Forest Ecology and Management 262 (2011) 188–194

Contents lists available at ScienceDirect

Forest Ecology and Management

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

Comparison of Forest Inventory and Analysis surveys, basal area models, and fitting methods for the aspen forest type in Minnesota

John M. Zobel*, Alan R. Ek, Thomas E. Burk

Department of Forest Resources, University of Minnesota, 115 Green Hall, 1530 Cleveland Ave. N., St. Paul, MN 55108-6112, USA

ARTICLE INFO

Article history:
Received 22 December 2010
Received in revised form 7 March 2011
Accepted 16 March 2011
Available online 13 April 2011

Keywords:
Basal area
Forest Inventory and Analysis (FIA)
Dataset comparison
Yield model
Fitting method
Populus tremuloides Michx

ABSTRACT

The Forest Inventory and Analysis (FIA) unit of the U.S. Forest Service has collected, compiled, and made available plot data from three measurement periods (identified as 1977, 1990, and 2003, respectively) within Minnesota. Yet little if any research has compared the relative utility of these datasets for developing empirical yield models. This paper compares these and other subdatasets in the context of fitting a basal area (B) yield model to plot data from the aspen (*Populus tremuloides* Michx.) forest type. In addition, several models and fitting methods are compared for their applicability and stability over time. Results suggest that the three parent datasets, along with their subdatasets, provide very similar three parameter B yield model prediction capability, but as model complexity increases, variability in coefficient estimates increases between datasets. The absence of data for older aspen stands and the inherent noise within B data prevented the exact determination of an overall best model. However, the model $B = b_1 S^{b_2}(1 - \exp(-b_3 A))$ with site index (S) and stand age (A) as predictors was found consistently among the highest in precision and stability. Additionally, nonlinear least squares and nonlinear mixed-effects fitting procedures produced similar model fits, but the latter is preferred for its potential to improve model projections. The results indicate little practical difference between datasets from different time periods and different sizes when used for fitting the models. Additionally, these results will likely extend to other states or regions with similar remeasurement data on aspen and other forest types, thus facilitating the development of other ecological models focused on forest management.

1. Introduction

For many years the U.S. Forest Service has collected inventory data on Minnesota's forests via the Forest Inventory and Analysis (FIA) program. This data has been made available online and includes the last three completed surveys, corresponding to the years 1977, 1990, and 2003, respectively. Each survey required 4–5 years of field work, with the survey date indicating the year of completion. Through the years, the various datasets have provided researchers with a source of forestry data representative of all Minnesota. In particular, Walters and Ek (1993) utilized the 1977 survey data to develop a system of linked yield models for basal area, density, and merchantable volume, among others, for 14 forest types in Minnesota.

As more years of data become available, questions arise as to the utility, similarity, and compatibility of the datasets for yield model building. For example, the sampling methodology was revised before each subsequent survey in an attempt to improve data quality, usefulness, and efficiency. In addition to methodological changes, possible physical differences between the three surveys include the size, representative quality, and scope of each dataset; the inherent weather during the years prior to measurement; and the stand treatments conducted since the last survey. Typical users often utilize only the most recent available data, for obvious reasons. However, in the context of model building, the methodological changes should theoretically have little effect on parameter estimates. Still, the physical differences may produce varying model fits. Therefore, this paper asks three questions related to the FIA data available for Minnesota: (1) which dataset (or subdataset or combination of datasets), if any, has the most utility/reliability for building yield models, specifically an aspen (*Populus tremuloides* Michx.) basal area (B) yield model; (2) which B model fits the best with respect to the available data (including all three surveys collectively) and best extrapolates per the usual assumptions; and (3) which fitting method yields the most credible parameter estimates across measurement periods. The answers to these questions will likely have relevance for other states or regions with similar remeasurement data on aspen and other forest types, thus facilitating the development of other ecological models focused on forest management.

Criteria for determining the optimal dataset, model form, and fitting method involves comparing coefficient estimates, fit statis-

* Corresponding author. Tel.: +1 612 624 4280; fax: +1 612 625 5212.
E-mail address: zobe0025@umn.edu (J.M. Zobel).

0378-1127/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.foreco.2011.03.022

Table 1
Aspen datasets considered in this study and their sample sizes.

Constraint[a]	Survey completion year			Total[b]
	1977	1990	2003	
CON1	378	378	378	1134
CON2	585	585	585	1755
CON3	652	706	653	n/a
CON4	3417	4410	1564	n/a

[a] CON1, remeasured plots that remain aspen across all periods; CON2, remeasured plots that start aspen, but may have changed type over time; CON3, aspen plots determined independently within each survey from among remeasured plots; CON4, aspen plots determined independently within each survey from all plots.
[b] Total, all three surveys combined.

tics (root mean square error (RMSE) and R^2), plotted curves, and/or residual plots from the following: (1) the same model fit to each dataset; (2) each model fit using the same dataset; and (3) the same model derived via each fitting method and the same dataset. Determining the best model form also includes evaluating the theoretical properties (both statistical and ecological) of each form.

2. Methods

2.1. Data

The parent datasets compared in this study were all obtained from the FIA online database (see http://199.128.173.17/fiadb4-downloads/datamart.html). Due to constant improvements in the sampling scheme and information collected, several differences exist between the 1977/1990 and 2003 surveys. For example, the first two surveys used a variable radius, 10-point cluster plot design, sampled on a periodic basis (each decade) (Leatherberry et al., 1995). The survey completed in 2003 used a fixed radius, 4-subplot design, with 20% of the plots in the State measured per year, with all the plots being measured in five years (FIA, 2008). Other changes in 2003 included the further breakdown of plots into conditions and updates to the forest type determination algorithm to include more types. However, across all three surveys, the plot layout encompassed roughly the same area, approximately one acre (see LaBau et al. (2007) for graphical representations of both designs). Many plots have been revisited during all three measurement periods, but many have also been either retired or introduced as new plots in the subsequent surveys. In 1990, numerous undisturbed plots were not actually revisited, but characterized via projection with the STEMS growth model (Belcher et al., 1982; Leatherberry et al., 1995).

At the time of this study, the online database included all inventory information for Minnesota dating back to the 1977 periodic inventory through the annual measurements in 2007. The database contained three completed surveys (i.e. a complete measurement of all FIA plots across Minnesota) and 80% of the most recent survey (completed in 2008, but not yet fully processed). This study focused on the three completed surveys and subsets of these datasets (subdatasets) as described below (CON1–CON4).

2.1.1. CON1

Through considerable assistance from the U.S. Forest Service Northern Research Station's FIA staff, 5141 plots were identified that had been measured in all three surveys (excluding projected plots but including nonforest plots), of which 378 maintained an aspen forest type across periods (see Table 1). This dataset is referred to as CON1 ("constraint one").

However, even with the availability of the remeasurement data in CON1, B yield instead of B growth was modeled for several reasons. First, FIA remeasurement data represents a broad range of stand conditions, from undisturbed to heavily managed, and differs considerably from research plot data. Additionally, the majority of FIA plots show signs of some interruption in the usual growth patterns. Thus most plots have modest value for modeling growth, but are useful for describing yield. Second, the aspen forest type has by far the most observations (plots) in every FIA survey in Minnesota. However, extending this research to the other forest types considered in Walters and Ek (1993) will encounter much smaller sample sizes, often too sparse to create precise yield (or growth) models without grouping data from various measurement periods. Finally, yield measurements/estimates often provide starting conditions for growth models, and thus developing yield models will aid the implementation of growth models if and when they become available.

2.1.2. CON2–CON4

We defined additional datasets based on alternative constraints: CON2 – the set of remeasured plots that start as aspen in 1977, but were allowed to maintain or change their forest type in one or both of the next two surveys; CON3 – the set of aspen plots that were determined independently within each survey from among the collection of *remeasured* plots; CON4 – the set of aspen plots that were determined independently within each survey from among the collection of *all* plots. In other words, we followed individual plots through time in CON1 and CON2, with each constraint defining a dataset with the same number of plots in each survey year (see Table 1). For CON3 and CON4, we applied a cross sectional approach, with each constraint defining a unique (and different) number of plots in each survey year.

Table 1 provides all sample sizes and Table 2 gives basic descriptions of the datasets under the first and last constraints (the summary statistics for CON2 and CON3 roughly resemble those from CON1). Note that without disturbance, the sequential survey datasets under CON1 should have mean ages roughly 10 years apart, a fairly constant mean site index, and increasing mean basal area. However, the summary statistics in Table 2 do not follow this pattern, suggesting the influence of management actions, natural disturbance, sampling distribution changes, or even measurement inconsistencies on the within plot values (e.g. annual rings in aspen trees are notoriously difficult to read, thus making age determination problematic). Note this observed influence relates to the first question of this study regarding the utility of the different datasets.

In an attempt to provide adequate comparison of the FIA surveys, subdatasets were defined as every combination of dataset (CON1–CON4) and measurement period (1977, 1990, and 2003), in addition to the CON1 and CON2 totals (i.e. all years combined), providing a total of 14 datasets for examination. These datasets were compared for their ability to produce similar B model performance (i.e. in terms of parameter values, fit statistics, and predictions). In addition, the diversity among the datasets allows us to investigate B model stability over time and under various conditions. Model stability (or lack of it) is important for model usage, since this characteristic embodies sampling error and changes in the environmental factors that influence growth and subsequently yield. However, this stability may only tell a partial story. Other influences on B yield, such as the age class distribution of plots, might change considerably. Also, the model used to estimate B may tend to force a stable model form, and thus the strength of conclusions drawn from consistent model fits may be limited.

2.2. Models

Two primary B yield models were considered in this study,

$$B = b_1 S^{b_2} A^{b_3} \quad (1)$$