REDRAWING THE BASELINE: A METHOD FOR ADJUSTING BIASED HISTORICAL FOREST ESTIMATES USING A SPATIAL AND TEMPORALLY REPRESENTATIVE PLOT NETWORK

Sara A. Goeking and Paul L. Patterson[1]

Abstract—Users of Forest Inventory and Analysis (FIA) data sometimes compare historic and current forest inventory estimates, despite warnings that such comparisons may be tenuous. The purpose of this study was to demonstrate a method for obtaining a more accurate and representative reference dataset using data collected at co-located plots (i.e., plots that were measured during both periodic and annual inventories). The approach described here uses co-located plot-level data to build linear regression models that relate annual inventory measurements to periodic inventory measurements. Separate models were constructed within each state, and wherever possible, for domains defined by factors that may affect forest attributes over time and that also affected the intensity of the periodic inventories (i.e., timber versus woodland forest types). We used these regressions to simulate periodic-era, plot-level response variables, on a per-acre basis, for annual plot locations that were not sampled during the periodic inventories. Because the extent of the resulting dataset coincides with the annual plot grid, the post-stratification procedures used to produce broad-scale annual inventory estimates can be applied to the simulated periodic dataset to produce periodic-era estimates of forest attributes. Construction of this simulated periodic-era dataset allows investigation of broad-scale trends in forest attributes, particularly as they vary across ownership group, reserved status, and forest type group due to disturbance and land management history.

In the eastern U.S., the Forest Inventory and Analysis (FIA) program has completed multiple inventory cycles and therefore provides assessments of trends in forest attributes such as volume, growth, mortality, biomass, and carbon over time. However, in the western U.S., the 10-year cycle length precludes long-term evaluations in states where only one cycle of data has been collected. In these areas, many users of FIA data rely on historical, periodic inventory estimates to serve as temporal reference conditions, and then directly compare them to annual estimates to quantify forest trends. Because the periodic plots did not representatively sample all forested locations, directly comparing the periodic and annual estimates can lead to erroneous conclusions (Goeking 2015). The purpose of this paper is to describe a methodological framework for obtaining a more accurate and representative reference dataset using data collected at co-located plots, or plots that were measured during both periodic and annual inventories, in states where direct comparisons of multiple inventories over time are tenuous.

STUDY AREA

The methods described below were applied to the eight states within the Interior West FIA region: Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, and Wyoming. The analysis was restricted to periodic plot measurements collected from 1993-2002 and annual plot measurements collected from 2004-2013.

[1] Biological Scientist (SAG), Rocky Mountain Research Station, USDA Forest Service, 507 25th Street, Ogden, UT 84401; and Statistician (PLP), Rocky Mountain Research Station, 2150-A Centre Ave, Suite 350, Fort Collins, CO 80526-8121. SAG is corresponding author: to contact, call (801) 625-5193 or e-mail at sgoeking@fs.fed.us.

METHODS

The response variables presented here include net volume of live trees, net volume of dead trees, and above-ground tree biomass at time 1 (periodic inventory), where the predictors are the values of these variables at time 2 (annual inventory). Tree-level volumes were obtained as the variable VOLCFNET in FIADB (O'Connell et al. 2015) and related to tree status (live or dead) to permit separate calculations of live and dead net volume. Biomass was queried from several variables that constitute the component ratio method, as described in O'Connell et al. (2015), and summed to a single above-ground metric. Because differences in periodic versus annual inventory plot designs preclude direct comparisons of total plot-level tree volume, these variables were calculated on a per-acre basis as described by Goeking (2015).

Based on the linear relationships evident between time 1 and time 2 plot-level volumes (Fig. 1), we adopted the approach of developing linear regression models where time 1 values were predicted based on time 2 values. Although this is contrary to typical regression modeling that seeks to predict future values based on current or previous measurements, in this situation the time 2 dataset is more complete and representative than any of the time 1 datasets.

Prior to building regression models, we identified domains for the purpose of developing separate regression models. Individual states formed the primary division into domains. Within each state, we considered that timber and woodland forest types might require separate regression models because their attributes may experience different rates of change, and also because this distinction undoubtedly affected the intensity of the periodic inventories (Goeking 2015). Thus, within each state, we tested whether timber and woodland forest types qualified as separate domains versus a single domain for the state. To qualify as a single domain, the regression models for timber and woodland plots within each state had to have slopes and intercepts that were not statistically different. We followed the procedure described by Zar (1996) for comparing two or more regression equations, which first tests for equal slopes and then if the slopes are not statistically different, tests for equal intercepts. Each response variable was considered separately, so the tests for equal slopes and intercepts were repeated for live and dead volume. An alpha of 0.05 was used to reject the null hypotheses that slopes and intercepts were equal between timber and woodland models.

Based on the results of the comparisons of slopes and intercepts in each state, we established domain-specific linear regression models relating the estimates made with the annual and periodic data at co-located plots. We used these relationships to estimate periodic-era, plot-level response variables for annual plot locations that were not sampled during the periodic inventory. Using plot-specific expansion factors obtained from the annual post-stratification estimation process, we then produced estimates of live volume, dead volume, and biomass.

RESULTS

Table 1 presents the results of tests for equal slopes and intercepts between regression models for timber and woodland plots within each state. Based on these results, timber and woodland domains were modeled separately in most states. Exceptions included Colorado, where a single model was used for each response variable (live volume, dead volume, and biomass); and Arizona and Wyoming, where each state had one model for biomass.

The relationships between above-ground biomass per acre at co-located plots, as measured at time 1 and time 2, for each modeling domain are shown in Figure 1. Adjusted r^2 values were generally lower for woodland models than for timber models.

Ongoing research and future papers will present the detailed calculation of statewide estimates based on the modeling approach described here and investigate trends in forest attributes such as volume, biomass, growth and mortality, particularly as they vary across ownership group, reserved status, and forest type group due to potential differences in disturbance and land management history. This plot-based approach will allow evaluation of changes in volume and

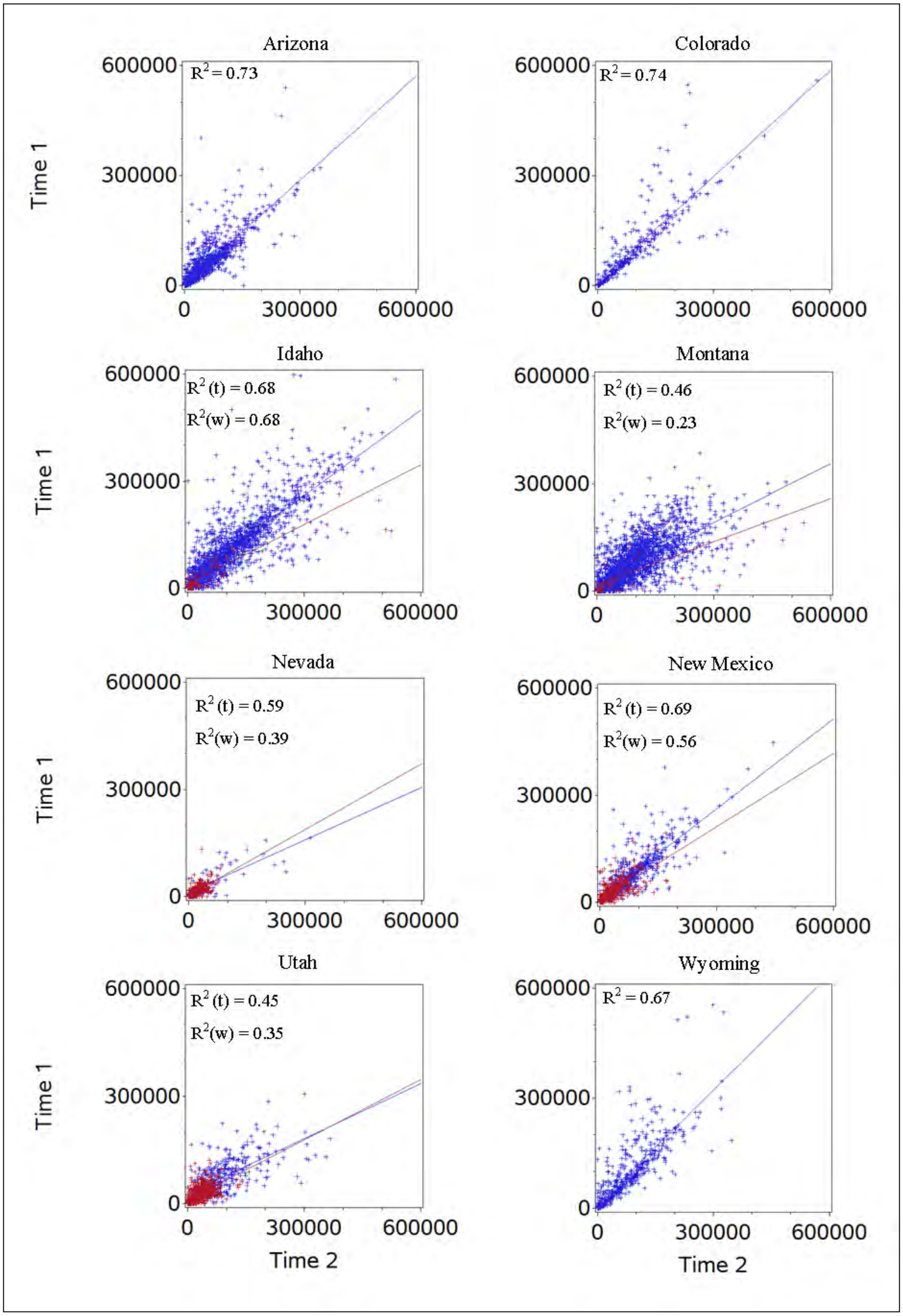

Figure 1—Scatter plots of time 1 (periodic) versus time 2 (annual) above-ground tree biomass, in dry tons per acre, by state. Within each state, each line represents a domain, where a regression model was developed for each domain. Five of the 8 states had separate domains for timber (blue markers) and woodland (red markers). In Arizona, Colorado, and Wyoming, timber and woodland were grouped into a single domain. Domains were defined based on the results of tests for equal slopes and intercepts of the timber and woodland regressions. Adjusted r2 values are shown for each modeling domain, where t indicates timber domains and w indicates woodland domains.

Table 1—Results of tests used for identifying domains for regression models, by state. Tests for equal slopes and equal intercepts follow Zar's (1996) methods for comparing two or more regression equations. Where "Number of domains"=2, timber and woodland domains were modeled separately.

		Test for equal slopes	Test for equal intercepts			Number of domains
State	Variable	p-value[a]	t[b]	df[b]	t*	
Arizona	Live volume	0.003	3.064	1635	1.962	2
	Dead volume	<.0001	4.606	1635	1.962	2
	Biomass	0.169	0.373	1635	1.962	1
Colorado	Live volume	0.365	1.589	333	1.968	1
	Dead volume	0.105	1.159	333	1.968	1
	Biomass	0.337	1.445	333	1.968	1
Idaho	Live volume	0.197	4.889	1760	1.962	2
	Dead volume	0.037	1.755	1760	1.962	2
	Biomass	0.008	4.109	1760	1.962	2
Montana	Live volume	0.437	2.264	2118	1.962	2
	Dead volume	0.665	2.068	2118	1.962	2
	Biomass	0.195	2.389	2118	1.962	2
Nevada	Live volume	<.0001	4.777	449	1.962	2
	Dead volume	0.073	2.791	449	1.962	2
	Biomass	<.0001	3.925	449	1.962	2
New Mexico	Live volume	0.015	1.420	1372	1.966	2
	Dead volume	0.732	4.438	1372	1.966	2
	Biomass	0.034	1.276	1372	1.966	2
Utah	Live volume	0.042	9.421	1218	1.962	2
	Dead volume	0.001	7.388	1218	1.962	2
	Biomass	0.258	6.037	1218	1.962	2
Wyoming	Live volume	0.383	3.302	509	1.965	2
	Dead volume	0.017	1.938	509	1.965	2
	Biomass	0.303	1.359	509	1.965	1

[a]P-values for tests of equal slopes were produced using a contrast statement in Proc GLM (SAS Institute, Inc. 2009).
[b]Values of t-statistics and df (degrees of freedom) were calculated as prescribed by Zar (1996) for testing equal elevations of regression models, and compared to critical values (t*). Where values of t are greater than t*, the null hypothesis that the intercepts are equal was rejected.

biomass by categories such as ownership group and reserved status, and the estimated variance of our volume and biomass estimates; the estimated variance will need to include the error associated with the simulated periodic dataset.

DISCUSSION

The modeling approach described here generates a spatially balanced dataset of periodic-era plot-level variables, to which the annual inventory's post-stratification estimation process can be applied to produce broad-scale periodic-era estimates. The fundamental advantage of FIA's annual inventory design over previous periodic inventories is that it provides spatially and temporally representative estimates of forest attributes (Bechtold and Patterson 2005). This advantage is especially pronounced in regions such as the Interior West, where periodic inventories were decidedly non-representative not only throughout space and time, but also relative to tabular attributes such as ownership and forest type. This paper describes an approach for using co-located plot data to produce more representative baseline estimates for the periodic inventories of the 1990s. Although the development of this approach is focused on the Interior West, it could be applied to other states or regions where (a) periodic datasets are known to be incongruous with annual inventory datasets, and (b) sufficient co-located plots exist to build regression models that allow prediction of time 1 (periodic inventory) values based on time 2 (annual inventory) values.

ACKNOWLEDGMENT

Mike T. Thompson provided invaluable conceptual and programming feedback, and Mark Brown, Jennifer Bakken, and L. Scott Baggett provided constructive reviews of this paper. This work was supported by the Rocky Mountain Research Station's Inventory & Monitoring Program.

LITERATURE CITED

Bechtold, W.A.; Patterson, P.L., eds. 2005. *The enhanced forest inventory and analysis program—National sampling design and estimation procedures.* USDA For. Serv., Gen. Tech. Rep. SRS-80, Southern Research Station, Asheville, NC. 85 p.

Goeking, S.A. In press. Disentangling forest change from forest inventory change: A case study from the U.S. Interior West. J. of Forestry, http://dx.doi.org/10.5849/jof.14-088.

O'Connell, B.M.; LaPoint, E.B.; Turner, J.A.; Ridley, T.; Boyer, D.; Wilson, A.M.; Waddell, K.L.; Pugh, S.A.; Conkling, B.L. 2015. *The Forest Inventory and Analysis database: Database description and user guide version 6.0 for phase 2 (version 6.0.2).* Available: http://www.fia.fs.fed.us/library/database-documentation/current/ver6.0/FIADB%20User%20Guide%20P2_6-0-2_final-opt.pdf. (Accessed June 12, 2015.)

SAS INSTITUTE, INC. 2009. *Base SAS 9.2 Procedures Guide.* SAS Institute, Inc., Cary, NC. 1705 p.

Zar, J.H. 1996. *Biostatistical analysis.* 3rd edition. Prentice Hall, Upper Saddle River, NJ. 662 p., plus appendices.