Testing for Change in Structural Elements of Forest Inventories

Melinda Vokoun, David Wear, and Robert Abt

Abstract: In this article we develop a methodology to test for changes in the underlying relationships between measures of forest productivity (structural elements) and site characteristics, herein referred to as structural changes, using standard forest inventories. Changes in measures of forest growing stock volume and number of trees for both hardwood and softwood on forestland in North Carolina are evaluated using plot-level data aggregated at both the state and survey unit level from the last three available completed Forest Inventory and Analysis surveys using exploratory data analysis and nonparametric statistics. When the survey data are aggregated at the state level, we accept the null hypothesis of no discernible between-survey differences in the means of the forest productivity measures for at least 90% of the plots in each of the four models. We also accept the null hypothesis of no discernible between-survey differences in the variance or higher moments of these forest productivity measures for at least 82% of comparisons. At a finer scale, our results show that structural stability is questionable in the Coastal Plain units of North Carolina. Overall, results provide evidence of some structural change in the forests of North Carolina but do not address the causes of such changes. The systematic comparison of forest inventories conducted in this article constitutes a new approach to testing for structural changes in forest relationships, one that can be implemented as a monitoring protocol within standard repeated forest inventories. FOR. SCI. 55(5):455–466.

Keywords: regression tree analysis, Forest Inventory and Analysis, inventory comparison, hypothesis testing

A PRINCIPLE OBJECTIVE of broad scale forest assessments is to evaluate and anticipate changes in forest conditions over time. An especially important element of such an evaluation is the ability to distinguish between changes that are the result of the orderly development of forests, e.g., aging, natural and management disturbances, and natural succession, and those that might arise from structural changes in forest production relationships or conditions (Loehle and LeBlanc 1996, Franklin et al. 2002). In particular, productivity measures of forest stock such as standing volume and number of trees per hectare can be viewed as structural elements of the forest, and changes in these structural elements or structural changes, conditioned on underlying landscape characteristics, serve to indicate potential changes in forest and ecosystem processes. Structural changes could arise, e.g., from climate alteration (Dixon et al. 1994) or nitrogen deposition (Magill et al. 2000). They may provide important signals regarding the long-term sustainability of forested ecosystems, possibly indicating changes in the fundamental biogeochemical interactions that underlie ecosystem structure and resilience and important implications for future management options (Franklin et al. 2002).

Assessing the future of forest conditions necessarily depends on examining the structure and dynamics of the ecosystem. An undisturbed ecosystem would naturally age over time with well documented changes in age structure, species mix, and associated understory (Shugart 2003). Disturbance may be incorporated into known dynamics of an ecosystem, e.g., fire in the longleaf pine (*Pinus palustris* M.) type. Other changes may be due to episodic events, anthropomorphic influences, or long-wave cycles that are not necessarily part of observed ecosystem development. An understanding of the nature of forest ecosystem dynamics is important in assessment of likely future conditions. In this article we attempt to empirically identify changes in forests over time and distinguish between those that appear to be consistent with historical patterns and those that may be associated with a structural change in the underlying dynamics.

The objectives of this article were to develop and apply tests for structural changes in forest conditions using plot-based surveys of forests, such as those conducted by the US Forest Service Forest Inventory and Analysis Program (FIA) (Miles et al. 2001). FIA surveys are designed to construct temporally consistent estimates of a variety of forest conditions for substate regions based on measurements of site and tree variables at hundreds of plots within each region. We chose to base our analysis on the FIA data system because it is the only comprehensive survey of forest vegetation conducted in the United States and provides a sample with a large number of observations. The FIA database contains individual state inventories that combine photo-based classifications of land use and ground plot measures. The ground samples are used to adjust photo-based classifications for changes, provide adjustment to photo-based classifications, and provide estimates of attributes that cannot be determined from remotely acquired

Melinda Vokoun, University of Wisconsin-Stevens Point, College of Natural Resources, 800 Reserve St., Stevens Point, WI 54481—Phone: 715-342-5161; Fax: 715-346-4554; mvokoun@uwsp.edu. David Wear, Project Leader, Forest Economics and Policy Research, Southern Research Station, US Forest Service, Research Triangle Park, NC—dwear@fs.fed.us. Robert Abt, North Carolina State University—bob_abt@ncsu.edu.

Acknowledgments: We thank William Smith, Robert Huggett, Ray Sheffield, and Pat Miles for assistance with FIA data and associated algorithms. We also thank the associate editor and three anonymous reviewers for their thoughtful comments on this manuscript. The standard qualifications apply. This study was funded through a cooperative research agreement with the US Forest Service Southern Research Station Forest Economics and Policy Work Unit in Research Triangle Park, NC.

Manuscript received November 20, 2006, accepted June 29, 2009

means (Miles et al. 2001). Because the FIA is a comprehensive and relatively stable program, tests for structural change that can be implemented within the FIA framework could provide a general approach for periodically evaluating changes in the forests of the United States. Because similar forest survey systems have been used in many other countries, these tests may have broader application.

Methods to test for structural change can provide a very useful adjunct to the monitoring role of forest inventories. Structural change analysis may indeed support important meta-scale sustainability monitoring and forest health detection within the standard reporting framework. Current inventory reports focus on trends without tests for significant change in biophysical relationships between surveys (e.g., Conner and Hartsell 2002).

The general premise of our approach is that although forests are highly diverse, complex, and dynamic terrestrial ecosystems, index or structural element variables can be used to characterize their important attributes. For example, volume per hectare may serve as a useful indicator of forest productivity. Here, we were not interested so much in whether forest density measures have changed over time as such changes are often explained by the simple aging of forests. Rather, our objective was to investigate whether the distribution of forest density measures associated with different forest conditions, including age, has changed. That is, we ask whether the distribution of density measures, such as volume per hectare, conditional on-site characteristics, has changed.

To develop our tests for structural change we distinguish between two types of variables in the FIA database. FIA ground plots have been located randomly within a 6,000-acre hexagon (Bechtold and Scott 2005). Before 2000 these ground plots were inventoried periodically using a variable-radius layout, whereas more recent inventories used a standardized fixed-radius plot layout for sample tree selection (Miles et al. 2001). The ground plots provide observations on individual trees that enable predictions of tree's volume, growth rate, and other qualities, which are summarized in the condition and individual tree records within the FIA database. The remote-sensed classification and area estimates are used to determine expansion factors for the ground plots. State variables are those that define the condition of each forest plot and include categorical variables such as physiographic region and broad management class as well as continuous variables such as stand age and site index. The other set of variables includes structural variables that we seek to evaluate for change, e.g., the volume of growing stock of hardwood and softwood trees or the number of trees per hectare; these variables comprise structural elements of the forest landscape. Our tests for structural change ask whether or not the probability distribution of a structural variable, conditioned on the state variables, changes between surveys. This allows us to account for orderly changes in the inventory related to the aging of forests and natural disturbances (i.e., changes in the state variables between surveys) within the context of our tests.

Constructing tests requires that we identify and summarize the key relationships between state variables and structural variables across surveys. These relationships determine the subset of plots or groups of plots (i.e., bins of plots) that are referred to in the computer science literature as information sets. Then, the subsets of plots comprising a modeled relationship between a structural variable and state variables over time are further compared using statistical evaluations. Our null hypothesis is that the bins of plots for the aggregate inventories describe the observed data from an individual survey; that is, the plots for each survey are equivalent within bins, or the distribution of each structural variable modeled over a landscape does not change between repeated surveys. Our alternative hypothesis is that the relationships between state and structural variables for an individual survey are different; that is, a structural change has occurred. We examine the distribution of each structural variable over the landscape over time, as opposed to multivariate distributions of the structural variables, as the former does not present the degree of difficulty in testing and the expectation of change over time that the latter would.

Because we have little a priori knowledge of the complex relationship between structural and state variables, we use a highly flexible exploratory data analysis approach and nonparametric statistics to test hypotheses. We define the bins of plots for a given structural variable over time by segmenting plots into homogeneous groups or bins determined by values of the state variables using a regression tree. A regression tree splits data into homogeneous groups or bins without any a priori definition of functional form. This technique is well-suited for our situation in which functional relationships are uncertain, complex, nonlinear, and interactive, and it supports compact tests of our hypotheses (Vayssières et al. 2000). A traditional model would test for change in the value of regression coefficients between surveys within the parametric model. Using the bins defined by the regression tree for all inventories, we compare the distributions of the structural variable in each bin for an individual inventory t, $t + 1$, and $t + 2$ using Kolmogorov-Smirnov (K-S) tests. The nonparametric approach may reduce the power of the resulting tests, but it is robust in the face of little knowledge regarding the appropriate functional form, i.e., a traditional statistical model would probably introduce specification error to these tests.

In the remainder of this article we develop and apply tests for structural changes using three inventories (1984, 1990, and 2002) for the State of North Carolina. North Carolina is a good test case because it has experienced a high rate of change in land use and many forest attributes over this period compared with other states in the southeastern United States (Conner and Hartsell 2002, Wear 2002). Forest planting and intensive management as well as the conversion of forest and agricultural land to developed uses have altered forest landscapes. First, we describe the data and the methods used to summarize forest inventories for this study. Second, the regression tree approach is discussed. Third, we present results for four structural variables explored in this case study. Fourth, we discuss how our nonparametric tests for structural change provide indicators of change and conclude with additional hypotheses regarding causative factors.

Data

In the FIA, plots are systematically located across the landscape in an area-frame survey. These plots are permanently established and repeatedly measured over time. The FIA data are used to assess forest conditions for substate regions (called survey units) based on measurements of tree, vegetation, and site characteristics for hundreds of plots within each region. FIA forest inventories were conducted on a periodic basis from the 1930s to the 1990s. Recently, the US Forest Service has begun to phase in a continuous inventory approach for most states in the United States that surveys a portion of permanent plots every year. We use the final three fully completed inventories for North Carolina for our analysis. Permanent plots from three inventories of the State of North Carolina (1984, 1990, and 2002) were used to evaluate our methods. The survey design changed between 1990 and 2002 in North Carolina, providing another motivation for testing whether or not changes in structural elements are revealed. By using our tests on permanent sample plots, we can evaluate structural change over time on forestland. Although the sampling methodology has changed, the goal of developing these tests is to examine change over time, which will eventually prove useful in the new annual inventory framework.

FIA data are stored in tables, three of which are used for our analysis (Miles et al. 2001). The plot, condition, and tree tables provide information on the overall plot characteristics, discrete landscape features, and measures associated with individual trees larger than 1 inch in diameter, respectively (Miles et al. 2001, p. 7). These data are extensive, and thus some manipulations were performed on the raw data to obtain values for the structural and state variables. The analysis was completed using data from FIA Database Versions 1.7 and 2.1. The most current FIA Database, Version 3, was made available in June 2007; the variables used in the subsequent analysis were compared and found to be consistent among database versions. Data were summarized at the plot level and then permanent plots were identified that spanned each of the three inventory years.

For each plot, four structural variables, volume of growing stock trees, and number of growing stock trees for both softwoods and hardwoods, were calculated on a per hectare basis for forested plots to account for changes in survey methodology over the years using algorithms derived from those suggested in Miles et al. (2001). For example, plots in the 2002 inventory were divided into condition class components, that is, each plot may contain one or more condition classes (based on changes in land use, owner, and forest type), and these needed to be expanded to define plot values (Bechtold and Scott 2005). As a validation step we used expansion factors to generate total values for survey units, which could then be compared with published reports (see Sheffield and Knight 1986, Johnson 1991, Brown 2004) and confirm accuracy of the algorithms used in this analysis. Growing stock volume and trees per hectare were delineated by broad species type, i.e., softwood and hardwood, using the species group variable recorded in the FIA database (Table 1).

Table 1. Descriptive statistics of variables used in regression tree models and hypothesis tests analyses of North Carolina 1984, 1990, and 2002 FIA inventories

Name	Variable definition	Mean	SD
State variables			
STDAGE	Average age of the trees in the predominant stand-size class of the condition	42.223	28.049
NPINE	= 1 if broad management class is natural pine, 0 otherwise	0.209	
PPINE	= 1 if broad management class is planted pine, 0 otherwise	0.112	
MPINE	= 1 if broad management class is mixed pine, 0 otherwise	0.144	
LHDWD	= 1 if broad management class is lowland hardwood, 0 otherwise	0.141	
UHDWD	= 1 if broad management class is upland hardwood, 0 otherwise	0.393	
SLOPE	Average percent slope of the condition	12.297	18.342
ASPECT	Predominant drainage direction (in degrees) from magnetic north for condition	108.267	113.416
XER	= 1 if physiographic condition class is xeric, 0 otherwise	0.073	
MES	= 1 if physiographic condition class is mesic, 0 otherwise	0.799	
HYD	= 1 if physiographic condition class is hydric, 0 otherwise	0.128	
PUBLIC	= 1 if condition ownership group is public, 0 otherwise	0.146	
PRIV	= 1 if condition ownership group is private, 0 otherwise	0.854	
SIQ1	= 1 for lowest site index quartile, 0 otherwise	0.276	
SIQ2	= 1 for second lowest site index quartile, 0 otherwise	0.294	
SIQ3	= 1 for second highest site index quartile, 0 otherwise	0.222	
SIQ4	= 1 for highest site index quartile, 0 otherwise	0.208	
D84	= 1 if assigned 1984 inventory plot attributes, 0 otherwise	0.342	
D90	= 1 if assigned 1990 inventory plot attributes, 0 otherwise	0.322	
D02	= 1 if assigned 2002 inventory plot attributes, 0 otherwise	0.335	
SCP	= 1 if survey unit is Southern coastal plain, 0 otherwise	0.269	
NCP	= 1 if survey unit is Northern coastal plain, 0 otherwise	0.250	
PIED	= 1 if survey unit is Piedmont, 0 otherwise	0.280	
MTN	= 1 if survey unit is Mountain, 0 otherwise	0.200	
Structural variables			
VSGSPH	Volume of softwood growing stock per hectare ($m^3 \times 1,000$)	0.046	0.068
VHGSPH	Volume of hardwood growing stock per hectare ($m^3 \times 1,000$)	0.076	0.099
TPHSW	Number of softwood growing stock trees per hectare ($\times$ 1000)	1.129	1.064
TPHHW	Number of hardwood growing stock trees per hectare ($\times$1000)	0.368	0.669

The possible state variables that resulted from combining information from the plot, condition, and tree tables in the FIA database were extensive and are summarized as follows (Table 1). Forest type and stand origin were combined to create a binary broad management class variable coinciding with the definition in published reports (e.g., see Sheffield and Knight 1986). These broad management classes are based on forest type, stand origin, and stocking which must be a minimum of 10% for all classes (Brown 2004). Planted pine (PPINE) and natural pine (NPINE) are both classed as pine or other softwood forest type and differ in that the former is artificially regenerated by planting or direct seeding, whereas the latter have not been artificially regenerated. Mixed pine (MPINE) types are classed under the Oak/Pine forest type group (Miles et al. 2001, p. 114). Upland hardwood (UHDWD) types are classed under the Oak/Hickory or Maple/Beech/Birch forest type groups. Last, lowland hardwood (LHDWD) types are classed under the Oak/Gum/Cypress, Elm/Ash/Cottonwood, Tropical Hardwoods, or Exotic Hardwoods forest type groups. The physiographic condition class code variable was similarly converted into binary dummy variables, which included xeric (XER), mesic (MES), and hydric (HYD). The four substate survey units were also used as state variables in the state-level models and included binary variables for the presence or absence in the Southern Coastal Plain (SCP), Northern Coastal Plain (NCP), Piedmont (PIED), and Mountain (MTN).

Because of variability in the recorded values of the site index variable among the survey cycles, we determined quartiles of the site index measure for the respective original inventory year and assigned binary variables associated with occurrence in a specific site index quartile; from lowest to highest these are SIQ1, SIQ2, SIQ3, and SIQ4. This variability in the recorded values of the site index variable among survey cycles was a result of recording practices and not a result of changes in base age used for determining this site productivity measure. The 1984 site index is recorded with a 10-unit difference in levels ranging from 35 to 125; thus, the quartiles from lowest to highest contained the following ranges: 35–65, 75, 85, and 95–125. The 1990 site index also exhibits a 10-unit difference in recorded levels ranging from 0 to 99, and the quartiles contain ranges of 0–60, 70, 80, and 90–99 from lowest to highest, respectively. The site index in the 2002 survey is recorded on a discrete basis ranging from 37 to 99, and the quartiles from lowest to highest include the ranges of 37–68, 69–79, 80–88, and 89–99. The aspect (ASPECT) variable also exhibited differences in recorded values between surveys. Aspect was given a value of 400 for flat land in 2002 (Ali Conner, pers. comm., US Forest Service–FIA, Oct. 20, 2005), and therefore we reassigned these occurrences a value of zero to coincide with the prior two surveys. The ownership group condition class variable was used to determine whether the plot was publicly owned (PUBLIC) or privately owned (PRIV), where the former included all Forest Service, other federal, and state and local ownerships. The average percent slope (SLOPE) and total age of the trees in the predominant stand-size class of the condition (STDAGE) were also included as state variables.

To preserve temporal independence among samples and avoid potential upward bias on condition sizes, the permanent plots with records in each of the inventories were summarized and selected for use in our analysis in the following fashion. A single condition was randomly selected for those permanent plots with multiple conditions in the 2002 inventory. The resulting matching set of permanent plots was composed of 2,845 plots and plot/conditions. Plots ranged from approximately 17 to 3,676 hectares in size. This set of permanent plots was then randomly assigned to an inventory year resulting in 974, 917, and 954 plots or plot/conditions on land delineated as forestland to be used for analysis of the 1984 (D84), 1990 (D90), and 2002 (D02) North Carolina surveys, respectively, using a binary coded variable for inventory year. Each resulting permanent plot was assigned characteristics associated with its representative year to be used in both the regression tree construction and subsequent testing of higher distributional moments. This assignment ensures temporal independence among the three surveys—i.e., no permanent plot's attributes are derived from or associated with more than one survey—and thus supports the underlying assumptions of independent and identical distributions of the K-S test.

Methods

Regression Trees

We used regression tree analysis to construct the set of bins of plots for each structural variable. This is a compact description of how the probability distribution of the structural variable is influenced by the set of state variables. An alternative approach to constructing this set of bins of plots might be to construct regression equations for the structural variables of interest. However, the complex relationship between state and structural variables probably includes discontinuities, nonlinearities, and complex variable interactions that would give rise to specification problems for the regression approach—i.e., we have no a priori guidance for specifying the regression equations. Because it does not impose a functional form a priori we used regression trees, an exploratory modeling technique that has been used in ecological analysis (Michaelsen et al. 1994, Andersen et al. 2000, De'ath and Fabricius 2000) and in analysis of forest conditions (Byler et al. 1990, Iverson and Prasad 1998), to construct a set of bins of plots for each structural variable over time.

Regression tree methods use a recursive partitioning algorithm to split a data set into homogeneous subgroups using splitting rules (Therneau and Atkinson 1997). The data set, which is referred to in tree terminology as the root node, is initially split into two subsets based on the most successful separation of high and low values of the dependent variable using one of the explanatory (state) variables (Breiman et al. 1984). The same methodology is applied to each individual subset, recursively, until an ending condition is satisfied. Output consists of a summary of the relationship between the dependent structural variable and the independent state variables, which can be presented in a graphic format as a tree structure with branches and terminal nodes.

The partitioning of data into various subgroups is referred to as splitting. The decision rules defining a split can be based on either a categorical variable (e.g., $x = 0$ or $x = 1$) or a continuous variable ($x < k$ or $x \geq k$, where the threshold k is selected by the algorithm) and are selected so that the resulting subsets are as homogeneous as possible in terms of the response, the mean of the dependent variable (De'ath and Fabricius 2000). Regression trees use splitting criteria based on the sum of squares about the group means (Therneau and Atkinson 1997). Define the sum of squares for the entire data set as

$$SS_{\mathrm{T}} = \sum_i (y_i - \bar{y})^2. \tag{1}$$

The sum of squares for the daughter nodes, splits to the right and left side of the tree, are defined as:

$$SS_{\mathrm{R}} = \sum_i (y_{\mathrm{R}i} - \bar{y}_{\mathrm{R}})^2, \tag{2a}$$

$$SS_{\mathrm{L}} = \sum_i (y_{\mathrm{L}i} - \bar{y}_{\mathrm{L}})^2, \tag{2b}$$

where y is the response, or dependent variable, which in this case is either the per hectare volume of growing stock or the number of growing stock trees per hectare. The split is defined to maximize the reduction in the sum of squares ($\max[SS_{\mathrm{T}} - (SS_{\mathrm{R}} + SS_{\mathrm{L}})]$). The same logic is then applied individually to the two resulting subsets on either side of the initial split. The data are split into subsequent subsets until the reduction in the sum of squares is minimal (see Breiman et al. 1984 for a more thorough description of the algorithm used).

The final regression tree defines groups of the survey plots (for ranges of the state variables) that have distributions of the dependent structural variable that minimizes the variance within each split (e.g., daughter node) while maximizing the variance between each split. Each subset is defined by a terminal node of the resulting regression tree (an example is Figure 1), constituting a bin of plots. The structure of the regression tree itself is informative. It identifies the variables and their respective ranges that explain variation in the structural variable. The degree of explanatory power or the share of reduction in total sum of squares for each split is indicated by the relative depth of the branch in the regression tree (e.g., Figure 1 shows that the binary variable for occurrence in the natural pine management class explains more variance in softwood growing stock volume than any one of the other variables).

Regression trees were generated in the open source R statistical programming platform using the RPART (Regression PARTitioning) package (R Development Core Team 2005, Therneau and Atkinson 2005). These methods were used to produce compact bins of plots describing the relationship between each of four dependent structural variables (volume of softwood or hardwood growing stock trees per hectare or the number of softwood or hardwood growing stock trees per hectare) and the independent state variables (public or private owner, physiographic type, broad management class, stand age, slope, aspect, site index quartile, survey unit, and survey year) using data from the three survey years (1984, 1990, and 2002), randomly assigned, as described previously. Regression tree models were generated at the state level and at the unit level, where the latter analysis excluded the survey unit variables, to examine how

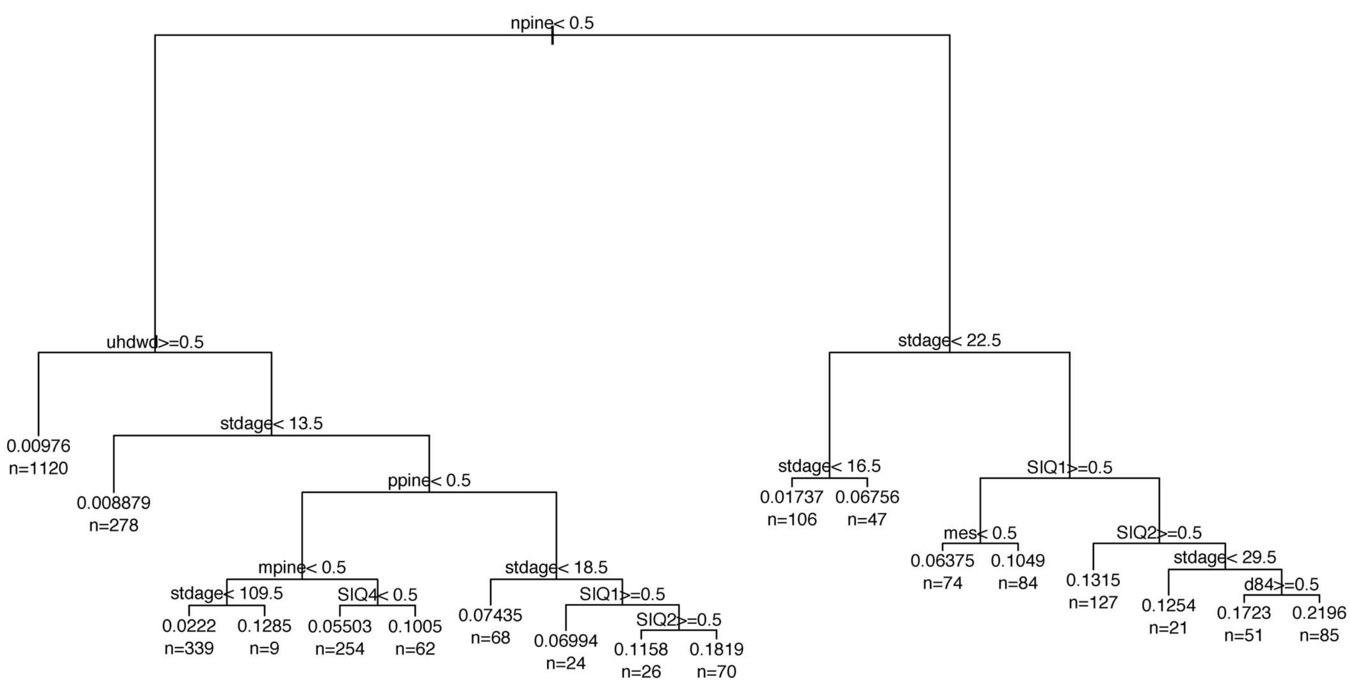

Figure 1. Regression tree relating softwood growing stock volume in North Carolina to various explanatory variables for permanent plots randomly assigned to an inventory year (1984, 1990, and 2002) obtained with stopping complexity parameter set to 0.005. Numbers at each terminal node are the mean cubic meters, in thousands, per hectare for plots in that bin and *n* indicates the number of plots included. Branches or splits in the tree are defined by the listed splitting rule, sending plots where rule is true to the left and where false to the right. For variable definitions see Table 1.

this level of aggregation affected the resulting set of bins of plots.

Comparing Bins of Plots

The bins of plots produced from regression tree models are defined by the rules regarding the limits of the state variables that define the terminal nodes—i.e., those that minimize within-bin plot variance while maximizing between-bin plot variance. Our objective in this analysis was to determine for each dependent (structural) variable whether the plots for one survey year are significantly different from those for the other survey years. Our null hypothesis in estimating the regression trees is that there is no discernible between-survey difference for each structural variable. The alternative hypothesis is that regression trees indicate a difference between survey years, i.e., a structural change has occurred. We construct hypothesis tests in two ways. The first is based on the structure of the regression trees. If a binary survey year variable is present in the regression tree as a splitting (state) variable, then this indicates a discernible difference in the distribution of the structural variable between surveys, i.e., rejection of the null hypothesis for at least a subset of the inventory. We accept the null hypothesis if survey year is not a splitting (explanatory) variable in the regression tree. However, the probability level of this hypothesis test is not determinant in the way that a *t* test for an ordinary least-squares regression would be. Rather it is related to the selection of the stopping rule for the regression tree algorithm. This stopping rule determines whether another split or further subsetting of the structural variable is worth pursuing and can be adjusted. We conduct some sensitivity analyses with the stopping rule in the regression tree models, but this remains a weakness of this hypothesis test. In the sensitivity analysis the stopping rule, or complexity parameter, was relaxed to 0.001 for each of the four structural variables modeled at the state level. The 0.004 reduction in the stopping complexity parameter resulted in a change in relative error, $(1 - R^2)$, for the regression tree models of volume of softwood and hardwood growing stock trees per hectare of 0.07 and 0.09, respectively. This same reduction in stopping complexity parameter changed the relative error for the number of softwood and hardwood growing stock trees per hectare models by 0.06 and 0.15, respectively.

Our second test involves evaluating the data within each final subset, bin of plots, produced from the estimated regression tree. Recall that each data point (plot) falls within a terminal node (bin of plots) and if a binary survey year variable does not split the data, then each resulting bin contains data (plots) from all 3 survey years. At each bin of plots, we test the null hypothesis that there is no discernible difference in the probability distribution of the structural variables for a given survey compared with the distribution of remaining surveys (alternate survey years). We use the nonparametric K-S test, with a Bonferroni adjusted α according to the number of comparisons made per plot bin, e.g., α/number of comparisons (Rao 1998). The null hypothesis is that the structural variable data for an individual survey year within a bin of plots were generated from the same distribution as that for the alternate inventory data set. For each structural variable and for each inventory we conduct a K-S test and use these results as an additional gauge of the overall comparability of survey years. For example, in the volume of softwood growing stock model we compare 1984 versus 1990 and 2002, 1990 versus 2002 and 1984, and 2002 versus 1984 and 1990, for each bin of plots. For each resulting K-S test, the probabilities were compared with 0.0167, the adjusted α, unless there were only two surveys present in a bin of plots. We compared the proportion of bin comparisons and the proportion of the total structural variable measured in a survey year for which the null hypotheses of identical distributions are rejected for each model.

This is a more general test than presence of a survey year variable as a splitting variable in the estimated regression trees. The regression tree algorithm maximizes the sum of square difference between bins of plots while minimizing the within-bin sum of squares difference in the structural variable modeled. Thus, a survey year is used as a splitting variable when these criteria are met. The K-S test is then used to address correspondence of the distributions of these data within each bin of plots; thereby identifying cases for which the means may be comparable across the survey years, but for which variances (or higher order moments) may be discernibly different. It should be noted that the following results are based on a specific subset of measures of forest conditions, specifically growing stock trees, and different results may be obtained using alternate measures of forest condition.

Results

Permanent plots present in the three most recent completed surveys of forest conditions for North Carolina (1984, 1990, and 2002) conducted by the US Forest Service FIA program were analyzed using regression trees. The data were summarized at the plot level, and plots were analyzed first for the state as a whole and then for each of the four survey units in North Carolina.

Regression Trees and Resulting Bins of Plots

For the state as a whole, regression tree estimates indicated that 66% of the variation in permanent plot measures of softwood growing stock volume (1,000 m^3 ha^{-1}) is explained by differences in (1) broad management classes, (2) stand age groups, (3) site quality as measured by site index quartiles, (4) physiographic type, and (5) survey year (Figure 1). Regression trees were estimated with a stopping complexity parameter of 0.005. Whether or not the plot is in a natural pine management class is the first splitting variable for the data set and the length of the branch indicates that this variable explains a majority of the variation in softwood growing stock volume. The next level of splits is defined by other management classes or stand age. Natural pine plots (the right split in Figure 1) are initially split into three age groups ($<$17, 17–22, and $\geq$23). The natural pine plots older than 23 years are then split according to site quality (SIQ =

1, SIQ = 2, and SIQ = 3, 4). For all other broad management classes, i.e., not natural pine, the left split in Figure 1, the regression tree, separates plots based on other management classes, age, and site index. Upland hardwood forest plots are initially split off on the left (Figure 1). The remaining lowland hardwood, mixed pine, and planted pine classes are split by age. Younger plots (≤13 years) of these classes have the lowest average volume of softwood growing stock per hectare. Mixed pine and planted pine plots are further split by age. Lowland hardwood plots are further split according to site quality (SIQ = or ≠ 4).

For hardwood growing stock volume, differences in (1) stand age, (2) broad management classes, (3) site quality, (4) aspect, and (5) survey year explain 48% of the variation in plot measures of the dependent variable. Stand age explains the most variation in hardwood growing stock volume, indicated by its position and relative length of the branch in the regression tree (Figure 2). The next level of splits is defined by management class or stand age. After the right split in Figure 2, whether or not the plot is in a natural pine management class is the second splitting variable for older stands (≥43). Hardwood, mixed pine, and planted pine plots are split according to site quality (SIQ = 1, 2, 3 or SIQ = 4). Hardwood plots with stand ages from 55 to 71 years on mid-quality sites with northeast aspects in the 2002 inventory have the highest average growing stock per hectare of the bins in this regression tree model. After the left split in Figure 2, the plots are further split according to age. Plots with stand age ≤26 years on nonprime quality sites (SIQ ≠ 4) have the lowest average volume of hardwood growing stock per hectare. Plots aged 27–42 years are split according to management class (pine/mixed pine, lowland hardwood, and upland hardwood).

With the number of growing stock trees (thousands) per hectare as the dependent measure of forest structure, the regression trees have starting points similar to those determined for the volume of growing stock trees per hectare (Figures 3 and 4). Also, when the splits among the two structural measures for softwood and hardwood (growing stock volume and number of trees) are compared, stand age appears as a splitting variable an equal number of times for each broad species class measure. However, splits in the regression trees modeling the number of trees per hectare seem to be less balanced than those conducted on the measures of growing stock volume, in that splits seem to move more to the right. Also, one variable does not dominate in the explanation of variance when the number of growing stock trees per hectare is measured. For instance, upland hardwood and lowland hardwood management classes explain an amount of variation similar to that of the natural pine management class when one is comparing the regression tree analysis of the number and volume of softwood growing stock trees per hectare, respectively.

For the number of softwood growing stock trees per hectare as the dependent structural variable, the upland hardwood and lowland hardwood management classes explain relatively similar proportions of variation, as indicated by the length of the splits (Figure 3). Overall, differences in (1) management class, (2) stand age, (3) site quality, (4) survey year, (5) survey unit, (6) slope, and (7) aspect explain approximately 49% of the variation in this state variable. Pine plots are split by differences in all of the

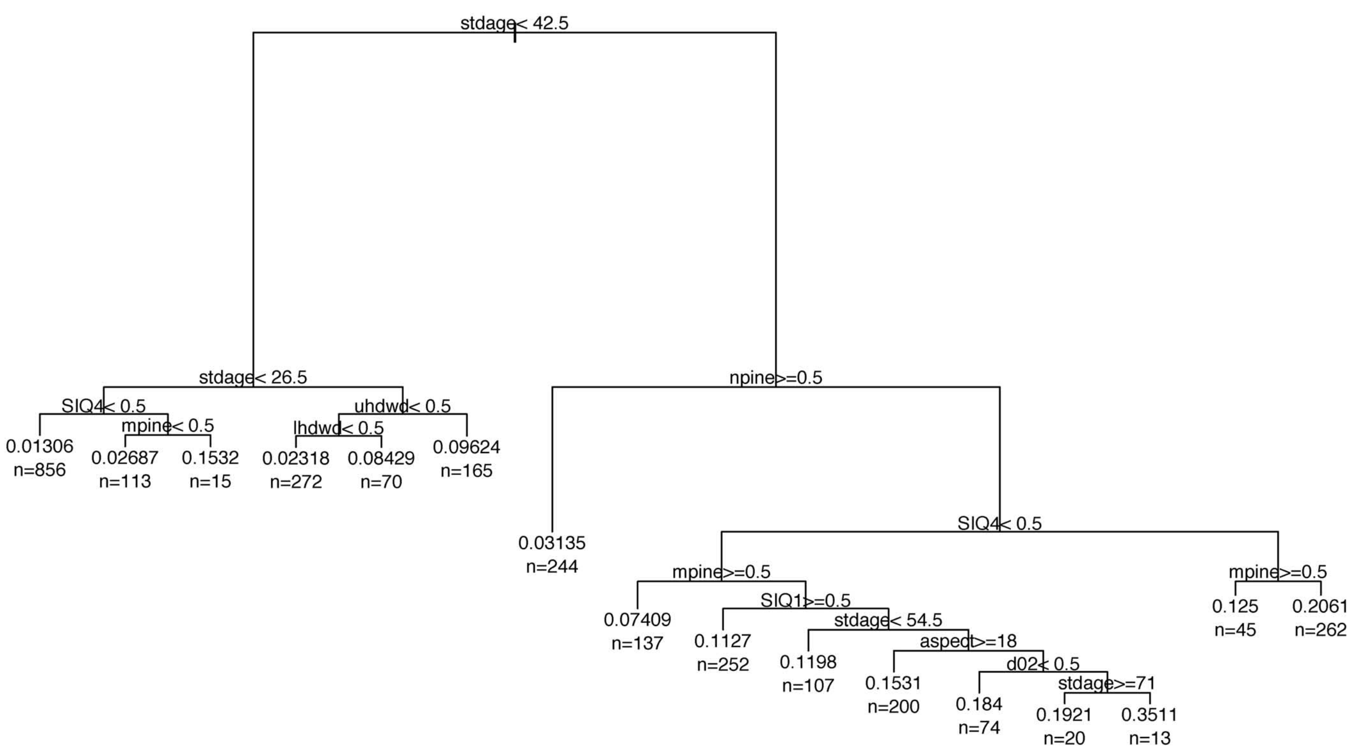

Figure 2. Regression tree relating hardwood growing stock volume (cubic meters, in thousands) per hectare in North Carolina to various explanatory variables for permanent plots randomly assigned to an inventory year (1984, 1990, and 2002) obtained with stopping complexity parameter set to 0.005. Variable definitions are found in Table 1.

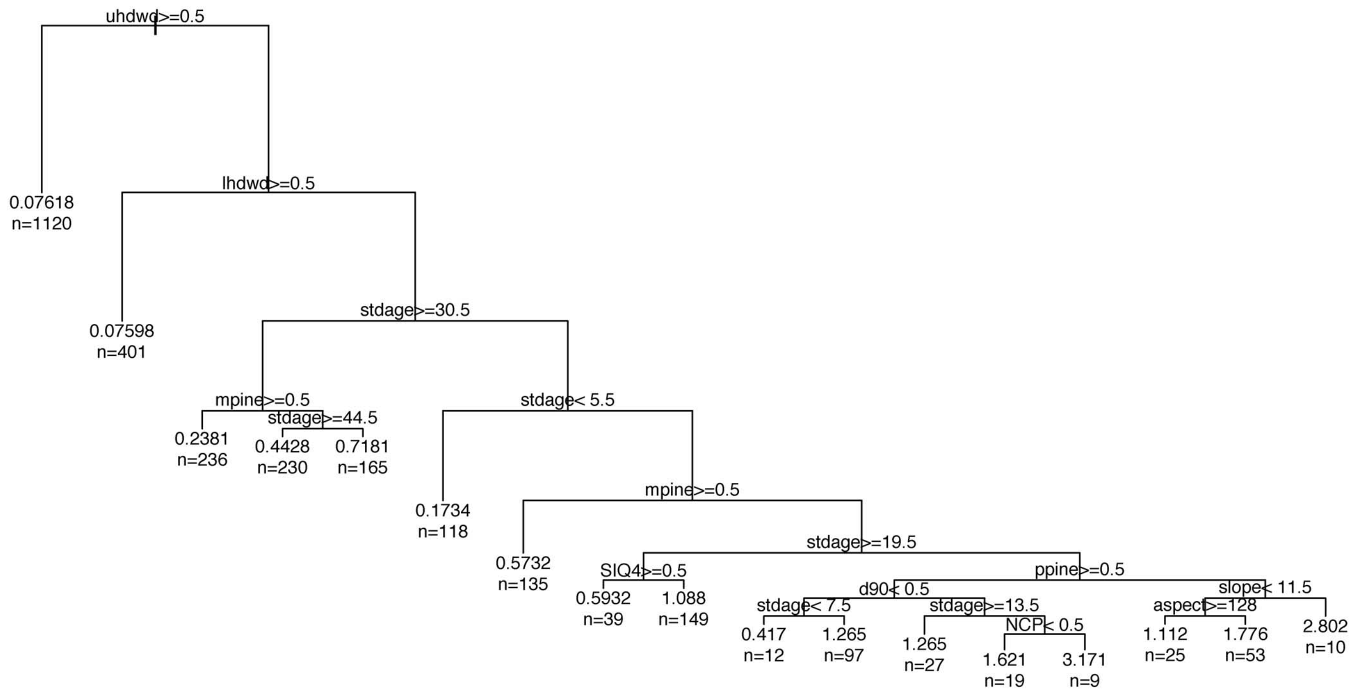

Figure 3. Regression tree relating number of softwood growing stock trees (in thousands) per hectare in North Carolina to various explanatory variables for permanent plots randomly assigned to an inventory year (1984, 1990, and 2002) obtained with stopping complexity parameter set to 0.005. Branches or splits in the tree are defined by the listed splitting rule, sending plots where rule is true to the left and where false to the right. Variable definitions are found in Table 1.

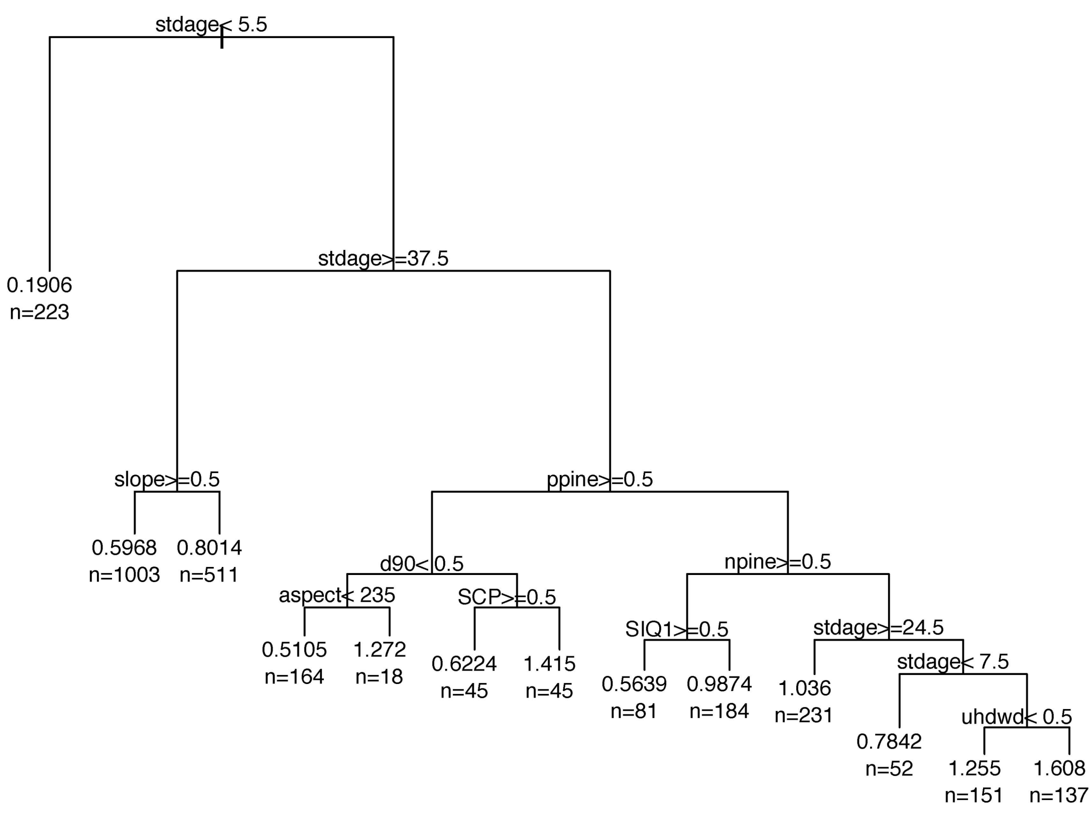

Figure 4. Regression tree relating number of hardwood growing stock trees (in thousands) per hectare in North Carolina to various explanatory variables for permanent plots randomly assigned to an inventory year (1984, 1990, and 2002) obtained with stopping complexity parameter set to 0.005. For variable definitions see Table 1.

above characteristics, whereas hardwood types are split only according to management class. Mixed pine, natural pine, and planted pine plots are split into three, five, and seven age groups, respectively. The bins with the lowest and highest average number of trees per hectare in this regression tree model are found on lowland hardwood plots and 1990 Northern Coastal Plain planted pine plots aged 14–19 years, respectively.

For the number of hardwood growing stock trees per hectare, differences in (1) stand age, (2) management class, (3) slope, (4) site quality, and (5) survey unit explained only 17% of plot variation measures of the dependent variable. Very young (≤5 years) and very old plots (≥38 years) are not subsequently split according to management class; however, the oldest plots are split according to slope (Figure 4). Plots with stand ages in the 5–37 year range are split according to type (planted pine or natural pine), and mixed pine, upland, and lowland hardwood classes are split into three distinct age groups (6–7, 8–24, and 25–37 years), with the middle age range split on management class. Upland hardwood plots in the 8–24 year age range possess the highest average number of trees per hectare in this regression tree model.

Tests for Structural Change

Survey year appears as a splitting variable in all of the four state-level regression trees, and in all cases a small portion (<10%) of additional variation is explained by the inclusion of survey year as a splitting variable. For the volume of softwood growing stock, the variable indicating presence or absence of a plot in the 1984 inventory occurs toward the bottom of the regression tree and indicates a structural change for only 5% (136 of 2,845) of the plots (Figure 1). Similarly, for volume of hardwood growing stock a structural change is indicated for only 4% (107 of 2,845) of the plots by the presence of the 2002 inventory as a splitting variable toward the bottom of the regression tree (Figure 2). For the number of softwood growing stock trees, the presence or absence of plot attributes resulting from the 1990 inventory indicates a structural change for 6% of the plots (164 of 2,845) (Figure 3). Last, for the number of hardwood growing stock trees, a structural change is indicated for 10% of the plots (272 of 2,845) by the presence of the 1990 survey year dummy as a splitting variable in the resulting regression tree (Figure 4). We therefore accept the null hypothesis of no structural change based on the regression tree results for 95 and 96% of plots for the volume of softwood and hardwood growing stock per hectare. For the number of softwood and hardwood growing stock trees, we accept the null hypothesis for 94 and 90% of the plots.

Based on the K-S test at the 5% significance level adjusted according to the Bonferroni correction using the number of comparisons made per bin of plots, the hypothesis of equivalent softwood growing stock volume distributions is rejected for approximately 2% of the bin comparisons using the structure of the state-level regression tree (Table 2). For the 1984 inventory, plot distribution in 1 of the 18 resulting bins containing older mixed pine plots on lower quality sites is significantly different from the set comprising the two alternate inventories. This is the only case at the state level of significant differences among survey years within bins of plots for the model of the structural variable volume of softwood growing stock, whereas for hardwood growing stock volume, the null hypothesis of equivalent information sets is rejected for 7% of bin comparisons. For both the 1990 and 2002 inventories, plots in at least 1 of the 16 bins are significantly different from the multiyear set of alternate inventories (Table 3). For both, significant differences in the distribution of plots were found for the youngest plots of all forest types on lower quality sites. The 2002 inventory also has significantly different plot level distributions occurring in an additional bin representing older hardwood plots on the lowest quality sites.

For the number of softwood growing stock trees, we accept the null hypothesis of equivalent distributions for all of the possible survey-year comparisons at the state level (Table 4). For the number of hardwood growing stock trees, approximately 18% of the possible comparisons are rejected for equivalent distributions within bins of plots using the K-S tests. For the 2002 inventory, distributions in 3 of 13 bins are significantly different from their respective multiyear sets (Table 5). These span all forest types and slopes, with stand age splits at 25 and 38 years. In contrast, there were no significantly different distributions for plots in the 1990 inventory when the resulting bins were compared with their respective multiyear distributions of plots. For the 1984 inventory, distributions in 3 of 12 bins of plots are significantly different. These include young hardwood or mixed pine plots, and both low and higher sloped plots of all management classes where stand age is ≥38 years.

Unit Level Regression Tree Models

We repeated the analysis of the four indicator variables at the individual survey unit level with somewhat different results and will summarize those data here (full results are available from the authors). All survey units have an inventory year as a splitting variable in at least three of the four

Table 2. Average softwood growing stock volume and regression tree split defining variables for multiyear inventory sets where the null hypothesis of equivalent distributions among inventory years within bins of plots is rejected at least once using permanent plots spanning three North Carolina FIA inventories

	Average volume per ha (m^3 in thousands)			
Model	1984[a]	1990[a]	2002[a]	Split defining variables
State [18]	0.057 (80)[b]	0.053 (89)	0.055 (85)	Stand age ≥14 year; mixed pine management class; SIQ = 1, 2, 3

Numbers in parentheses indicate numbers of observations in sample; number in brackets indicates number of bins per model.
[a] Randomly assigned.
[b] Significantly different from other inventories using K-S tests (Bonferroni adjusted $\alpha = 0.0167$).

Table 3. Average hardwood growing stock volume and regression tree split defining variables for multiyear inventory sets where the null hypothesis of equivalent distributions among inventory years within bins of plots is rejected at least once using permanent plots spanning three North Carolina FIA inventories

	Average volume per ha (m^3 in thousands)			
Model	1984[a]	1990[a]	2002[a]	Split defining variables
State [16]	0.011 (280)	0.018[b] (293)	0.010[b] (283)	Stand age ≤26 years; SIG = 1, 2, 3
	0.097 (62)	0.105 (73)	0.126[b] (117)	Stand age ≥43 years; hardwood management classes; SIQ = 1
Unit 1 [16]	0.007 (121)	0.011 (131)	0.010[b] (111)	Stand age ≤46 years; pine or mixed pine management classes
Unit 2 [12]	0.189 (26)	0.189 (36)	0.163[b] (45)	Stand age ≥50 years; hardwood management classes; SIQ = 1, 2, 4
Unit 3 [14]	0.015 (102)	0.023[b] (89)	0.010[b] (90)	Stand age ≤27 years

Numbers in parentheses indicate numbers of observations in sample; number in brackets indicates number of bins per model.
[a] Randomly assigned.
[b] Significantly different from other inventories using K-S tests (Bonferroni adjusted $\alpha = 0.0167$).

Table 4. Average number of softwood growing stock trees and regression tree split defining variables for multiyear inventory sets where the null hypothesis of equivalent distributions among inventory years within bins of plots is rejected at least once using permanent plots spanning three North Carolina FIA inventories

	Average no. trees (in thousands) per hectare			
Model	1984[a]	1990[a]	2002[a]	Split defining variables
Unit 2 [14]	1.115[b] (9)	n/a	0.572[b] (8)	Stand age 20–25 years; planted pine management class; survey year = 1984 or 2002
Unit 4 [7]	0.424[b] (28)	0.256 (24)	0.280 (15)	Mixed pine or lowland hardwood management classes

Numbers in parentheses indicate numbers of observations in sample; number in brackets indicates number of bins per model.
[a] Randomly assigned.
[b] Significantly different from other inventories using K-S tests (Bonferroni adjusted $\alpha = 0.0167$).

Table 5. Average number of hardwood growing stock trees per ha and regression tree split defining variables for multiyear inventory sets where the null hypothesis of equivalent distributions among inventory years within bins of plots is rejected at least once using permanent plots spanning three North Carolina FIA inventories

	Average no. trees (in thousands) per ha			
Model	1984[a]	1990[a]	2002[a]	Split defining variables
State [13]	0.639[b] (362)	0.623 (315)	0.524[b] (326)	Stand age ≥38 years; slope ≥0.5
	0.914[b] (180)	0.777 (157)	0.707[b] (174)	Stand age ≥38 years; slope <0.5
	1.074 (86)	1.155 (71)	0.876[b] (74)	Stand age 25–37 years; mixed pine or hardwood management classes
	1.164[b] (15)	0.715 (20)	0.530 (17)	Stand age 6–7 years; mixed pine or hardwood management classes
Unit 1 [22]	1.050 (22)	0.886 (13)	0.599[b] (19)	Stand age 35–56 years; mixed pine or hardwood management classes; mesic or hydric physiographic type; SIQ = 1, 3, 4

Numbers in parentheses indicate numbers of observations in sample; number in brackets indicates number of bins per model.
[a] Randomly assigned.
[b] Significantly different from other inventories using K-S tests (Bonferroni adjusted $\alpha = 0.0167$).

structural models. The Mountain survey unit has the lowest occurrence of inventory years as a splitting variable: only once in the model of the number of hardwood growing stock trees. The Northern and Southern Coastal Plain survey units have five and four occurrences, respectively, with generally one occurrence of an inventory year as a splitting variable per survey unit per model. The Piedmont survey unit also has four occurrences of survey year as a splitting variable, with one occurrence in each of the models of volume of softwood growing stock and number of softwood growing stock trees and two occurrences in the model of number of hardwood growing stock trees.

We applied the K-S test to evaluate the null hypothesis of equivalent distributions for bins of plots that are not survey year-specific, i.e., a survey year indicator is not a splitting variable on the path to reach the resulting bin or if it is there are at least two survey years present to be compared (Tables 2–5). For all models, number of softwood and hardwood growing stock trees and the volume of softwood and hardwood growing stock, the hypothesis of equivalent distributions, is rejected for <2.6% of the total comparisons (respectively, 3 of 132 tests for softwood trees, 1 of 172 tests for hardwood trees, 0 of 174 tests for softwood volume, and 4 of 159 tests for hardwood volume). There were five, one, and two rejections of the null hypothesis of no structural change among assigned survey years (for 2002, 1990, and 1984, respectively) for the indicator models of hardwood growing stock volume and number of both softwood and hardwood growing stock trees (Tables 3, 4, and 5). Also, rejections of the null hypothesis were concentrated in the Coastal Plain survey units for these models (five of eight), with only two rejections occurring in the Piedmont survey unit for the model of number of hardwood growing stock trees and one rejection occurring in the Mountain survey unit for the number of softwood growing stock trees.

Discussion

Our compendium of tests conducted at both the state and survey unit levels do not provide a single unambiguous answer regarding structural change between North Carolina's forest inventories. Taken together, they do provide useful insights into changes in the forest inventory, assessed using volume of growing stock and number of trees as structural indicators of forest conditions. At the state level we found the following:

- ➤ Survey year does explain differences in hardwood and softwood volumes of growing stock and numbers of softwood and hardwood trees (more precisely, survey year does split data in the regression trees). However, in general, survey year serves to define plot bins for these structural variables for <10% of plots. Based on this test we accept the null hypothesis that the three inventories are comparable with respect to these variables in characterizing these specific attributes of the forest for 90% of the plots.
- ➤ We do find significant differences in the distributions of hardwood variables at some bins of plots using the K-S test, for at least 7% of the distribution analyses. For the models of softwood variables we find low (2%) or no significant differences in the distributions of these variables in their resulting bins of plots among survey years.

The regression tree results, based on a total sum of squares minimizing algorithm, suggest that the means of the four structural variables in each resulting bin of plots are only marginally different across the surveys. Regression tree results indicated that the survey year variable explained an additional 3–7% of the variation in the means of the plots for the four structural models. The K-S test is a much more general test, comparing the entire distribution of plots in each resulting bin. Coupled with the regression tree results, the rejection of the null hypothesis using the K-S test suggests that the variances (or higher moments) of the dependent (structural) variable are significantly different for these subsets of the data. The variance of the four structural variables for each bin of plots generally increases from 1984 to 2002 and from 1990 to 2002. The exception is in the model of number of hardwood growing stock trees, for which the difference in variance of the bins of plots in the 1984 and 2002 inventories is equal in a comparison among the bins.

There seems to be no generalizable pattern regarding significant differences across the three survey years. Given the change in survey design between 1990 and 2002, we might have expected the results of the 2002 inventory to be significantly different from those of the earlier inventories. However, our findings suggest that there was no change in the likelihood of observing structural differences after the change in inventory methodology. This result suggests that the results from the 2002 inventory are as comparable to the 1990 inventory results as the 1990 inventory results were comparable to the 1984 inventory results.

Application of these models at a finer scale (i.e., the four FIA survey units in North Carolina) yields somewhat different results. In 12 of the 16 estimated trees, one of the survey years appears as a splitting variable. When survey year does occur as a splitting variable, it typically contributes a larger share of the explained variance: an average of at least 10% of additional variance is explained by inclusion of the survey year variable.

Nine of the 14 cases for which survey year appears in the finer scale regression tree models are in the Northern Coastal Plain and the Southern Coastal Plain survey units. The remaining cases arise in the Piedmont and Mountain survey units. These findings focus attention on the Coastal Plain, suggesting that a significant shift in forest conditions has occurred in this subregion. The Coastal Plain contains the largest share of intensively managed pine forests and active management. These findings suggest a need for further research into the details of forest dynamics in this subregion and what these changes might portend for forest management and production in the future.

Conclusions

Forest assessments are undertaken to evaluate changes in forest conditions over time. We set out to use forest inventories to distinguish between changes that can be attributed to general development cycles of the resource and those that arise from structural changes in production relationships. With the increasing focus on changing climate and the role that forests play in mitigating carbon emissions, it is important to investigate whether forests are undergoing structural changes and what, if anything, these changes may imply for the long-term sustainability of forest ecosystems.

These methods provide a useful initial approach for a systematic comparison of forest inventories. We found through our initial modeling efforts that regression trees clearly flag where key variables were coded or measured differently between inventories (e.g., site index). The collection of results indicates where in the State of North Carolina structural stability between inventories is less tenable: the Coastal Plain survey units. Overall the results show no strong evidence of a structural shift in measured inventories with the change from the plot design used for periodic inventories and that used for the new continuous inventory and therefore provide support for the comparability of inventories generated by the two designs.

Structural variables were used to characterize important attributes of the dynamic forest ecosystems, and results provide insight into what variables are important to track when these productivity measures are projected: broad age groupings, forest management classes, and site quality. We test for structural changes in the forest inventory over time using an approach that does not impose an a priori functional form on the data. As a result, our findings of structural change do not point to causative factors; they simply indicate shifts in the distributions of certain characteristics. This approach can therefore be used to define a set of research questions that deserve additional scrutiny to test hypotheses regarding the factors driving these changes.

The strength of this approach to testing for structural changes is that it can be readily applied to any two inventories, without additional specification. In addition, it provides useful insights into the state variables that explain variation in a set of structural variables without appealing to a restrictive functional form. It therefore provides insights into what variables are most important for explaining change and predicting change in forest inventories.

Our approach, although it does not yield a single measure of structural change, does provide a suite of indicators of where and for which forest conditions structural change might be indicated. This information could be used to design a more detailed (and highly specified) research strategy for isolating the causes of change.

The choice of structural variables used in this type of analysis is critical. Our tests pivot off of a definition of structural change based on structural variables describing growing stock measures of two attributes of the forest state, i.e., volume and number of trees. Other perspectives are possible, e.g., those that more directly depict change in forest conditions such as growth, so the preceding comparisons of inventories apply only to the structural variables selected for this analysis. Analysis of alternate structural measures is a potential next step in this process. Furthermore, data ancillary to the FIA inventories, such as temperature and precipitation, might prove useful for explaining additional variance in the structural variables.

Literature Cited

ANDERSEN, M.C., J.M. WATTS, J.E. FREILICH, S. R. YOOL, G.I. WAKEFIELD, J.F. MCCAULEY, AND P.B. FAHNESTOCK. 2000. Regression-tree modeling of desert tortoise habitat in the central Mojave desert. *Ecol. Applic.* 10(3):890–900.

BECHTOLD, W.A., AND C.T. SCOTT. 2005. *The forest inventory and analysis plot design*. US For. Serv. Gen. Tech. Rep. SRS-80. South. Res. Stn., Asheville, NC. 37–52.

BREIMAN, L., J.H. FRIEDMAN, R.A. OLSHEN, AND C. J. STONE. 1984. *Classification and regression trees.* Wadsworth International Group: Belmont, CA. 358 p.

BROWN, M.J. 2004. *Forest statistics for North Carolina, 2002.* Resour. Bull. SRS-88. US Forest Service South. Res. Stn., Asheville, NC. 78 p.

BYLER, J.W. M.A. MARSDEN, AND S.K. HAGLE. 1990. The probability of root disease on the Lolo National Forest, Montana. *Can. J. For. Res.* 20:987–994.

CONNER, R., AND A. J. HARTSELL. 2002. Forest area and conditions. In Southern forest resource assessment, Wear, D. N. and Greis, J.G. (eds.). US For. Serv. Gen. Tech. Rep. SRS-53. South. Res. Stn., Asheville, NC. 635 p.

DE'ATH, G., AND K.E. FABRICIUS. 2000. Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology* 81(11):3178–3192.

DIXON, R.K., S. BROWN, R.A. HOUGHTON, A.M. SOLOMON, M.C. TREXLER, AND J. WISNIEWSKI. 1994. Carbon pools and flux of global forest ecosystems. *Science* 263(14):185–190.

FRANKLIN, J.F., T.A. SPIES, R. VAN PELT, A.B. CAREY, D.A. THORNBURGH, D. RAE BERG, D.B. LINDENMAYER, M.E. HARMON, W.S. KEETON, D.C. SHAW, K. BIBLE, AND J. CHEN. 2002. Disturbances and structural development of natural forest ecosystems with silvicultural implications, using Douglas-fir forests as an example. *For. Ecol. Manag.* 155:399–423.

IVERSON, L.R., AND A.M. PRASAD. 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. *Ecol. Monogr.* 68(4):465–485.

JOHNSON, T.G. 1991. *Forest statistics for North Carolina, 1990.* US For. Serv. Resour. Bull. SE-120. South. Res. Stn., Asheville, NC. 63 p.

LOEHLE, C., AND D. LEBLANC. 1996. Model-based assessments of climate change effects on forests: A critical review. *Ecol. Model.* 90:1–31.

MAGILL, A.H., J.D. ABER, G.M. BERNTSON, W.H. MCDOWELL, K.J. NADELHOFFER, J.M. MELILLO, AND P. STEUDLER. 2000. Long-term nitrogen additions and nitrogen saturation in two temperate forests. *Ecosystems* 3(3):238–253.

MICHAELSEN, J., D.S. SCHIMEL, M.A. FRIEDL, F.W. DAVIS, AND R.C. DUBAYAH. 1994. Regression tree analysis of satellite and terrain data to guide vegetation sampling and surveys. *J. Veg. Sci.* 5:673–686.

MILES, P.D., G.J. BRAND, C. L. ALERICH, L. F. BEDNAR, S.W. WOUDENBERG, J.F. GLOVER, AND E.N. EZZELL. 2001. *The forest inventory and analysis database: Database description and users manual version 1.0.* US For. Serv. Gen. Tech. Rep. NC-218. North Central Res. Stn., St. Paul, MN. 130 p.

R DEVELOPMENT CORE TEAM. 2005. r: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at www.R-project.org; last accessed Jun. 8, 2009.

RAO, P.V. 1998. *Statistical research methods in the life sciences.* Duxbury Press, North Scituate, MA. 889 p.

SHEFFIELD, R.M., AND H.A. KNIGHT. 1986. *North Carolina's forests.* US For. Serv. Resour. Bull. SE-88. South. Res. Stn., Asheville, NC. 97 p.

SHUGART, H.H. 2003. *A theory of forest dynamics: The ecological implications of forest succession models.* Blackburn Press, Caldwell, NJ. 296 p.

THERNEAU, T.M., AND E. J. ATKINSON. 1997. *An introduction to recursive partitioning using RPART routines.* Tech. Rep. Ser. no. 61. Department of Health Science Research, Mayo Clinic, Rochester, MN. 52 p.

THERNEAU, T.M., AND E.J. ATKINSON.2005. *R port by Ripley, B. The rpart package*. R package version 3.1-23. R Foundation for Statistical Computing, Vienna, Austria.

VAYSSIÈRES, M.P., R.E. PLANT, AND B.H. ALLEN-DIAZ. 2000. Classification trees: An alternative non-parametric approach for predicting species distributions. *J. Veg. Sci.* 11:679–694.

WEAR, D.N. 2002. Land use. In *Southern forest resource assessment,* Wear, D.N., and J.G. Greis (eds.). US For. Serv. Gen. Tech. Rep. SRS-53. South. Res. Stn., Asheville, NC. 635 p.