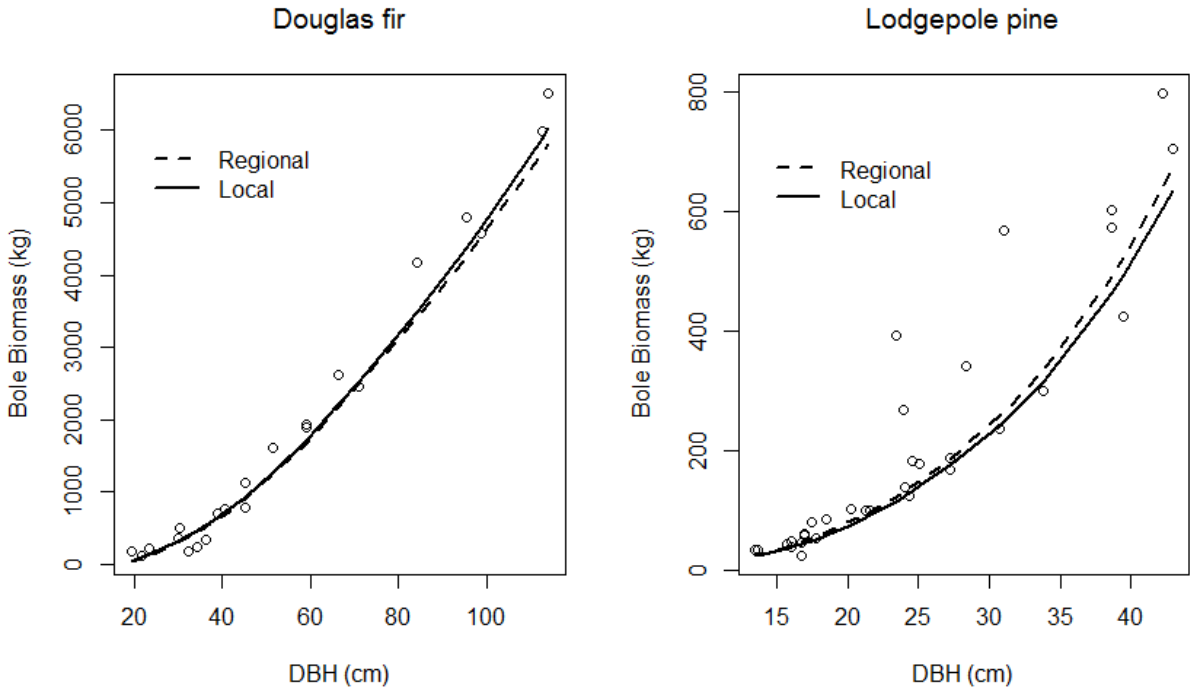

Figure 4.3. Graphical comparison of inside-bark bole biomass estimates obtained from regional model coefficients and locally refitted model of the same form.

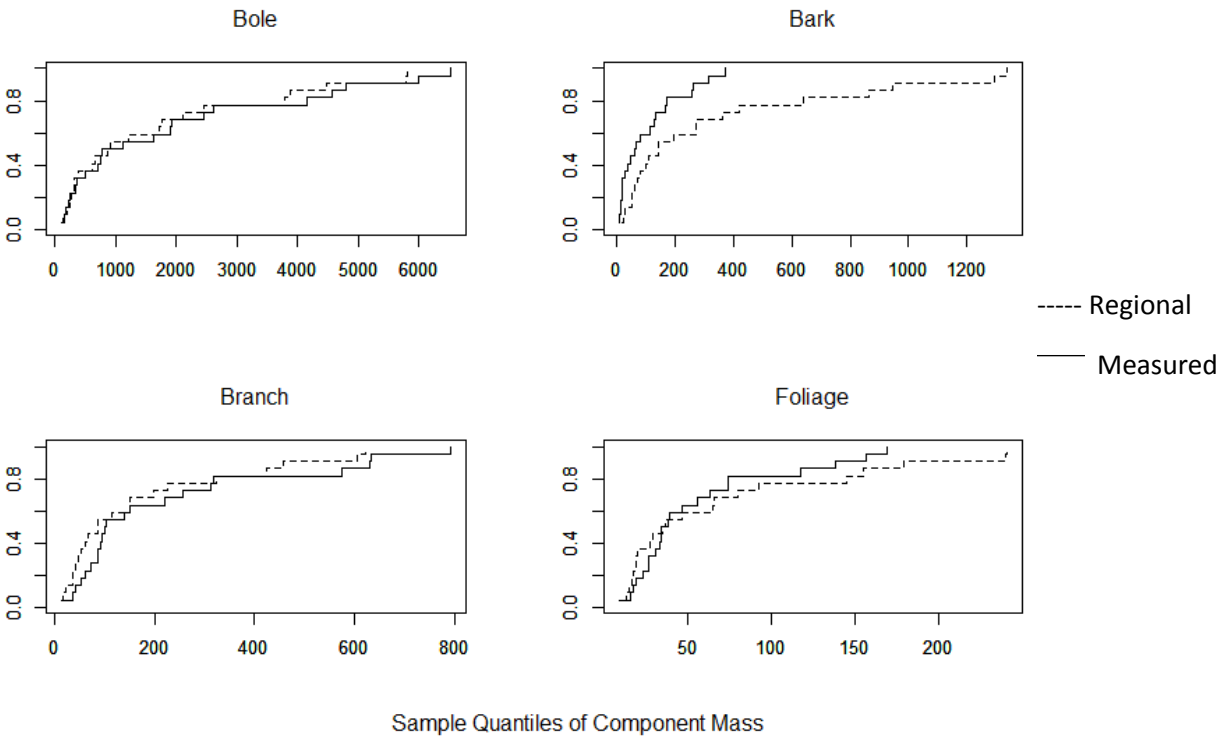

Figure 4.4. Comparison of ECDFs of Douglas-fir component biomass estimates obtained from regional equations (dotted line) and measured (solid line) values.

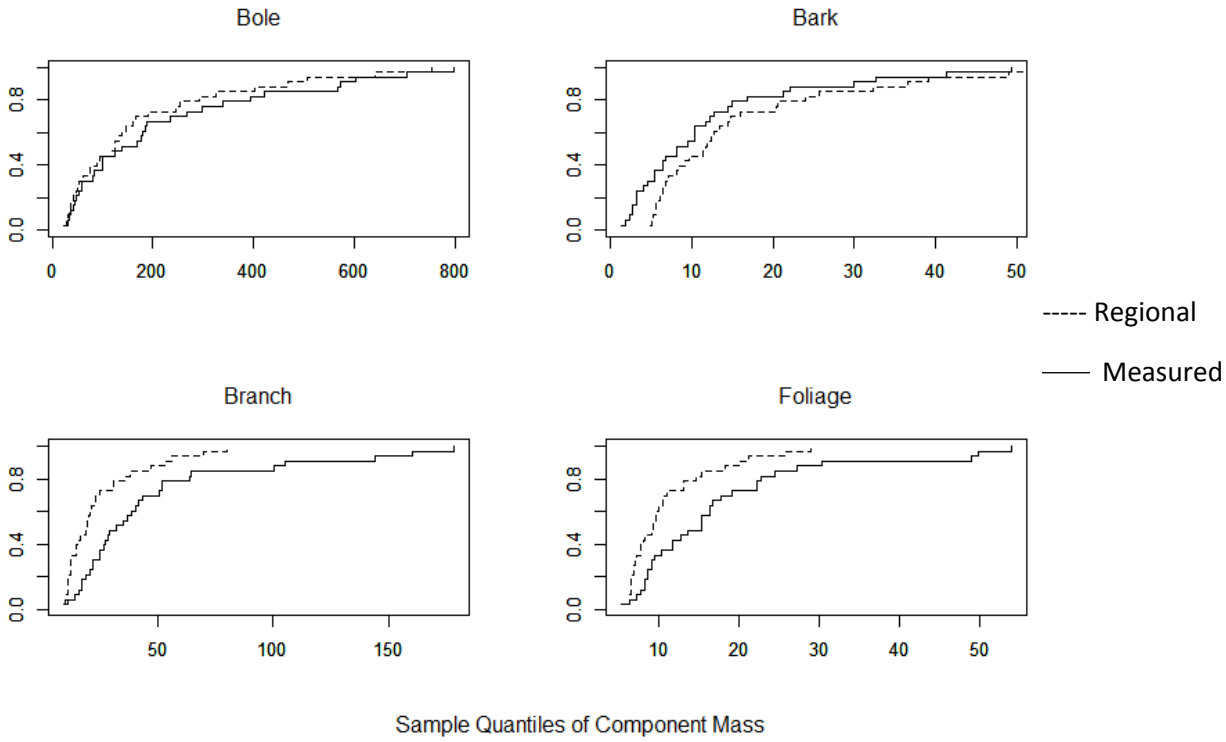

Figure 4.5. Comparison of ECDFs of lodgepole pine component biomass estimates obtained from regional equations (dotted line) and measured (solid line) values.

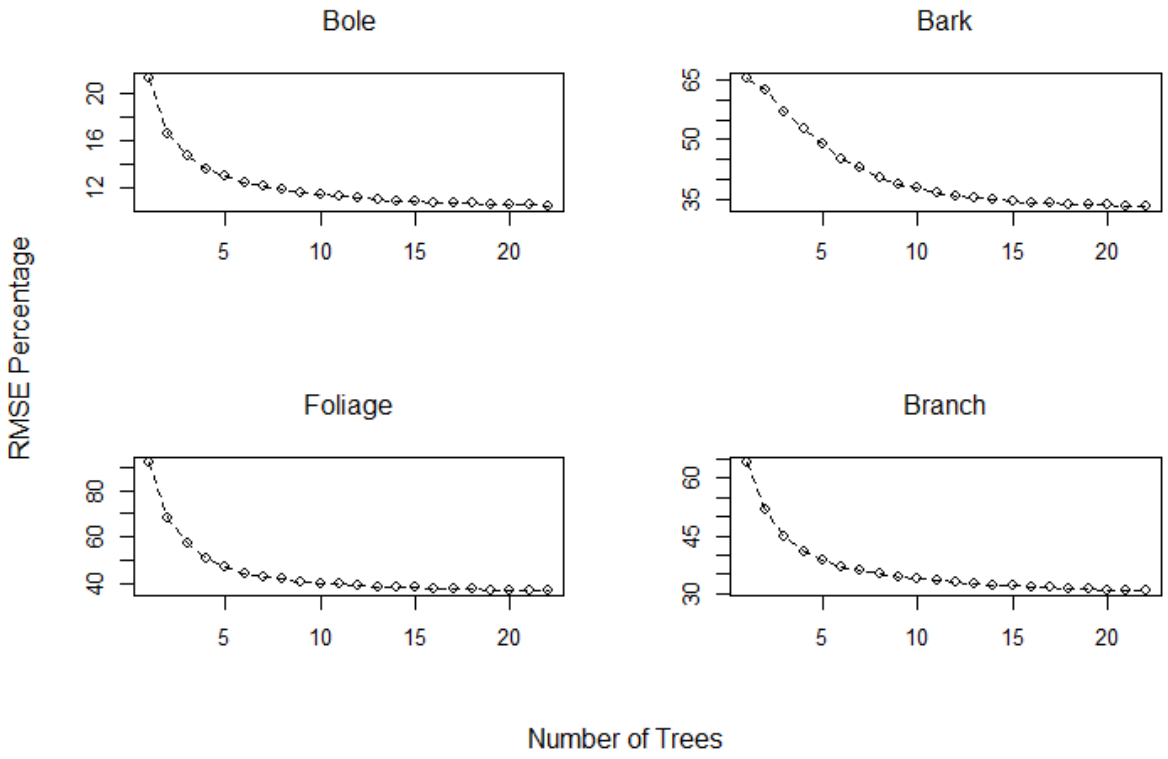

Figure 4.6. Trend in RMSE percent in estimating component biomass from regional models adjusted with OLS-RTO method calculated from randomly selected m Douglas-fir trees. The RMSE percent from unadjusted regional equations were 16.3, 361.6, 66.2, and 38.7 percent for bole, bark, foliage, and branch biomass respectively.

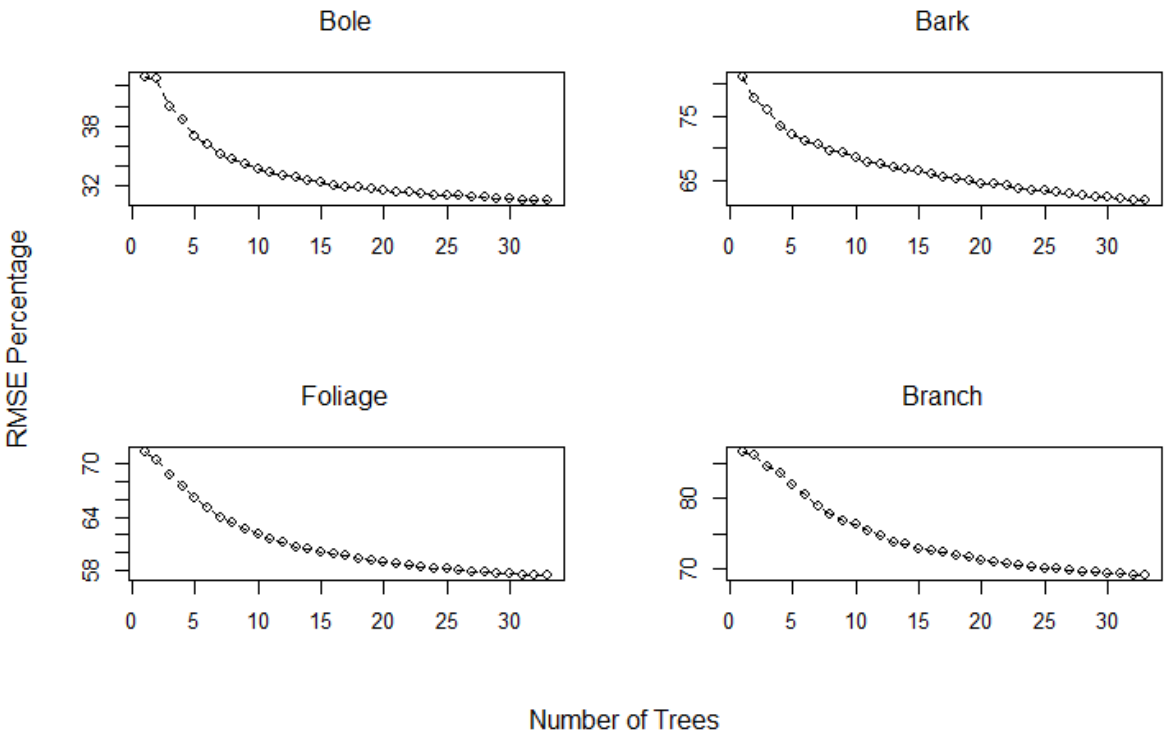

Figure 4.7. Trend in RMSE percent in estimating component biomass from regional models adjusted with OLS-RTO method calculated from randomly selected m lodgepole pine trees. The RMSE percent from unadjusted regional equations were 35.8, 78.1, 68.1, and 85.4 percent for bole, bark, foliage, and branch biomass respectively.

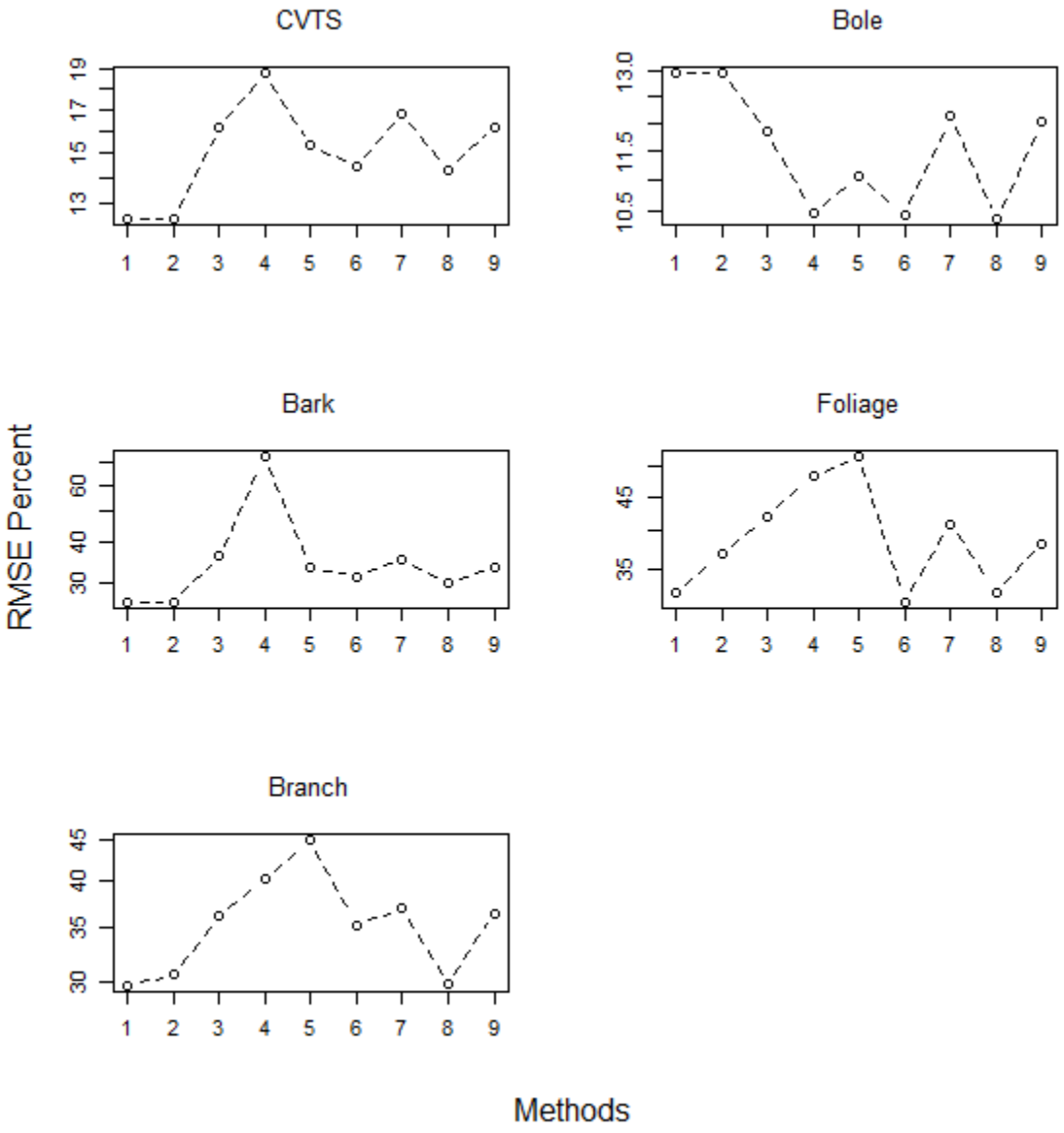

Figure 4.8. Comparison of RMSE percentage, for Douglas-fir, obtained from different methods: (1) Local model, (2) OLS-RTO based on all data, (3) OLS-RTO (leave one out), (4) OLS-RTO based on dominant tree, (5) OLS-RTO based on tree with maximum DBH, (6) OLS-WI based on all data, (7) OLS-WI (leave one out), (8) Inverse approach based on all data, and (9) Inverse approach (leave one out).

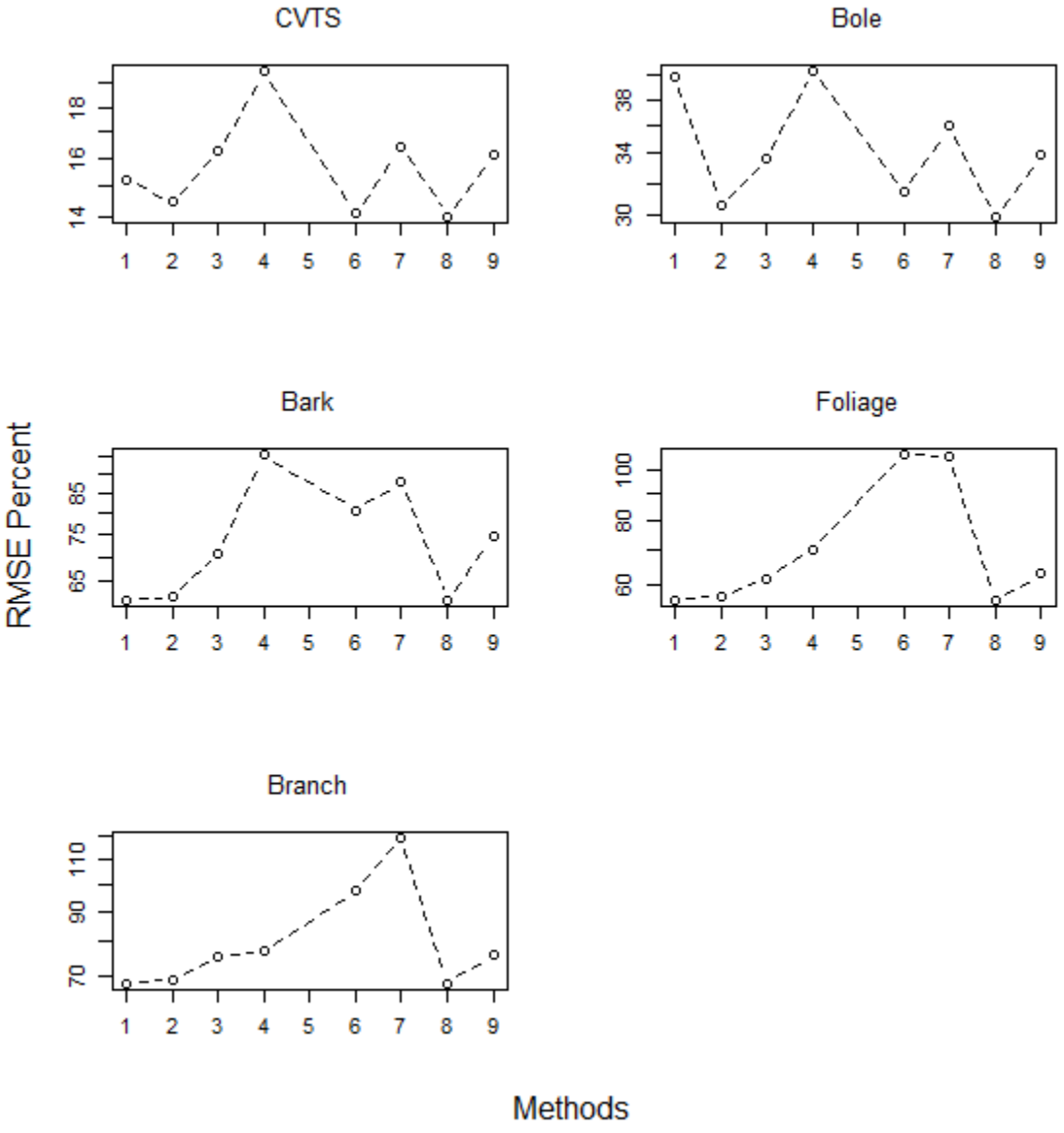

Figure 4.9. Comparison of RMSE percentage, for lodgepole pine, obtained from different methods: (1) Local model, (2) OLS-RTO based on all data, (3) OLS-RTO (leave one out), (4) OLS-RTO based on dominant tree, (5) OLS-RTO based on tree with maximum DBH, (6) OLS-WI based on all data, (7) OLS-WI (leave one out), (8) Inverse approach based on all data, and (9) Inverse approach (leave one out).

Table 4.1. Volume equations used by FIA-PNW to estimate cubic volume including top and stump (CVTS).

Species	Volume Equation	
Douglas-fir	$CVTSL = \begin{matrix} -3.21809 + 0.04948 \times \log(HT) \times \log(DBH) - 0.15664 \times (\log(DBH))^2 \\ +2.02132 \times \log(DBH) + 1.63408 \times \log(HT) - 0.1685 \times (\log(HT))\text{^}2 \end{matrix}$ $CVTS = 10^{CVTSL}$	(4)
Lodgepole Pine	$CVTSL = -2.615591 + 1.847504 \times \log(DBH) + 1.085772 \times \log(HT)$ $CVTS = 10^{CVTSL}$	(5)

(Base of the logarithm is 10, HT = total tree height, and DBH = diameter at breast height)

Table 4.2. Biomass equations used by the FIA-PNW to estimate component and total aboveground biomass for Douglas-fir and lodgepole pine trees.

Species	Component	Biomass Equation	
Douglas Fir	Bole	$CVTS \times WD$	(6)
	Foliage	$10.3 + 3.9 \times DBH^2 \times HT$	(7)
	Bark	$3.1 + 15.6 \times DBH^2 \times HT$	(8)
	Branch	$12.6 + 23.5 \times DBH^2 \times HT$	(9)
Lodgepole Pine	Bole	$CVTS \times WD$	
	Foliage	$5.5 + 4.0 \times DBH^2 \times HT$	(10)
	Bark	$3.2 + 9.1 \times DBH^2 \times HT$	(11)
	Branch	$7.8 + 12.3 \times DBH^2 \times HT$	(12)

(WD is wood density which is 28.7 and 23.71 pounds per cubic foot for Douglas-fir and lodgepole pine respectively)

Table 4.3. Evaluation statistics of inside bark diameter predictions for Douglas-fir and lodgepole pine trees using different fitting techniques: (I) nonlinear fixed effects model, (II) nonlinear mixed effects model, (III) nonlinear mixed effects model with first order continuous autoregressive correlation, and (IV) nonlinear mixed effects model with second order continuous autoregressive correlation structure.

Species	Method	Bias (cm)	Bias Percent	RMSE (cm)	RMSE Percent
Douglas-fir	I	-0.01	-0.03	3.10	8.31
	II	0.01	0.04	2.57	6.90
	III	0.06	0.15	2.61	6.99
	IV	0.08	0.20	2.65	7.10
Lodgepole pine	I	-0.01	-0.03	1.79	9.40
	II	0.02	0.11	1.56	8.20
	III	0.05	0.25	1.59	8.38
	IV	0.10	0.51	1.65	8.69

Table 4.4. Evaluation statistics produced by different methods in inside bark CVTS predictions for Douglas-fir and lodgepole pine trees.

Species	Method	Bias (m^3)	Bias Percent	RMSE (m^3)	RMSE Percent
Douglas fir	Local	0.01	0.30	0.50	12.48
	PNW	0.13	3.22	0.59	14.83
	Taper	-0.40	-10.14	0.70	17.53
Lodgepole pine	Local	0.00	-0.34	0.07	15.24
	PNW	-0.03	-6.66	0.09	19.52
	Taper	-0.19	-41.39	0.27	60.02

Table 4.5. The parameter estimates and associated chi-squared p-values for testing a joint null hypothesis H0: (a, b) = (1, 1) for the nonlinear model ($y = a * y_R^b$) describing measured volume and component biomass as function of FIA-PNW estimates of the same quantity are (1, 1).

Variable	Parameter	Douglas-fir		Lodgepole pine	
		Estimate	Pr > ChiSq	Estimate	Pr > ChiSq
Volume	a	1.1976	0.7493	0.9241	<.0001
	b	0.9735		0.9274	
Bole	a	1.5405	0.4292	2.4334	0.0001
	b	0.9627		0.8783	
Bark	a	1.8849	<.0001	1.2527	0.0003
	b	0.7494		0.8403	
Foliage	a	3.0913	<.0001	2.9366	0.4139
	b	0.7462		0.7598	
Branch	a	2.3883	0.0647	4.1689	0.3482
	b	0.9007		0.7749	

Table 4.6. Evaluation statistics in volume and component biomass estimation obtained from regional and locally refitted volume and component biomass equations. Units for bias and RMSE are cubic meter for CVTS and kg for component biomasses, respectively.

Species	Variable	Local				Regional			
		Bias	Bias Percent	RMSE	RMSE Percent	Bias	Bias Percent	RMSE	RMSE Percent
DF	CVTS	0.01	0.30	0.50	12.48	0.13	3.22	0.59	14.83
	Bole	130.49	6.81	248.06	12.94	182.83	9.54	311.68	16.26
	Bark	-2.49	-2.36	27.82	26.32	-234.16	-221.52	382.25	361.60
	Foliage	0.01	0.02	18.41	32.38	-14.89	-26.18	37.66	66.22
	Branch	0.94	0.42	66.11	29.82	47.05	21.23	85.86	38.73
LP	CVTS	0.00	-0.34	0.07	15.26	-0.03	-6.66	0.09	19.52
	Bole	49.58	22.52	87.58	39.78	38.83	17.64	78.71	35.75
	Bark	0.00	-0.01	7.38	61.47	-3.87	-32.21	9.39	78.14
	Foliage	0.00	0.01	9.96	56.40	6.59	37.31	12.02	68.07
	Branch	-0.01	-0.02	32.42	68.10	22.67	47.61	40.65	85.38

Table 4.7. Evaluation statistics in CVTS and component biomass estimates obtained from OLS-RTO adjustment of the regional equations using all data. Units for bias and RMSE are cubic meter for CVTS and kg for component biomasses, respectively.

Species	Variable	Bias	Bias Percent	RMSE	RMSE Percent
DF	CVTS	0.03	0.63	0.57	12.48
	Bole	17.41	0.91	200.17	12.94
	Bark	11.62	10.99	35.23	26.32
	Foliage	7.25	12.76	21.10	37.10
	Branch	11.16	5.03	68.01	30.68
LP	CVTS	0.01	2.61	0.06	14.48
	Bole	9.95	4.52	67.25	30.54
	Bark	0.56	4.68	7.44	61.91
	Foliage	0.83	4.68	10.12	57.32
	Branch	3.12	6.56	32.87	69.05

Table 4.8. Evaluation statistics in CVTS and component biomass estimates obtained from OLS-RTO adjustment of the regional equations: results from leave one out cross validation. Units for bias and RMSE are cubic meter for CVTS and kg for component biomasses, respectively.

Species	Variable	*k*	Bias	Bias Percent	RMSE	RMSE Percent
DF	CVTS	1.0268	0.02	0.52	0.64	16.11
	Bole	1.0955	15.15	0.79	227.18	11.85
	Bark	0.2770	10.78	10.20	38.30	36.24
	Foliage	0.6920	6.85	12.04	23.97	42.15
	Branch	1.2061	10.26	4.63	80.37	36.25
LP	CVTS	0.9134	0.01	2.32	0.07	16.26
	Bole	1.1598	8.88	4.03	74.05	33.63
	Bark	0.7215	0.47	3.92	8.49	70.71
	Foliage	1.5212	0.74	4.17	10.94	61.91
	Branch	1.7849	2.75	5.77	35.86	75.32

Table 4.9. Evaluation statistics obtained by using OLS-RTO method based on dominant tree in calibrating regional volume and component biomass equations. Units for bias and RMSE are cubic meter for CVTS and kg for component biomasses, respectively.

Species	Variable	Bias	Bias Percent	RMSE	RMSE Percent
DF	CVTS	0.36	9.02	0.75	18.78
	Bole	10.67	0.56	200.41	10.45
	Bark	-32.69	-30.92	77.13	72.97
	Foliage	19.83	34.87	27.55	48.44
	Branch	50.46	22.76	88.98	40.14
LP	CVTS	0.05	11.84	0.09	19.52
	Bole	50.62	22.99	88.54	40.22
	Bark	7.30	60.80	11.46	95.36
	Foliage	7.28	41.20	12.46	70.53
	Branch	16.39	34.43	36.66	77.01

Table 4.10. Evaluation statistics obtained by using OLS-RTO method based on tree with maximum DBH in calibrating regional volume and component biomass equations for Douglas-fir. Units for bias and RMSE are cubic meter for CVTS and kg for component biomasses, respectively.

Variable	Bias	Bias Percent	RMSE	RMSE Percent
CVTS	0.17	4.39	0.61	15.34
Bole	-31.67	-1.65	212.35	11.08
Bark	11.18	10.58	35.23	33.33
Foliage	21.81	38.36	29.43	51.74
Branch	60.85	27.45	99.44	44.85

Table 4.11. RMSE and RMSE percentage obtained by using a correction factor based on OLS-WI method in calibrating regional volume and component biomass equations using all data. Units for RMSE are cubic meter for CVTS and kg for component biomasses, respectively.

Variable	Douglas-fir		Lodgepole pine	
	RMSE	RMSE Percent	RMSE	RMSE Percent
CVTS	0.58	14.46	0.06	14.14
Bole	199.73	10.42	69.20	31.43
Bark	33.29	31.49	9.69	80.68
Foliage	69.38	31.30	50.77	106.64
Branch	20.11	35.36	17.19	97.30

Table 4.12. Evaluation statistics obtained by using a correction factor based on OLS-WI method in calibrating regional volume and component biomass equations from leave one out cross validation. Units for bias and RMSE are cubic meter for CVTS and kg for component biomasses, respectively.

Species	Variable	*a*	*b*	Bias	Bias Percent	RMSE	RMSE Percent
Douglas-fir	CVTS	0.0268	0.9611	-0.01	-0.16	0.67	16.78
	Bole	-12.8586	0.9113	-3.07	-0.16	232.28	12.12
	Bark	-41.4628	3.6063	-1.22	-1.15	37.46	35.44
	Foliage	-9.2455	1.4245	-0.26	-0.47	23.28	40.93
	Branch	-1.374	0.794	-0.66	-0.30	81.94	36.96
Lodgepole pine	CVTS	-0.0154	1.1008	0.00	-0.40	0.07	16.42
	Bole	1.3692	0.8173	-1.96	-0.89	79.16	35.95
	Bark	5.4361	0.8699	-0.10	-0.84	10.55	87.81
	Foliage	6.281	0.2715	-0.11	-0.63	18.76	106.20
	Branch	12.2349	0.2671	-0.65	-1.37	56.82	119.36

Table 4.13. RMSE and RMSE percentage obtained using inverse approach in calibrating regional volume and component biomass equations using all data. Units for RMSE are cubic meter for CVTS and kg for component biomasses, respectively.

Variable	Douglas-fir		Lodgepole pine	
	RMSE	RMSE Percent	RMSE	RMSE Percent
CVTS	0.57	14.32	0.06	14.01
Bole	198.71	10.37	65.76	29.87
Bark	31.77	30.06	7.38	61.47
Foliage	18.41	32.38	9.96	56.40
Branch	66.26	29.89	32.42	68.10

Table 4.14. Evaluation statistics obtained by using inverse approach to calibrate regional volume and component biomass equations, results from leave one out cross validation. Units for bias and RMSE are cubic meter for CVTS and kg for component biomasses, respectively.

Species	Variable	*c*	*d*	Bias	Bias Percent	RMSE	RMSE Percent
Douglas-fir	CVTS	0.048	1.0209	-0.01	-0.15	0.64	16.14
	Bole	33.3182	1.0863	-2.98	-0.16	230.22	12.01
	Bark	19.925	0.2525	-0.69	-0.65	35.26	33.36
	Foliage	14.6011	0.5892	-0.20	-0.35	21.80	38.33
	Branch	20.9354	1.1498	-0.83	-0.37	80.86	36.48
Lodgepole pine	CVTS	0.0231	0.8893	0.00	-0.38	0.07	16.11
	Bole	19.7394	1.1055	-1.45	-0.66	74.58	33.87
	Bark	1.3806	0.6698	-0.18	-1.51	8.97	74.69
	Foliage	3.8887	1.2446	-0.26	-1.47	11.16	63.16
	Branch	9.3362	1.5354	-0.74	-1.56	36.12	75.88

5. General Conclusion

The overall objective of this dissertation was to explore different methods for sampling and estimating aboveground tree biomass. The specific objectives were to (1) evaluate different sampling strategies to estimate crown biomass, (2) develop methods for estimating aboveground biomass and its components, and (3) calibrate regional volume and component biomass. These objectives were assessed using data collected destructively from sampled trees that belonged to five different species namely Douglas-fir, grand fir, lodgepole pine, western hemlock, and red alder. These species account for approximately 50 percent of gross live volume in the Pacific Northwest.

In Chapter 2, eleven different strategies for crown biomass sampling were evaluated based on their performance in terms of 6 different evaluation statistics. These strategies belonged to three major categories: simple random sampling, systematic sampling, and stratified sampling. We also evaluated their performance when different numbers of branches (3, 6, 9, and 12) are selected from each tree. The stratified sampling method with proportional to size produced better results when 3 or 6 branches per tree were sampled. When at least nine branches per tree are sampled, the systematic sampling with ratio estimation technique provided the smallest RMSE compared to all other methods. Under the stratified sampling strategy, selecting 4, 3, and 2 branches from lower, middle, and upper section of the tree produced approximately similar results to simple random sampling, but it further decreased RMSE when information on branch diameter is used in the design and estimation phases. Measuring more branches per tree would reduce the bias and RMSE, but our results showed sampling 9 branches per tree to be reasonably efficient and limit the amount of fieldwork in crown biomass estimation.

In chapter 3, total aboveground biomass and/or its components were estimated using (a) analytical approach with existing biomass equations, (b) simultaneous fitting of a system of biomass equations based on DBH only and DBH and other explanatory variables using a seemingly unrelated

regression (SUR), and (c) three regression methods (beta regression, Dirichlet regression, and multinomial log linear) that predicted proportion of biomass in main stem, stem bark, bare branch, and foliage components. The analytical approaches were the component ratio method used by the USDA Forest Service, Forest Inventory and Analysis (FIA), approach used by the Pacific Northwest unit of FIA, and the approach described by Jenkins et al. (2003). Biomass estimates obtained from the analytical approaches were highly biased for our dataset, especially in component biomass estimation. The need for biomass equations that include height or crown length in them was justified by the reduction in RMSE by such models in our analysis. The methods that predicted proportions were unbiased and produced RMSEs smaller than that provided by analytical methods. These methods produced smaller RMSEs for 85% of the species component combinations. The component proportions are simultaneously fitted in the Dirichlet regression thus ensuring that the predicted component proportions sum to 1. If desired, the component proportion models within Dirichlet system can have separate explanatory variables. Thus this approach of estimating component biomass might have greater value than the existing methods.

In chapter 4, methods for adjusting regional volume and component biomass equations were explored. Performance of a variable-exponent taper function in estimating upper stem diameters and inside bark cubic volume was also assessed. The regional volume equations provided more accurate volume estimates than that obtained by using the Smalian's formula and the predicted inside bark diameters from the fitted taper equation. Three methods for adjusting regional volume and component biomass equations namely (a) use of a correction factor based on OLS regression through origin (OLS-RTO method), (b) use of a correction factor based on OLS with intercept (OLS-WI method), and (c) an inverse approach were applied. These adjustment methods were able to improve the performance of regional equations in terms of root mean squared error (RMSE). The results from leave one out cross validations showed that the adjusted models produced better estimates (low RMSE) than provided by

the regional models. For some components, the error in prediction was same as obtained from the locally refitted model of the same form. Among these three adjustment methods, the OLS-RTO and inverse approach provided similar results that were superior to that of the OLS-WI method.

5.1. Future Directions

In this study we did not evaluate the performance of RBS compared to other sampling strategies. The RBS is a special case of multistage sampling. Even though, this method was originally used in estimating fruit count in individual tree, it has been used in estimation branch and leaf/foliage biomass (e.g. Valentine et al. 1984; Williams 1989; Gregoire et al. 1995; Schlecht and Affleck 2014). One advantage of the RBS is that, samples are sequentially collected along the path thus counting the total number of branches beforehand is not necessary. Schlecht and Affleck (2014) found the performance of RBS to be intermediate between the probability proportional to size sampling and the simple random sampling in estimating green crown mass in Douglas-fir and western larch. There is increasing interest in the use of this method as a nationally consistent method for crown sampling. Therefore, it is necessary to assess the performance of this method compared to the traditional simple random sampling, systematic sampling, and stratified random sampling.

With the available technology, upper stem diameters can be obtained without destructive sampling. This makes it possible to calibrate the taper equations to obtain better upper stem diameter estimates. For example, Cao (2009) found optimal gain in upper stem diameter prediction when the taper function is calibrated using diameter at midpoint between breast height and the tree tip. The density data can be obtained from small tree cores with minimal field work. Because, biomass is the function of volume and density, thus integrating these two variables could provide efficient method for estimating aboveground biomass.

The data in this study came from trees selected subjectively, thus the scope of inference could be limited. Therefore, future works should consider data collection that is safeguarded by theory of random sampling. Even though the destructive nature of biomass sampling creates logistic problem, it is essential that the biomass data be collected in the same FIA plots that are used to measure tree for volumetric information.

There have been significant researches in biomass mapping using the remotely sensed data. Combining remote sensed data with the data from ground measurement can improve the accuracy of remote sensed based biomass estimates. This approach would, additionally, provide opportunity to validate the models based on remotely sensed dataset. Furthermore, the biomass models based on remotely sensed data can be calibrated using small sample of the ground measured data.

6. Bibliography

Barney, R.J., K. VanCleve, and R. Schlentner. 1978. Biomass distribution and crown characteristics in two Alaskan Picea mariana ecosystems. *Can. J. For. Res.* 8(1): 36-41.

Beaucham, J.J., and J.S. Olson. 1973. Corrections for bias in regression estimates after logarithmic transformation. *Ecology*. 54(6): 1403-1407.

Bluhm, A. A., S.M. Garber, and D.E. Hibbs. 2007. Taper equation and volume tables for plantation-grown red alder. *Gen. Tech. Rep. PNW-GTR-735*. Portland, OR: U.S. Department of Agriculture, Forest Service, p. 74.

Boudewyn, P., X. Song, S. Magnussen, and M. Gillis. 2007. Model-based, volume-to-biomass conversion for forested and vegetated land in Canada. Canadian Forest Service, Victoria, Canada. Information Report BC-X-411.

Brackett, M. 1973. Notes on TARIF tree-volume computation. Resources Management Report No. 24. Washington Department of Natural Resources, Olympia, WA, p 26.

Brown, I.F., L.A. Martinelli, W.W. Thomas, M.Z. Moreira, C.A. Cid Ferreira, and R.A. Victoria. 1995. Uncertainty in the biomass of Amazonian forests: An example from Rondônia, Brazil. *For. Ecol. Manage.* 75(1): 175-189.

Brown, S. 1997. Estimating biomass and biomass change of tropical forests: a primer. Rome: FAO Forestry Paper 134. Food and Agriculture Org.

Brown, S. 2002. Measuring carbon in forests: current status and future challenges. *Environmental pollution*. 116(3): 363-372.

Cao, Q.V. 2009. Calibrating a segmented taper equation with two diameter measurements. *South. J. Appl. For.* 33(2): 58-61.

Catchpole, W., and C. Wheeler. 1992. Estimating plant biomass: a review of techniques. *Aust. J. Ecol.* 17(2): 121-131.

Chirici, G., N. Puletti, R. Salvati, F. Arbi, C. Zolli, and P. Corona. 2014. Is randomized branch sampling suitable to assess wood volume of temperate broadleaved old-growth forests? *For. Ecol. Manage.* 312: 225-230.

Cribari-Neto, F., and A. Zeileis. 2010. Beta Regression in R. *J. Stat. Softw.* 34(2): 1-24.

Czapowskyj, M.M., D.J. Robison, R.D. Briggs, and E.H. White. 1985. Component biomass equations for Black Spruce in Maine. *Res. Pap. NE-564.* Broomal, PA: U.S. Department of Agriculture, Forest Service, p. 7.

De Paola, M., V. Scoppa, and R. Lombardo. 2010. Can gender quotas break down negative stereotypes? Evidence from changes in electoral rules. *Journal of Public Economics*. 94(5): 344-353.

de-Miguel, S., L. Mehtätalo, and A. Durkaya. 2014. Developing generalized, calibratable, mixed-effects meta-models for large-scale biomass prediction. *Can. J. For. Res.* 44(6): 648-656.

de-Miguel, S., T. Pukkala, N. Assaf, and Z. Shater. 2014. Intra-specific difference in allometric equations for aboveground biomass of eastern Mediterranean Pinus brutia. *Ann. For. Sci.* 71(1): 101-112.

Devine, W.D., P.W. Footen, R.B. Harrison, T.A. Terry, C.A. Harrington, S.M. Holub, and P.J. Gould. 2013. Estimating tree biomass, carbon, and nitrogen in two vegetation control treatments in an 11-year-old Douglas-fir plantation on a highly productive site. *Res. Pap. PNW-RP-591*. Portland, OR: U.S. Department of Agriculture, Forest Service, p 29.

Draper, N.R., and H. Smith. 1998. Applied regression analysis. 3rd ed. John Wiley & Sons, New York.

Eisenhauer, J. G. 2003. Regression through the origin. *Teaching Statistics*. 25(3): 76-80.

Eskelson, B.N., L. Madsen, J.C. Hagar, and H. Temesgen. 2011. Estimating Riparian understory vegetation cover with Beta regression and copula models. *For. Sci.* 57(3): 212-221.

Ferrari, S., and F. Cribari-Neto. 2004. Beta regression for modelling rates and proportions. *J. Appl. Stat.* 31(7): 799-815.

Flewelling, J.W., and L.V. Pienaar. 1981. Multiplicative regression with lognormal errors. *For. Sci.* 27(2): 281-289.

Fonseca, W., F.E. Alice, and J.M. Rey-Benayas. 2012. Carbon accumulation in aboveground and belowground biomass and soil of different age native forest plantations in the humid tropical lowlands of Costa Rica. *New Forests.* 43(2): 197-211.

Garber, S.M., H. Temesgen, V.J. Monleon, and D.W. Hann. 2009. Effects of height imputation strategies on stand volume estimation. *Can. J. For. Res.* 39(3): 681-690.

Good, N.M., M. Paterson, C. Brack, and K. Mengersen. 2001. Estimating tree component biomass using variable probability sampling methods. *J. Agric. Biol. Environ. Stat.* 6(2): 258-267.

Goodman, R.C., O.L. Phillips, and T.R. Baker. 2014. The importance of crown dimensions to improve tropical tree biomass estimates. *Ecol. Appl.* 24(4): 680-698.

Gregoire, T.G., H.T. Valentine, and G.M. Furnival. 1995. Sampling methods to estimate foliage and other characteristics of individual trees. *Ecology*. 76(4): 1181-1194.

Hann, D.W. 2006. ORGANON user's manual eidtion 8.2 [computer-manual]. Oregon State University, Department of Forest Resources, Corvallis, Ore., USA.

Hansen, M. 2002. Volume and biomass estimation in FIA: national consistency vs. regional accuracy. P. 109 –120 in Proc. of the Third annual forest inventory and analysis symposium, McRoberts, R.E., G.A. Reams, P.C. Van Deusen, and J.W. Moser (eds.). *US For. Serv. Gen. Tech. Rep. NC-230*. p 216.

Hanus, M.L., D.W. Hann, and D.D. Marshall. 1999. Predicting height for undamaged and damaged trees in southwest Oregon. Oregon State Univ. For. Res. Lab. Res. Cont. 27.

Harrison, R.B., T.A. Terry, C.W. Licata, B.L. Flaming, R. Meade, I.A. Guerrini, B.D. Strahm, D. Xue, M.R. Lolley, A.R. Sidell, G.L. Wagoner, D. Briggs, and E.C. Turnblom. 2009. Biomass and stand characteristics of a highly productive mixed Douglas-fir and Western Hemlock plantation in Coastal Washington. *West. J. Appl. For.* 24(4): 180-186.

He, Q., E. Chen, R. An, and Y. Li. 2013. Above-ground biomass and biomass components estimation using LiDAR data in a coniferous forest. *Forests*. 4(4): 984-1002.

Heath, L.S., M. Hansen, J.E. Smith, P.D. Miles, and B.W. Smith. 2008. Investigation into calculating tree biomass and carbon in the FIADB using a biomass expansion factor approach. *McWilliams, WH, Moisen, GG & Czaplewski, RL (eds.)*.

Henry, M., N. Picard, C. Trotta, R.J. Manlay, R. Valentini, M. Bernoux, and L. Saint-André. 2011. Estimating tree biomass of sub-Saharan African forests: a review of available allometric equations. *Silva Fenn.* 45(3B): 477-569.

Hepp, T.E., and G.H. Brister. 1982. Estimating crown biomass in loblolly pine plantations in the Carolina Flatwoods. *For. Sci.* 28(1): 115-127.

Hollinger, D.Y. 2008. Defining a Landscape-Scale Monitoring Tier for the North American Carbon Program. In *Field Measurements for Forest Carbon Monitoring.* C. Hoover (ed.), Springer Netherlands, p 3-16.

Hoover, C.M. 2008. Field measurements for forest carbon monitoring: A landscape-scale approach. Springer: New York, NY, USA, 2008.

Horvitz, D.G., and D.J. Thompson. 1952. A generalization of sampling without replacement from a finite universe. *J. Am. Stat. Assoc.* 47: 663-685.

Houghton, R.A. 2005. Aboveground forest biomass and the global carbon balance. *Glob. Change Biol.* 11(6): 945-958.

Hubben, G.A.A., D. Bishai, P. Pechlivanoglou, A.M. Cattelan, R. Grisetti, C. Facchin, F. Compostella, J. Bos, M. Postma, and A. Tramarin. 2008. The societal burden of HIV/AIDS in Northern Italy: An analysis of costs and quality of life. *AIDS care*. 20(4): 449-455.

Ishii, H., and M.E. Wilson. 2001. Crown structure of old-growth Douglas-fir in the western Cascade Range, Washington. *Can. J. For. Res.* 31(7): 1250-1261.

Ishii, H., and N. McDowell. 2001. Age-related development of crown structure in coastal Douglas-fir trees. *For. Ecol. Manage.* 169: 257-270.

Jenkins, C.J., D.C. Chojnacky, L.S. Heath, and R.A. Birdsey. 2003. National-scale biomass estimators for United States tree species. *For. Sci.* 49(1): 12-35.

Jessen, R. J. 1955. Determining the fruit count on a tree by randomized branch sampling. *Biometrics.* 11(1):99-109.

Junior, L.R.N., V.L. Engel, J.A. Parrotta, A.C.G.d. Melo, and D.S. Ré. 2014. Allometric equations for estimating tree biomass in restored mixed-species Atlantic Forest stands. *Biota Neotropica*. 14(2): 1-9.

Kershaw Jr, J.A., and D.A. Maguire. 1995. Crown structure in western hemlock, Douglas-fir, and grand fir in western Washington: trends in branch-level mass and leaf area. *Can. J. For. Res.* 25(12): 1897-1912.

Ketterings, Q.M., R. Coe, M. van Noordwijk, Y. Ambagau, and C.A. Palm. 2001. Reducing uncertainty in the use of allometric biomass equations for predicting above-ground tree biomass in mixed secondary forests. *For. Ecol. Manage.* 146(1): 199-209.

Korhonen, L., K.T. Korhonen, P. Stenberg, M. Maltamo, and M. Rautiainen. 2007. Local models for forest canopy cover with beta regression. *Silva Fennica*. 41(4): 671.

Kozak, A. 2004. My last words on taper equations. *For. Chron.* 80(4): 507-514.

Krutchkoff, R. G. 1967. Classical and inverse regression methods of calibration. *Technometrics*. 9(3): 425-439.

Kutner, M.H, C. J. Nachtsheim, and J. Neter. 2004. Applied linear regression models. 4th ed. McGraw-Hill Irwin, Boston.

Kuyaha, S., J. Dietz, C. Muthuri, M.V. Noordwijk, and H. Neufeldt. 2013. Allometry and partitioning of above- and below-ground biomass in farmed eucalyptus species dominant in Western Kenyan agricultural landscapes. *Biomass and Bioenergy.* 55: 276-284.

Kuznetsova, T., M. Tilk, H. Parn, A. Lukjanova, and M. Mandre. 2011. Growth, aboveground biomass, and nutrient concentration of young Scots pine and lodgepole pine in oil shale post-mining landscapes in Estonia. *Environ. Monit. Assess.* 183(1-4): 341-350.

Lambert, M.C., C.H. Ung, and F. Raulier. 2005. Canadian national tree aboveground biomass equations. *Can. J. For. Res.* 35(8): 1996-2018.

Li, R., A. Weiskittel, A. R. Dick, J. A. Kershaw Jr., and R. S. Seymour. 2012. Regional stem taper equations for eleven conifer species in the Acadian Region of North America: development and assessment. *North. J. Appl. For.* 29(1): 5-14.

Liang, J.J., M. Zhou, D.L. Verbyla, L.J. Zhang, A.L. Springsteen, and T. Malone. 2011. Mapping forest dynamics under climate change: A matrix model. *For. Ecol. Manage.* 262(12): 2250-2262.

Lu, D. 2005. Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon. *Int. J. Remote Sens.* 26(12): 509-2525.

Lu, D. 2006. The potential and challenge of remote sensing-based biomass estimation. *Int. J. Remote Sens.* 27(7): 1297-1328.

Lu, D., P. Mausel, E. Brondizio, and E. Moran. 2002. Above-ground biomass estimation of successional and mature forests using TM images in the Amazon Basin. In *Advances in Spatial Data Handling*, Springer, p 183-196.

Luo, Y., X. Wang, X. Zhang, Y. Ren, and H. Poorter. 2013. Variation in biomass expansion factors for China's forests in relation to forest type, climate, and stand development. *Ann. For. Sci.* 70(6): 589-599.

Maier, M.J. 2014. DirichletReg: Dirichlet Regression for Compositional Data in R. Research Report Series / Department of Statistics and Mathematics, 125. WU Vienna University of Economics and Business, Vienna.

Maier, M.J. 2014a. DirichletReg: Dirichlet Regression in R. R package version 0.5-2.

Melson, S.L., M.E. Harmon, J.S. Fried, and J.B. Domingo. 2011. Estimates of live-tree carbon stores in the Pacific Northwest are sensitive to model selection. *Carbon Balance Manag.* 6(2): 1-16.

Miles, P.D., and W.B. Smith. 2009. Specific gravity and other properties of wood and bark for 156 tree species found in North America. *Res. Note. NRS-38*. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, p 35.

Naidu, S.L., E.H. DeLucia, and R.B. Thomas. 1998. Contrasting patterns of biomass allocation in dominant and suppressed loblolly pine. *Can. J. For. Res.* 28(8): 1116-1124.

Nunes, L., D. Lopes, F.C. Rego, and S.T. Gower. 2013. Aboveground biomass and net primary production of pine, oak and mixed pine-oak forests on the Vila Real district, Portugal. *For. Ecol. Manage.* 305: 38-47.

Omule, S.A.Y., V.E. Fletcherand, and K.R. Polsson. 1987. Total and merchantable volume equations for small coastal Douglas-fir. Economic & Regional Develop. Agree., B.C. Min. For. and Lands, Res. Br., FRDA Rep. 010, Victoria, B.C.

Paladinic, E., D. Vuletic, I. Martinic, H. Marjanovic, K. Indir, M. Benko, and V. Novotny. 2009. Forest biomass and sequestered carbon estimation according to main tree components on the forest stand scale. *Periodicum Biologorum.* 111(4): 459-466.

Parresol, B.R. 1999. Assessing tree and stand biomass: A review with examples and critical comparisons. *For. Sci.* 45(4): 573-593.

Parresol, B.R. 2001. Additivity of nonlinear biomass equations. *Can. J. For. Res.-Rev. Can. Rech. For.* 31(5): 865-878.

Peduzzi, P., J. Concato, E. Kemper, T.R. Holford, and A.R. Feinstein. 1996. A simulation study of the number of events per variable in logistic regression analysis. *Journal of clinical epidemiology*. 49(12): 1373-1379.

Pflugmacher, D., W. Cohen, R. Kennedy, and M. Lefsky. 2008. Regional Applicability of Forest Height and Aboveground Biomass Models for the Geoscience Laser Altimeter System. *For. Sci.* 54(6): 647-657.

Pooreter, H., K.J. Niklas, P.B. Reich, J. Oleksyn, P. Poot, and L. Mommer. 2012. Biomass allocation to leaves, stems and roots: meta-analyses of interspecific variation and environmental control. *New Phytologist*. 193: 30-50.

R Core Team. 2014. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Raile, G.K. 1982. Estimating stump volume. *Res. Pap. NC-224.* St Paul, MN: U.S. Department of Agriculture, Forest Service, p 4.

Räty, M., and A. Kangas. 2008. Localizing general models with classification and regression trees. *Scand. J. Forest Res.* 23(5): 419-430.

Repola, J. 2006. Models for vertical wood density of Scots pine, Norway spruce and birch stems, and their application to determine average wood density. *Silva Fennica*. 40(4): 673-685.

Ritchie, M.W., J. Zhang, and T.A. Hamilton. 2013. Aboveground Tree Biomass for Pinus ponderosa in Northeastern California. *Forests*. 4(1): 179-196.

Rojo, A., X. Perales, F. Sanchez-Rodriguez, J.G. Alvarez-Gonzalez, and K. von Gadow. 2005. Stem taper functions for maritime pine (*Pinus pinaster* Ait.) in Galicia (Northwestern Spain). *Eur. J. For. Res.* 124(3): 177-186.

Saatchi, S., K. Halligan, D. G. Despain, and, R.L. Crabtree. 2007. Estimation of forest fuel load from radar remote sensing. *IEEE Trans. Geosci. Remote Sens.* 45(6): 1726-1740.

SAS Institute Inc. 2013. Base SAS® 9.4 Procedures Guide: Statistical Procedures, Second Edition. Cary, NC: SAS Institute Inc.

Schlecht, R., and D.L. Affleck. 2014. Branch aggregation and crown allometry condition the precision of randomized branch sampling estimators of conifer crown mass. *Can. J. For. Res.* 44(5): 499-508.

Sedjo, R., and B. Sohngen. 2012. Carbon sequestration in forests and soils. *Annu. Rev. Resour. Econ.* 4: 127-144.

Smithson, M., and J. Verkuilen. 2006. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological methods*. 11(1): 54.

Snowdon, P. 1986. Sampling strategies and methods of estimating the biomass of crown components in individual trees of *Pinus radiata* D. Don. *Aust. For. Res.* 16(1): 63-72.

Snowdon, P. 1991. A ratio estimator for bias correction in logarithmic regressions. *Can. J. For. Res.* 21(5): 720-724.

Standish, J.T., G.H. Manning, and J.P. Demaerschalk. 1985. Development of biomass equations for British Columbia tree species. Can. For. Serv., Info. Rep. BC-X-264, Pacific Forest Resource Center, Victoria, BC, Canada. p 47.

Swank, W. T., and H. T. Schreuder. 1974. Comparison of three methods of estimating surface area and biomass for a forest of young eastern white pine. *For. Sci.* 20(1): 91-100.

Taylor, D.C.A., V. Pawar, D. T. Kruzikas, K. E. Gilmore, M. Sanon, and M. C. Weinstein. 2012. Incorporating calibrated model parameters into sensitivity analyses deterministic and probabilistic approaches. *Pharmacoeconomics.* 30 (2): 119-126.

Temesgen, H. 2003. Evaluation of sampling alternatives to quantify tree leaf area. *Can. J. For. Res.* 33(1): 82-95.

Temesgen, H., D. Affleck, K. Poudel, A. Gray, and J. Sessions. 2015. A review of the challenges and opportunities in estimating above ground forest biomass using tree-level models. *Scand. J. Forest Res.* 30(4): 326-335.

Temesgen, H., V. Monleon, A. Weiskittel, and D. Wilson. 2011. Sampling strategies for efficient estimation of tree foliage biomass. *For. Sci.* 57(2): 153-163.

Temesgen, H., V.J. Monleon, and D.W. Hann. 2008. Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. *Can. J. For. Res.* 38: 553-565.

Thompson, S.K. 2002. Sampling. 2nd edn. John Wiley & Sons, Inc.: New York.

Tumwebaze, S. B., E. Bevilacqua, R. Briggs, and T. Volk. 2013. Allometric biomass equations for tree species used in agroforestry systems in Uganda. *Agroforest Syst.* 87(4): 781-795.

U.S. Environmental Protection Agency. 2015. Inventory of U.S. greenhouse gas emissions and sinks: 1990–2013. EPA 430-R-15-004., U.S. Environmental Protection Act, Washington, DC: U.S.

Valentine, H. T., and S. J. Hilton. 1977. Sampling oak foliage by the randomized-branch method. *Can. J. For. Res.* 7: 295-298.

Valentine, H.T., L.M. Tritton, and G.M. Furnival. 1984. Subsampling trees for biomass, volume, or mineral content. *For. Sci.* 30(3): 673-681.

Vashum, K.T., and S. Jayakumar. 2012. Methods to Estimate Above-Ground Biomass and Carbon Stock in Natural Forests-A Review. *J. Ecosyst. Ecogr.* 2(4): 1-7.

Venables, W.N., and B.D. Ripley. 2002. Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0.

Westfall, J.A., D.W. MacFarlane, and A.R. Weiskittel. 2012. Biomass measurement and modeling challenges for hardwood species in the northern region. *In Moving from status to trends: Forest Inventory and Analysis (FIA) symposium 2012. December 4-6, 2012, Baltimore, MD*.

Williams, R.A. 1989. Use of randomized branch and importance sampling to estimate loblolly pine biomass. *South. J. Appl. For.* 13(4): 181-184.

Woodall, C., L.S. Heath, G.M. Domke, and M.C. Nichols. 2011. Methods and equations for estimating aboveground volume, biomass, and carbon for trees in the US forest inventory, 2010. Gen. Tech. Rep. NRS-GTR-88. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, p 30.

Xiao, C.W., J.C. Yuste, I.A. Janssens, P. Roskams, L. Nachtergale, A. Carrara, B.Y. Sanchez, and R. Ceulemans. 2003. Above- and belowground biomass and net primary production in a 73-year-old Scots pine forest. *Tree Physiology*. 23(8): 505-516.

Zeng, W.S., H.R. Zhang, and S.Z. Tang. 2011. Using the dummy variable model approach to construct compatible single-tree biomass equations at different scales - a case study for Masson pine (Pinus massoniana) in southern China. *Can. J. For. Res.* 41(7): 1547-1554.

Zhang, Y.J., B.E. Borders, R.E. Will, and H.D. Posadas. 2004. A model for foliage and branch biomass prediction for intensively managed fast growing loblolly pine. *For. Sci.* 50(1): 65-80.

Zhou, X., and M.A. Hemstrom. 2009. Estimating aboveground tree biomass on forest land in the Pacific Northwest: a comparison of approaches. *Res. Pap. PNW-RP-584*. Portland, OR: U.S. Department of Agriculture, Forest Service. Pacific Northwest Research Station, p. 18.

Zhou, X., and M.A. Hemstrom. 2010. Timber volume and aboveground live tree biomass estimations for landscape analyses in the Pacific Northwest. *Gen. Tech. Rep. PNW-GTR-819*. Portland, OR: U.S. Department of Agriculture, Forest Service, p. 31.

Zianis, D., P. Muukkonen , R. Mäkipää, and M. Mencuccini. 2005. Biomass and stem volume equations for tree species in Europe. *Silva Fennica Monographs.* 4. p 63.