Forest Ecology and Management 315 (2014) 112–120

Contents lists available at ScienceDirect

Forest Ecology and Management

journal homepage: www.elsevier.com/locate/foreco

Strategies to compensate for the effects of nonresponse on forest carbon baseline estimates from the national forest inventory of the United States

Grant M. Domke *, Christopher W. Woodall, Brian F. Walters, Ronald E. McRoberts, Mark A. Hatfield

USDA Forest Service, Northern Research Station, 1992 Folwell Ave., St. Paul, MN 55108, USA

ARTICLE INFO

Article history:
Received 6 September 2013
Received in revised form 19 December 2013
Accepted 23 December 2013
Available online 18 January 2014

Keywords:
Imputation
Missing data
Greenhouse gas reporting
Forest Inventory and Analysis (FIA)

ABSTRACT

Forest ecosystem carbon (C) stocks and stock change in the United States (US) have been documented using Intergovernmental Panel on Climate Change (IPCC) procedures and guidance with 1990 as a baseline reference for all United Nations Framework Convention on Climate Change reports. In the US, estimates of forest C stocks and stock change are obtained from data collected and maintained by the Forest Inventory and Analysis (FIA) program of the US Forest Service. Over the course of the IPCC monitoring period, the FIA program made a transition from state-by-state multiyear periodic inventories selected on a rotating basis – with reporting standards largely tailored to regional requirements – to nationally consistent, annual inventories (where a proportion of plots is measured in each state each year) designed for large-scale strategic requirements. Lack of measurements on all forest land during the periodic inventories, along with plot access difficulties and misidentification of forest plots as nonforest due to poor aerial imagery, have resulted in missing data (i.e., nonresponse) throughout the FIA database. Nonresponse, which in some US states is greater than 20%, may lead to differences in estimates of forest C stock change due to the procedural transition from periodic to annual inventories. As an initial step towards rectifying the differences in estimates, we examined several strategies to compensate for missing observations using the most recent annual inventory data from the Lake States region of the US. Results varied by state in the study but given the annual reporting cycle and requirements to compile national estimates of forest C, it was deemed that techniques, where non-observed samples are removed from estimation procedures, provided the optimal combination of statistical performance and efficiency. While the initial analysis focused on the Lake States region, several compensation strategies described may be useful in bridging the gap between national C flux estimates from periodic and annual forest inventories.

Published by Elsevier B.V.

1. Introduction

As signatories to the United Nations Framework Convention on Climate Change (UNFCCC), the United States (US) has been providing annual estimates of forest ecosystem carbon (C) stocks and stock change (USEPA, 2013) to the UNFCCC in accordance with IPCC Good Practice Guidance (IPCC, 2006). Carbon stocks and stock change are estimated by pool each year across a defined reporting period that runs from 1990 to the present (IPCC, 2006). In the US, C stocks and stock change are estimated from data collected and maintained by the Forest Inventory and Analysis (FIA) program of the US Forest Service, which conducts the national forest inventory (NFI) of the US (USDA Forest Service, 2012). Over the course of the IPCC monitoring period, the FIA program made a transition from state-by-state periodic inventories – with reporting standards largely tailored to regional requirements (Gillespie, 1999) – to nationally consistent, annual inventories designed for large-scale strategic requirements (McRoberts et al., 2010). Lack of measurements on all forest land during the periodic inventories, along with plot access difficulties and misidentification of forest plots as nonforest due to poor aerial imagery, have resulted in missing data throughout the FIA database (McRoberts, 2003; Birdsey, 2004; Patterson et al., 2012; Goeking and Patterson, 2013). These data gaps may contribute to structural differences in estimates of forest C stock change between periodic and annual inventories that are procedural artifacts as opposed to changes in natural resources. Because the US's forest carbon baseline (USEPA, 2013) serves to inform policy as well as carbon science, improving the accuracy and scientific rigor of the baseline is paramount (Woodall, 2012).

* Corresponding author. Tel.: +1 651 649 5138; fax: +1 651 649 5140.
E-mail addresses: gmdomke@fs.fed.us (G.M. Domke), cwoodall@fs.fed.us (C.W. Woodall), bfwalters@fs.fed.us (B.F. Walters), rmcroberts@fs.fed.us (R.E. McRoberts), mahatfield@fs.fed.us (M.A. Hatfield).

0378-1127/$ - see front matter Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.foreco.2013.12.031

Fundamentally, the lack of appropriate data to estimate an annual forest C baseline can be viewed as a missing observation research problem (Van Deusen, 1997; Little and Rubin, 2002). The issue of missing observations (i.e., nonresponse) has been persistent through the course of forest inventories around the world (McRoberts, 2003; Eskelson et al., 2009; Tomppo et al., 2010; Beets et al., 2011; Barrett and Maltamo, 2012; Patterson et al., 2012; Goeking and Patterson, 2013) whereas forest C baselines suffer from the same dilemma: how should missing observations be accommodated when using large scale forest inventories to estimate forest resources whether C or sawtimber volume? The extent to which nonresponse is a serious problem depends to some degree on the underlying reasons for nonresponse. In the US, the most common reason for nonresponse is that private forest land owners refuse field crews access to their land (Patterson et al., 2012). However, for most countries, individual property rights are not nearly as strong as in the US. Thus, contrary to the situation in the US, NFI field crews in only very few countries must obtain permission of landowners to access plots on private forestland. For these countries, two results follow; first, the reasons for nonresponse are limited to natural causes such as floods and hazardous terrain, and second, the proportion of nonresponses is so small that the problem is generally ignored. In the southern hemisphere, for example, NFIs are sufficiently rare and of such recent initiation that the degree to which nonresponse is a problem has not been rigorously evaluated.

As a surrogate for examining techniques to mitigate nonresponse by year for C baselines, data that exist from forest inventories collected over multiple years may be used to inform annual C baseline strategies. Within annual forest inventories, missing observations are optimally estimated with data of interest for all sample units that were observed. Unfortunately, nonresponse is inevitable in most large inventories which has led to a well-defined order of operations for determining strategies to compensate for missing observations (Sande, 1982; Lemeshow, 1985; Rubin, 1987; Särndal et al., 1992; Lesser, 2001; Little and Rubin, 2002; McRoberts, 2003; Eskelson et al., 2009; Patterson et al., 2012). In the annual FIA inventory, nonresponse is largely the result of denied access on private forest lands and, to a lesser extent, hazardous areas (McRoberts, 2003; Patterson et al., 2012; Goeking and Patterson, 2013). On private forest lands, FIA field crews typically make a single attempt to obtain permission to gain access to the plot via letter, phone call, or location visit. Landowners may deny the field crews access to measure plots on their private forest land and in those cases plots are listed as denied access. Plot locations on forest land deemed hazardous (e.g., flooded) may be revisited later in the field season when conditions allow access but this is not always possible. Once an attempt has been made and the plot or portion of the plot (i.e., condition – area classification mapped on each plot using discrete variables such as land use, forest type, and/or ownership group to enable division of forest into various domains of interest, USDA Forest Service, 2013) – hereafter collectively referred to as plot – cannot be observed and measured, it is recorded as nonsampled and given a reason code in the FIA database (USDA Forest Service, 2013). Identifying plausible strategies to compensate for nonresponse first requires quantification of nonresponse, followed by an assessment of the properties of the nonresponse elements and an understanding of the nonresponse mechanisms (Little and Rubin, 2002). Because missing data is an issue for both annual forest inventories and UNFCCC forest C baselines, examining potential strategies for nonresponse compensation is warranted.

We examined several approaches that compensate for missing observations with respect to the accuracy and precision of estimates of C stocks per unit area using data from the FIA annual inventory in the Lake States region (Michigan, Minnesota, and Wisconsin) of the US (Fig. 1). The specific objectives of the study were to: (1) quantify nonresponse in the annual FIA data, (2) describe the properties of the missing observations, (3) identify strategies to compensate for nonresponse, (4) describe the process of incorporating compensation strategies into the FIA sampling framework, (5) assess each compensation strategy under increasing levels of simulated nonresponse, and (6) describe how the selected strategies may be employed to compensate for missing observations in the periodic inventories dating back to the 1990 C baseline used in greenhouse gas reporting (e.g., UNFCCC).

2. Methods

Because this study is an initial step toward rectifying the differences in forest C stock change due to the procedural transition from periodic to annual inventories, we examined several well-established strategies to compensate for missing observations using the most recent annual inventory data. We chose to restrict our initial assessment to the annual inventory because it includes a nationally consistent sampling frame and plot design so the methodologies established for compensating for missing observations could be applied nationally without substantial modification. Furthermore, causes of nonresponse were thought to be more consistent across regions of the US in the annual inventory, thus providing a parsimonious approach to identifying potential nonresponse patterns and selection of strategies to compensate for nonresponse in the data.

2.1. Data

The FIA program employs a three phase inventory, with each phase contributing to the subsequent phase. Phase 1 is a variance reduction step where satellite imagery is used to assign Phase 2 plots to strata (Bechtold and Patterson, 2005). A stratum is a defined geographic area (e.g., state or estimation unit) that includes plots with similar attributes; in the Lake States region, strata are defined by predicted percent canopy cover. Data in this study came from Phase 2 plots measured in each of the two most recent annual inventory cycles (2002–2006 and 2007–2011) in the Lake States region of the US. Phase 2 plots are distributed approximately every 2428 hectares across the 48 conterminous states of the US (Fig. 1). Each Phase 2 permanent ground plot comprises a series of smaller plots (i.e., subplots) where tree- and site-level attributes – such as diameter at breast height (dbh) and tree height – are measured at regular temporal intervals (USDA Forest Service, 2013).

In the Lake States region, gross tree volume is estimated using a model with tree dbh, site index (as a proxy for tree height), and basal area as explanatory variables (Woodall et al., 2011). Gross volume estimates are adjusted to account for volume loss due to rotten and missing cull defect and the sound volume estimates are converted to oven-dry biomass using the component ratio method (Heath et al., 2009; Woodall et al., 2011). Tree biomass is then multiplied by 0.5 to convert to C and live tree-level C estimates are summed within the plot and then converted to a per unit area basis. The plot-level estimate obtained as the sum of tree-level estimates is assumed to be an observation without error (McRoberts and Westfall, 2014).

2.2. Stratification

Because the precision standards (USDA Forest Service, 1970) established by the FIA program may not be satisfied for estimates of some parameters, the estimation process is enhanced through stratification. Stratification is used to reduce the variance of estimates of parameters such as C stocks, by partitioning the popula-