For. Sci. 60(2):206–213
http://dx.doi.org/10.5849/forsci.12-072

biometrics

A Comparison of Carbon Stock Estimates and Projections for the Northeastern United States

Richard G. MacLean, Mark J. Ducey, and Coeli M. Hoover

We conducted a comparison of carbon stock estimates produced by three different methods using regional data from the USDA Forest Service Forest Inventory and Analysis (FIA). Two methods incorporated by the Forest Vegetation Simulator (FVS) were compared to each other and to the current FIA component ratio method. We also examined the uncalibrated performance of FVS growth simulations for predicting net carbon accumulation in live trees. In general, the three carbon stock estimation approaches do not produce estimates that are either equivalent or are simply convertible. A strong spatial pattern of relationships between estimates was associated with regional variation in stand top height. Uncalibrated growth projections gave downwardly biased results that were also poorly correlated with observed carbon accumulation rates, yielding little improvement in root mean square error over the use of a simple regional average. These results reinforce the need for managers and scientists to be careful in choosing methods and reporting carbon stock estimates and to use appropriate model calibration methods in projecting future carbon accumulation.

Keywords: biomass, growth and yield, model validation, carbon markets

The role of forests in the global carbon cycle and the potential of forests as a greenhouse gas mitigation tool have been topics of increasing attention in recent years, both at the national and international levels. Nongovernment organizations, national governments, and corporations are some of the entities implementing forest carbon projects around the world; the Forest Carbon Portal provides a searchable index of planned and active projects and currently lists 412 projects representing 7.9 million ha (Forest Carbon Portal 2012). Maintaining and increasing forest carbon stocks can also result in a number of cobenefits, including wildlife habitat, soil quality, and water quality (for example, Imai et al. 2009). The value of carbon sequestration as an ecosystem service is perhaps best illustrated by the REDD+ Programme of the United Nations (Reducing Emissions from Deforestation and Forest Degradation), which recognizes the role of forests as a tool for mitigating both greenhouse gas emissions and poverty (UN-REDD 2012).

In the United States, forests sequestered enough carbon to offset 13.5% of national greenhouse gas emissions in 2010, the most recent reporting year for which data are publicly available. (US Environmental Protection Agency 2012). Recognizing the importance of forests and their interactions with climate, and the role of national forests, the USDA Forest Service has published a roadmap for responding to climate change and has directed each national forest and grassland to use a 10-point scorecard to assess and report progress in implementing the agency's climate change strategy (USDA Forest Service 2011a). One of the elements of the scorecard relates to carbon assessment (USDA Forest Service 2011b) and asks the following: "Does the unit have a baseline assessment of carbon stocks and the influence of disturbance and management activities on these stocks? Is the unit integrating carbon stewardship with the management of other benefits being provided by the unit?" The increased emphasis on carbon in forests presents managers with a need for tools to estimate forest carbon stocks, assess the implications of management actions on those stocks, and investigate the tradeoffs between carbon sequestration and other forest management objectives.

Because managers were receiving an increasing amount of inquiries about the carbon consequences of planned management actions, in 2006 carbon estimation was added to the Fire and Fuels Extension (FFE; Rebain 2010) of the Forest Vegetation Simulator (FVS; Crookston and Dixon 2005). This functionality allows managers who are familiar with FVS to quickly generate carbon estimates along with standard FVS output as part of routine simulations. The carbon reports in the FFE-FVS are widely used by researchers,

Manuscript received June 20, 2012; accepted February 14, 2013; published online June 27, 2013.

Affiliations: *Richard G. MacLean (rich.maclean@unh.edu), University of New Hampshire. Mark J. Ducey (mjducey@unh.edu), University of New Hampshire, Department of Natural Resources, Durham, NH. Coeli M. Hoover (choover@fs.fed.us), USDA Forest Service.*

Acknowledgments: *This project was supported by Research Joint Venture Agreement RJVA 09-JV-11242302110 between the USDA Forest Service and the University of New Hampshire. We gratefully acknowledge the constructive comments of three reviewers and the associate editor.*

public and private forest managers, and forestry consultants to generate carbon estimates and examine the carbon outcomes of management alternatives and are also recommended as a means of addressing the carbon assessment element of the Climate Scorecard for National Forests and Grasslands. The carbon reports available in the FFE-FVS offer a choice of two calculation methods for live aboveground tree biomass: the FFE default method (FFE), based on regional volume equations, and the Jenkins et al. (2003) national biomass equations (Jenkins). Due to the increasing use of the carbon reports for a range of objectives, we are conducting a region-by-region comparison of the aboveground carbon estimation methods. Prior to 2009 the Forest Inventory and Analysis Program (FIA) calculated aboveground biomass using the Jenkins approach but currently implements the component ratio method (CRM), which uses FIA regional volume equations combined with the tree component rations from the Jenkins approach (Woodall et al. 2011). Since FIA now provides the official US forest carbon stock estimates it is important to assess how the FVS carbon report values compare to the estimates resulting from the FIA CRM approach.

Here, we present a regional comparison carbon estimates for the Northeast variant of the FVS growth-and-yield model. Specific objectives of this paper are:

1. To test whether the estimates produced by the FFE, Jenkins, and FIA CRM approaches can be regarded as operationally equivalent and, if not, to determine whether simple conversion factors or other straightforward mathematical adjustments can be used to translate results from one system to another.
2. To describe between-plot and geographic variation in the relationships between estimates, and to examine what factors predict observed differences.
3. To test whether use of FIA data with the FVS growth-and-yield model run without calibration to local conditions can produce realistic estimates of carbon stock changes at the plot or regional level.

Materials and Methods

Field Data and Growth Projections

We used data from the USDA Forest Service, Forest Inventory and Analysis (FIA) to evaluate differences in estimated carbon stocks and growth projections. We used all available remeasured phase 2 plot data from the current annualized design for those states appropriate to the Northeast variant of FVS. Remeasured plot data were available for Connecticut (CT), Maine (ME), Massachusetts (MA), New Hampshire (NH), New York (NY), Ohio (OH), Pennsylvania (PA), and Vermont (VT). Data were obtained from the FIA Database version 4.0.[1] The phase 2 plots consist of a cluster of four subplots, with one phase 2 plot per 2,428 ha arranged on a systematic national grid. Within each subplot, all trees larger than 12.7 cm DBH are measured on a 7.32 m radius fixed-area sample, while trees less than 12.7 cm DBH are measured on a 2.07 m radius sample. In principle, all plots are remeasured on a 5-year interval, though occasional slight deviations occur. Additional details on FIA protocols and associated estimators can be found in Bechtold and Patterson (2005).

Each state-level database was processed with Structures Query Language (SQL) scripts to isolate remeasured plots and their associated data for conversion to FVS-ready formats. Remeasured plots showing any harvested trees during the remeasurement interval were eliminated. We then used the FIA2FVS program[2] (Vandendriesche 2012) to convert each state-level database into an FVS-ready database. All data were processed through the Northeast variant of FVS (v. 6.21) with the same set of simulation parameters. For each set of input data, carbon stocks were estimated using either the FFE or Jenkins method generated by the FFE-FVS carbon reports. We summarize both methods briefly here; further detail on the development of carbon reporting within FVS can be found in Hoover and Rebain (2008), Rebain (2010), and Hoover and Rebain (2011).

- The *default FFE method* for calculating aboveground live tree carbon begins with estimates of merchantable cubic volume (in the Northeast variant merchantable trees are defined as ≥ 12.7 cm DBH for softwoods and ≥ 12.7–20.3 cm DBH for hardwoods, depending on location), calculated using variant-specific volume equations that depend on DBH and tree height. If measured tree heights are available, these are used; if not, they are predicted internally within FVS. The volume estimates are converted to biomass using species-specific wood density (Forest Products Laboratory 1999). In the Northeast variant, biomass of crowns and nonmerchantable stems is estimated using Jenkins et al. (2003), added to the merchantable biomass, and converted to carbon assuming that half of all woody biomass is carbon. Bark is not calculated as part of biomass in the FFE and so is not included in the aboveground tree carbon estimate, although this may change as the tool is updated over time.
- The Jenkins method predicts aboveground live tree biomass directly using a series of nationally averaged biomass allometric equations from Jenkins et al. (2003), which lump tree species into nine taxonomic and structural groups. These allometric equations are simple power functions of DBH for trees greater than 2.5 cm DBH (and include bark on the bole). For trees less than 2.5 cm DBH, prediction of biomass is by linear interpolation (assuming biomass is 0 when DBH is 0). For all trees, carbon is assumed to be one-half of total aboveground biomass.

For comparison purposes, we also calculated aboveground live tree carbon using the protocols recently adopted by the FIA program for official reporting purposes (Heath et al. 2009, Woodall et al. 2011). Until reporting results from the 2010 inventories, FIA had used protocols essentially identical to those used in the Jenkins method described above. However, FIA now uses a volume conversion factor approach (component ratio method, CRM) that is similar to that of the FFE method but with some key differences. In the northeast, for all trees greater than 12.7 cm DBH, gross cubic volume per tree is predicted using region- and species-specific volume equations that depend on DBH and measurements of either total or merchantable height (depending on the equation). In general, the volume equations used by FIA (Woodall et al. 2011) are different from those incorporated in FVS; for the Northeast variant, these can be found in Miles and Hill (2010). Then, gross volume is converted to sound volume after deduction for any observed rot or voids and converted to bole biomass using species-specific wood density (Miles and Smith 2009). Finally, bole biomass is converted to total aboveground carbon using the volume conversion factors from Jenkins et al. (2003), an adjustment factor, and a carbon fraction of 0.5 (for further details, see Woodall et al. 2011). For smaller trees, the procedure is different. Trees with DBH greater than 2.5 cm and less than 2.7 cm the whole-tree allometric equations of Jenkins et al.

(2003) are used to predict biomass (again with a carbon fraction of 0.5). Trees with DBH less than 2.5 cm are not included in the FIA carbon estimates, and foliage is excluded for all trees.

In all cases, we used total aboveground live carbon, computed as described above, as the metric for comparisons. Calculation methods and components included in the aboveground live tree carbon pool vary by approach; the purpose of this study is to compare the output of the tools as they are currently configured, and so no adjustments were made to harmonize carbon pool components. Because the live tree tally from FIA plots serves as the common shared data for all the approaches used here and aboveground live tree carbon represents the largest pool of carbon (soil carbon is not included in FFE-FVS) reported, it is the focus of this study. Growth was modeled in FVS using 1-year intervals from the initial measurement year to the remeasurement year for each plot. Although extended modeling over long durations with 1-year intervals is not typically advisable using FVS (Wykoff et al. 1982, Hoover and Rebain 2011), the remeasurement interval of the FIA data is substantially less than the default 10-year projection interval of the Northeast variant of FVS (Dixon and Keyser 2008). No local calibration was used in the growth simulations, as we wished to assess the impact of using FIA data "as translated" using the FIA2FVS tool without additional supplementary information. For all comparisons, differences between carbon estimates calculated using the FIA data at the initial measurement and remeasurement periods were treated as observed net growth, while differences between estimates based on projected stand conditions at the remeasurement period and those from the initial measurement were treated as modeled net growth.

Statistical Analysis

To test whether the estimates produced by FFE, Jenkins, and FIA can be regarded as operationally equivalent, we used an equivalence testing framework (e.g., Robinson and Froese 2004, Robinson et al. 2005) with the estimates produced at the initial measurement of all plots. Equivalence testing reverses the usual "burden of proof" of statistical hypothesis testing, in which "no difference" is treated as the null hypothesis and the data must demonstrate that a difference actually exists. Under the equivalence testing framework, two sets of estimates are assumed not to be equivalent unless the data demonstrate convincingly that the estimates are similar to within a predefined tolerance interval. We initially considered a regression-based equivalence test (Robinson et al. 2005). However recognizing that for a plot without any trees all three approaches would return an estimate of 0 t/ha, there is no basis for asserting a shift in intercept. Instead, we used "two one-sided tests" (TOST; Berger and Hsu 1996) on the ratios between methods to determine if the ratio between estimates could be demonstrated to be sufficiently close to one that the two methods should be regarded as equivalent. Because some estimates were very close to zero the distribution of ratios could be very long-tailed, so we conducted equivalence testing on both the mean and median of the ratios. We defined two sets of equivalence criteria. For tight equivalence, we required the data to demonstrate that the mean or median ratio of estimates between two methods was within a tolerance interval from 0.95 to 1.05. For rough equivalence, we required the mean or median ratio to be within a tolerance interval from 0.80 to 1.20. Although criteria and discounting of uncertainty continue to evolve (Olander and Haugen-Kozyra 2011), these limits correspond loosely to those suggested under some greenhouse gas registries. For example, the California Climate Action Registry (2008) indicates that for credits based on forest inventories, confidence limits of $< 5\%$ of the mean are needed for an estimate to be accepted without discounting, while confidence limits of $> 20\%$ of the mean require 100% discounting (i.e., receive no credit). Under the TOST approach, using a nominal $\alpha = 0.05$, equivalence is demonstrated if both the upper and lower 90% confidence limits on the mean (or median) fall within the tolerance interval (Berger and Hsu 1996, Wellek 2003). Confidence limits for the mean were calculated using the conventional parametric approach; confidence limits for the median were calculated nonparametrically following Conover (1999, p. 143–144).

To describe factors driving regional differences in carbon estimates, we used a combination of geographic visualization and predictive modeling techniques. Mean ratios were mapped at the county scale across the study region. We also used Random Forest regression (Breiman 2001) to predict the ratio between methods at the plot level. The Random Forest algorithm is a sophisticated "bagging" algorithm that constructs multiple regression trees from resampled versions of the original data set. The resulting regression trees are then averaged to produce a final prediction. We used 500 trees, with several plot-level candidate predictors: the "fuzzed" plot coordinates[3] (latitude and longitude), plot elevation, stand age, trees per ha, basal area per ha, quadratic mean diameter, crown competition factor, Reineke's (1933) stand density index, and top height. Modeling was implemented using the Random Forest extension (Liaw and Wiener 2002) of the R statistics package (R Development Core Team 2010).

To evaluate FVS predictions of net carbon gain, we used simple descriptive statistics along with visualization using the ggplot2 extension (Wickham 2009) of R. Our original intent had been to use the regression-based equivalence testing approach of Robinson et al. (2005) to assess whether predicted and observed growth fell sufficiently close to a 1:1 line, but the outcomes of modeling proved that to be largely unnecessary. We reemphasize that the growth projections used here were performed without any local calibration; our intent was to evaluate the quality of projections that would result from users utilizing the FIA2FVS tool in an automated fashion using only the data that would be imported directly from the FIA database, not to evaluate the full potential for FVS to produce accurate short-term projections when calibrated thoughtfully using additional site- or project-specific auxiliary data.

Results

Overall, 4,341 remeasured, nonharvested FIA plots were available for analysis within the geographic scope of the FVS Northeast variant. Results of the comparison of total aboveground live tree carbon stocks are shown in Figure 1. There is an overall pattern of the Jenkins method providing the highest estimates, while the FIA approach gives the lowest, but there is also a great deal of variability at the individual plot level.

This overall pattern is confirmed by the equivalence tests for both means and medians of ratios (Table 1). None of the methods showed tight equivalence to any other (defined here as being within 5% of each other on average or on a typical basis, i.e., using the median). The FFE method did show rough equivalence (within 20% on an average or typical basis) to both the Jenkins and FIA methods, but those two methods were not equivalent to each other. The spread in plot-level ratios as depicted in Figure 1 indicates that no single conversion factor will provide a consistent and reliable

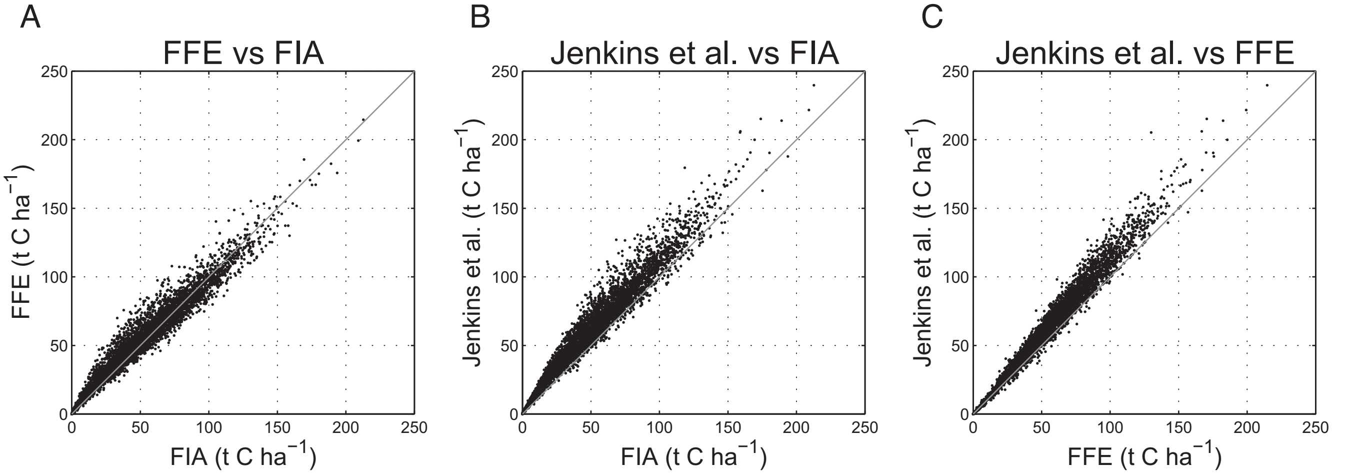

Figure 1. Comparison of aboveground live tree carbon stock estimates obtained using (A) FFE and FIA, (B) Jenkins and FIA, and (C) FFE and Jenkins. Solid line indicates 1:1 relationship.

translation between methods across the region. For example, considering the ratio of FFE to FIA, which was closest to one on average among all the pairs (Table 1), 25% of the individual plots had values less than 0.969, while 25% had values in excess of 1.222; only 50% of individual plots were contained within this interval. To capture 90% of plots would require an interval from 0.879 to 1.793. The spread for the other ratios was similarly broad.

Maps of the mean ratios at the county level are shown in Figure 2. A strong geographic trend in all three sets of ratios is readily apparent. For example, while plot-level estimates calculated using the default FFE method average 13.7% higher than those calculated using the new FIA procedures, FFE estimates can be nearly 30% higher for counties in northern ME, and can actually be lower than FIA estimates in OH and PA. Similar trends can be seen for the relationship between Jenkins and FIA estimates. These results further highlight the challenge of developing a conversion factor, as using a single conversion factor to translate between methods would lead to underestimates in part of the region and overestimates in another. The geographic patterns for maps of the median ratios, or for ratios-of-means, are similar (though the exact numerical values differ) and are not shown.

Random Forest analysis showed that while ratios between estimates could be predicted at the plot level, both stand-level variables and geographic location contributed to prediction accuracy. Moreover, predictions were associated with considerable residual variability. The Random Forest model predicting the ratio of FFE to FIA estimates explained 71.6% of the variance, with a root mean squared residual of 0.171. Similarly, the model predicting the ratio of Jenkins to FIA estimates explained 71.7% of the variance, with a root mean squared residual of 0.154. The model predicting the ratio of Jenkins to FFE explained a smaller fraction of the variance, at 48.2%, but because that ratio was less variable in the data the root mean squared residual was also smaller at 0.077. Variable importance scores for the Random Forest modeling of the ratios are shown in Table 2. Although interpretation of variable importance scores in Random Forest can be problematic (Strobl et al. 2007), the results highlight the influence of stand structure (especially top height) on the ratio of either FFE or Jenkins to the FIA estimates of carbon stocks. In some ways, this is not surprising since the Jenkins method uses single-entry allometric equations, while the FIA method employs volume equations that also use measured height information, therefore, it is natural to expect that the ratio of Jenkins:FIA would be lower in areas where measured heights are taller than average and higher where measured heights are shorter. However, the same general pattern is observed for the FFE estimates, which also use measured heights in a volume conversion factor framework.

Comparison of observed and predicted net carbon growth rates, when FFE-FVS is used without local calibration, reveal performance that is not encouraging (Figure 3). Using the FFE approach to calculate carbon for the observed and modeled tree lists, mean observed net carbon accumulation in aboveground live trees for plots in the study region is 0.978 ± 0.045 t/ha/year (mean plus or minus one standard error). Without calibration, FFE-FVS tended to underpredict net carbon accumulation, with a mean of 0.806 ± 0.008 t/ha/yr. Moreover, while correlation between observed and modeled net carbon accumulation was statistically significant, it was also poor (Spearman's $\rho = 0.18$, $P < 0.0001$). As a result of the combination of bias and variability, the uncalibrated FVS projections had a large root mean square error (RMSE) of 2.961 t/ha/yr, nearly identical to that which would be obtained by simply using the mean regional

Table 1. Results of the equivalence tests comparing estimates of the total aboveground live tree carbon stocks using three different methods.

Comparison	Mean of ratios (95% confidence limits)	Equivalence[1]	Median of ratios (95% confidence limits)	Equivalence
FFE:FIA	1.137 (1.128–1.146)	Rough	1.053 (1.047–1.060)	Rough
Jenkins:FIA	1.281 (1.274–1.289)	None	1.214 (1.208–1.219)	None
Jenkins:FFE	1.140 (1.137–1.143)	Rough	1.137 (1.133–1.140)	Rough

[1]Tight equivalence: mean or median ratio of estimates between two methods was within a tolerance interval from 0.95 to 1.05. Rough equivalence: mean or median ratio within a tolerance interval from 0.80 to 1.20.

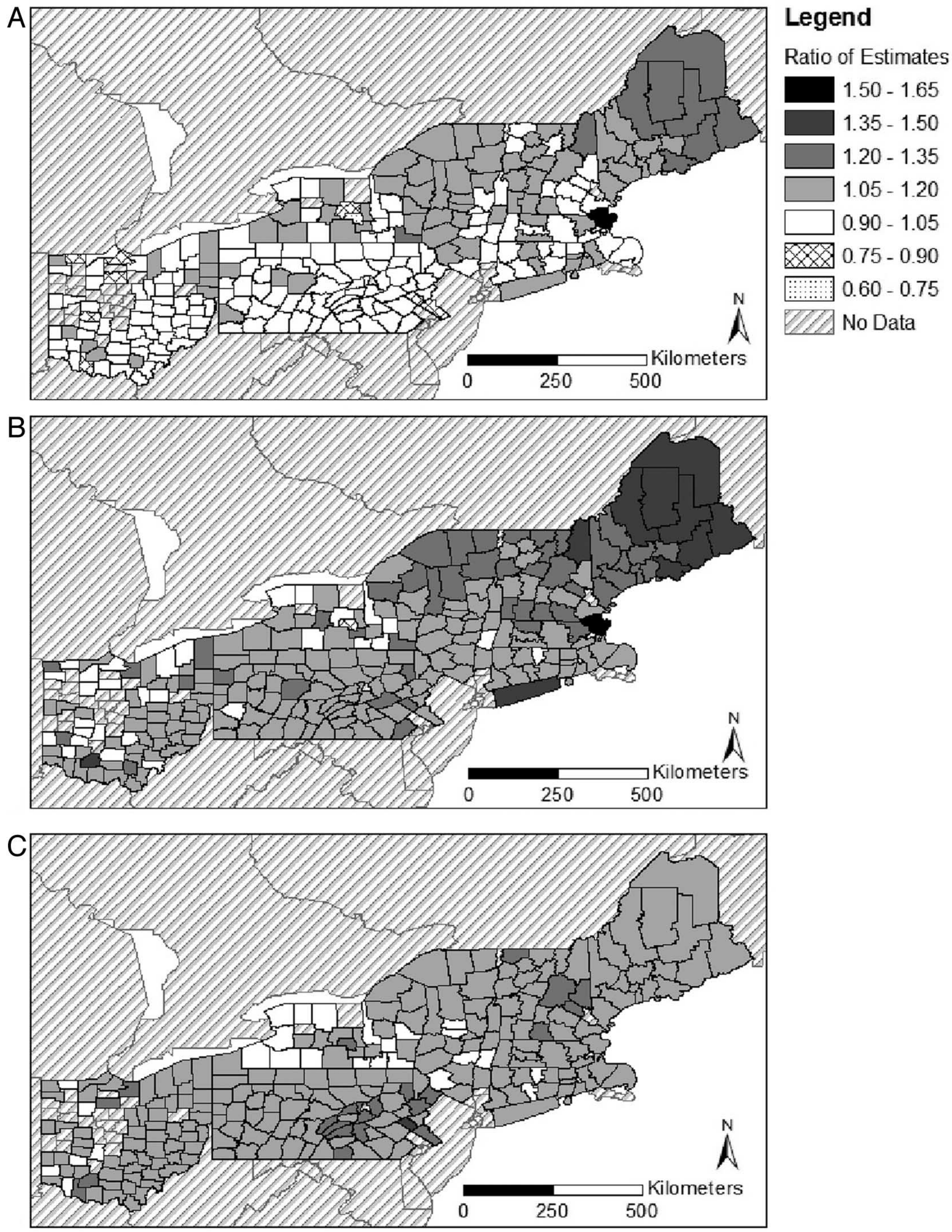

Figure 2. County-level means of the ratios between aboveground live tree carbon stock estimates obtained using (A) FFE and FIA, (B) Jenkins and FIA, and (C) FFE and Jenkins.

growth rate, or 2.972 t/ha/yr. An examination of the raw data suggested, however, that there were numerous plots showing large, unexplained declines in observed live tree carbon. Because plots with a history of harvest had been screened from the data, we supposed that these declines might be due to disturbance that is not possible for FVS to predict. Moreover, some plots showed substantial carbon increases due to ingrowth, and the Northeast variant of FVS does not model regeneration (except from postharvest sprouting, and harvested plots were excluded in this study); many of these plots had no modeled net carbon accumulation. Therefore, we reanalyzed the data, retaining only those plots (n = 3,278) that had positive observed and modeled net carbon accumulation, expecting that this would improve metrics of model performance. The actual results were worse. Observed carbon accumulation within the restricted set of plots was 1.974 ± 0.022 t/ha/yr, while modeled carbon accumulation remained low at 0.0878 ± 0.009 t/ha/yr. Correlation between observed and modeled carbon accumulation was barely affected (Spearman's $\rho = 0.18$, $P < 0.0001$). The elimination of plots with unusual carbon trajectories did improve model performance as measured using RMSE, but because of the increase in bias, the RMSE of modeled carbon accumulation was actually greater than that which would be obtained by simply using the regional mean

Table 2. Importance measures for Random Forest prediction of the ratios between aboveground live tree carbon stock estimates obtained using three different methods. Variables are listed in order of importance.

FFE:FIA			Jenkins:FIA			FFE:Jenkins		
Variable	Increase MSE[1]	Increase purity	Variable	Increase MSE	Increase purity	Variable	Increase MSE	Increase purity
Top ht (m)	36.1	95.8	Top ht (m)	37.4	63.7	CCF	84.8	8.1
Latitude	25.9	42.6	Latitude	28.1	51.8	Top ht (m)	71.7	7.7
Longitude	24.6	46.5	Longitude	23.4	43.2	Latitude	42.0	4.2
CCF[2]	24.1	38.1	Trees/ha	17.5	26.0	Age	41.3	5.1
QMD[2]	18.8	57.9	QMD	17.1	42.5	SDI	39.5	4.1
Trees/ha	18.6	30.9	Basal area	14.5	33.7	Longitude	38.8	4.0
Elevation	15.8	16.2	CCF	13.5	26.0	Basal area	37.3	4.7
Basal area	15.5	36.6	SDI	12.2	25.2	QMD	29.3	3.5
Age	13.0	42.5	Elevation	10.8	13.2	Trees/ha	27.9	2.8
SDI[2]	12.7	25.4	Age	7.6	23.6	Elevation	26.3	3.1

[1]Increase MSE = increase in mean squared error following permutation of the variable; Increase Purity = averaged change in residual sum of squares due to splitting on the variable.
[2]CCF = crown competition factor; QMD = quadratic mean diameter (cm); SDI = Reineke's Stand Density Index.

(1.658 versus 1.253 t/ha/yr). We conducted identical analyses using the carbon estimates produced by the Jenkins method and while the results were not numerically identical, the general outcomes and practical implications were.

Discussion and Conclusions

Carbon stocks calculated using the FFE, Jenkins, and FIA methods in the northeast United States are not equivalent (unless one is willing to tolerate mean differences approaching or exceeding 20%), and not simply convertible. This has implications for carbon markets, comparability of scientific studies, and policy analyses. Researchers and managers conducting forest carbon stock assessments need to be aware of the differences between the methods when selecting an approach, should clearly state which method is in use, and must either use that method consistently through time or recalculate prior estimates if a change in approach occurs and document the changes.

The same base data result in considerable variation in carbon stock estimates at the plot and regional scales, with strong spatial patterning of the differences. These results are not unusual; several investigators have found substantial differences in carbon stock estimates produced by various methods. Guo et al. (2010) compared carbon stock estimates for Chinese forests by three approaches: mean biomass, mean ratio (using a constant biomass expansion factor), and continuous biomass (biomass expansion factors based on volume), and found that the mean biomass approach produced carbon stock estimates for young stands that were 300% higher than those from the continuous biomass approach, while stocks in older stands were underestimated. The reverse was true for the mean ratio method relative to the continuous biomass method, though the difference were smaller. Net increment was similar for all three approaches. Petersson et al. (2012) compared carbon stock estimates calculated using individual tree biomass equations and volume to carbon conversions (the biomass expansion factor approach) for Swedish forests and found that estimates from the volume approach were about 30% higher than those calculated using biomass equations, the opposite result from this study. Westfall (2012) conducted a comparison for trees in the northeast United States, comparing dry biomass estimates for 15 common tree species using the "old" Northeast FIA biomass equations, the Jenkins et al. (2003) equations, the FIA CRM method, and the Canadian national biomass equations. While tree component results differed, aboveground biomass estimates for the Jenkins and Northeast FIA methods were not significantly different from each other and were higher than estimates from the FIA CRM and Canadian approaches, which were also not significantly different from each other. The differences between estimates varied with DBH. In Sweden, Jalkanen et al. (2005) compared biomass estimates developed from equations applied to tree-level national inventory data to those from age-dependent biomass expansion factors applied to aggregated stand-level volumes. At the national level, estimates from the expansion factor approach were about 7% lower than those developed using biomass equations; although differences varied by species and region. In general, relative differences in biomass estimates were greatest in the intermediate age classes, with better agreement between methods in the younger and older classes. Finally, Domke et al. (2012) investigated the effects of the change in estimation approach from Jenkins to FIA on United States forest carbon stock estimates using 20 common species and report a result similar to ours: for nearly all of the species studied, estimates produced using the FIA method were lower than those from the Jenkins equations. For the northeast United States, carbon stocks of the most common species decreased by 11% when using the FIA method. Domke et al. (2012) also report that at the tree level, the differences in carbon stocks decreased as DBH increased.

Regional scale or project scale carbon analyses with FVS using either the FFE or Jenkins calculation method will not provide FIA equivalents for the Northeast variant because the available methods in FVS are not readily convertible to FIA equivalents. The lack of consistent differences between estimates produced by the different methods suggests that multiple variables may affect the outcomes of each calculation approach, which is reflected in the differences reported by other investigators. Guo et al. (2010) found that age class influenced results strongly in their analysis of biomass estimation in China, while Westfall (2012) and Domke et al. (2012) report that differences in estimates at the individual tree level varied with DBH. Our examination using Random Forest regression indicates that top height had the largest impact on the ratio of the estimates. This suggests that if comparability to FIA is important for national forest managers, and for other users of the FVS tool, additional work is needed to discern the mathematical relationship between the estimates, and to ascertain whether a sufficiently reliable conversion

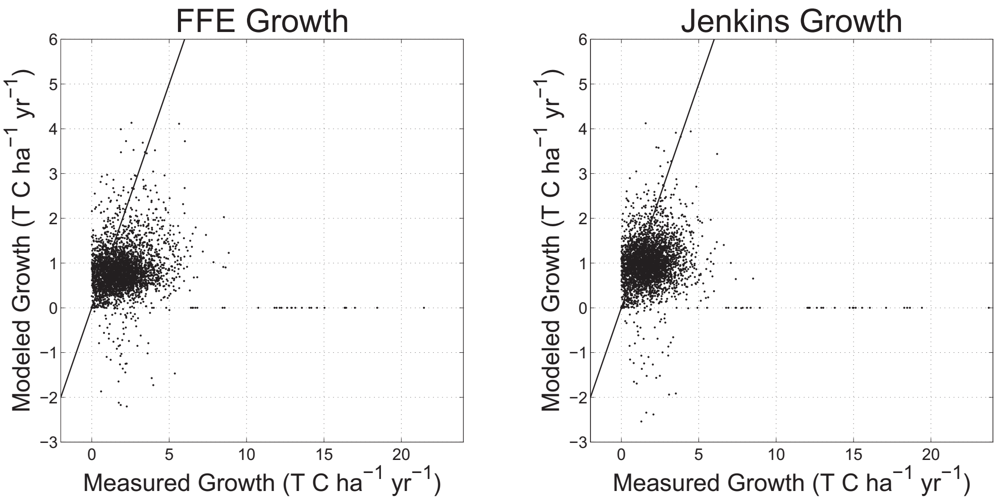

Figure 3. Comparison of observed with predicted net annual growth of carbon (t/ha), where observed is based directly on remeasured FIA plot data, modeled is based on uncalibrated FVS projections, and estimates are calculated using either the FFE or Jenkins methods. Plots with declines in observed carbon stocks are not shown to facilitate comparison; solid line indicates 1:1 relationship.

equation can be developed. Alternatively, it may be necessary to incorporate the FIA CRM approach directly within FVS.

Efforts have begun to develop a set of national biomass equations from destructive harvesting, and when completed, these equations will be the standard by which FIA computes biomass; however, this is a long-term project and results will not be available for some time. Until then, there remains a need to understand the factors that influence each of the calculation approaches, to provide a context in which to interpret results and a basis for developing a method to harmonize the estimates. There are multiple considerations when selecting a calculation approach, and users should not simply select the method that will provide the largest number but should weigh the strengths and weaknesses of each for the region, spatial scale, and forest type of interest. Zhou and Hemstrom (2009) present a comparison of the regional volume, Jenkins, and FIA methods for the Pacific Northwest and include a thoughtful discussion of the tradeoffs between methods, particularly with regard to spatial scale. It is important to note that regardless of the differences between biomass estimates produced by the difference calculation methods, values for net changes will not be affected as long as the same method is consistently applied.

Our results should also serve as a cautionary note against the naive use of FVS for forest growth projection without appropriate attention to calibration. The need for calibration has long been understood by the FVS community (Hamilton 1994, Vandendriesche and Haugen 2008, Ray et al. 2009). In this study, we deliberately avoided local calibration to evaluate model performance when plot data were imported directly using the new FIA2FVS tool. The resulting predictions of net carbon accumulation were no better than, and in some ways worse than, simple use of a regional average. Some of the poor performance may be due to difficulties in modeling gross productivity, some due to the inherent challenge in modeling mortality, and some due to disturbance. Partitioning these effects would be of substantial interest for understanding what calibration factors are most important for improving regional scale prediction, but such analysis is outside the scope of this study. In our view, the biased and weak relationship between observed and predicted net change in C stocks in the absence of calibration is of potential concern. The growing interest in carbon accounting and the emergence of carbon markets has led to new and wider uses of FVS (Hoover and Rebain 2011). While we laud the straightforward access to data provided by tools such as FIA2FVS, ease of access can also facilitate use by those who lack adequate background, training, and familiarity with the modeling tools. Specifically, our results provide a stark warning against the potential practice of downloading local or regional FIA data and then using FVS for forward simulation without calibration to set baselines and expectations for management or for carbon projects. Without local calibration, the projections will be unreliable at plot, project, and regional scales and are no more informative than regional averages.

Endnotes

1. The FIA database is available online at apps.fs.fed.us/fiadb-downloads/datamart.html.
2. The FIA2FVS program is available online at www.fs.fed.us/fmsc/fvs/software/data.shtml.
3. FIA does not make exact plot coordinates publicly available to protect the privacy of the landowner. Publicly available plot coordinates are "fuzzed" by 0.8–1.6 km.

Literature Cited

Bechtold, W.A., and P.L. Patterson, eds. 2005. The enhanced Forest Inventory and Analysis Program—National sampling design and estimation procedures. USDA For. Serv., Gen. Tech. Rep. SRS-80. 85 p.

Berger, R.L., and J.C. Hsu. 1996. Bioequivalence trials, intersection tests and equivalence confidence sets. *Stat. Sci.* 11:283–319.

Breiman, L. 2001. Random forests. *Mach. Learn.* 45:5–32.

California Climate Action Registry. 2008. Urban forest project reporting protocol, version 1.0. California Climate Action Registry, San Francisco, CA. 100 p. Available online at www.climateaction-reserve.org/how/protocols/urban-forest/dev/#version1; last accessed Jan. 28, 2013.

Conover, W.J. 1999. *Practical nonparametric statistics.* Wiley, New York. 584 p.

Crookston, N.L., and G.E. Dixon. 2005. The forest vegetation simulator: A review of its structure, content, and applications. *Comput. Electron. Agr.* 49:60–80.

Dixon, G.E., and C.E. Keyser (comps.). 2008 (revised March 16, 2012). *Northeast (NE) variant overview—Forest Vegetation Simulator.* USDA For. Serv., Forest Management Service Center, Internal Rep., Fort

Collins, CO. 40 p. Available online at www.fs.fed.us/fmsc/fvs/documents/userguides.shtml; last accessed June 4, 2012.

Domke, G.M., C.W. Woodall, J.E. Smith, J.A. Westfall, and R.E. McRoberts. 2012. Consequences of alternative tree-level biomass estimation procedures on US forest carbon stock estimates. *For. Ecol. Manage.* 270:108–116.

Forest Carbon Portal. 2011. Available online at www.forestcarbonportal.com/projects; last accessed May 30, 2012.

Forest Products Laboratory. 1999. *Wood handbook—Wood as an engineering material.* USDA For. Serv., Gen. Tech. Rep. FPL-GTR-113, Forest Products Laboratory, Madison, WI. 463 p. Available online at www.fpl.fs.fed.us/documnts/FPLGTR/fplgtr113/fplgtr113.htm; last accessed May 30, 2012.

Guo, Z., J. Fang, Y. Pan, and R. Birdsey. 2010. Inventory-based estimates of forest biomass carbon stocks in China: A comparison of three methods. *For. Ecol. Manage.* 159:1225–1231.

Hamilton, D.A. Jr. 1994. *Uses and abuses of multipliers in the Stand Prognosis Model.* USDA For. Serv., Gen. Tech. Rep. INT-GTR-310. 9 p.

Heath, L.S., M.H. Hanson, J.E. Smith, W.B. Smith, and P.D. Miles. 2009. Investigation into calculating tree biomass and C in the FIADB using a biomass expansion factor approach in *Forest Inventory and Analysis (FIA) symposium 2008*, McWilliams, W., G. Moisen, and R. Czaplewski (eds.). USDA For. Serv., Proc. RMRS-P-56CD. 26 p.

Hoover, C.M., and S.A. Rebain. 2008. The Kane Experimental Forest carbon inventory: Carbon reporting with FVS. P. 17–22 in *Third Forest Vegetation Simulator conference 2007*, Havis, R.N., and N.L. Crookston (comps.). USDA For. Serv., Proc. RMRS-P-54.

Hoover, C.M., and S.A. Rebain. 2011. Forest carbon estimation using the Forest Vegetation Simulator: Seven things you need to know. USDA For. Serv., Gen. Tech. Rep. NRS-77. 16 p.

Imai, N., H. Samejima, A. Langner, R.C. Ong, S. Kita, J. Titin, A.Y.C. Chung, P. Lagan, Y.F. Lee, and K. Kitayama. 2009. Co-benefits of sustainable forest management in biodiversity conservation and carbon sequestration. *PLoS ONE* 4(12):e8267.

Jalkanen, A., R. Mäkipää, G. Ståhl, A. Lehtonen, and H. Petersson. 2005. Estimation of the biomass stock of trees in Sweden: Comparison of biomass equations and age-dependent biomass expansion factors. *Ann. For. Sci.* 62:845–851.

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.

Liaw, A., and M. Wiener. 2002. Classification and regression by randomForest. *R. News* 2(3):18–22.

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. USDA For. Serv., Res. Note NRS-38. 35 p.

Miles, P.D., and A.D. Hill. 2010. Volume equations for the Northern Research Station's Forest Inventory and Analysis program as of 2010. USDA For. Serv., Gen. Tech. Rep. NRS-74. 50 p.

Olander, L., and K. Haugen-Kozyra. 2011. Using biogeochemical process models to quantify greenhouse gas mitigation from agricultural management projects. Nicholas Institute for Environmental Policy Solutions, Rep. NI-R-11–03, Duke University, Durham, NC. 34 p.

Petersson, H., S. Holm, G. Stahl, D. Aldger, J. Fridman, A. Lehtonen, A. Lundström, and R. Mäkipää. 2012. Individual tree biomass equations or biomass expansion factors for assessment of carbon stock changes in living biomass–A comparative study. *For. Ecol. Manage.* 270:78–84.

R Development Core Team. 2010. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

Ray, D.G., M.R. Saunders, and R.S. Seymour. 2009. Recent changes to the northeast variant of the Forest Vegetation Simulator and some basic strategies for improving model outputs. *North. J. Appl. For.* 26:31–34.

Rebain, S.A. (comp.). 2010 (revised Mar. 20, 2012). *The Fire and Fuels Extension to the Forest Vegetation Simulator: Updated model documentation.* USDA For. Serv., Forest Management Service Center, Internal Rep., Fort Collins, CO. 397 p. Available online at www.fs.fed.us/fmsc/fvs/documents/userguides.shtml; last accessed June 1, 2012.

Reineke, L.H. 1933. Perfecting a stand density index for even-aged forests. *J. Agr. Res.* 46:627–638.

Robinson, A.P., and R.E. Froese. 2004. Model validation using equivalence tests. *Ecol. Model.* 176:349–358.

Robinson, A.P., R.A. Duursma, and J.D. Marshall. 2005. A regression-based equivalence test for model validation: Shifting the burden of proof. *Tree Physiol.* 25:903–913.

Strobl, C., A.-L. Boulesteix, A. Zeileis, and T. Hothorn. 2007. Bias in randomForest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics* 8:25, doi: 10.1186/1471-2105-8-25.

UN REDD Programme. 2012. Available online at www.un-redd.org/; last accessed May 29, 2012.

USDA Forest Service. 2011a. *National roadmap for responding to climate change*, FS-957-b. Available online at www.fs.fed.us/climatechange/pdf/Roadmapfinal.pdf; last accessed May 30, 2012.

USDA Forest Service. 2011b. *Navigating the climate change performance scorecard: A guide for national forests and grasslands* (Version 2, August 2011). Available online at www.fs.fed.us/climatechange/advisor/scorecard/scorecard-guidance-08-2011.pdf; last accessed June 4, 2012.

US Environmental Protection Agency. 2011. *Inventory of US greenhouse gas emissions and sinks: 1990–2010.* EPA 430-R-12–001. Available online at www.epa.gov/climatechange/emissions/usinventoryreport.html; last accessed Dec. 10, 2012.

Vandendriesche, D.A. 2012. Advanced FVS tools for landscape planning. USDA For. Serv., Forest Management Service Center, Interoffice Publ., Fort Collins, CO. 197 p.

Vandendriesche, D., and L. Haugen. 2008. Comparison of FVS projection of oak decline on the Mark Twain National Forest to actual growth and mortality as measured over three FIA inventory cycles. P. 68–80 in *Third Forest Vegetation Simulator conference, February 13–15, 2007*, Havis, R.N., and N.L. Crookston (comps.). Proc. RMRS-P-54, Fort Collins, CO.

Westfall, J.A. 2012. A comparison of above-ground dry-biomass estimators for trees in the Northeastern United States. *North. J. Appl. For.* 29(1):26–34.

Wickham, H. 2009. *ggplot2: Elegant graphics for data analysis.* Springer, New York.

Woodall, C.W., L.S. Heath, G.M. Domke, M. Nichols, and C. Oswalt. 2011. *Methods and equations for estimating volume, biomass, and C for forest trees in the US national inventory, 2010.* USDA For. Serv., Gen Tech. Rep. NRS-88. 30 p.

Wykoff, W.R., N.L. Crookston, and A.R. Stage. 1982. *User's guide to the Stand Prognosis Model.* USDA For. Serv., Gen. Tech. Rep. INT-133. 112 p.

Zhou, Z., and M.A. Hemstrom. 2009. *Estimating aboveground tree biomass on forest land in the Pacific Northwest: A comparison of approaches.* USDA For. Serv., Res. Pap. PNW-RP-584. 18 p.