METHODS FOR ESTIMATING ABOVEGROUND BIOMASS AND ITS COMPONENTS FOR FIVE PACIFIC NORTHWEST TREE SPECIES

Krishna P. Poudel[1] and Temesgen Hailemariam[2]

Abstract—Performance of three groups of methods to estimate total and/or component aboveground biomass was evaluated using the data collected from destructively sampled trees in different parts of Oregon. 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 was superior to group I methods. The third group of methods predicted the proportions of biomass in each component using beta, Dirichlet, and multinomial logistic regression (MLR). The MLR approach produced smaller root mean squared error (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.

Forest carbon reporting requires information on tree measurements, forest area estimates, and methods to estimate forest biomass. Tree measurements and forest area estimates for the official U.S. forest carbon reporting are obtained from the U.S. Forest Service's Forest Inventory and Analysis (FIA). Forest biomass estimates until 2009 were based on the equations developed by Jenkins and others (2003) but after 2009 these estimates are obtained using the component ratio method of the FIA (FIA-CRM). These methods were developed for national scale application but are commonly used to estimate biomass at the local scale. Therefore these methods may be biased at the local scale if there is spatial variation in the tree form due to one or more unknown predictors in the sub-region or subarea.

DATA

A detailed biomass data collection was carried out by destructively sampling 90 trees in different forests within the state of Oregon. 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).* 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 stand edges. Trees that were forked below breast height and with damaged tops were also not included in sampling. The average D.B.H. ranged from 24.6 cm to 54.9 cm and average height ranged from 17 m to 33 m. Volume of 5.18 m bole 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 section masses. Individual branch wood and foliage biomass was obtained by fitting species specific log linear regression as a function of branch diameter.

METHODS

Methods for estimating aboveground biomass (AGB) used in this study belonged to three major groups. The first group of methods used analytical approach based on existing equations. The analytical methods are the FIA-CRM, the equations used by the FIA-PNW and the equations developed by Jenkins and others (2003). The equations used in FIA-CRM and FIA-PNW methods can be found in Woodall and others (2011)

[1] PhD Candidate, Department of Forest Engineering, Resources and Management, Oregon State University, Corvallis, OR, 97333. Tel: (541) 737-2687, Krishna.Poudel@oregonstate.edu.

[2] Professor, Department of Forest Engineering, Resources and Management, Oregon State University, Corvallis, OR, 97333. Tel: (541) 737-8548 Hailemariam.Temesgen@oregonstate.edu.

and Zhou and Hemstrom (2010), respectively. The second group is the locally fitted systems of equations using a seemingly unrelated regression. We fitted two systems of equations: a DBH based single entry system (simple SUR) and a multiple entry system that included DBH and other explanatory variables (extended SUR) to estimate component and total AGB. The systems were constrained such that the sum of the predicted biomass from component equations is equal to the total AGB obtained from the equation for total AGB. Third group of methods predicted proportion of total AGB in different components using beta, Dirichlet, and multinomial logistic regression (MLR). The predicted proportions were then multiplied by observed total AGB to obtain predicted biomass estimates in different components. These generalized regression models assume that the errors of component biomass equations are beta, Dirichlet, and multinomial distributed, respectively. The beta and Dirichlet regressions have been described by Ferrari and Cribari-Neto (2004) and Maier (2010), respectively. Performance of all the methods was evaluated based on bias and root mean squared error (RMSE).

RESULTS AND DISCUSSION

The FIA-CRM, FIA-PNW, and the Jenkins methods were biased and produced the highest RMSE values among the methods used in the study. The average bias and RMSE produced by these methods are given in Table 1. These methods produced similar estimates for total AGB except for Douglas-fir. The Jenkins method for Douglas-fir produced total AGB 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 AGB, these methods showed inconsistent discrepancies in component biomass estimates (Fig. 1).

It is important to note, however, that the component biomass estimates obtained from these methods were similar for lodgepole pine and red alder, trees with smaller D.B.H. in our study. Indeed, these methods were more sensitive to tree size compared to other methods. For example, in Douglas-fir, the RMSE percent for total AGB dropped from 57.7 to 11.1 percent, 10.2 to 7.1 percent, and 16.3 to 8.5 percent when Jenkins, FIA-CRM, and FIA-PNW equations, respectively, were applied to trees with less than 94 cm D.B.H.

The average bias and RMSE produced by simple and extended SUR approaches are shown in Figure 2. These methods consistently provided smaller RMSE compared to FIA-CRM, FIA-PNW, and Jenkins methods. Including additional explanatory variables than just D.B.H. in the SUR models 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 AGB 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 tree with same D.B.H. but different height which would not be accounted for by D.B.H. only models.

However, it should be noted that even though the RMSE for total AGB is decreased by using the extended SUR, the RMSE for some component biomass increased (Fig. 2). 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 MLR unbiasedly predicted the proportions of biomass in different components. 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 (Table 2). There was no clear winner within this group of methods.

The beta regression produced smaller RMSEs compared to the simple SUR models except for grand fir foliage and bark biomass and Douglas-fir branch

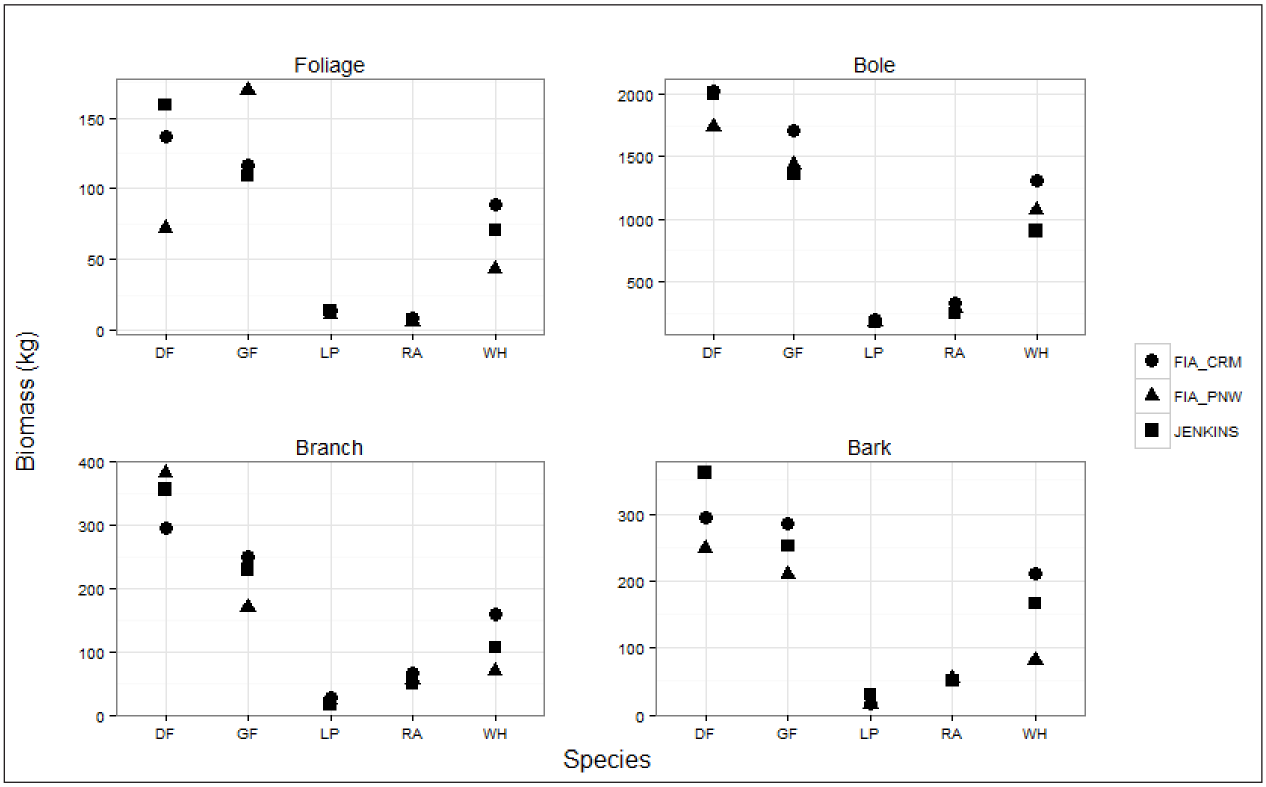

Figure 1—Average component biomass estimates produced by the FIA-CRM, FIA-PNW, and Jenkins methods in different species (Species are DF = Douglas-fir; GF = Grand Fir; LP = Lodgepole Pine; RA = Red Alder; WH = Western Hemlock).

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 compared to the simple and extended SUR methods respectively.

The Dirichlet regression also produced smaller RMSEs compared to 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 simple and extended SUR methods respectively. In the case of conifers, it produced smaller RMSEs compared to simple 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.

The MLR consistently produced smaller RMSEs compared to simple SUR methods for all species and all components. It also produced smaller RMSEs 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 and others (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 for grand fir.

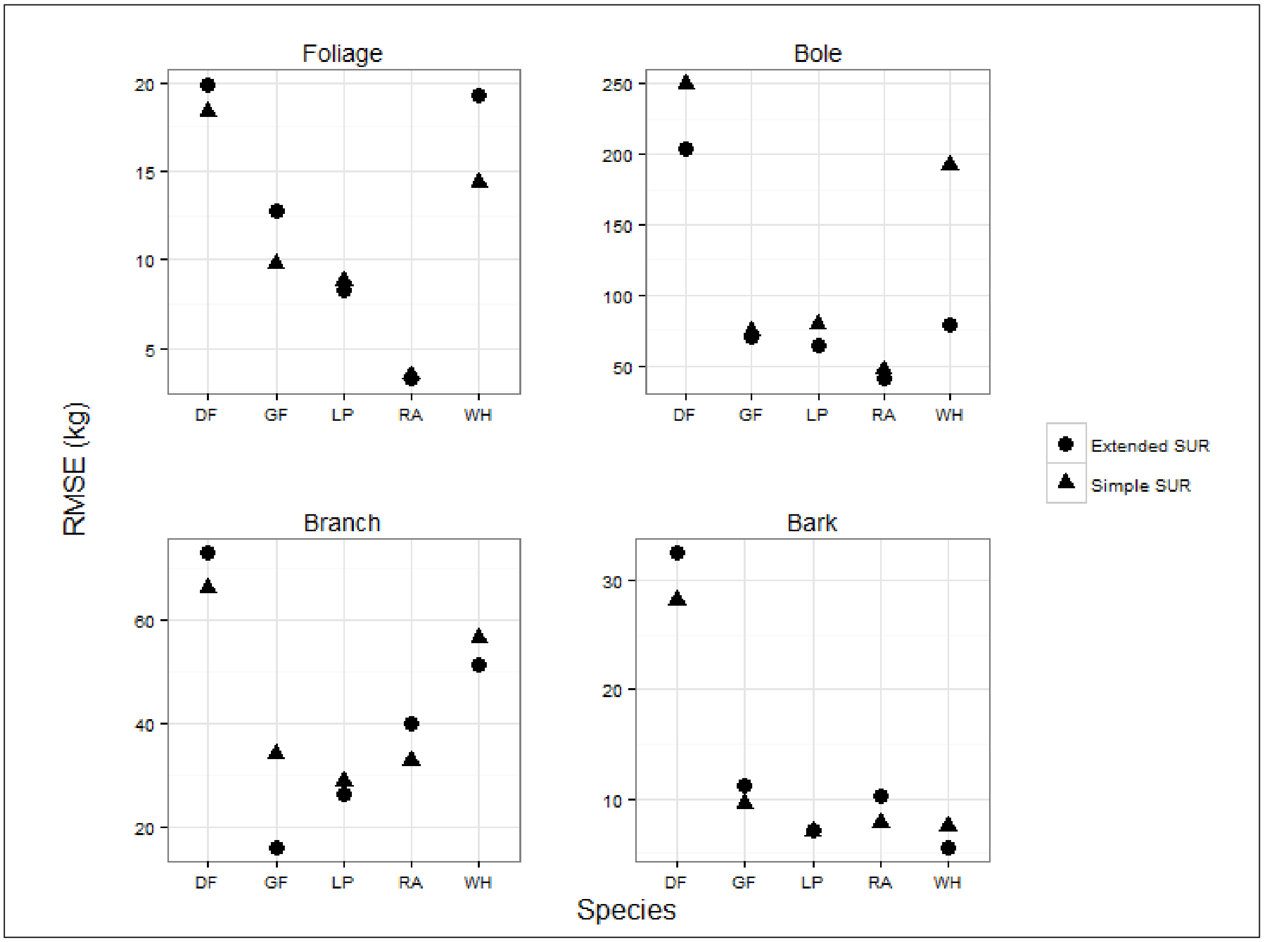

Figure 2—RMSE produced by simple and extended SUR approaches in estimating component biomass for different species (Species are DF = Douglas-fir; GF = Grand Fir; LP = Lodgepole Pine; RA = Red Alder; WH = Western Hemlock).

Even though the methods or models that are capable of predicting biomass at large scale are desired, the use of such models without local calibration could lead to serious bias due to the differences in scale of development and application of such models. Findings of this study provide information on the efficiency of selected methods in quantifying component and total AGB. Methods to predict proportions are promising to apportion total AGB to different components. One advantage of using Dirichlet regression and MLR over beta regression is that these regressions allow simultaneous fitting of the component proportions and therefore the predicted proportions sum to 1. Application of the methods to predict component proportions for other species and locations and with the larger dataset would further validate their accuracy.

ACKNOWLEDGMENT

The authors thank Drs. Lisa Madsen and Glen Murphy, both professors at Oregon State University and Dr. Andrew Gray, Research Ecologist at USDA Forest Service, Pacific Northwest Research Station for their comments and suggestions during the manuscript preparation.

Table 1—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 2—Average bias and RMSE of component biomass produced by the beta, multinomial logistic, 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.

		Bias (kg)				**RMSE (kg)**			
Method	**Species**	**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

LITERATURE CITED

Ferrari S. L. P.; Cribari-Neto F. (2004). Beta regression for modelling rates and proportions. Journal of Applied Statistics. 31: 799–815.

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

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.

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

Woodall, C.W.; Heath L.S.; Domke G.M.; Nichols M.C. 2011. Methods and equations for estimating aboveground volume, biomass, and carbon for trees in the US forest inventory. USDA For. Serv., Gen. Tech. Rep. NRS-88, Northern Research Station, Newtown Square, PA. 30 p.

Zhou, X.; Hemstrom M.A. 2010. Timber volume and aboveground live tree biomass estimations for landscape analyses in the Pacific Northwest. USDA For. Serv., Gen. Tech. Rep. PNW-GTR-819. Pacific Northwest Research Station, Portland, OR. 31 p.