Estimating Current Forest Attributes from Paneled Inventory Data Using Plot-Level Imputation: A Study from the Pacific Northwest

Bianca N.I. Eskelson, Hailemariam Temesgen, and Tara M. Barrett

Abstract: Information on current forest condition is essential to assess and characterize resources and to support resource management and policy decisions. The 1998 Farm Bill mandates the US Forest Service to conduct annual inventories to provide annual updates of each state's forest. In annual inventories, the sample size of 1 year (panel) is only a portion of the full sample and therefore the precision of the estimations for any given year is low. To achieve higher precision, the Forest Inventory and Analysis program uses a moving average (MA), which combines the data of multiple panels, as default estimator. The MA can result in biased estimates of current conditions and alternative methods are sought. Alternatives to MA have not yet been explored in the Pacific Northwest. Data from Oregon and Washington national forests were used to examine a weighted moving average (WMA) and three imputation approaches: most similar neighbor, gradient nearest neighbor, and randomForest (RF). Using the most recent measurements of the variables of interest as ancillary variables, RF provided almost unbiased estimates that were comparable to those of the MA and WMA estimators in terms of root mean square error. FOR. SCI. 55(1):64-71.

Keywords: moving average, weighted moving average, nearest neighbor imputation, missing panels

INITIATED BY THE Agricultural Research, Extension, and Education Reform Act of 1998 (PL 105-185), the Forest Inventory and Analysis (FIA) program of the US Forest Service has switched from periodic inventories that varied from state to state to a consistent nationwide annual inventory. A portion of the inventory of the nation's forests is now conducted annually within each state. The fraction of the plots measured annually is 10% in the western United States and 20% in the eastern United States.

The precision of the estimates of current status and changes in the forest resources using only data from the panel of plots measured in the current year has been found to be unacceptable because of the small annual sample size (Mckoberts and Hansen 1999). There have been efforts to combine data of multiple panels to achieve a higher precision. The current FIA default estimator is a moving average (MA), which is operationally convenient and requires few assumptions (Gartner and Reams 2001). The MA approach can improve the precision of the estimates by using data from the panels measured in the most recent years. However, MA reflects an average of conditions over the past 10 years rather than current forest conditions, resulting in a bias of the current year's population parameter (McRoberts 2000, Johnson et al. 2003). The MA estimates can be improved with a weighted moving average (WMA), which weighs panels that were measured more recently more heavily than those measured earlier (Roesch and Reams 1999). Other approaches to combine data from all panels include updating unmeasured panel data to the current year using growth models (Lessard et al. 2001, McRoberts 2001), time series models (Johnson et al. 2003) or mixed estimation (Van Deusen 1996, 1999, 2002, Scott et al. 1999); filling in missing panel data using tree- and plot-level imputation techniques (Gartner and Reams 2001, 2002, McRoberts 2001); or modifying the annual inventory of interpenetrating, nonoverlapping panels to an inventory system with balanced annual partial remeasurements so that estimators based on sampling with partial replacement can be used (Scott et al. 1999, Arner et al. 2(04).

There is a need to develop new methods that will be included in the annual inventory system according to their performance (Reams et al. 1999). Because spatial, temporal, and forest characteristics differ within and among regions. it is unclear whether any single technique will work for all regions (Patterson and Reams 2005), and it is necessary to evaluate different methods in all regions. Studies comparing different alternatives to the MA approach for estimating current forest attributes in the Pacific Northwest (PNW) are still lacking, whereas a variety of methods have been tested in the other regions of the United States (Van Deusen 1996, 1997, 1999, 2002, Lessard et al. 2001, McRoberts 2001, Arner et al. 2004).

The imputation and modeling approaches examined by McRoberts (2001) asserted that model development requires a greater resource investment than development of an imputation procedure. As the difference in the estimation results was negligible, it is reasonable to focus on investigating and improving the imputation techniques. McRoberts

Bianca N.I. Eskelson, Oregon State University Department of Forest Engineering, Resources and Management, 204 Peavy Hall, Corvallis, OR 97331-5706-Phone: (541) 737-4457; Fax: (54 I) 737-4316; bianca.eskelson@oregonstate.edu Hailemariam Temesgen, Oregon State University Department of Forest Engineering, Resources and Management. Corvallis, Oregon---hailemariam.temesgen@oregonstate.edu. Tara M. Barrett, US Forest Service, Pacific Northwest Research Station Forestry Sciences Laboratory, Anchorage, Alaska-tbarrett@fs.fed.us.

Acknowledgments: We thank Janet Ohmann and Matt Gregory for sharing their ancillary data from the National Land Cover Database and DAYMET with us and for their insights on the data. We appreciate the help that Jim Alegria, Carol Apple, Bob Brown, and Melinda Moeur provided in obtaining national forest data and thank Kurt Campbell for assistance with volume and biomass equations. We also thank David Hann, three anonymous reviewers, and the associate editor for their helpful review comments.

Manuscript received April 23, 2008, accepted October 3, 2008