Canadian Journal of Forest Research
Revue canadienne de recherche forestière

Methods for Estimating Aboveground Biomass and its Components for Douglas-fir and lodgepole pine trees

Journal:	<i>Canadian Journal of Forest Research</i>
Manuscript ID	cjfr-2015-0256.R1
Manuscript Type:	Article
Date Submitted by the Author:	21-Sep-2015
Complete List of Authors:	Poudel, Krishna; Oregon State Univeristy, Forest Engineering, Resources, and Management Temesgen, Hailemariam; Oregon State University
Keyword:	Component Ratio Method, Seemingly Unrelated Regression, Multinomial Loglinear Regression, Beta Regression, Dirichlet Regression

Methods for Estimating Aboveground Biomass and its Components for Douglas-fir and lodgepole pine trees

Poudel, K.P. and H. Temesgen

Estimating aboveground biomass and its components requires sound statistical formulation and evaluation. Using data collected from 55 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 predicted the proportion of biomass in each component using beta, Dirichlet, and multinomial loglinear regression (MLR). The predicted proportions were then applied to the total aboveground biomass to obtain amount of biomass in each component. The MLR approach consistently produced smaller RMSEs compared to both SUR approaches. The beta and Dirichlet regressions were superior to both SUR methods except for Douglas-fir branch biomass for which the simple SUR produced smaller RMSE compared to the beta and Dirichlet regressions.

Keywords: Component Ratio Method, Seemingly Unrelated Regression, Multinomial Loglinear Regression, Beta Regression, Dirichlet Regression

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 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 dioxide. 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 preferred to regional models in their current form when predicting biomass at the local scale. de-Miguel et al. (2014) found that the generalized mixed-effects meta-models calibrated with location specific small sample using a best unbiased linear predictor performed better than fitting location specific equations based on much larger samples.

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, different tree components are used for different purposes requiring separate estimates of component biomasses. Primarily, the main bole is used for timber production, crown biomass can provide information on fuel load and wildfire assessment, and small branches and leaves are useful in bioenergy production. Thus the information on biomass in each component is required to determine the associated cost of transportation and processing biomass. However, on the other hand, if the objective is to process whole trees through cull removal or removal of small diameter trees for whole tree chipping at the bioenergy plant, the component biomass estimates may not be as essential.

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 allocation of biomass to different components is influenced by differences in management practices (Tumwebaze et al. 2013; de-Miguel et al. 2014a), age and stand structure (de-Miguel et al. 2014a). 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 without local calibration. Weighing trees in the field would be the most accurate method for estimating tree biomass. However, it is time consuming and very expensive making the use of allometric equations inevitable. 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. Their methods of obtaining total and component aboveground biomass include 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 the equations developed by Jenkins et al. (2003). 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. The stem biomass for 20-, 40-, and 60-year old eastern Mediterranean *Pinus brutia* estimated using intrinsically linear models was found by de-Miguel et al. (2014a) to be 79.8, 80.5, and 80.6 percent of total aboveground biomass while the crown biomass was 20.2, 19.5, and 19.4 percent respectively. They found the proportion of stem and crown biomass for uneven-aged stands to be 61.8 and 38.1 percent 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 account for approximately 50 percent of the gross live volume in the Pacific Northwest. In this study, we developed and evaluated different approaches for estimating aboveground biomass and its components for Douglas-fir and lodgepole pine trees using the data collected destructively from sampled trees. The ideal method for estimating biomass for a population would randomly select sample plots, harvest all trees in that plot to estimate plot biomass and then to extrapolate to the population of interest. However, the design based biomass estimation of a population is prohibitively expensive. Therefore, commonly the sample tree measurement and forest area estimates are obtained from independent forest inventory. The individual tree biomass equations are then applied to these inventory data to obtain biomass of the target population.

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 of methods (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) predicted 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.

Materials and Methods

Data and Study Area

A detailed biomass data collection was carried out by destructively sampling 22 Douglas-fir (*Pseudotsuga menziesii* (Mirbel) Franco) and 33 lodgepole pine (*Pinus contorta*) trees in different forests within the state of Oregon (Figure 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 trees with damaged tops were also not included in sampling. 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 2975 and 4458 branches from Douglas-fir and lodgepole pine respectively. The average diameter at breast height was 54.9 cm and 24.6 cm and average height was 33 m and 17 m for Douglas-fir and lodgepole pine trees respectively (Table 1).

The crown of the sample tree was divided in to three equal length strata. For all first order branches i. e. for the branches that are directly attached to the main bole, 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. Once on the ground, the bole of the sampled tree was sectioned into 5.18 m long sections. Three to five centimeter thick disks were removed from the top of the stump and every 5.18 meter; diameter inside bark at both ends of the 5.18 m sections were also obtained. The inside-bark cubic volume of these sections was calculated using the Smalian's formula which was then converted to inside bark section biomass by multiplying it by the average density of the disks taken from two ends. Inside bark volume of stump was calculated as cylinder while the volume of top section was calculated as cone. Stump volume and volume of top section was multiplied by density of the lowest and top disks to obtain the stump and top biomass respectively. Total inside bark bole biomass was obtained by summing the stump, sections, and top masses. Outside bark bole volume was obtained in the same way as to obtain total inside bark stem volume using outside bark diameters in place of inside bark diameters. Total stem bark volume was then obtained by subtracting inside bark stem volume from the outside bark stem volume. The density of bark samples was obtained as the ratio of dry weight to its volume calculated assuming a rectangular

shape. Then, the method of bark biomass calculation was same as the method for inside bark bole biomass calculation except that the bark sample densities were used instead of disk density.

Individual branch wood and foliage biomass was obtained by fitting species specific log linear model of the following form:

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

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; β_{ij} 's are regression parameters to be estimated from the data; $\ln(\cdot)$ is the natural logarithm; and ε_{ij} 's are the random error. The scatterplot of logarithm of branch wood and foliage biomasses against the logarithm of branch basal diameter showed that a linear model was appropriate for modeling these relationships. Similar model form has been used by Temesgen et al. (2011) to model foliar biomass for Douglas-fir and ponderosa pine and by Poudel et al. (2015) to estimate crown biomass in Douglas-fir. When logarithmic regressions are used, there is inherent negative bias due to the fact that the dependent variable is transformed prior to estimation (Snowdon, 1993). The commonly used remedy to this is to multiply the back transformed results by a correction factor $[\exp(\frac{MSE}{2})]$, proposed by Baskerville (1972), 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. The correction factors for Douglas-fir branch wood and foliage biomass models were 1.06 and 1.14 respectively and for lodgepole pine the correction factors were 1.12 and 1.16 for branch wood and foliage biomass models respectively. 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 being present in the main bole (Figure 2).

Methods for Estimating Aboveground Biomass

Group I Methods

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 to bole biomass 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.

Group II Methods

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:

$$\text{Bole} = \exp(a_{11} + a_{12} * X_1 + a_{13} * X_2 + a_{14} * X_3) \quad (2)$$

$$\text{Bark} = \exp(a_{21} + a_{22} * X_1 + a_{23} * X_2 + a_{24} * X_3) \quad (3)$$

$$\text{Branch} = \exp(a_{31} + a_{32} * X_1 + a_{33} * X_2 + a_{34} * X_3) \quad (4)$$

$$\text{Foliage} = \exp(a_{41} + a_{42} * X_1 + a_{43} * X_2 + a_{44} * X_3) \quad (5)$$

$$\text{Total} = \sum_{i,j} \exp(a_{ij} + a_{ij} * X_1 + a_{ij} * X_2 + a_{ij} * X_3) \quad (6)$$

Where a_{ij} , ($i = 1, 2, 3, 4$, and $j = 1, 2, 3, 4$) are parameters to be estimated from the data and X_1 , X_2 , and X_3 are natural logarithms of DBH, total tree height and crown length respectively. Note that for the simple SUR, only X_1 was used as explanatory variable. The dependent variables in these equations were component or total biomass (kg) in original scale not in the logarithmic scale. The choice of model form was primarily guided by past literature and the diagnostic plots did not show severe problem with the model form. The studentized residuals for the extended SUR models for bole biomass estimation in Douglas-fir and lodgepole pine trees are shown in Figure 3. 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.

Group III Methods

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.

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:

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

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:

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

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

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

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

Where, $g(\cdot)$ is strictly increasing and double differentiable link function that maps $(0, 1)$ in to the real line $\mathbb{R}$, $\mathbf{x}_i = (x_{i1}, \dots, x_{ik})^T$ is a vector of k explanatory variables, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_k)^T$ is a vector of unknown k unknown regression parameters ($k < n$), and $\boldsymbol{\eta}_i$ is a linear predictor (i.e. $\boldsymbol{\eta}_i = \beta_1 x_{i1} + \dots + \beta_k x_{ik}$, 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). The predictor variables in beta regression were DBH and total tree height for both species.

Dirichlet Regression

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

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

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. $B(\boldsymbol{\alpha}) = \frac{\prod_{c=1}^C \Gamma(\alpha_c)}{\Gamma(\sum_{c=1}^C \alpha_c)}$ 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:

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

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:

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

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 2014). The predictor variables used in Dirichlet regression were also the DBH and total tree height for both species

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:

$$p_{Bole} = \frac{1}{1+e^{(a1+a2\times X_1+a3\times X_2)}+e^{(b1+b2\times X_1+b3\times X_2)}+e^{(c1+c2\times X_1+c3\times X_2)}} \tag{13}$$

$$p_{Bark} = \frac{e^{(a1+a2\times X_1+a3\times X_2)}}{1+e^{(a1+a2\times X_1+a3\times X_2)}+e^{(b1+b2\times X_1+b3\times X_2)}+e^{(c1+c2\times X_1+c3\times X_2)}} \tag{14}$$

$$p_{Foliage} = \frac{e^{(b_1+b_2 \times X_1+b_3 \times X_2)}}{1+e^{(a_1+a_2 \times X_1+a_3 \times X_2)}+e^{(b_1+b_2 \times X_1+b_3 \times X_2)}+e^{(c_1+c_2 \times X_1+c_3 \times X_2)}} \quad (15)$$

$$p_{Branch} = \frac{e^{(c_1+c_2 \times X_1+c_3 \times X_2)}}{1+e^{(a_1+a_2 \times X_1+a_3 \times X_2)}+e^{(b_1+b_2 \times X_1+b_3 \times X_2)}+e^{(c_1+c_2 \times X_1+c_3 \times X_2)}} \quad (16)$$

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 a_i, b_i, c_i ($i = 1, 2, 3$) 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.

The methods used to estimate aboveground biomass and/or its components in this study were applied to Douglas-fir and lodgepole pine trees. Performance of all the methods was evaluated based on the bias and RMSE 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 the predicted value of the variable of interest.

Results and Discussion

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 2. These methods produced similar estimates for total aboveground biomass for lodgepole pine but for the Douglas fir trees these methods provided differing estimates. 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. These methods provided the component biomass estimates for lodgepole pine that were closer to each other than the component biomass estimates for Douglas-fir. However, none of these methods were consistent in over- or under-predicting the component masses (Figure 4). Previous studies have also shown great discrepancies in biomass estimates obtained from the FIA-CRM, FIA-PNW, and Jenkins methods. For example, Nay and Bormann (2014) found Jenkins equation to predict up to 190% of the site specific foliage biomass equation for Douglas-fir.

The lodgepole pine trees in our sample were the trees with smaller DBH (average DBH 24.6 cm) compared to the Douglas-fir trees (average DBH 54.9 cm). Indeed, the Jenkins, FIA-CRM, and FIA-PNW 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 are presented in Table 3. The average bias and RMSE produced by simple and extended SUR approaches are presented in Table 4. These locally fitted SUR models (both simple and extended) provided component and total aboveground biomass estimates that were more accurate (lower RMSE) than the estimates produced by the FIA-CRM, FIA-PNW, and Jenkins methods. Including additional explanatory variables than just DBH in the SUR models resulted in the decrease in RMSE percent from 10.7 to 8.3 for Douglas fir and from 22.8 to 20.5 for lodgepole pine total aboveground biomass. The RMSE for bole biomass estimation was reduced by 2.3 and 6.9 percent for Douglas fir and lodgepole

pine respectively by using the extended SUR approach instead of the simple SUR models. 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 slightly increased (Figure 5). 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. In practice, it is unreasonable to assume that information on actual/measured total AGB would be available. In that case, the AGB should be replaced with $\widehat{AGB}$ i.e. the predicted total aboveground biomass. The model given in equation (6) can be used to obtain $\widehat{AGB}$. The parameter estimates for these models are available on request. The bias and RMSE produced by these methods are given in Table 5. These methods unbiasedly predicted component proportions for all species. These methods also provided component biomass estimates that were more accurate (lower RMSE) than the estimates produced by the FIA-CRM, FIA-PNW, and Jenkins methods.

The beta regression produced smaller RMSEs compared to the simple SUR models except for Douglas-fir branch biomass while it produced smaller RMSEs than the extended SUR models for both Douglas-fir and lodgepole pine component biomasses. The beta regression produced 2.8 percent higher RMSE for Douglas-fir branch biomass but for foliage, bark, and bole biomass, it decreased RMSE by 5.4,

7.8 and 8.8 percent compared to the simple SUR models. In case of lodgepole pine, the beta regression decreased RMSEs by 16.5, 4.2, 13.7, and 24.6 percent for foliage, bark, branch, and bole biomass compare to the simple SUR models. Comparing the performance of beta regression against the extended SUR models, it decreased RMSEs by 8.1, 11.8, 0.2, and 6.5 for Douglas-fir and by 13.6, 4.2, 8.4, and 17.7 percent for lodgepole pine foliage, bark, branch, and bole biomasses respectively. 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).

Similarly, the Dirichlet regression also produced smaller RMSE compared to the simple SUR methods except for Douglas-fir branch biomass but it produced smaller RMSEs than the extended SUR models for both Douglas-fir and lodgepole pine component biomasses. The Dirichlet regression produced 2.1 percent higher RMSE for Douglas-fir branch biomass but for foliage, bark, and bole biomass, it decreased RMSE by 1.6, 3.7, and 8.7 percent compared to the simple SUR models. In case of lodgepole pine, the Dirichlet regression decreased RMSEs by 7.4, 0.8, 14.7, and 23.2 percent for foliage, bark, branch, and bole biomass compare to the simple SUR models. Comparing the performance of Dirichlet regression against the extended SUR models, it decreased RMSEs by 4.2, 7.8, 0.9, and 6.3 for Douglas-fir and by 4.5, 0.8, 9.5, and 16.3 percent for lodgepole pine foliage, bark, branch, and bole biomasses respectively. 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 better performance of Dirichlet regression models could be attributed to their capability to capture the variance in proportions that sum to a constant exhibiting skewness and heteroscedasticity.

The MLR consistently produced smaller RMSE compared to both SUR methods for both species and all components. The MLR decreased RMSEs by 2.1, 4.7, 1.1, and 9.1 percent for Douglas-fir and by

10.2, 5.0, 14.9, and 23.5 percent for lodgepole pine foliage, bark, branch, and bole biomasses respectively compared to the simple SUR models. The MLR produced RMSEs that were 4.7, 8.8, 4.2, and 6.8 percent smaller for Douglas-fir and 7.4, 5.0, 9.7, and 16.6 percent smaller for lodgepole pine foliage, bark, branch, and bole biomasses, respectively, compared to the RMSEs produced by the extended SUR models.

There was no single method to predict proportion that was best for all both species and component biomass. The beta regression produced smallest RMSEs for foliage and bark biomass estimation while the MLR produced the smallest RMSEs for branch and bole biomass estimation for Douglas-fir. In the case of lodgepole pine, the beta regression produced smallest RMSEs for foliage and bole biomass but the MLR produced smallest RMSEs for bark and branch biomass estimation. It should be noted, however, that the RMSE percent produced by these methods were within 4.2 percent of each other except for the lodgepole pine foliage estimation in which case the Dirichlet regression produced 9.1 percent higher RMSE compared to the beta regression.

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. 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 3.4 and 2.6 times higher average bark biomass estimates. The bole biomass estimates were within 10 percent for Douglas fir and within 22 percent for lodgepole pine. These methods were very inconsistent for branch and foliage biomass, over -estimating for one species-component combination 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.

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, it did not improve prediction for some 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-component combinations. One desired property in the component biomass estimation is the property of additivity which can be attained by simultaneous fitting of

component proportions in the Dirichlet regression and MLR. These methods ensure 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. When the large scale models are calibrated, these models could perform even better than the locally fitted equations (e.g. de-Miguel et al. 2014). 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.

482 Literature Cited

- 483 Baskerville, G. 1972. Use of logarithmic regression in the estimation of plant biomass. *Can. J. For. Res.*
 484 2(1): 49-53.
- 485 Beaucham, J.J., and J.S. Olson. 1973. Corrections for bias in regression estimates after logarithmic
 486 transformation. *Ecology*. 54(6): 1403-1407.
- 487 Boudewyn, P., X. Song, S. Magnussen, and M. Gillis. 2007. Model-based, volume-to-biomass conversion
 488 for forested and vegetated land in Canada. Canadian Forest Service, Victoria, Canada.
 489 Information Report BC-X-411.
- 490 Brown, S. 2002. Measuring carbon in forests: current status and future challenges. *Environmental*
 491 *pollution*. 116(3): 363-372.
- 492 Cribari-Neto, F., and A. Zeileis. 2010. Beta Regression in R. *Journal of Statistical Software* 34(2), 1-24.
- 493 Czapowskyj, M.M., D.J. Robison, R.D. Briggs, and E.H. White. 1985. Component biomass equations for
 494 Black Spruce in Maine. *Res. Pap. NE-564*. Broomal, PA: U.S. Department of Agriculture, Forest
 495 Service, Northeastern Research Station, p. 7.
- 496 de-Miguel, S., L. Mehtatalo, and A. Durkaya. 2014. Developing generalized, calibratable, mixed-effects
 497 meta-models for large-scale biomass prediction. *Can. J. For. Res.* 44(6): 648-656.
- 498 de-Miguel, S., T. Pukkala, N. Assaf, and Z. Shater. 2014a. Intra-specific differences in allometric equations
 499 for aboveground biomass of eastern Mediterranean *Pinus brutia*. *Ann. For. Sci.* 71(1): 101-112.
- 500 De Paola, M., V. Scoppa, and R. Lombardo. 2010. Can gender quotas break down negative stereotypes?
 501 Evidence from changes in electoral rules. *J. Pub. Econ.* 94(5): 344-353.
- 502 Eskelson, B.N., L. Madsen, J.C. Hagar, and H. Temesgen. 2011. Estimating Riparian understory vegetation
 503 cover with Beta regression and copula models. *For. Sci.* 57(3): 212-221.
- 504 Ferrari, S., and F. Cribari-Neto. 2004. Beta regression for modelling rates and proportions. *J. Appl. Stat.*
 505 31(7): 799-815.
- 506 Flewelling, J.W., and L.V. Pienaar. 1981. Multiplicative regressions with lognormal errors. *For. Sci.* 27(2):
 507 281-289.
- 508 Harrison, R.B., T.A. Terry, C.W. Licata, B.L. Flaming, R. Meade, I.A. Guerrini, B.D. Strahm *et al.* 2009.
 509 Biomass and Stand Characteristics of a Highly Productive Mixed Douglas-Fir and Western
 510 Hemlock Plantation in Coastal Washington. *West. J. Appl. For.* 24(4): 180-186.
- 511 He, Q.S., E.X. Chen, R. An, and Y. Li. 2013. Above-Ground Biomass and Biomass Components Estimation
 512 Using LiDAR Data in a Coniferous Forest. *Forests*. 4(4): 984-1002.
- 513 Heath, L.S., M. Hansen, J.E. Smith, P.D. Miles, and B.W. Smith. 2008. Investigation into calculating tree
 514 biomass and carbon in the FIADB using a biomass expansion factor approach. *McWilliams, WH,*
 515 *Moisen, GG & Czaplewski, RL (eds.).*
- 516 Hollinger, D.Y. 2008. Defining a Landscape-Scale Monitoring Tier for the North American Carbon
 517 Program. In *Field Measurements for Forest Carbon Monitoring*. C. Hoover (ed.), Springer
 518 Netherlands, p. 3-16.
- 519 Hoover, C.M. 2008. Field measurements for forest carbon monitoring: A landscape-scale approach.
 520 Springer : New York, NY, USA, 2008.
- 521 Hubben, G.A.A., D. Bishai, P. Pechlivanoglou, A.M. Cattelan, R. Grisetti, C. Facchin, F. Compostella, J. Bos,
 522 M. Postma, and A. Tramarin. 2008. The societal burden of HIV/AIDS in Northern Italy: An
 523 analysis of costs and quality of life. *AIDS care*. 20(4): 449-455.
- 524 Jenkins, J.C., D.C. Chojnacky, L.S. Heath, and R.A. Birdsey. 2003. National-scale biomass estimators for
 525 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, pp. 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. 2014. 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.
- Nay, S.M., and B.T. Bormann. 2014. Site-specific Douglas-fir biomass equations from the Siskiyou Mountains, Oregon, compared with others from the Pacific Northwest. *For. Sci.* 60(6): 1140-1147.
- 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.
- Poudel, K.P., H. Temesgen, and A.N. Gray. 2015. Evaluation of sampling strategies to estimate crown biomass. *For. Ecosyst.* 2:1 p11.
- 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.
- 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.

- 573 Temesgen, H., V. Monleon, A. Weiskittel, and D. Wilson. 2011. Sampling strategies of efficient
574 estimation of tree foliage biomass. *For. Sci.* 57(2): 153-163.
- 575 Tumwebaze, S.B., E. Bevilacqua, R. Briggs, and T. Volk. 2013. Allometric biomass equations for tree
576 species used in agroforestry systems in Uganda. *Agroforestry Systems*. 87(4): 781-795.
- 577 Vashum, K.T., and S. Jayakumar. 2012. Methods to estimate above-ground biomass and carbon stock in
578 natural forests-a review. *J. Ecosyst. Ecogr.* 2(4): 1-7.
- 579 Venables, W.N., and B.D. Ripley. 2002. *Modern Applied Statistics with S*. Fourth Edition. Springer, New
580 York. ISBN 0-387-95457-0.
- 581 Westfall, J.A., D.W. MacFarlane, and A.R. Weiskittel. 2012. Biomass measurement and modeling
582 challenges for hardwood species in the northern region. *In Moving from status to trends: Forest
583 Inventory and Analysis (FIA) symposium 2012. December 4-6, 2012, Baltimore, MD.*
- 584 Xiao, C.W., J.C. Yuste, I.A. Janssens, P. Roskams, L. Nachtergale, A. Carrara, B.Y. Sanchez, and R.
585 Ceulemans. 2003. Above- and belowground biomass and net primary production in a 73-year-
586 old Scots pine forest. *Tree Physiology*. 23(8): 505-516.
- 587 Zhang, Y.J., B.E. Borders, R.E. Will, and H.D. Posadas. 2004. A model for foliage and branch biomass
588 prediction for intensively managed fast growing loblolly pine. *For. Sci.* 50(1): 65-80.
- 589 Zhou, X., and M.A. Hemstrom. 2009. Estimating aboveground tree biomass on forest land in the Pacific
590 Northwest: a comparison of approaches. *Res. Pap. PNW-RP-584*. Portland, OR: U.S. Department
591 of Agriculture, Forest Service, p. 18.
- 592 Zhou, X., and M.A. Hemstrom. 2010. Timber volume and aboveground live tree biomass estimations for
593 landscape analyses in the Pacific Northwest. *Gen. Tech. Rep. PNW-GTR-819*. Portland, OR: U.S.
594 Department of Agriculture, Forest Service, p. 31.

List of Figures:

Figure 1. Locations of study sites. Data collection was carried out in the summer of 2012 and 2013.

Figure 2. Aboveground biomass distribution in different tree components (percentage of total aboveground biomass) in different species (Species are DF = Douglas-fir and LP = Lodgepole Pine).

Figure 3. Studentized residuals for the extended SUR models for bole biomass estimation in a) Douglas-fir and b) lodgepole pine trees.

Figure 4. Average component biomass estimates produced by the FIA-CRM, FIA-PNW, and Jenkins methods in Douglas-fir and lodgepole pine trees.

Figure 5. RMSE (kg) produced by simple and extended SUR approaches in estimating component biomass in Douglas-fir and lodgepole pine trees.

List of Tables:

Table 1. Summary statistics for tree data sampled. Min, max, and CV are minimum, maximum, and coefficient of variation of the variables, respectively.

Species	Variable	Min	Mean	Max	CV
Douglas-fir	Foliage (kg)	10.0	56.91	169.00	0.82
	Branch (kg)	19.0	221.7	791.0	1.03
	Bark (kg)	6.0	105.6	372.0	1.03
	Bole (kg)	123.0	1916.9	6515.0	1.05
	Total (kg)	187.0	2301.2	7578.0	1.04
	DBH (cm)	19.3	54.9	114.0	0.54
	Total Height (m)	16.6	33.0	48.8	0.31
Lodgepole pine	Foliage (kg)	5.0	17.61	54.00	0.70
	Branch (kg)	10.0	47.6	178.0	0.90
	Bark (kg)	1.0	11.9	49.0	0.97
	Bole (kg)	24.0	220.2	798.0	0.97
	Total (kg)	49.0	297.5	899.0	0.87
	DBH (cm)	13.5	24.6	42.9	0.35
	Total Height (m)	9.2	17.0	31.9	0.35

Table 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 and LP = Lodgepole pine).

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
	LP	3.2	-4.8	19.4	25.1	72.4	12.7	12.5	37.9	70.5	101.8
FIA-PNW	DF	-14.9	-143.2	-161.2	182.8	-136.6	37.7	232.5	280.3	311.7	376.1
	LP	6.6	-3.9	22.7	38.8	64.2	12.0	9.4	40.6	78.7	99.7
Jenkins	DF	-102.1	-255.5	-134.8	-93.8	-586.2	176.0	418.7	246.5	597.6	1327.7
	LP	3.5	-18.7	29.7	49.1	63.6	9.9	26.7	44.6	102.5	104.2

Table 3. Parameter estimates and their approximate standard errors of the simple and extended SUR models for Douglas-fir and lodgepole pine.

Method	Parameter	Douglas-fir		Lodgepole pine	
		Estimate	Approximate Standard Error	Estimate	Approximate Standard Error
Simple SUR	a11	-0.349	0.096	-2.374	0.334
	a12	1.925	0.021	2.373	0.093
	a21	-3.247	0.701	-4.401	1.150
	a22	1.919	0.154	2.105	0.324
	a31	-2.604	0.822	-3.384	1.137
	a32	1.942	0.180	2.202	0.318
	a41	-3.242	0.896	-2.472	0.855
	a42	1.758	0.197	1.637	0.245
Extended SUR	a11	-1.993	0.062	-1.622	0.174
	a12	1.597	0.012	1.067	0.069
	a13	0.827	0.009	1.234	0.077
	a21	-4.459	0.806	-4.602	1.165
	a22	2.169	0.176	1.842	0.643
	a23			0.361	0.606
	a31	-3.856	0.454	-6.345	0.921
	a32	1.795	0.093	4.678	0.318
	a33	0.611	0.046	-1.865	0.283
	a41	-4.517	0.564	-5.134	0.776
	a42	1.497	0.147	3.843	0.324
	a43	0.779	0.171	-1.647	0.325

Table 4. Average bias and RMSE for component and total aboveground biomass produced by the simple and extended SUR approaches (Species are DF = Douglas-fir and LP = Lodgepole pine).

Method	Species	Bias (kg)					RMSE (kg)				
		Foliage	Bark	Branch	Bole	Total	Foliage	Bark	Branch	Bole	Total
Simple	DF	4.0	0.4	0.3	-44.1	-39.3	18.4	28.2	66.2	249.2	245.8
SUR	LP	0.7	0.1	2.3	-2.7	0.5	8.8	7.1	28.7	79.7	67.9
Extended	DF	-0.7	12.7	-16.6	-16.3	-20.9	19.9	32.5	72.9	204.3	190.9
SUR	LP	2.5	0.2	6.8	-10.9	-1.2	8.3	7.1	26.2	64.4	60.9

Table 5. Average bias and RMSE of component biomass produced by the beta, multinomial loglinear, and Dirichlet regression approaches (Species are DF = Douglas-fir and LP = Lodgepole pine). 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
	LP	-0.122	0.105	-0.111	-0.766	5.9	6.6	22.2	25.5
MLR	DF	0.001	-0.005	-0.004	-0.083	17.2	23.2	63.7	74.3
	LP	-0.007	-0.006	-0.025	-0.114	7.0	6.5	21.6	27.9
Dirichlet	DF	-1.948	-2.971	6.807	-1.979	17.5	24.3	70.8	83.0
	LP	-1.515	-1.356	0.729	1.990	7.5	7.0	21.7	28.6