## Proceedings of the

# 2001 Southern Mensurationists' Conference

*Edited by*
Paul F. Doruska
Don C. Bragg

*Sponsored by*
Arkansas Forest Resources Center
Society of American Foresters Biometrics Working Group
University of Arkansas, School of Forest Resources
University of Georgia, D.B. Warnell School of Forest Resources
USDA Forest Service, Southern Research Station

# Proceedings of the 2001 Southern Mensurationists' Conference

**November 4-6, 2001**
**Chattanooga, Tennessee**

***Edited by***
Paul F. Doruska
Don C. Bragg

***Hosted by***
Arkansas Forest Resources Center
University of Arkansas, School of Forest Resources
USDA Forest Service, Southern Research Station

***Sponsored by***
Arkansas Forest Resources Center
Society of American Foresters Biometrics Working Group
University of Arkansas, School of Forest Resources
University of Georgia, D.B. Warnell School of Forest Resources
USDA Forest Service, Southern Research Station

***Published by***
Arkansas Forest Resources Center
Monticello, Arkansas
March 2002

## FOREWORD

The 2001 Southern Mensurationists' Conference was held in scenic Chattanooga, Tennessee at the historical Chattanooga Choo Choo Holiday Inn. This conference was the latest in the series of annual gatherings of southern biometricians, and attracted speakers and participants from Texas to Virginia. A variety of papers were presented and the quaint atmosphere, as always, led to many fruitful discussions.

Further details about the Southern Mensurationists and other related organizations can be found at the following URL:

http://www.mensurationists.com


## ACKNOWLEDGMENTS

We would like to personally thank the Conference Steering Committee for their support during the development of the meeting, with special recognition going to Chris Cieszewski and Mike Strub. The Arkansas Forest Resources Center, the University of Georgia Warnell School of Forest Resources, and the USDA Forest Service Southern Research Station graciously provided the equipment used during the meeting.



Paul F. Doruska
Arkansas Forest Resources Center
Monticello, Arkansas

Don C. Bragg
USDA Forest Service
Monticello, Arkansas

## TABLE OF CONTENTS

# SOMA 2001 Agenda

**November 4-6, 2001**
**Chattanooga Choo Choo Holiday Inn**
**Chattanooga, TN**

**SUNDAY, NOVEMBER 4, 2001**

| | |
|---|---|
| 6:30pm - 9:00pm | Ice Breaker Reception |

**MONDAY, NOVEMBER 5, 2001**

| | |
|---|---|
| 7:00am – 8:15am | Continental Breakfast |
| 8:15am – 8: 20am | Welcome: Paul Doruska, University of Arkansas, Chair SOMA 2001 |

**Session 1: Moderator, Chris Cieszewski, University of Georgia**
**Location: Town Hall Theatre**

| | |
|---|---|
| 8:20am – 9:00am | Revised Base-Age Invariant Height Growth Model for Loblolly Pine Based on Intensively Managed Plantation Data. William M. Harrison*, and Chris Cieszewski, University of Georgia |
| 9:00am – 9:40am | Fitting Criterion for Site Index Curves. Mike Strub, Weyerhaeuser Company, Hot Springs, AR |
| 9:40am – 10:00am | Break |
| 10:00am – 10:40am | Preliminary Evaluation of Tree Indicator Species for Classifying Forest Site Productivity. Henry McNab*, F. Thomas Lloyd, and David L. Loftis, US Forest Service, Southern Research Station, Asheville, NC |
| 10:40am – 11:20am | Feasability of Miniature Plantations as Experimental Tools. Ralph Amateis*, Mahadev Sharma, and Harold Burkhart, Virginia Polytechnic Institute and State University |
| 11:30am – 1:00pm | Luncheon |

**Session 2: Moderator, Paul Doruska, University of Arkansas**
**Location: Town Hall Theatre**

| | |
|---|---|
| 1:00pm – 1:40pm | Long-term Sustainability Analysis of Fiber Supply in Georgia. Chris Cieszewski*, Michal Zasada, Roger Lowe, Bruce Borders, Mike Clutter, and Dick Daniels, University of Georgia |
| 1:40pm – 2:20pm | Large-Scale Estate Modeling in Georgia. M. Zasada*, C.J. Cieszewski, R. Lowe, B.E. Borders, M.L. Clutter, R.F. Daniels, University of Georgia |
| 2:20pm – 2:40pm | Break |

| | |
|---|---|
| 2:40pm – 3:20pm | Current Status of Southern Annual Inventories. Greg Reams, USDA Forest Service Southern Research Station, Asheville, NC |
| 3:20pm – 4:00pm | Modifying the Daniels index to indicate when competition begins, Philip Radtke, James Westfall, and Harold Burkhart, Virginia Polytechnic Institute and State University |
| 4:00pm – 4:30 pm | Business Meeting |
| 5:00pm – 6:30pm | Dinner |

**TUESDAY, NOVEMBER 6, 2001**

| | |
|---|---|
| 7:00am – 8:15am | Breakfast |

**Session 3: Moderator, Mike Strub, Weyerhaeuser Company**

**Location: Town Hall Theatre**

| | |
|---|---|
| 8:20am – 9:00am | Planting Rectangularity: How Much Is Too Much? Mahadev Sharma* and Ralph A. Amateis, Virginia Polytechnic Institute and State University |
| 9:00am – 9:40am | Using Nonparametric Regression to Examine Taper of Irregular Logs. Paul Doruska*, University of Arkansas |
| 9:40am – 10:00am | Break (refreshments in or around the Town Hall Theatre) |
| 10:00am – 10:40am | Fuzzy Set Classification of Old-Growth Southern Pine. Don Bragg*, USDA Forest Service Southern Research Station, Monticello, AR |
| 10:40am – 11:30am | Roundtable Discussion: Southern Forest Mensuration Needs/Issues<br>Moderators: Chris Cieszewski, University of Georgia<br>Paul Doruska, University of Arkansas<br>Mike Strub, Weyerhaeuser Company |
| 11:30am – 1:00pm | Luncheon |
| 1:00pm | Conference adjourned |

# 2001 Southern Mensurationists' Conference

**190 year old shortleaf pine cut in 1943 in Ashley County, Arkansas**
**Butt log was 50 inches in diameter and scaled 1,350 board feet**
(Picture courtesy of the Crossett Public Library)

# Papers and Abstracts

# Updated Base Age Invariant Height Growth Models for Loblolly Pine (*Pinus taeda* L.) Based on Intensively Managed Plantation Data

by

**W.M. Harrison[1], C.J. Cieszewski[2], S.W. Martin[3], M. Zasada[4, 6], and B.E. Borders[5]**

**Abstract**

This is an updated report on an earlier study that has been revisited after additional measurements were taken on the considered permanent sample plots. Thus, much of the material in this report replicates the original publication intended to demonstrate statistically robust methods for site-dependent height model development. The choice of form of the model is treated in this study as a secondary issue, while the main focus is on the conceptual aspect of parameter estimation for self-referencing models, such as site index models. The data considered came from the Consortium for Accelerated Pine Production Studies (CAPPS) at the Warnell School of Forest Resources. In this paper, we present two practical approaches to base-age invariant and non-biased parameter estimation and compare the results to those obtained from base-age dependent methods. Unlike the traditional methods, the base age invariant techniques do not violate regression assumptions and produce identical results regardless of the choice of base age or applied algorithm.

## INTRODUCTION

Heights of dominant and codominant trees of southern pines are little affected by a wide range of stand densities. Therefore, site quality estimation procedures based on stand height data are the most commonly used techniques for evaluating site productivity (Clutter et al. 1983). Most of these height-based site quality evaluation techniques involve the development of site index curves. Each site index curve defines an expected height/age relationship referenced by the expected height at a specified base age.

In general, the most desirable data for the development of site index curves come from remeasured permanent sample plots or from stem analysis. Both of these data sources provide a number of observed height/age pairs for a given location and allow the ultimate flexibility in model forms and estimation techniques. Remeasurements or sectioned stem data are usually combined by plot, resulting in an average height/age relationship for a given location. In order to reference the resulting curves by the height at a given age (site index), a choice of base age must then be made. This approach requires that site index be included in the model prior to estimation and makes the parameter estimates unique to the preselected base age (Bailey and Clutter 1974).

Having identified this problem, Bailey and Clutter (1974) proceeded to describe a base age invariant site index estimation technique using the following general model:

$$\log (H) = a_i + b(1/A)^c \quad (i = 1, 2, \ldots, m)$$

where:

$a_I$ = a parameter specific to the $i^{th}$ site,
$b$ = a common slope parameter, and
$c > 0$ is a linearization parameter.

Traditional methods for fitting site index equations involve definition of the site-specific parameter, $a_i$, in terms of the height at a given base age (site index). The estimate of the site-specific parameter, therefore, is based on a single height/age pair for a given plot. This causes the measurement error associated with that data point to be heavily weighted in the fitting process. If, however, we actually fit the site-specific parameter to all of the height/age data for each plot, the resulting model will be unbiased and the parameter estimates will be stable regardless of the choice of base age. Bailey and Clutter recognized this fact, but the computing resources necessary to carry out this fitting technique on intrinsically nonlinear model forms were unavailable in 1974.

---

[1]Program Coordinator, [2]Assistant Professor, [4]Postdoctoral Fellow, [5]Professor, D.B. Warnell School of Forest Resources, University of Georgia, Athens, GA 30602, USA
[3]Biometrician/SAS Programmer, Dyntel Contractor, Centers for Disease Control and Prevention, Data Management Division/SAB, Atlanta, GA 30333, USA
[6]Assistant Professor, Faculty of Forestry, Agricultural University, 02-528 Warsaw, Poland

*Citation for proceedings:* P.F. Doruska and D.C. Bragg (editors). 2002. Proceedings of the Southern Mensurationists' Conference, Chattanooga, TN, November 4-6, 2001. Arkansas Forest Resources Center, Monticello, AR. 43 p.

The purpose of this short report is to describe two methods for fitting linear or nonlinear, base age invariant site index equations using PC-SAS. Both approaches were evaluated using three different choices of base age. The resulting models were compared to site index models fit using traditional techniques.

## DATA AND METHODS

The data for this study came from a series of loblolly pine (*Pinus taeda* L.) growth and yield plots established at six locations across the state of Georgia. The main objective of the study was to provide real growth series data for loblolly pine plantations managed under various silvicultural treatment regimes. Study sites were located near weather monitoring stations so that the effects of climate and atmospheric pollution factors on pine growth could be evaluated and modeled where appropriate (Borders and Bailey 1997). The original study plan called for two complete blocks, each containing four 3/8-acre treatment plots. One of the following treatments was applied to each plot:

- H – Use herbicides to control all competing vegetation throughout the life of the study.

- F – Fertilize as follows: First two growing seasons, 250 lbs/ac DAP + 100 lbs/ac KCL in the spring and 50 lbs/ac of ammonium nitrate in mid summer. During each subsequent growing season, 50 lbs/ac of ammonium nitrate were applied in early to mid summer.

- HF – Both H and F treatments.

- C – Control treatment, no treatment following initial site preparation.

In addition, the treatment plots at each location were replicated at various points in time. The first series of plots was established near Waycross, GA in 1987. Additional installations were established at the same location in 1989 and 1993. All plots have been measured annually since the first growing season.

To develop the methodology to fit base age invariant site index equations, a single treatment was chosen from the four. This eliminated the need to model any treatment effects in addition to the height/age relationship. The vegetation control treatment (H) was chosen for this exercise. As of 2001, this resulted in 26 plots and 253 total observations.

The Chapman-Richards equation has been successfully used to model height/age relationships for southern pines (Newberry and Pienaar 1978, Pienaar and Shiver 1980). Two forms of the model were used in this study:

<u>Projection</u>

$$Hd_2 = Hd_1\left[\frac{1-\exp(-\alpha A_2)}{1-\exp(-\alpha A_1)}\right]^{\beta} \qquad (1)$$

<u>Site Index</u>

$$Hd = S\left[\frac{1-\exp(-\alpha A)}{1-\exp(-\alpha A_b)}\right]^{\beta} \qquad (2)$$

where:

$Hd_i$ = average dominant/codominant height at time $i$,
$A_i$ = age at time $i$,
$S$ = site index, defined as $Hd$ at any given base age $A_b$,
$\alpha,\beta$ = parameters estimated from data.

## MODELING APPROACH I

The first approach to fitting the base age invariant site index equation uses an iterative procedure to fit the global and site-specific parameters. The procedure begins by fitting the global parameters (α and β) in equation (2), using the observed site index ($S$) for each plot. The site index in this case is defined as the average dominant/codominant height at any given base age. In the second step, estimates of the global parameters are used as constants and the site-specific parameter ($S$) is estimated for each plot. The observed $S$ values for each plot are used as starting values for the fitting procedure. Next, the estimated $S$ values for each plot become the "observed" values and the global parameters are refit. This procedure is repeated until the global parameters stabilize. Figure 1 illustrates the iterative procedure with a flow chart.

The initial data setup to implement the iterative procedure is straightforward. First, the base age must be chosen and the average dominant/codominant height for a given plot at the base age is defined as the site index ($S_i$). The site index value must be placed on each observation for each plot. The rest of the data consists of a plot identifier, age and the average dominant/codominant height at that age. Table 1 shows an example dataset for the iterative procedure.

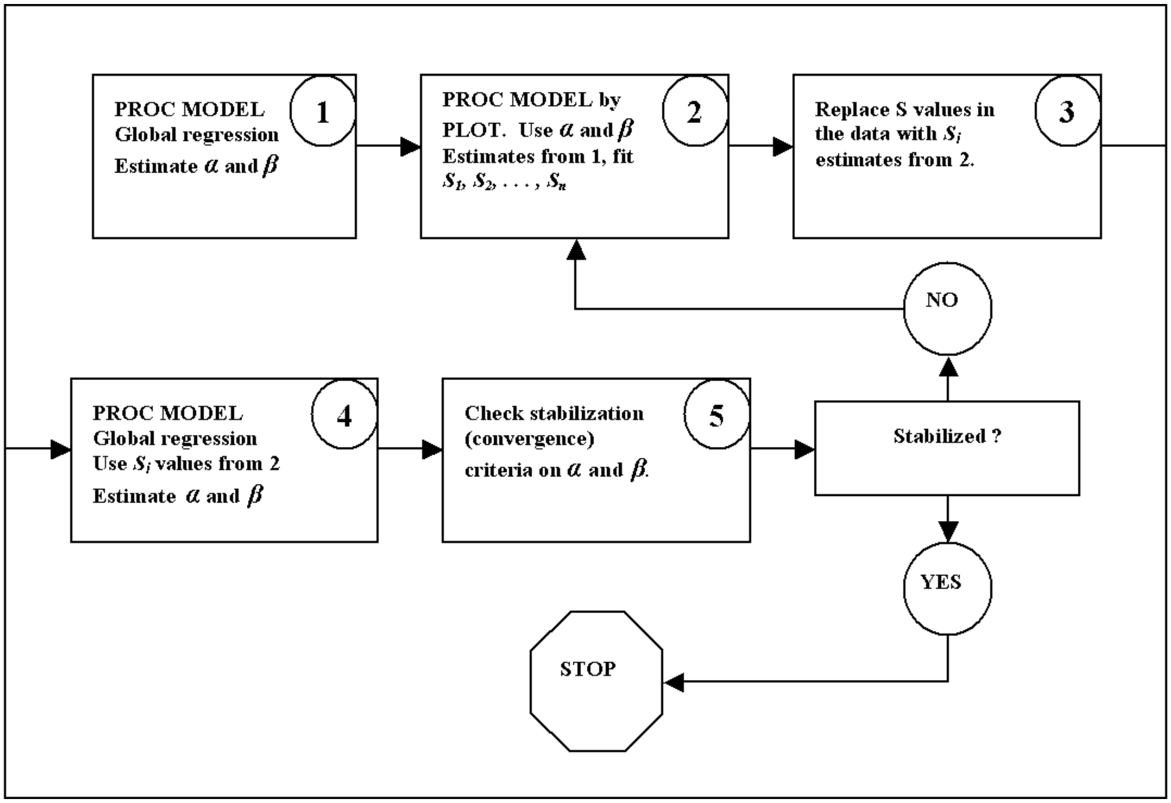



**Figure 1.** Iterative procedure for fitting base age invariant site index equation using PC-SAS.

**Table 1.** Example data setup for the iterative fitting procedure.

| Plot | Age | Hd | S |
|---|---|---|---|
| 1 | 3 | 14 | 49 |
| 1 | 4 | 18 | 49 |
| 1 | 5 | 24 | 49 |
| 1 | 6 | 28 | 49 |
| . | . | . | . |
| . | . | . | . |
| 1 | 10 | 49 | 49 |
| 2 | 3 | 12 | 46 |
| 2 | 4 | 17 | 46 |
| 2 | 5 | 22 | 46 |
| . | . | . | . |
| . | . | . | . |
| 2 | 10 | 46 | 46 |

## MODELING APPROACH II

The second approach to fitting a base age invariant site index equation is a dummy variable procedure that runs in a single PROC MODEL (SAS Institute 1993) step. The following model was fit:

$$Hd = (S_1 p_1 + S_2 p_2 + \ldots + S_n p_n)\left[\frac{1-\exp(-\alpha A)}{1-\exp(-\alpha A_b)}\right]^{\beta} \qquad (3)$$

where: $S_i$, = site-specific parameter for plot *i*,
$p_i$ = dummy variable; = 1 for plot i;
= 0 otherwise,
$A_b$ = site index base age,
all others previously defined.

The first term, containing a site-specific parameter and a dummy variable for each plot, collapses into a single parameter in the fitting process. The observed site index values at the specified base age are used as starting values for the $S_i$ parameters for each plot. The starting values are loaded via the ESTDATA dataset in PROC MODEL. This dataset consists of a single observation containing the starting values for each parameter. The order of the elements in the dataset is determined by order of parameters in the PARMS statement in PROC MODEL. Table 2 shows an example of the input data setup for the dummy variable procedure. Table 3 illustrates the ESTDATA dataset.

**Table 2.** Example data setup for the dummy variable fitting procedure.

| Plot | Age | Hd | $p_1$ | $p_2$ | $p_3$ . . | $p_n$ |
|---|---|---|---|---|---|---|
| 1 | 3 | 14 | 1 | 0 | 0 | 0 |
| 1 | 4 | 18 | 1 | 0 | 0 | 0 |
| 1 | 5 | 24 | 1 | 0 | 0 | 0 |
| 1 | 6 | 28 | 1 | 0 | 0 | 0 |
| . | . | . | . | . | . | . |
| . | . | . | . | . | . | . |
| 1 | 10 | 49 | 1 | 0 | 0 | 0 |
| 2 | 3 | 12 | 0 | 1 | 0 | 0 |
| 2 | 4 | 17 | 0 | 1 | 0 | 0 |
| 2 | 5 | 22 | 0 | 1 | 0 | 0 |
| . | . | . | . | . | . | . |
| . | . | . | . | . | . | . |
| 2 | 10 | 46 | 0 | 1 | 0 | 0 |

**Table 3.** Example setup for the ESTDATA dataset for the dummy variable fitting procedure.

| $\alpha$ | $\beta$ | $S_1$ | $S_2$ | $S_3$ . . . | $S_n$ |
|---|---|---|---|---|---|
| 0.07 | 1.4 | 49 | 46 | 52 . . . | 44 |

**RESULTS**

To ensure the consistency of results, the convergence criteria in PROC MODEL (FIT statement) was changed to 0.0000001. This setting increases the run time for both procedures, but helps avoid converging to a local minimum. Using both the iterative and dummy variable procedures with base age site indexes at 3, 5 and 10 years, the results were identical in all cases. The following model was obtained:

$$Hd = S\left[\frac{1-\exp(-0.10372A)}{1-\exp(-0.10372A_b)}\right]^{1.51558} \qquad (4)$$

$$\text{MSE} = 5.19;\ R^2 = 0.97$$

Figure 2 shows the height growth curve implied by equation (4) with a base age 10 site index of 45.47 feet. The observed average dominant and codominant heights are also shown.

For comparison purposes, equation (2) was fit to the same 253 observations using the ordinary, nonlinear least squares technique in PROC MODEL. The following models were obtained with site indexes defined at base ages 10, 5, and 3 years:

<u>Base age 10</u>

$$Hd = S\left[\frac{1-\exp(-0.105578A)}{1-\exp(-0.105578A_b)}\right]^{1.522201} \qquad (5)$$

$$\text{MSE} = 5.75;\ R^2 = 0.96$$

<u>Base age 5</u>

$$Hd = S\left[\frac{1-\exp(-0.103756A)}{1-\exp(-0.103756A_b)}\right]^{1.293266} \qquad (6)$$

$$\text{MSE} = 15.38;\ R^2 = 0.90$$

<u>Base age 3</u>

$$Hd = S\left[\frac{1-\exp(-0.014909A)}{1-\exp(-0.014909A_b)}\right]^{1.100903} \qquad (7)$$

$$\text{MSE} = 26.90;\ R^2 = 0.83$$

Figure 3 shows the height growth curves implied by equations (4), (5), (6) and (7) with a base age of 10 and a site index of 46 feet.

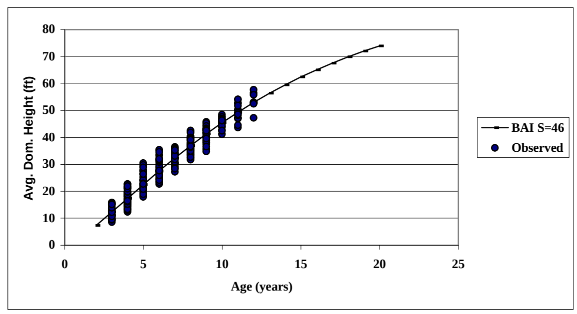



**Figure 2.** Predicted average dominant and codominant height growth curve implied by equation (4) and observed height measurements.

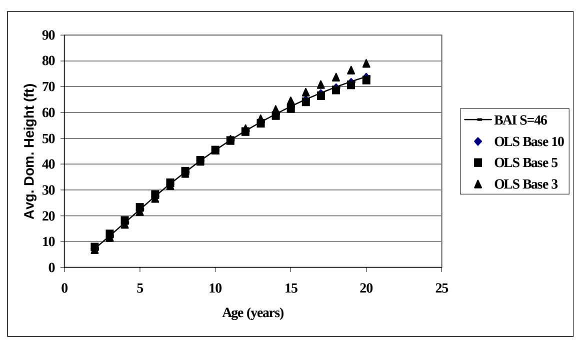



**Figure 3.** Predicted average dominant and codominant height growth curves for models fit with the base age invariant technique and the ordinary least squares technique for three different base ages.

**CONCLUSIONS**

The base age invariant site index equation, as first identified by Bailey and Clutter (1974), can be successfully fit with sufficient data and fairly straightforward programming with PROC MODEL and the SAS Macro Language (SAS Institute 1990). The data must consist of repeated average dominant/codominant height measurements on some number of plots. It is unclear as to how many measurements are required but, with this method, all of the data can be utilized. There is no need to make any arbitrary choice regarding measurement intervals. In this exercise, we chose plots that received the same silvicultural treatment regime to avoid the need to model treatment response in addition to the basic height growth pattern.

Compared to site index equations fit with ordinary, nonlinear least squares, the base age invariant model is more consistent. The traditional technique requires an arbitrary choice of base age prior to fitting the model. The model is then forced through the chosen height/age point. In the base age invariant technique, we recognize that each measurement is made with error and, therefore, it seems unreasonable to force the model through any given measurement. Instead, the curve is fit to the observed trend in the data.

This work is preliminary in that the choice of data may not have been ideal. We would like to evaluate the method on data with older ages and more inherent site variation. We would also like to evaluate the fitting technique with other dependent variables such as per-acre basal area, survival and per-acre yield. The ultimate application may involve using the technique to fit a system of seemingly unrelated equations with more than one site-specific parameter.

**LITERATURE CITED**

Bailey, R.L. and J.L. Clutter. 1974. Base-age invariant polymorphic site curves. For. Sci. 20(2):155-159.

Borders, B.E. and R.L. Bailey. 1997. Loblolly pine - pushing the limits of growth. Univ. of Ga. Consortium on Accelerated Pine Production Studies. Tech. Rep. CAPPS 1997-1.

Newberry, J.D. and L.V. Pienaar. 1978. Dominant height growth models and site index curves for site-prepared slash pine plantations in the lower coastal plain of Georgia and North Florida. Univ. of Ga. Plantation Man. Res. Coop. Res. Paper No. 4.

Pienaar, L.V. and B.D. Shiver. 1980. Dominant height growth and site index curves for loblolly pine plantations in the Carolina flatwoods. So. J. Appl. For. 4:55-59.

SAS Institute. 1990. SAS Guide to Macro Processing, Version 6, Second Edition. SAS Institute, Inc. Cary, NC.

SAS Institute. 1993. SAS/ETS User’s Guide, Version 6, Second Edition. SAS Institute, Inc. Cary, NC.

**Fitting Criterion for Site Index Curves**

by

**Mike Strub**[1]

Weyerhaeuser Company, Hot Springs, AR

**Abstract**

Three criterion for fitting site index curves are examined. The criterion involve minimizing error sums of squares in height, height growth, and site index. The impact on residuals and curve shape is examined for each fitting criterion. A mixture of site and height criterion gives curves that are unbiased over a wide range of ages.

[1] *Citation for proceedings:* P.F. Doruska and D.C. Bragg (editors). 2002. Proceedings of the Southern Mensurationists' Conference, Chattanooga, TN, November 4-6, 2001. Arkansas Forest Resource Center, Monticello, AR. 43 p.

**Preliminary Evaluation of Methods for Classifying Forest Site Productivity Based on Species Composition in Western North Carolina**

by
**W. Henry McNab[1], F. Thomas Lloyd[1], and David L. Loftis[2]**

**Abstract**

The species indicator approach to forest site classification was evaluated for 210 relatively undisturbed plots established by the USDA Forest Service Forest Inventory and Analysis unit (FIA) in western North Carolina. Plots were classified by low, medium, and high levels of productivity based on 10-year individual tree basal area increment data standardized for initial stocking. Chi-square analysis of contingency tables indicated that productivity classes were not independent ($P < 0.05$) of the frequencies of occurrence for 4 of 27 common tree species. Multiple logistic regression of a binary variable formed by the high productivity class compared to the combined low and medium classes resulted in a model consisting of elevation and seven significant ($P < 0.05$) species that produced a classification accuracy of 85 percent; a similar model based on the low productivity class resulted in classification accuracy of 70 percent. A multinomial logistic regression model indicated that elevation and six species were significantly ($P < 0.05$) associated with the three productivity classes, but overall classification accuracy dropped to 61 percent, mainly due to the poor predictability of low productivity classes. Chestnut oak (*Quercus prinus*) and serviceberry (*Amelanchier* spp.) were the most consistent indicator species. Results of this exploratory study suggest that using indicator species for site classification shows promise in hardwood stands by avoiding problems associated with conventional methods based on site index.

## INTRODUCTION

Forest productivity evaluation based on indicator species -- where the presence of certain vegetative species is associated with the rate of tree growth on forestland -- has received relatively little attention in the United States (Daubenmire 1976). Indicator species integrate the complex array of forest environmental components important for tree growth and their presence can used as a phytometer to conveniently assess productivity (Kimmins 1987). Vegetative composition is the basis of the habitat type approach to site classification in much of the arid western United States (Daubenmire 1976), but elsewhere other methods are generally used to evaluate forest productivity (Carmean 1975).

Site index, the most widely used method of evaluating forest productivity, is also based on the phytometer premise (Carmean 1975, Spurr and Barnes 1984), but requires the acceptance of a number of underlying assumptions that are typically unknown for the subject stand (Beck and Trousdel 1973). Using site index is problematic in hardwood stands, which tend to be many-aged and consist of mixed species. Determining the age and height of sample trees in these stands is laborious and prone to measurement error; in addition, site index relationships are typically based on simple guide-curve relationships that may be inaccurate (Carmean 1975). Adopting indicator species for site classification could prove particularly useful in growth and yield equations that use tree lists to drive the models, such as the forest vegetation simulator (Teck et al. 1996), because plot inventory data could also be used for site quality determination.

Several problems with using indicator species for site quality assessment soon become apparent. Among the most important of these is the paucity of quantitative information on the physiological requirements of different tree species. A considerable body of qualitative information exists on vegetation and environmental associations in the southern Appalachian Mountains, and occurrence of species has been frequently used to classify vegetative associations on forest sites (Whittaker 1966, Mowbray and Oosting 1968, Golden 1974, Callaway et al. 1989). Lacking, however, are quantitative relationships -- based on the presence (or absence) of multiple tree species -- that could be used to assess site productivity for management purposes.

Additional site classification problems include how to determine measurement units of forest productivity and the number of categories to use in

[1] Research Forester, USDA Forest Service, Southern Forest Experiment Station, Bent Creek Experimental Forest, Asheville, NC 28806
[2] Project Leader, USDA Forest Service, Southern Forest Experiment Station, Bent Creek Experimental Forest, Asheville, NC 28806

*Citation for proceedings:* P.F. Doruska and D.C. Bragg (editors). 2002. Proceedings of the Southern Mensurationists' Conference, Chattanooga, TN, November 4-6, 2001. Arkansas Forest Resources Center, Monticello, AR. 43 p.

assessing productivity. Data from periodic inventories of sample plots probably includes only tree diameter and height by species, which allows calculation of change in tree size as a measure of productivity. Because productivity depends on the number of trees present, plots must be adjusted for variation in stocking levels. Comparison of measured plot increments across many sites is facilitated by using a measure of stocking at the beginning of the inventory period as a method of correcting for uncontrolled variation. Associated with the question of stocking levels is the related question of using a two-dimensional (basal area) or three-dimensional metric (bole volume) as a measure of site productivity.

Last, but equally important, is the question of the appropriate number of productivity classes to use. Two broad categories of site quality (e.g. good or poor) is the smallest number that can be recognized, and depending on limits imposed can provide information needed for management decisions (Beck and Della-Bianca 1981). For example, instead of using the median of the population of sample plot productivities to divide good sites from poor, we could identify only the prime sites for management, those in the upper 25 percentile. Conversely, the lower 25 percentile of sites from the population of plots could be identified as cull and excluded from management practices. Alternatively, we could identify three classes of productivity; low, medium, or high. This logic could be extended to identify perhaps five or more classes, but at some point the ability to use species to discriminate among productivity levels would be reached with corresponding increase of classification accuracy.

Because the information available on the three components of the indicator species problem discussed above is sparse, we decided to conduct a pilot study to collect preliminary data. Our study focused on finding tree species that showed fidelity and constancy with site quality that would have high value as indicators of forest productivity. The primary objective of our study was to investigate how well species composition and selected environmental variables could distinguish among broad categories of forest site productivity. We also planned to evaluate an experimental metric for controlling variation in stocking levels among plots and to examine classification accuracy with varying numbers of productivity categories. The scope of our study was limited to sites in western North Carolina.

**METHODS**

Using the Eastwide Forest Inventory Data Base, we obtained FIA data for 1974 and 1984 for the 21 predominantly mountainous counties in western North Carolina that we used as our study area (Hansen et al. 1992). The eastern border of the study area was formed by the following counties, moving from northeast to southwest: Alleghany, Wilkes, Caldwell, Burke, McDowell, Buncombe, and Henderson. Ecoregions of the border counties include the foothills of the Appalachian piedmont, the Blue Ridge escarpment, and Blue Ridge Mountains. Elevations range from about 1000 ft in the Yadkin River valley to over 6600 ft in the Black Mountains.

The region's climate is primarily oceanic, is characterized by short, mild winters and long, warm summers: temperature averages range from 36$^{o}$F in January to 72$^{o}$F in July. The growing season averages 180 days, with average annual precipitation ranging from <40 to >90 in. Precipitation is well distributed annually and varies locally as a result of orographic influences and proximity to the escarpment. Geologic formations are predominantly highly weathered gneisses and schists of Precambrian age, but include localized rock units of horneblende gneiss that weather to form soils of higher pH and increased fertility. Soils are typically deep (>40 in) and predominately Ultisols in areas of gently sloping low hills and broad ridges; Inceptisols occur on steep mountain slopes and in the colluvia of coves. The forest canopy below about 5000 ft consists primarily of deciduous hardwoods, and is dominated by six species of oaks. More than 30 other tree species may be present, ranging in abundance from sparse to common depending on disturbance regimes and site conditions (Whittaker 1966, Mowbray and Oosting 1968, Golden 1974, Callaway et al. 1989).

We were able to obtain tree and plot level data for analysis. Tree data included initial and final periodic diameter at breast height (dbh) of arborescent species = 1-in dbh. Plot data consisted of four environmental variables: elevation (nearest 10 feet), aspect (classes of 45$^{o}$ azimuth), slope gradient (nearest percent), and an index of solar radiation (Golden 1974, 205). We started out by identifying 489 plots for analysis but we omitted 279 because they had been disturbed by logging, insects, or disease during the previous 10-year interval or were dominated by relatively uncommon species (i.e. red spruce, *Picea rubens*; yellow buckeye, *Aesculus octandra*). Unfortunately, shrub species -- previously shown to work well as indicator species (Spurr and Barnes 1984) -- were not included in the FIA dataset. Also, species had been pooled for three genera: (1) birches, sweet (*Betula lenta*); and yellow (*B. alleghaniensis*), (2) hickories, bitternut (*Carya cordiformis*); pignut (*C. glabra*); shagbark (*C. ovata*); and mockernut (*C. tomentosa*), and (3) serviceberries, downy (*Amelanchier arborea*) and Allegheny (*A.*

*laevis*). Taxonomic nomenclature follows Little (1953) and is presented for other species in Table 1.

Productivity (the net periodic change in the dimension of trees occupying an area) for each of the 210 plots is an index ($I_p$) calculated as periodic basal area increment of survivor trees per unit of 100 linear feet of cumulative tree circumference at 4.5 ft of all trees that are greater than 1 in dbh at the beginning of the period, expressed as

$$I_p = 0.1736(\sum_N (d_f^2 - d_i^2) / \sum_N d_i),$$

where $d_i$ and $d_f$ are diameters at the beginning and end of the period, respectively. The index is similar to that used by Lexen (1943) who found cumulative bole area a useful indicator of potential stand growth.

We arbitrarily subdivided the population of forest site productivities into three classes: low, medium, or high. The 25 percentile of the total 210 plots with the lowest levels of productivity were assigned to the low class, the 25 percentile of plots with highest productivity to the high class, with the remaining 50 percent of the plots assigned to the medium class. We used this method of allocating plots to three discrete classes of productivity to ensure that we had adequate numbers of samples in each class to do the analysis.

We also investigated a second method of analyzing productivity by recombining the three classes to form only two categories of site quality - poor and good. Poor sites consisted of plots in the low productivity class compared with others (i.e. pooled medium and high classes). Good sites consisted of sites in the high class compared with others (i.e. pooled medium and low classes). Stratification of productivity classes in this manner provided information for comparing the relative value of xerophytic and mesophytic species for classifying sites, and could provide additional information to managers for broadscale land management planning. For example, it might be desirable for a manager to expend limited resources for intensive practices only on good sites and practice custodial management of all other sites.

We used contingency tables and logistic regression to develop quantitative relationships of dependent variables -- productivity classes and site quality categories --with environmental variables and tree species. Independent variables consisted of the presence or absence of common species and the four environmental variables. To reduce the confounding influence of disturbance associated with certain species (e.g. yellow-poplar, red maple, sweet birch), we did not use a measure of abundance -- such as number of individuals of each species present or crown area by species -- as the independent variable. (Beck and Della-Bianca 1981, Golden 1974). Contingency tables were used to determine if species were independent of site productivity (assuming uniform expected frequency of species occurrence among the three classes). We used multiple logistic regression to evaluate the relationship of multiple independent variables with the two site quality categories: (1) poor versus others and (2) good versus others. We developed models so that the dependent variable resulted in a positive outcome, which affected interpretation of the sign of the coefficients. Logistic models were formulated using a stepwise, backward elimination method with a probability level of 0.05 required for retention of variables in the model. Classification tables of model performances were based on a cutpoint of $P_{(Y=1)} = 0.5$. We evaluated significance of the model coefficients with the Wald test at the $P = 0.05$ level (Hosmer and Lemeshow 2000). This is the conventional use of logistic regression, to predict two possible model outcomes.

We used multinomial logistic regression to simultaneously examine the relationship of species and site variables with the three classes of site productivity (i.e. low, medium, and high). Because STATA, our statistical software package (StataCorp 1999), did not allow a stepwise procedure for efficiently determining a suitable model with this type of analysis, we began initial trial formulations with influential species from the contingency table analysis and site quality models. We followed the rationale of Hosmer and Lemeshow (2000) to develop and interpret significant, parsimonious, multinomial models. The advantage of using multinomial logistic regression for three or more possible outcomes is that probabilities of multiple class membership sum to one. This is not possible when simply applying two separate multiple logistic models (e.g. one for poor sites and another for good sites) because their different formulations can allow non-mutually exclusive membership of sample plots. STATA uses maximum likelihood methods for developing logistic regression models.

**RESULTS**

Periodic 10-yr productivity ranged from 0.134 $ft^2$/ac/in to 1.218 $ft^2$/ac/in on the 210 sample plots and averaged 0.405 $ft^2$/ac/in (±0.188 s.d.) (Figure 1). The median of the distribution of plot productivities was 0.3661 $ft^2$/ac/in. A Shapiro-Wilk test of normality indicated that the distribution of sample plots was not normally distributed ($P < 0.001$), but was positively skewed ($P < 0.0001$) toward sites of high productivity. Periodic productivity was below average on 60 percent (127) of the 210 plots. Only 15 (7 percent) of the plots represented the upper half of the range in productivity (i.e. >0.405 $ft^2$/ac/in). The plot of lowest productivity (0.134 $ft^2$/ac/in) was occupied by pitch pine, red maple, black gum, and scarlet and chestnut oaks; species on the plot of highest productivity (1.218 $ft^2$/ac/in) consisted of yellow-poplar and black locust.

Elevations of the 210 plots averaged 2933 ft and ranged from 1000 ft to 5320 ft. Plot aspects were well distributed with 96 plots (46 percent) having a northerly exposure. Slope gradients averaged 43 percent and ranged from 5 to 95 percent. Correlations between productivity and the two continuous variables were small, but significant ($P < 0.05$) for elevation ($r = -0.17$) and gradient ($r = -0.18$).

**Single species association with site productivity**

The FIA data set included 59 species: 27 of these occurred on =10 plots and were retained for the analysis (Table 1). The hickory, birch, and serviceberry tree groups each consisted of more than one species. The overall (27 rows x 3 columns) contingency analysis indicated that species frequency on plots was not independent of classification by productivity ($P = 0.0001$). Chi-square tests of hypothesized ratios of species occurrence of 1:2:1 for sites of low, medium, and high productivity respectively indicated that chestnut oak, scarlet oak, and hickory were associated ($P < 0.05$) with low sites. Serviceberry was unusual among the 27 species in being associated ($P < 0.05$) with sites of both low and high productivities; however, it occurred on relatively few plots (6 percent). Twelve species were present on <10 percent of the plots. Red maple occurred on the highest proportion (52 percent) of the 210 plots, followed by chestnut oak (42 percent), yellow-poplar (38 percent), and flowering dogwood (36 percent).

**Table 1.** Observed frequency of species by site productivity classes on 210 FIA plots in western North Carolina.

| Species | Productivity class: Low | Medium | High | Chi-square probability | Total Plots | Percent of total |
|---|---|---|---|---|---|---|
| Chestnut oak (*Quercus prinus*) | 31[a] | 49[b] | 9[a] | 0.003 | 89 | 41.8 |
| Scarlet oak (*Quercus coccinea*) | 15 | 25 | 3 | 0.020 | 43 | 20.2 |
| Hickory (*Carya* spp.) | 17 | 41 | 7 | 0.023 | 65 | 30.5 |
| Serviceberry (*Amelanchier* spp.) | 6 | 2 | 4 | 0.045 | 12 | 5.6 |
| Black locust (*Robinia pseudoacacia*) | 7 | 24 | 19 | 0.054 | 50 | 23.5 |
| Yellow-poplar (*Liriodendron tulipifera*) | 13 | 39 | 28 | 0.059 | 80 | 37.6 |
| Red maple (*Acer rubrum*) | 26 | 65 | 20 | 0.142 | 111 | 52.1 |
| Black oak (*Quercus velutina*) | 9 | 22 | 4 | 0.154 | 35 | 16.4 |
| Sourwood (*Oxydendrum arboreum*) | 18 | 43 | 8 | 0.181 | 69 | 32.4 |
| White ash (*Fraxinus americana*) | 2 | 7 | 7 | 0.185 | 16 | 7.5 |
| Carolina silverbell (*Halesia carolina*) | 6 | 4 | 3 | 0.191 | 13 | 6.1 |
| Blackgum (*Nyssa sylvatica*) | 14 | 20 | 6 | 0.202 | 40 | 18.8 |
| Northern red oak (*Quercus rubra*) | 15 | 37 | 10 | 0.209 | 62 | 29.1 |
| Virginia pine (*Pinus virginiana*) | 2 | 10 | 7 | 0.261 | 19 | 8.9 |
| Cucumbertree (*Magnolia acuminata*) | 3 | 4 | 5 | 0.287 | 12 | 5.6 |
| Flowering dogwood (*Cornus florida*) | 22 | 40 | 14 | 0.388 | 76 | 35.7 |
| Pitch pine (*Pinus rigida*) | 9 | 11 | 5 | 0.440 | 25 | 11.7 |
| Sassafras (*Sassafras albidum*) | 3 | 9 | 2 | 0.526 | 14 | 6.5 |
| American basswood (*Tilia americana*) | 4 | 4 | 2 | 0.549 | 10 | 4.7 |
| Eastern white pine (*Pinus strobus*) | 8 | 18 | 12 | 0.632 | 38 | 17.8 |
| American beech (*Fagus grandifolia*) | 5 | 8 | 3 | 0.779 | 16 | 7.5 |
| Southern red oak (*Quercus falcata*) | 2 | 6 | 2 | 0.819 | 10 | 4.7 |
| Sugar maple (*Acer saccharum*) | 5 | 7 | 4 | 0.829 | 16 | 7.5 |
| Eastern hemlock (*Tsuga canadensis*) | 5 | 10 | 6 | 0.931 | 21 | 9.8 |
| Shortleaf pine (*Pinus echinata*) | 4 | 7 | 3 | 0.931 | 14 | 6.6 |
| White oak (*Quercus alba*) | 13 | 27 | 12 | 0.944 | 52 | 24.4 |
| Birch (*Betula* spp.) | 11 | 21 | 10 | 0.976 | 42 | 19.7 |

[a] Expected frequencies for the low and high productivity classes = total plots / 4.

[b] Expected frequency for the medium productivity class = total plots / 2.

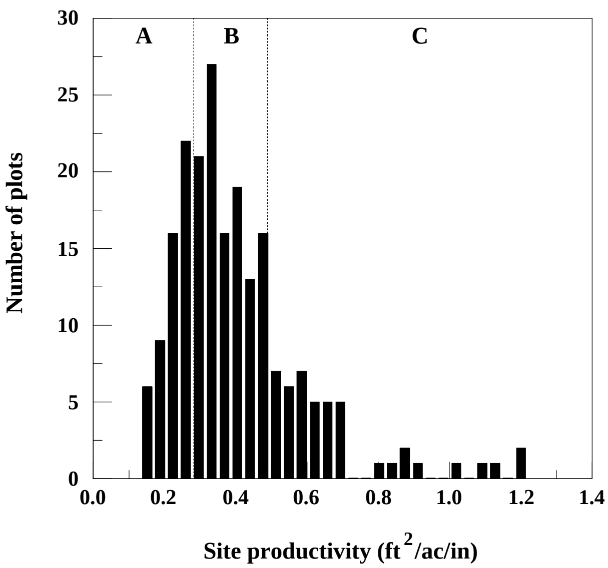



**Figure 1.** Distribution of the 210 sample plots by low (A), medium (B), and high (C) productivity classes. The low (<0.280 $ft^2$/ac/in) and high (>0.489 $ft^2$/ac/in) classes each consist of the lower and upper 25 percentile of plots ranked by productivity; the medium (=0.280 to =0.489 $ft^2$/ac/in) class consists of the remaining 50 percent of plots.

**Multiple species and two site quality categories**

Multiple logistic regression analysis indicated that the two categories of site quality (e.g. poor versus others, good versus others) were associated with elevation and the presence or absence of a number of species (Table 2). Both models were highly significant ($P < 0.0001$). The model for predicting sites of poor quality included five species and elevation. The presence of two species, black locust and yellow-poplar, reduced the probability that a site was of poor quality, while the presence of the other three species increased the probability. Application of this model to the analysis data resulted in classification accuracy of 70 percent.

The logistic model for predicting sites of good quality included three species from the poor quality sites (flowering dogwood, chestnut oak, and black locust) along with four other species. In this model, the presence of black locust increased the probability that sites were of good quality; however, the presence of the other six species decreased that probability. Similar to the model for poor sites, the presence of black locust increased the probability of membership in the good quality class. However, the absence of the other species also increased the probability that a site was predicted to be good quality. This model resulted in a classification accuracy of 85 percent when applied to the analysis.

**Table 2.** Best formulation of logistic models for predicting either poor or good categories of site quality on 210 FIA plots in western North Carolina.

| Site quality category[a] | Independent variable | Regression coefficient | P > \|z\| |
|---|---|---|---|
| Poor | Constant | -2.999 | 0.001 |
| | Elevation | 0.00055 | 0.014 |
| | Black locust | -0.980 | 0.039 |
| | Chestnut oak | 0.719 | 0.040 |
| | Flowering dogwood | 0.903 | 0.021 |
| | Serviceberry | 1.439 | 0.034 |
| | Yellow-poplar | -0.844 | 0.033 |
| Good | Constant | 3.112 | 0.001 |
| | Elevation | -0.0008 | 0.001 |
| | Black locust | 0.968 | 0.025 |
| | Chestnut oak | -1.175 | 0.009 |
| | Flowering dogwood | -1.025 | 0.026 |
| | Hickory spp. | -1.284 | 0.011 |
| | Red maple | -0.830 | 0.039 |
| | Scarlet oak | -1.881 | 0.007 |
| | Sourwood | -1.723 | 0.001 |

[a] The two site quality categories were modeled as positive outcomes of the logistic regression. Each category consists of the lower or upper 25 percentile of the total number of plots ranked by productivity.

We made an unplanned analysis to explore effects of classifying sites in two equal-size groups (inferior or superior) defined by the mean productivity of the 210 sample plots. Results of this analysis were consistent with the two planned tests: lower elevations, yellow-poplar and black locust indicated superior sites while higher elevations, scarlet oak and chestnut oaks indicated inferior sites. The classification accuracy of this model was only 68 percent. Although the indicator effect of yellow-poplar changed from the models for productivity class (where its absence was associated with the model for poor sites), this analysis identified a small set of core variables that would likely be significant in models using other methods of defining classes of productivity or categories of quality.

Elevation was the only significant ($P < 0.01$) environmental variable included in the multiple logistic models for predicting site quality. Although the sign of the coefficient for elevation differed between the models, the effects on site quality were consistent: plots at higher elevations increased the probability that the site was of poor quality. None of the other environmental variables or solar radiation explained significant variation associated with site quality.

**Multiple species and three site productivity classes**
The multinomial analysis indicated that elevation and six species were significantly associated with the three productivity classes of the 210 sample plots (Table 3). Species significant in predicting the low productivity class of sites included only the presence of serviceberry, although chestnut oak was almost significant (P = 0.059). On sites of high productivity, however, the presence of serviceberry and the absence of hickory, sourwood, scarlet oak, red maple, and chestnut oak were significant independent variables. The high productivity class was associated with lower elevations. The overall classification accuracy of this model was 61 percent (Table 4). The largest source of model inaccuracy occurred in the low productivity class, where many plots were misclassified as medium productivity.

**Table 3.** Logistic model for predicting productivity classes of 210 FIA plots in western North Carolina.

| Logistic model[a] | Independent variable | Regression Coefficient | P > \|z\| |
|---|---|---|---|
| Low | Constant | -1.460 | 0.100 |
| | Elevation | 0.0003 | 0.251 |
| | Chestnut oak | 0.690 | 0.059 |
| | Hickory spp. | -0.494 | 0.197 |
| | Red maple | -0.563 | 0.121 |
| | Scarlet oak | 0.276 | 0.496 |
| | Serviceberry | 1.914 | 0.027 |
| | Sourwood | -0.250 | 0.535 |
| High | Constant | 2.694 | 0.001 |
| | Elevation | -0.0006 | 0.019 |
| | Chestnut oak | -0.962 | 0.036 |
| | Hickory spp. | -1.660 | 0.001 |
| | Red maple | -0.855 | 0.036 |
| | Scarlet oak | -1.647 | 0.016 |
| | Serviceberry | 2.036 | 0.032 |
| | Sourwood | -1.914 | 0.001 |

Prediction of site quality class requires three calculations: (1) determine probability using the model for the low productivity class, (2) determine probability using the model for the high productivity class, and (3) subtraction of the low and high probabilities from 1.0 to obtain the probability of the medium productivity class. The predicted productivity class is the probability of greatest magnitude.

**Table 4.** Observed and predicted productivity classes of 210 FIA plots in western North Carolina.

| Observed class | Predicted class: Low | Medium | High | Totals |
|---|---|---|---|---|
| Low | 7 | 39 | 8 | 54 |
| Medium | 5 | 88 | 12 | 105 |
| High | 2 | 16 | 33 | 51 |
| Totals | 14 | 143 | 53 | 210 |

As with the two-category models, signs of the coefficients reveal their effects on behavior of the models. Coefficient signs of three species (red maple, hickory, and sourwood) were negative in both models, indicating their presence would decrease membership in the lower and higher classes and increase the probability of inclusion in the medium productivity class. Chestnut oak and scarlet oak coefficient signs were positive in the low model and negative in the high model, indicating their presence was associated with sites of low productivity but not with sites of high productivity. Interestingly, serviceberry was the only species with two positive signs, suggesting the ambiguous result that its presence decreases the probability that a site is of medium productivity, with about equal indication of membership in both the low or high classes as shown by similar magnitudes of the coefficients. Caution should be used, however, in interpreting meaning of insignificant regression coefficients.

## DISCUSSION

The relationship of species to site productivity in our study was consistent with that reported by others in the southern Appalachians (Whittaker 1966, Callaway et al. 1989, Golden 1974, Mowbray and Oosting 1968). Our results indicated that sites of high productivity tended to be occupied by species considered as mesophytic: black locust, and yellow-poplar (Burns and Honkala 1990, Mowbray and Oosting 1968, Whittaker 1966, Golden 1974, Callaway et al. 1989). Low quality sites were associated with sites where xerophytic species, such as oaks, dominated. As indicated by the lack of significance in the contingency table analysis, several species occurred across the range of site qualities, including red maple, sassafras, white oak, and hemlock. Northern red oak, a species typically associated with highly productive sites (Loftis 1990, Burns and Honkala 1990), was present on poor and good sites, but with less than expected frequency on the latter, a condition that could have resulted from past management practices and lack of advance regeneration (Loftis 1990).

One species, flowering dogwood seemed to occur illogically in relation to site quality. Flowering dogwood is generally associated with mesic, high quality sites in the southern Appalachians (Beck and Della-Bianca 1981, Mowbray and Oosting 1968, Chellemi et al. 1992), although it can occur over a range of moisture regimes (Burns and Honkala 1990). In our study, however, flowering dogwood occurred more commonly on sites of low productivity than on sites of high productivity. One explanation of the lower than expected frequency of flowering dogwood on high sites could be with the presence of anthracnose disease (*Discula destructive*), which causes high mortality rates in this species on mesic sites of northern aspects and lower slope positions (Chellemi et al. 1992).

Our index of site productivity was based on tree circumference, an unconventional measure of tree and stand increment, but it appeared to perform in a satisfactory manner for most plots. However, in evaluations of error analysis for some two-class models, evidence (not presented) suggested that the index behaved illogically for some plots, particularly those with a low level of initial basal area stocking. Performance of the index should be examined in greater detail and compared with conventional measures of productivity such as periodic basal area or bole volume increment.

Our arbitrary definition of site productivity classes and categories likely had a small but important influence on results of our study, particularly the significance of certain species in the model. For example, if we had defined the good category as consisting of the 33.3 percentile of the plots with highest productivities (with the remaining 66.7 percent in the not-good category), the significant species in the model, in decreasing order of importance (with sign of coefficient), were chestnut oak (-), white ash (+), sourwood (-), and yellow-poplar (+). Inclusion of white ash as a significant indicator of good site quality would agree with our collective field observations and evidence of others (Burns and Honkala 1990). Callaway et al. (1989) ranked white ash about midway in the range of species productivity reported in the Great Smoky Mountains but did not indicate if it was associated with stands of higher than average productivity.

Callaway et al. (1989) also reported stands containing yellow-poplar as among the most productive. Our study based on data from a broader geographic area confirmed their findings. Also, consistent with our findings, Callaway et al. (1989) reported that stands with Virginia pine, scarlet oak, or pitch pine were the least productive, and tended to be associated with perceived xeric sites (i.e. ridges and south facing slopes). Callaway et al. (1989) reported that other species associated with dry sites included sourwood, black oak, and blackgum.

Nine species, considered individually and in combination with others, were significantly associated with classes or categories of site productivity. Many significant indicator species were notably absent from sites. In the bipartite model of good and poor sites, for example, only black locust had a positive coefficient. As a component of all models, chestnut oak was the most consistent indicator species. We were surprised to find that serviceberry, a relatively minor species in the southern Appalachians, was among the most important indicator species in this data set. Serviceberry proved unique among all species as a significant positive indicator of both low and high productivity classes in the tripartite analysis. Whittaker (1966) found Allegheny serviceberry on high elevation south-slope and ridge (presumably dry) sites, where it was a minor component of moderately productive stands dominated by chestnut oak and beech. Little information on the ecology of serviceberry is available in the literature to explain this apparent conflict of indicator values -- a conflict that might be an artifact of the FIA data set resulting from pooling of two or more species. On high quality sites, other significant species included scarlet oak and sourwood; red maple was least important. Except for serviceberry, none of the species we used proved a significant indicator of poor sites. Results of our study, particularly for poor sites, might have been influenced by our decision to arbitrarily omit from the analysis species that occurred on less than 10 plots.

Grouping species in three genera probably also affected the sensitivity of our predictive models, particularly for serviceberry. The serviceberry genus occurring in the study area consists primarily of two species: downy (present throughout, but mainly in the piedmont) and Allegheny (restricted mainly to the mountains), which occur on sites of differing environmental characteristics associated probably with elevation. Pooling the two species might have caused the contradictory effects associated with productivity observed for serviceberry. A similar situation exists for species of hickory and birch.

Significant environmental variables were limited to elevation in our study. We found that site quality categories and productivity classes were negatively correlated with elevation ($r = -0.17$), consistent with results of Whittaker (1966) and Callaway et al. (1989) in the Great Smoky Mountains. Elevation has been consistently associated with species occurrence (Whittaker 1966, Golden 1974) and productivity (Callaway et al. 1989). Although slope gradient was also significantly correlated (negatively) with productivity (r = -0.18), elevation was the only environmental variable included in the regression models. The sign of the elevation coefficient differed between the two bipartite models of site quality classes, but the effects of elevation were consistent. Aspect affects species composition in the southern Appalachians (Mowbray and Oosting 1968, Golden 1974), but not productivity (Whittaker 1966, Callaway et al. 1989) -- as results of our study confirm. We found, as did Whittaker (1966), that solar radiation was not associated with site productivity.

In summary, the results of this study suggest that the indicator species method of assessing site productivity has promise for use in the southern Appalachian Mountains. Our study in the mountains of western North Carolina indicated that good quality sites could be predicted with acceptable results (85 percent accuracy) based on the presence or absence of six tree species. However, results were less satisfactory (71 percent accuracy) when identifying the poor category of site quality. Results were less acceptable (60 percent accuracy) when predicting membership of sites into one of three classes of productivity based on species. Elevation was a significant environmental variable associated with all classes of productivity, and accounted for about 5 percent of the variation on good sites, but only about 1 percent on poor sites. Golden (1974) states however, that "an important limitation to this approach lies in the reality that plant species have widely differing ecological amplitudes or tolerances, thus having variable value as indicators."

Several areas might be fruitful to pursue in future studies of the indicator species approach to site classification. These include: (1) expanding the area of analysis by including plots from similar mountainous areas in adjoining states, (2) increasing the number of site classes to four or five, and (3) the reformulation of the measure of productivity to include stand height. Other multivariate methods, such as principal components analysis, might be also be explored, with the caveat that the interpretation and application of results would be more complicated. Model validation, beyond the scope of this exploratory study, is particularly important area for future evaluation -- as is the effect of omitting plots that showed evidence of disturbance. Because of this decision we omitted pines -- a group of species that generally requires disturbed sites -- from our models.

Another possibility for future work involves restricting the analysis to species represented by two or more individuals per plot. Our experience in southern Appalachian forests suggests that any species can occur on any site as a result of chance, a reality that tends to confound species indicator relationships. Inclusion of an "off-site" species in the analysis would be less likely if we required a minimum of two occurrences per plot for inclusion in the data set. This strategy, however, would have eliminated serviceberry -- a particularly valuable indicator species -- from our analysis. Inclusion of shrub species in the analysis would also be desirable.

## ACKNOWLEDGMENTS

We gratefully acknowledge the assistance by Joseph F. Glover and Linda S. Heatherly, USDA Forest Service, Southern Research Station, Forest Inventory and Analysis, for providing -- with little advance notice -- plot elevations, an environmental variable that proved important in our analysis. We thank Bruce Baldwin, Mike Rauscher, and Ray Sheffield for helpful comments on an earlier version of our manuscript.

## LITERATURE CITED

Beck, D.E. and K.B. Trousdell. 1973. Site index: Accuracy of prediction. Research Paper 108. U.S. Department of Agriculture. Forest Service. Southeastern Forest Experiment Station. Asheville, North Carolina. 7 p.

Beck, D.E. and L. Della-Bianca. 1981. Yellow-poplar: characteristics and management. Agricultural Handbook 583. U.S. Department of Agriculture. Forest Service. Washington, DC. 91 p.

Burns, R.M. and B.H. Honkala. 1990. Silvics of North America, Volume 2, Hardwoods. Agricultural Handbook 654. U.S. Department of Agriculture. Forest Service. Washington, DC. 877 p.

Callaway, R.M., E.E.C. Clebsch, and P.S. White. 1989. Predicting wood production by canopy trees in forest communities in the western Great Smoky Mountains. Forest Science 35:338-348.

Carmean, W.H. 1975. Forest site quality evaluation in the United States. Adv. in Agronomy 27:209-269.

Chellemi, D.O., K.O. Britton, and W.T. Swank. 1992. Influence of site factors on dogwood anthracnose in the Nantahala Mountain range of western North Carolina. Plant Disease 76(9):915-918.

Daubenmire, R. 1976. The use of vegetation in assessing the productivity of forest lands. The Botanical Review 42:115-143.

Golden, M.S. 1974. Forest vegetation and site relationships in the central portion of the Great Smoky Mountains National Park. Dissertation. University of Tennessee. 275 p.

Hansen, M.H., T. Frieswyk, T., J.F. Glover, and J.F. Kelly. 1992. The eastwide forest inventory data base: users manual. General Technical Report NC-151. U.S. Department of Agriculture. Forest Service. North Central Forest Experiment Station. St. Paul, Minnesota. 56 p.

Hosmer, D.W. and S. Lemeshow. 2000. Applied logistic regression, Second Edition. John Wiley & Sons, Inc. New York. 373 p.

Kimmins, J.P. 1987. Forest Ecology Edition 1. Macmillian Publishing Company, New York. 531 p.

Lexen, B. 1943. Bole area as an expression of growing stock. Journal of Forestry 41:883-885.

Little, Jr., E.L. 1953. Check list of native and naturalized trees of the United States. Agriculture Handbook 41. U.S. Department of Agriculture. Forest Service. Washington, DC. 472 p.

Loftis, D.L. 1990. Predicting post-harvest performance of advance red oak reproduction in the southern Appalachians. Forest Science 36(4):908-916.

Mowbray, T.B. and H.J. Oosting. 1968. Vegetation gradients in relation to environment and phenology in a southern Blue Ridge gorge. Ecological Monographs 38:309-344.

Spurr, S.H. and B.V. Barnes. 1984. Forest ecology. Edition. 2. The Ronald Press Company. New York. 571 p.

StataCorp. 1999. Stata Statistical Software: Release 6.0. Stata Corporation. College Station, Texas. 384 p.

Teck, R., M. Moeur, and B. Eav. 1996. Forecasting ecosystems with the forest vegetation simulator. Journal of Forestry 94(12):7-10.

Whittaker, R.H. 1966. Forest dimensions and production in the Great Smoky Mountains. Ecology 47:103-121.

**Long-term Sustainability Analysis of Fiber Supply in Georgia**

by

**C.J. Cieszewski[1], M. Zasada[2, 5], B.E. Borders[4], M.L. Clutter[4], R.F. Daniels[4], R.I. Elle[6], and R. Lowe[3]**

**Abstract**

The general scope of a new cooperative study at the School of Forest Resources, University of Georgia, in analysis of long-term fiber supply in Georgia is presented and described. Essentially, this study is an applied research effort into effective techniques for annual forest inventory using Landsat TM imagery, GPS, and GIS. The overriding objective is to develop a methodology for realistic fiber supply assessment on a short and long term basis. The resulting methodology will be used to reassess sustainability of Georgia's forest resources on an annual basis by rerunning the long-term simulations with annually updated forest inventory data. Our approach is similar to the approach adapted by British Columbia for determination of annual allowable cut in national forests, whereby the levels of allowable cuts every year are determined by long-term sustainability analyses based on 200-year simulations of the existing inventory under assumed levels of utilization. Currently, the analysis is based on basic inventory summaries generated using the databases provided by the USDA Forest Service Forest Inventory and Analysis unit. However, even this simplified preliminary approach to the analysis is very comprehensive.

## BACKGROUND

Collection and analysis of large-scale forest inventory data has a long tradition in the United States. As early as 1928 the McSweeney-McNary Forest research Act authorized the USDA Forest Service to conduct periodic national forest surveys. The Act called for “a determination of the present potential forest productivity of forestland” in the United States. Since this time the Forest Inventory Analysis section of the USDA Forest Service has conducted such an inventory service across the United States.

In the southeastern (SE) part of the US the first forest inventory was implemented in the 1930. Until recently the inventories have been carried out on what were to be 10-year cycles, but in truth, many of the SE inventories have taken some 15 years to complete. The SE region is currently being inventoried for the seventh time. Emphasis in past inventories was placed more on providing current estimates of forest resources than on providing estimates of growth and other changes over time.

The sampling units, the basic unit of measurement on the ground, have been changing over time. These have included fixed area plots and point sampling with variable area plots of different basal area factors for trees greater than five inches in DBH and fixed area plots for trees less than six inches DBH (Frayer and Furnival 1999).

The sampling design of the FIA plots was a combination of systematic and random sampling. Specifically, the plots have been selected through a random selection of grid points, and site productivity is described in terms of site class and site index, based on arbitrarily selected site-representative dominant tree-height measurements.

Measurements recorded in the FIA program were very extensive, labor consuming, and covered the collection of various and detailed information that were dimmed to be appealing to various interest groups. Some examples are information on dead trees and cavities, wildlife cover and habitat types, recreation indicators, presence of litter, and profiles of understory vegetation. Thus much of the effort was spent obtaining environmental and other non-tree-measurement data. In recent years increased demand for forest products, intensified forest management, population growth, accelerated immigration to the state of Georgia and continued urbanization, have all greatly changed Georgia’s forest landscape. These changes include an increase in the amount of forestland being converted away from Forest production. Forest management has intensified in order to produce increased yields from the remaining commercial land base (Daniels 1999). Such demands placed on the forest resources raise important questions concerning their long-term sustainability.

While intensive management has increased forest productivity, many individuals are asking if this will be enough to meet the demand for forest

---

[1]Assistant Professor, [2]Postdoctoral Fellow, [3]Information Analyst, [4]Professor, [6]Research Coordinator, D.B. Warnell School of Forest Resources, University of Georgia, Athens, GA 30602, USA

[5]Assistant Professor, Department of Forest Productivity, Faculty of Forestry, Warsaw Agricultural University, Rakowiecka 26/30, 02-528 Warsaw, Poland

*Citation for proceedings:* P.F. Doruska and D.C. Bragg (editors). 2002. Proceedings of the Southern Mensurationists' Conference, Chattanooga, TN, November 4-6, 2001. Arkansas Forest Resources Center, Monticello, AR. 43 p.

products given a shrinking land base? Even more important is whether or not Georgia's present resource needs and values can be met without jeopardizing those of future generations. Such uncertainties underscore the need for more extensive analysis of forest inventory data, and for advancements in analysis of long-term trends in changes of the natural resources.

The forest industry and, by extension, the state's economy, rely on accurate assessment of fiber supply and on the ability to predict changes in available inventory brought about by land use shifts, changes in fuel loading, balance of growth to harvest, and catastrophic events such as drought, fire and insect infestation. A desire for greater accuracy and timeliness of forest inventory has led to the re-organization of Forest Inventory Analysis (FIA) methods into the Southern Annual Inventory System (SAFIS).

The state of Georgia was one of the first southern states to initiate work on the southern inventory system in partnership with the USDA Forest Service Forest Inventory and Analysis program. The annual cost of implementing SAFIS will be many millions of dollars, and the forest products industry will rely to a great deal on the inferences made from the collected data. Yet, much of the analysis performed on this data may be fairly simplistic, and in turn inferences coming from this analysis cannot answer key questions raised above.

For example, reliable answers to the question of long-term forest sustainability cannot be answered from simple comparisons of current growth rates and removal levels. The growth of forests does not follow linear trends. In addition, growing stock is not recorded by the FIA inventory until the trees reach a certain minimum size. Sustainability related questions must therefore be addressed with much more complex analysis involving proper modeling of nonlinear changes over time. This includes using clear assumptions regarding future land use changes, explicit assumptions with respect to regeneration dynamics and considering forest product demand.

Without such rigorous methods, claims of threat to sustainability may be based on inadequate information, arbitrary assumptions and over-simplified analyses. This in turn may cause unrealistic expectations about the resource or unjustified concerns, either of which may lead to less than optimal resource management and utilization.

The proposed study includes basic modeling of supply and demand conditions, land use changes, including urban and suburban development, and explicit assumptions concerning regeneration and changes in growth dynamics. In addition, the study should produce a reliable and effective method for analyzing the impacts of specified environmental regulations and other forest management related regulatory constraints. The study will investigate impacts of such issues as maximum harvest areas, adjacency constraints and green-up regulations, and the effects of introducing road buffers and riparian zones, or water resource buffers (commonly referred to as streamside management zones, or SMZs).

## OBJECTIVES

The short-term objectives of this study are to illustrate a proof-of-concept for running statewide simulations based on FIA data and to conduct preliminary visibility analysis based on simplified assumptions and criteria of forest management.

The long-term objectives are to:

1. carry out spatially explicit, detailed simulations given various assumptions and regulatory scenarios.
2. identify those elements that will most strongly affect Georgia's long-term wood supply.
3. conduct sensitivity analysis on simulations with respect to assumptions of forest management practices.
4. estimate the effect of various management practices on future states of Georgia's forest resources.
5. estimate the effect of various regulatory scenarios on the state's wood supply.

Implicit in these objectives will be the methodology research using Landsat Thematic Mapper, GPS and GIS for supply assessment on a short and long term basis. The sustainability of Georgia's forest resources will be reassessed on an annual basis. This will determine levels of permissible cuts by long-term sustainability analysis based on 200-year simulations of the existing inventory under the assumed levels of inventory.

## DATA

The first stage of the study is based only on the FIA data from the USDA Forest Service. Presently all FIA data from the periodic inventories is available from the web site of the North Central Research station in St. Paul, MN (Miles et al. 2000). Included in this data are the Eastwide Database (EWDB) (Hansen et al. 1992) and Westwide Databases (WWDB) (Woudenberg and Farrenkopf 1992), which give results for the last two periodic inventories.

For a more detailed discussion of the data handling for this project, see Zasada et al. (2002) in this same proceedings publication, where the authors discuss species group definitions, yield table development, data sources, minimum data requirements, sample setup, etc.

## METHODS

The study is based on 200-year simulations for the whole state of Georgia with the same harvesting levels as what is currently observed in the state (i.e., 1,476 million cubic feet per annum) (Thompson 1998). Running of these simulations is based on various requirements and assumptions with regards to:

1. appropriate simulation tool;
2. management regimes;
3. simplified representation of species growing in Georgia;
4. maturity criteria for different species groups; and
5. harvesting priorities.

### Approach to Long-term Simulation

The first requirement was to select a comprehensive and spatially explicit forest estate model capable of a multi-landscape level simulation. After extensive research into various programming tools the decision was made to purchase the off-the-shelf OPTIONS software by DR Systems Inc., 1615 Bowen Road, Nanaimo, B.C. This software is a simulator with GIS functionality that is designed to test different forest management scenarios for large land areas. The tested scenarios are based on various variables in such areas as financial, industrial, policy decisions and sustainability analysis.

Optimization was not intended for this study because one cannot reasonably expect to optimize harvesting levels for 630,000 independent landowners who do not work towards a common goal or management plan. While many timber companies in the state may optimize forest management and apply harvest scheduling, it would be unreasonable to attempt this on a statewide level. At the same time, while choosing not to optimize may lead to bias by underestimating Georgia's forest growth capabilities, in this study such bias could be considered desirable to assure that the estimates are conservative instead of aggressive.

OPTIONS is an estate simulator that can handle highly complex scenarios but can also run basic simulations with a minimum of data requirements (area, species group, site index, age or year). Accordingly, the initial simulations are based on using only crude data summaries and basic assumptions.

The initial simulations use only two yield tables, three species groups, three management regimes, and five site index classes. The tables are those for planted loblolly pine (Harrison and Borders 1996) and for upland oaks (Gingrich, 1971). The three basic species groups are natural softwoods, planted softwoods, and hardwoods. The initial runs also included some basic assumptions on cover type transitions, and specified maturity criteria and harvesting priorities.

## RESULTS

The comparisons of inventory information based on the most recent (1997) periodic FIA data and the first panel SAFIS annual inventory data (1999) suggested a strong consistency in the resource assessment. This is a desirable outcome, because it supports the plans of continuing the long-term sustainability analysis in future years despite the frequent concerns of many individuals about the diminished accuracy of the much smaller sample in the annual inventory design.

The results of the preliminary analysis are in no way final or even conclusive for the obvious reason of its incompleteness. Yet, this simplified analysis of Georgia's forest management and harvesting scenarios based on current practices and public inventory data suggests a sustainable wood supply at Georgia's current harvesting level.

Sensitivity analysis conducted with respect to harvesting levels and adjustments of yield tables showed reasonable responses to both of these parameters. The earlier suggested that the current harvesting level is close to the level of maximum possible harvesting that could be sustainable by the forest resources in Georgia. The later suggested that lower yields within deviations that might be associated with errors in projected yields would not dramatically impact the results of the analysis. The later results suggested that yields increased by margins expected from forest management intensification would significantly increase the sustainable harvesting limits to the extent that could offset exclusions of a large portion of land due to urbanization, harvesting-free national forest policies, and potential land conversions from forestry into other uses. More details about the preliminary results can be found in Cieszewski et al. (2001).

## DISCUSSION

The current study is the first approximation based on greatly simplified data and assumptions. The concerns are limited in this study to biophysical factors such as forest management practices, timber growth rates and inventory definitions in terms of broad classes of species and productivity sites. The FIA inventory data and growth and yield information defined by the yield tables comprised the majority of the technical information. Future long-term sustainability analysis of forest resources also should consider other available inventory and GIS data, as well as any available social and economic information.

The data and assumptions used in this analysis were drastically simplified. It was assumed in the preliminary analysis that all lands can be harvested, while it is reasonable to expect that there will be constraints on some land uses. On the other hand the calculations of growth were underestimated by not including in the analysis any management intensification that could dramatically increase yields. With respect to the applied yield tables there was an observable small negative bias in the prediction of future yields. In addition, it is expected that the FIA data contain a small negative bias in estimation of growing stock, because FIA does not record trees smaller than 5 inches in diameter.

The future constraints on land uses are expected to include some of the forestland base that will be set aside for the purpose of urbanization and other uses. A certain amount of the land base may be given up as protective buffers, which might be significant as carbon storage but not relevant as a source of timber supply. It is assumed in this project that land conversions from forestry uses to agriculture will be offset by land conversions from agriculture into forestry uses. This assumption is supported by findings of the U.S. Forest Service (Wear and Greis 2001).

In addition to other factors, the future simulations will need to account for: species growth characteristics; the shrinking timber-supply land base; and the fact that intensively managed plantations grow much faster than unmanaged stands.

Phase 1 of the study included acquisition and processing of the FIA data, programming of the simulation software (OPTIONS), and running initial scenarios for the entire state of Georgia. The next steps in this phase include acquiring and processing GIS data on roads and urban and suburban developments, water and other data and linking to ground truth data where resolution permits. Other efforts will concentrate on refining yield tables and incorporating wood quality relationships into the analysis.

In phase 2 of this study we will include a semi-spatial sensitivity analysis including effects of riparian zones. It will begin with the review and validation of initial scenario results and comparison with maps and reports and will include considering reasonability and visibility. This will be followed by sensitivity analysis for the effects of water buffer zones and forest management options including riparian zones and other management options.

Phase 3 will involve a full spatially explicit analysis of those areas of the state where spatial information exists. This will include acquisition of spatial data from industry and other study participants, processing of the data to a common format, and screening of the data for quality assurance. Where resolution permits the data will be linked to ground data obtained from detailed information on roads, urban and suburban developments and water resource data and other sources.

The yield tables at this stage will have to be prepared for all cover types and scenarios and validated against ground truth data where possible. The final simulations, which include constructing final databases, will process detailed spatially explicit scenarios and generate tabular and graphical outputs. The results from all Phases will be published in the forms of journal publication and in some cases web-based dissemination.

## ACKNOWLEDGMENTS

This study has been sponsored by the Traditional Industry Program 3 (TIP$^3$) funding and the School of Forest Resources, University of Georgia. We are much indebted for this support. This paper was prepared for the proceedings of the 2001 Southern Mensurationists' Conference. Technical details associated with the project discussed here are described in the companion paper by Zasada et al. (2002) that is included in the same proceedings. Another description of this project, with much additional information on the results of the preliminary runs, has also been presented, and included in the Proceedings, of the IUFRO Conference held in Vancouver in 2001 (Cieszewski et al. 2001), and was submitted for publication in the Southern Journal of Applied Forestry.

## LITERATURE CITED

Chojnacky, D.C. 1998. Double sampling for stratification. A forest inventory application in the interior West. Research Paper RMRS-RP-7. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Cieszewski, C.J., D.P. Turner, and D.L. Phillips. 1996. Statistical analysis of error propagation in national level carbon budgets. Proceedings of the Second International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Fort Collins, Colorado.

Cieszewski, C.J., M. Zasada, R.F. Daniels, M.L. Clutter, and B.E. Borders. 2001. Long-term Sustainability of fiber supply in Georgia - a Preliminary Approach. In: Forest Modelling for Ecosystem Management, Forest Certification and Sustainable Management. Proceedings of the IUFRO Conference, Vancouver, BC, Canada. August 12-17. p. 173-182.

Daniels, R.F. 1999. Sustainability and forest industry: growing more wood than ever before. In: Proc. 1998 SAF Natl. Conv. Society of American Foresters, Bethesda, MD, USA. p. 75-80.

Frayer, W.E. and G.M. Furnival. 1999. Forest survey sampling designs: a history. Journal of Forestry 97(12):4-10.

Gingrich, S.F. 1971. Management of young and intermediate stands of upland hardwoods. U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station, Res. Pa. NE-195. 26 p.

Helms, J.A. 1998. The dictionary of forestry. Society of American Foresters, Bethesda, MD, USA. 210 p.

Hansen M.H., T. Frieswyk, J.F. Glover, and J.F. Kelly. 1992. The Eastwide forest inventory database: users manual. Gen. Tech. Rep. NC-151. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station. 48 p.

Harrison W.M. and B.E. Borders. 1996. Yield Prediction and Growth Projection For Site-prepared Loblolly Pine Plantations in the Carolinas, Georgia, Alabama and Florida. PMRC Tech. Rep. 1996-1. Plantation Management Research Cooperative. Daniel B. Warnell School of Forest Resources. University of Georgia. Athens, GA. 64 p.

Thompson, M.T. 1998. Forest Statistics for Georgia, 1997. Resource Bulletin SRS-36. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 92 p.

Wear, D.N. and J.G. Greis. 2001. Southern Forest Assessment, draft report. USDA Forest Service.

Zasada, M., C.J. Cieszewski, R. Lowe, B.E. Borders, M.L. Clutter, and R.F. Daniels. 2002. Selected technical aspects of a statewide forest resource modeling project in Georgia. In P.F. Doruska and D.C. Bragg (editors). 2002. Proceedings of the Southern Mensurationists' Conference, Chattanooga, TN, November 4-6, 2001.

# Selected Technical Aspects of a Statewide Forest Resource Modeling Project in Georgia

by

**M. Zasada[1, 6], C.J. Cieszewski[2], R. Lowe[3], B.E. Borders[4], M.L. Clutter[4], R.F. Daniels[4], and W.M. Harrison[5]**

**Abstract**

We describe here selected technical aspects of a large-scale sustainability analysis of wood supply in Georgia with particular focus on: i) available data; ii) species group definition; iii) yield tables. The study is conducted at the School of Forest Resources, University of Georgia, as a new cooperative effort financed by the Traditional Industry Program 3 and the D.W. Warnell School of Forest Resources, University of Georgia. We identify here the directions, resources and problems in such a large scale of modeling. At the current state the analysis is based on basic inventory summaries founded on the databases provided by the USDA Forest Service Forest Inventory Analysis unit. We explain our main assumptions based on available literature, and unpublished research on intensive plantation management.

## BACKGROUND

Forestry and the forest industry is an important part of Georgia's economy. It generates about 177,000 jobs and contributes $22 billion annually to the State's economy. Forests cover almost ¾ of Georgia's area and they are owned by about 630,000 landowners. American forestry, especially in the South, in recent years has change very fast. Production areas are decreasing due to such factors as urban expansion and political and regulatory constraints on harvesting. At the same time we have seen an increase in wood-product demand as well as non-timber related uses, such as tourism, recreation, fishing, and hunting. Given this multitude of interests it is no mystery why society is concerned about the sustainability of our state's forests.

It is possible to talk about sustainability in terms of harvesting decisions, especially the amount of timber available for cut, management practices, and different regulatory constraints having influence on wood available for utilization.

Because of the nonlinear nature of changes in forest resources and many factors that influence these changes, it is not enough to use existing inventory data for simply comparing growth and removals. The situation requires conducting long-term and large-scale research on the forest trends and future impacts.

## OBJECTIVES

Our objectives are to:

- conduct spatially explicit simulations of growth and harvest of forest resources in Georgia under various assumptions;
- identify major factors impacting the long-term wood supply in Georgia;
- conduct sensitivity analysis on simulations with respect to the assumptions and forest management practices;
- estimate the impact of various forest management practices on future forest resources in Georgia;
- estimate the effect of various regulatory scenarios on the state's wood supply.

## APPLIED SOFTWARE - OPTIONS

With our objectives we started looking for research tools that were capable of state-level simulations. We did not want to develop our own simulator from scratch knowing that this is a very long and resource-intensive process (e.g., the multi-year and multi-million dollar CLAMS project at Oregon State University). After extensive research we selected the estate simulation software called OPTIONS, from DR SYSTEMS Inc. as the most comprehensive spatially explicit forest estate model.

OPTIONS can be used to examine different forest management scenarios for land areas including financial, industrial and policy decisions and sustainability analysis. The simulator is based on

[1]Postdoctoral Fellow, [2]Assistant Professor, [3]Information Analyst, [4]Professor, [5]Program Coordinator, D.B. Warnell School of Forest Resources, University of Georgia, Athens, GA 30602, USA

[6]Assistant Professor, Faculty of Forestry, Agricultural University, 02-528 Warsaw, Poland

*Citation for proceedings:* P.F. Doruska and D.C. Bragg (editors). 2002. Proceedings of the Southern Mensurationists' Conference, Chattanooga, TN, November 4-6, 2001. Arkansas Forest Resources Center, Monticello, AR. 43 p.

forecasting information for individual polygons (database records). Each piece of information is processed annually, record by record.

We have chosen to conduct the study using OPTIONS software in spite of the fact that it lacks optimization capabilities since in truth it is impossible to optimize harvesting levels and times among hundreds of thousands of independent landowners who own most of Georgia's forestland. Even though many timber companies in Georgia do optimize their forest management on their properties and apply optimized harvest scheduling, their forests make up only 1/4 of total forested area in Georgia. While not optimizing harvesting in our simulations will likely result in underestimation of the growth potential in Georgia, we are willing to be conservative in this case.

The OPTIONS program is a simulator with GIS functionality. It keeps topology for polygons from up to 25 GIS layers. This information can be used with an external GIS system through links to particular polygons. Because it keeps polygon topology it is possible to define combinations of constraints and targets for every included spatial layer. GIS layers can be switched on and off according to the goals of the simulation.

For each basic polygon the simulator does all defined silvicultural treatments such as regeneration, pre-commercial thinning (PCT), fertilization etc., excluding commercial harvesting (thinning and final harvesting). After applying all planned treatments the simulator performs commercial harvesting.

The model starts with the current inventory data and simulates changes over time according to the defined rules and provided yield tables. Even though OPTIONS is a landscape simulator with the capability to handle very complex tasks it also has low minimum data requirements. Basic simulations in OPTIONS can be run using as little information as area, species groups, site index, age or year of establishment, and some basic assumptions on yield tables. For this reason we found OPTIONS particularly suitable for our study and decided to start the simulations using crude data summaries and only basic assumptions without giving up the perspective of increasing the complexity of our analysis to the highest desirable levels in the foreseeable future.

## DATA

All assumptions and settings are stored in the OPTIONS definition file. It requires information about the database (inventory data which can be combined with GIS information and satellite imagery), management and silvicultural practices and economics rules applied in the analyzed scenario.

The bulk of the data that we used for estate modeling in Georgia we acquired from the plot-level FIA database. We used the following FIA variables: EXPACR (area of the polygon), SI (site index) and STDAGE (stand age) directly (Hansen et al. 1992). A species group can't be used directly. It requires additional database processing according to simulation assumptions.

In the simplest case all other values are automatically filled in from the yield tables as the program starts a scenario. But it is possible to also use additional information from the inventory, such as: basal area, height, diameter, stocking, volume, year of thinning, current and future management regime, stand activity and GIS constraints.

As mentioned, we used in our study data collected by the USDA Forest Service's FIA program. We had access to at least 3 sources of this data. Until the year 2001, the primary FIA data site was handled by the Southern Research Station in Asheville, NC and provided by its web server as the Eastwide Database (EWDB) (Hansen et al. 1992), and Westwide Database (WWDB) (Woudenberg and Farrenkopf 1995). In 2000 the Forest Service began switching to the new system and now all FIA data from periodic inventories is available as the FIADB database from the website of the North Central Research Station in St. Paul, MN (Miles et al. 2000). We obtained MR (Master Record) data directly from Forest Service (Southern Research Station, Ashville, NC). In the future we plan to get available data from the forest industry and other sources to have as detailed and as current data as possible.

Each of purchased data sets has specific features, advantages and disadvantages. We describe below some of the characteristics of the Georgia data and also some of the differences in data content between data formats and county regions.

The most popular and widely used are the EWDB and WWDB databases (Hansen et al. 1992, Woudenberg and Farrenkopf 1995). They provide results from the last two periodic inventories for most of states. The databases for states often have their own set of unique characteristics. For example the database for the State of Georgia contains an additional variable, basal area, calculated from the tree level. Some limitations of the Georgia data include lack of recorded height measurements, site index variables rounded to the nearest 10 feet and the maximum site index value limited to 99 feet. It should also be kept in mind that data may be labeled with a certain year but was in fact collected over a range of years. For example, collection of data later labeled as 1997 data actually began in 1992.

Files in FIADB databases from the latest two periodic inventories were merged by conversion of EWDB and WWDB files to the same format. The structure was then extended to include space for the added variables in the new round of collected data. This new structure includes changes in the table design (7 tables instead of 3) and software for processing the data (Oracle ™ database server) and a higher level of detail (all collected field variables are included in the database). The FIADB databases are also connected to the system for tables and maps creation called the Forest Inventory Mapmaker:

**http://www.ncrs.fs.fed.us/4801/FIADB/index.htm**

Master record data can be obtained directly from the Forest Service. Its format provides tree height measurements, but they are differently defined in different inventory cycles. MR files have a "card" structure, which means that instead of separated files with data for each level (plot, tree etc) – this format has only one file with all data in it. This format has almost no documentation.

Forest industry data, which we believe we will be able to get, are usually very detailed data covering relatively large areas. Unfortunately they are very often confidential, provided in various formats, and collected using different procedures and designs. All these factors can make this data very difficult to apply.

We noticed also some differences between states and more frequently, regions, which makes use of the data on a regional scale and comparison of the data very difficult. First of all, start and end dates and the numbers of years spent collecting data may differ between regions and states. Another inconsistency is that in Georgia and Tennessee inventories very often one plot is divided into several conditions, e.g., forest-non-forest, young forest-old forest, coniferous-hardwood forest and so on. Data from other southeastern states contained only one plot condition per plot. We should also note that there is no site index data available for some states, as: Alabama, Arkansas, Kentucky, Louisiana, Mississippi, Oklahoma and Texas. However according to Forest Service information, in the annual inventory all states will have a structure like Georgia and Tennessee.

## OTHER ASSUMPTIONS

The next important scenario element is to define species groups. The meaning of "species group" can be unique to different scenarios – it is not always the same as e.g., "loblolly pine" or "softwood". Species groups are to be defined according to current forest management practices, simulation objectives (e.g., hardwoods-softwoods vs. detailed species list), and significant differences in growth and yield (caused by different regions or management practices).

To define any species group we have to answer several questions and decide if certain factors are important or not. Even for a given species, as loblolly pine, we have to know at least how it is regenerated (naturally of artificially), how it is managed (as traditional plantation, improved plantation, intensively managed plantation or CRP area), and in which region it grows. All the above factors have a significant impact on species growth and yield.

After the species group definition it is necessary to choose yield tables, i.e., data about a stand's growth and yield. They are divided by the species groups and site index classes. Their task is to fill in for data that is not available from the inventory and to provide information for growing the stands.

The OPTIONS simulator does not have built-in yield tables or growth models. Its construction allows us to use any data from existing yield tables or tables generated by any growth models. The provided yield tables should be built for stands without any treatments, because responses of stands on treatments are realized by using special factors (treatment response factors). It is also necessary to provide separated sets of yield tables for natural and planted stands.

Use of growth models or yield tables for large-scale simulations isn't as easy as for small objects. One issue is lack of yield tables for many species, especially hardwoods. Existing yield tables have different formats and assumptions (e.g., different site index base age) and very often are also only "partial" models, including e.g., only height growth, basal area or volume. In addition – all this information is spread through many publications, which makes its collection laborious.

Searching for yield tables and growth models gave us a lot of them, but most of them are relatively old, the oldest having been published more than 40 years old (e.g., Schumacher and Coile 1960, Nelson et al. 1961, Forbes 1961). Relatively few new research results are available. Those that are available are mainly for pine plantations. (e.g., Harrison and Borders 1996, Pienaar et al. 1996, Martin and Brister 1999). Very often there are no yield tables for stands of different origin and no sufficient information about growth and yield of some stands, such as stands established under the CRP (Conservation Resource Program) or intensively managed plantations. There is also insufficient information on the influence of treatments such as weed control, pre-commercial thinning, fertilization, thinning, or genetic improvements on growth rate.

Silvicultural treatments are all activities, which can be applied to a particular stand. These include regeneration, weed control, pre-commercial thinning, fertilization, commercial thinning, genetic improvement, and final harvesting. The treatments assigned for a particular species group, site index class and time comprise a management regime.

Yield tables in OPTIONS work for species groups managed according to a particular management regime. Stands within a specific management regime are adjusted through adjustment factors in yield tables assigned to the particular regime.

**SAMPLE SETUP**

Four examples of management regimes are given below. We assume here that we have defined 4 species groups: natural softwoods, planted softwoods, hardwoods, and additionally intensively managed pine plantations. These are of course only sample regimes and they can vary in different scenarios and for different assumptions.

- Management regime #1 is very extensive and intended for natural softwood stands. It includes natural regeneration immediately after harvesting, one thinning at the age of 20-25 (depends on site index) and final harvesting at the age of 30-40.

- Management regime #2 is a sample of treatments for planted softwood stands (traditional plantations). It includes artificial regeneration one year after harvesting with 1-year-old seedlings, one thinning at age of 14-16 with fertilization applied one year after thinning, and harvesting at the age of 20-30.

- Management regime #3 is intended for natural hardwood stands. It includes only natural regeneration and no additional treatments up to the final harvesting at the age of 40-50.

- Management regime #4 is intended for intensively managed plantations. This is a highly intensive management plan, which includes artificial regeneration with site preparation after harvesting using genetically improved 1-year-old seedlings, one herbicide treatment at the age of two, two fertilizations at age of five and 10, commercial thinning at age of 14, the next fertilization one year later, and the final harvesting at the age of 20-21.

Forestland use changes over time. Different species can be regenerated after harvesting, and different treatments can be applied, etc.. All these dynamic processes also have to be included in the definition of the large-scale simulation. OPTIONS realizes it is using a so- called “regime allocation table”, which means transition of management regimes between species groups over time. The regime allocation table shows what management regimes belong with particular species groups growing on a certain site, and the percentage of stands going from one regime to another. Sample tables are as follows:

**Table 1.** Sample regime allocation table (for regimes 1, 2 and 3 only).

| Species group | Current regime | % of allocation | Future regime |
|---|---|---|---|
| Natural softwoods | 1 | 10 | 1 |
| Natural softwoods | 1 | 90 | 2 |
| Planted softwoods | 2 | 100 | 2 |
| Hardwoods | 3 | 50 | 2 |
| Hardwoods | 3 | 50 | 3 |

**Table 2.** Sample regime allocation table (for regimes 1, 2, 3 and 4).

| Species group | Current regime | % of allocation | Future regime |
|---|---|---|---|
| Natural softwoods | 1 | 10 | 1 |
| Natural softwoods | 1 | 50 | 2 |
| Natural softwoods | 1 | 40 | 4 |
| Planted softwoods | 2 | 50 | 2 |
| Planted softwoods | 2 | 50 | 4 |
| Intensive plantations | 4 | 100 | 4 |
| Hardwoods | 3 | 30 | 2 |
| Hardwoods | 3 | 20 | 4 |
| Hardwoods | 3 | 50 | 3 |

These examples show that if we increase the number of species groups and include site index classes, such a table will be very long and complicated. The biggest challenge is to predict decisions of forest owners and changes caused primarily by economic factors (e.g., Abt et al. 2000).

**UP-TO-DATE PROGRESS**

We have started preliminary research in and conducted studies on available data, species definitions, yield tables, and management regimes. We have started the yield table calibration and adjustments, and did some preliminary GIS and remote sensing works and have tested several first runs of OPTIONS with the raw FIA database (Cieszewski et al. 2001, Cieszewski et al. 2002).

These experiences help us to estimate the necessary resources needed for execution of this type of project. Assuming that we use Pentium III class computers with at least 1 GB RAM a simple scenario run set to 200 years takes up to 36 hours. The size of a detailed database file for a simulation is over 300 MB in size. A single state-of-forest output file for any given year of the simulation is over 100 MB. The sum of output files can exceed several gigabytes per each considered scenario. Overall, this type of study needs powerful equipment to conduct the research within a reasonable timeframe.

**FUTURE PLANS**

The next stage of this study will use spatially explicit analysis on a level of resolution defined by the FIA subplots. Instead of statewide summaries, individual polygons of various cover types will be constructed based on the annual inventory data, Landsat TM satellite images, and state GIS data. The future efforts will concentrate on increasing the resolution of the analysis, including the use of spatially explicit data, and incorporating spatial information from various available GIS data and satellite images.

**ACKNOWLEDGMENTS**

This study has been sponsored by the School of Forest Resources, University of Georgia, and the Traditional Industry Program 3 ($TIP^3$) funding. We are much indebted for this support. This paper was prepared for the proceedings of the 2001 Southern Mensurationists' Conference. A more general description of this project is described in the companion paper by Cieszewski et al. (2002) that is included in the same proceedings. Another general description has been published in the Proceedings of the IUFRO Conference held in Vancouver, BC, Canada in 2001.

**LITERATURE CITED**

Abt, R.C., F.W. Cubbage, and Pacheco G., 2000. Southern Forest Resource Assessment using the Subregional Timber Supply (SRTS) model. For. Prod. Journal 50(4):25-33.

Cieszewski, C.J., M. Zasada, B.E. Borders, M.L. Clutter, R.F. Daniels, R.I. Elle, and R. Lowe. 2002. Long-term sustainability analysis of fiber supply in Georgia. In P.F. Doruska and D.C. Bragg (editors). 2002. Proceedings of the 2001 Southern Mensurationists' Conference, held in Chattanooga, TN, Nov. 4-6.

Cieszewski, C.J., M. Zasada, R.F. Daniels, M.L. Clutter, and B.E. Borders. 2001. Long-term Sustainability of fiber supply in Georgia - a Preliminary Approach. In: Forest Modelling for Ecosystem Management, Forest Certification and Sustainable Management. Proc. of the IUFRO Conf., Vancouver, BC, Canada. August 12-17. p. 173-182.

DR Systems Inc. OPTIONS V. Forest Estate Planning Model for Windows 95&NT. User's Guide.

Forbes, R.D. 1961. Forestry Handbook. The Ronald Press Company, New York.

Hansen, M.H., T. Frieswyk, J.F. Glover, and J.F. Kelly. 1992. The Eastwide forest inventory database: users manual. Gen. Tech. Rep. NC-151. St. Paul, MN: USDA Forest Service, North Central Forest Experiment Station. 48 p.

Harrison, W.M. and B.E. Borders. 1996. Yield Prediction and Growth Projection For Site-prepared Loblolly Pine Plantations in the Carolinas, Georgia, Alabama and Florida. PMRC Tech. Rep. 1996-1. Plantation Management Research Cooperative. Daniel B. Warnell School of Forest Resources. University of Georgia. Athens, GA. 64 p.

Martin, S.W. and G.H. Brister. 1999. A Growth and Yield Model Incorporating Hardwood Competition for Natural Loblolly Pine Stands in the Georgia Piedmont. So. J. Appl. For. 16(3):179-185.

Miles, P.D., G.J. Brand, C.L. Alerich, L.F. Bednar, S.W. Woudenberg, J.F. Glover, and E.N. Ezzell. 2000. The Forest Inventory and Analysis Database: Database Description and User's Manual Version 1.0. http://www.ncrs.fs.fed.us/4801/fiadb/fiadb_documentation/FIADB_DOCUMENTATION.htm

Nelson, T.C., J.L. Clutter, and L.E. Chaiken. 1961. Yield of Virginia Pine. Station Paper No. 124. Asheville, NC: USDA Forest Service, Southeastern Forest Experiment Station. 11 p.

Pienaar, L.V., B.D. Shiver, and J.W. Rheney. 1996. Yield Prediction for Mechanically Site-Prepared Slash Pine Plantations in The Southeastern Coastal Plain. PMRC Tech. Rep. 1996-3. Plantation Management Research Cooperative. Daniel B. Warnell School of Forest Resources. University of Georgia. Athens, GA. 57 p.

Schumacher, F.X. and T.S. Coile. 1960. Growth and Yields of Natural Stands of the Southern Pines. T.S. Coule Inc., Durham, N.C. 115 p.

Woudenberg, S.W. and T.O. Farrenkopf. 1995. The Westwide forest inventory data base: user's manual. USDA Forest Service, Intermountain Research Station, Gen. Tech. Rep. INT-GTR-317. 67 p.

# Current Status of Southern Annual Inventories

by
**Greg Reams**[1]
USDA Forest Service Southern Research Station, Asheville, NC

**Abstract**

The current status of research and operational components of the Southern FIA program are presented. Topics covered include the national commitment to use TM based classifications of forest cover from USGS, and database and estimation issues related to panel creep.

[1] *Citation for proceedings:* P.F. Doruska and D.C. Bragg (editors). 2002. Proceedings of the Southern Mensurationists' Conference, Chattanooga, TN, November 4-6, 2001. Arkansas Forest Resources Center, Monticello, AR. 43 p.

**Modifying the Daniels index to indicate when competition begins**

by

**Philip Radtke*, James Westfall, and Harold Burkhart**[1]

Virginia Polytechnic Institute and State University

**Abstract**

Data from a loblolly pine (*Pinus taeda* L.) spacing trial were used to investigate relationships between a distance dependent competition index (CI) and the inflection age of sigmoidal single-tree cumulative basal area growth curves. Inflection ages increased with increasing initial growing space, consistent with the hypothesis of maximum usable growing space. Competition intensity as measured by the CI was generally smallest at a given point in time for trees with large inflection ages, but the trend varied by planting density. CI values at the inflection age also varied with planting density. The CI was modified so that it gave a constant CI value at the inflection age, on average, across all planting densities. Effects of site quality were accounted for to a limited degree, but the range of sites was narrow in the spacing trial. The modified CI should be useful as an absolute measure of competition independent of spacing and, to a lesser degree, site quality.

[1] *Citation for proceedings:* P.F. Doruska and D.C. Bragg (editors). 2002. Proceedings of the Southern Mensurationists' Conference, Chattanooga, TN, November 4-6, 2001. Arkansas Forest Resources Center, Monticello, AR. 43 p.

**Planting Rectangularity: How Much Is Too Much?**

by
**Mahadev Sharma* and Ralph A. Amateis[1]**
Virginia Polytechnic Institute and State University

**Abstract**

The effect of rectangularity on tree growth and stand development was evaluated using annually measured tree data on a loblolly pine (*Pinus taeda* L.) spacing trial data through age 16. Spacings 1.83m x 2.44m and 2.44m x 1.83m with rectangularity 3:4 (or 4:3) and spacings 1.22m x 3.66m and 3.66m x 1.22m with rectangularity 1:3 (or 3:1) were used to evaluate the rectangularity effect.

[1] *Citation for proceedings:* P.F. Doruska and D.C. Bragg (editors). 2002. Proceedings of the Southern Mensurationists' Conference, Chattanooga, TN, November 4-6, 2001. Arkansas Forest Resources Center, Monticello, AR. 43 p.

# Using Nonparametric Regression to Examine Taper of Irregular Logs

by

**Paul F. Doruska[1] and David W. Patterson[2]**

**Abstract**

The volume of sawlogs can be calculated from the taper of the log. Many excellent taper equations have been developed over the years because of the importance of describing log (and tree taper). Occasionally, unusual logs are encountered, that is, anomalous logs with respect to taper. We examined the use of nonparametric regression to trace the taper of 11 such logs. The use of nonparametric regression and diagnostics used therein will be discussed in conjunction with describing the taper of these logs.

## INTRODUCTION

Parametric regression has long been a staple of forest research. Many researchers have published a myriad of papers using this tried and true technique. In many and perhaps most applications, parametric regression performs extremely well for describing the relationship between one or more variables or predicting the value of one variable based on the values of others.

Nonetheless, there are times when parametric regression does not perform adequately. In such cases, a researcher might consider nonparametric regression as an alternative. There are no functional forms to define in nonparametric regression, and no equation results. A curve is described, however, that defines the relationship between two variables. This paper will introduce the basics of nonparametric regression (kernel, local linear, and local quadratic) and apply said techniques to defining taper curves of logs with unusual curvatures present in the logs.

## BACKGROUND

*Kernel Regression*: Kernel regression traces a curve through the data according to the following formula (Härdle 1990).

$$\hat{y}_i = \sum_{j=1}^{n} w_{ij} y_j \qquad (1)$$

Any predicted response, $\hat{y}_i$, is obtained by multiplying each observation in the dataset, $y_i$ by a weight, $w_{ij}$. This formula is used to predict responses at the n points in the dataset as well as for any other potential point of prediction. The weights are determined by the combination of a kernel function, *K(u),* and a bandwidth, *h,* as follows (Nadaraya 1964, Nadaraya 1965, Watson 1964):

$$w_{ij} = K[(x_i - x_j)/h] / \sum_{i=1}^{n} K[(x_i - x_j)/h] \qquad (2)$$

The bandwidth determines the rate at which the respective weights of the observations go to 0 as one predicts at locations remote from the location of the respective observations. Härdle (1990) reported that the choice of bandwidth is more important than the choice of the kernel function. Hence, only the normal kernel function will be used herein.

$$K(u) = \exp(-u^2) \qquad (3)$$

Kernel regression is known to exhibit poorer fits at the boundaries of the data (Hastie and Loader 1993), and as exhibited in crown profile fits reported in Doruska and Mays (1998).

*Local Linear Regression:* Local linear regression is a form of nonparametric regression that first applies kernel regression to obtain weights to use in weighted least squares regression. Local linear regression applies these weights to an otherwise simple linear regression fit. Once such weights are determined, the parameter estimates for the local linear regression fit at each point of prediction *i* are found via:

$$\underline{\hat{\beta}}_i = \underline{x}_i \left( X' W_{x_i} X \right)^{-1} X' W_{x_i} \underline{y} \qquad (4)$$

See Cleveland (1979) and Hastie and Loader (1993) for expanded discussion on local linear regression.

*Local Quadratic Regression*: The mechanics of local quadratic regression mimic that of local linear regression. However, instead of applying the kernel regression found weights to a simple linear

[1] Assistant Professor, Biometrics, [2] Research Professor, Wood Science, Arkansas Forest Resources Center, School of Forest Resources, University of Arkansas, Monticello, AR 71656 USA

*Citation for proceedings:* P.F. Doruska and D.C. Bragg (editors). 2002. Proceedings of the Southern Mensurationists' Conference, Chattanooga, TN, November 4-6, 2001. Arkansas Forest Resources Center, Monticello, AR. 43 p.

regression, the weights are applied to a model that is polynomially quadratic with respect to x. See Cleveland (1979) and Hastie and Loader (1993) for further discussion and development of this form of nonparametric regression.

*Choice of Bandwidth*: There is some subjectivity involved with nonparametric regression, and that subjectivity manifests itself in bandwidth selection. The practitioner is free to select any bandwidth of his/her choosing. However, more objective criterion for bandwidth selection have been developed (see Härdle 1990). Also, a statistic similar to PRESS (Allen 1974), or cross validation (Rudemo 1982) can also be used. Based on past research (Doruska and Mays 1998), the PRESS* (Einsporn 1987) criterion will be used herein to select appropriate bandwidths.

*Past Use in Forestry*: Nonparametric regression garnered attention in the statistical literature in the mid to late 1980's. However, few applications to forestry have been found in the literature. See Doruska and Mays (1998) for a brief review of nonparametric regression use on forestry.

## DATA

The data used herein are a small subset of another dataset belonging to the second author. The larger dataset consisted of taper (diameter in inches) measurements made at intervals ranging from 2 inches to 1 foot along southern pine sawlogs. Eleven logs were noted to be quite irregular in form (see Figure 1 for an example) and not easily described by conventional taper functions, thus making them candidates for nonparametric regression.

Although there were eleven logs in the nonparametric regression dataset, only 5 will be shown herein. The results for the other six were similar.

## RESULTS

Table 1 contains the PRESS* selected bandwidth and the nonparametric regression estimate of $s^2$ (the nonparametric version of mean squared error) for the kernel, local linear, and local quadratic fits to each of the 5 logs depicted in Figures 2-6.

**Table 1.** Nonparametric fit statistics for the 5 logs depicted in Figures 2-6.

| Figure | Form | Bandwidth | $s^2$(in.$^2$) |
|---|---|---|---|
| 2 | kernel | 0.062 | 0.66 |
| | local linear | 0.128 | 0.69 |
| | local quadratic | 0.197 | 0.69 |
| 3 | kernel | 0.033 | 0.37 |
| | local linear | 0.060 | 0.40 |
| | local quadratic | 0.076 | 0.33 |
| 4 | kernel | 0.040 | 0.17 |
| | local linear | 0.043 | 0.17 |
| | local quadratic | 0.070 | 0.19 |
| 5 | kernel | 0.013 | 0.09 |
| | local linear | 0.092 | 0.91 |
| | local quadratic | 0.133 | 0.89 |
| 6 | kernel | 0.016 | 0.31 |
| | local linear | 0.054 | 1.21 |
| | local quadratic | 0.106 | 1.23 |

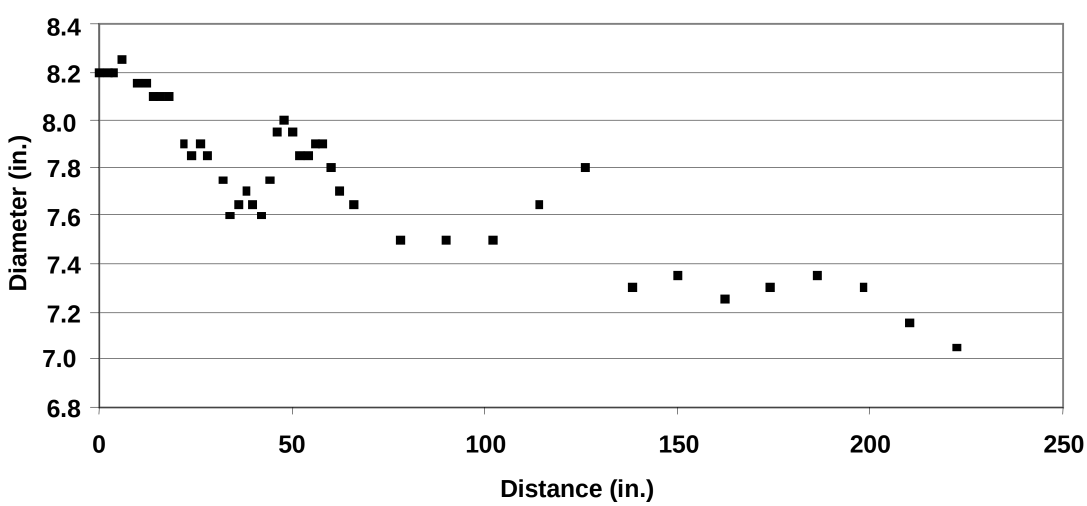



**Figure 1.** Unusual taper exhibited by one of the southern pine sawlogs.

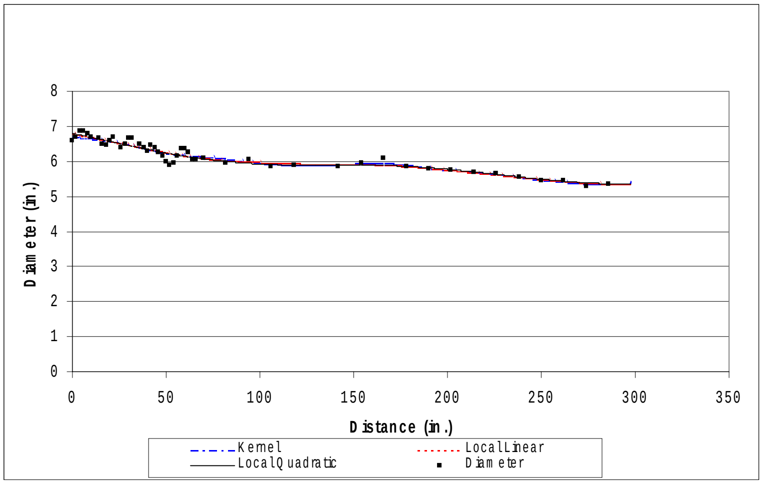



**Figure 2**. Nonparametric regression fits to one of the southern pine sawlogs.

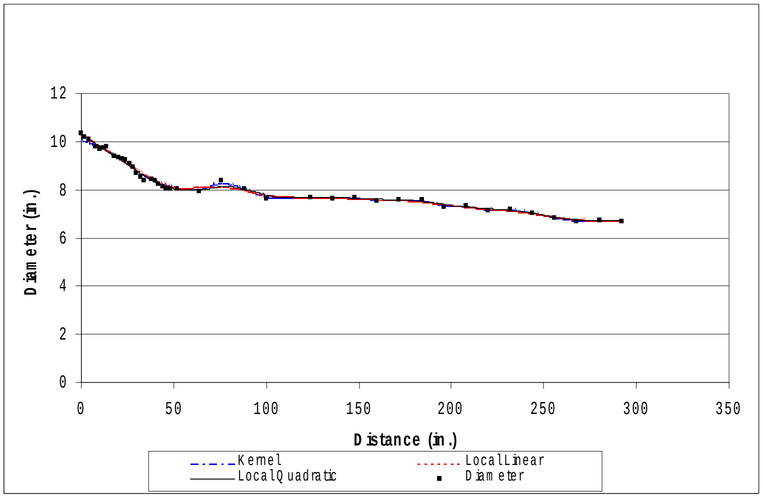



**Figure 3.** Nonparametric regression fits to one of the southern pine sawlogs.

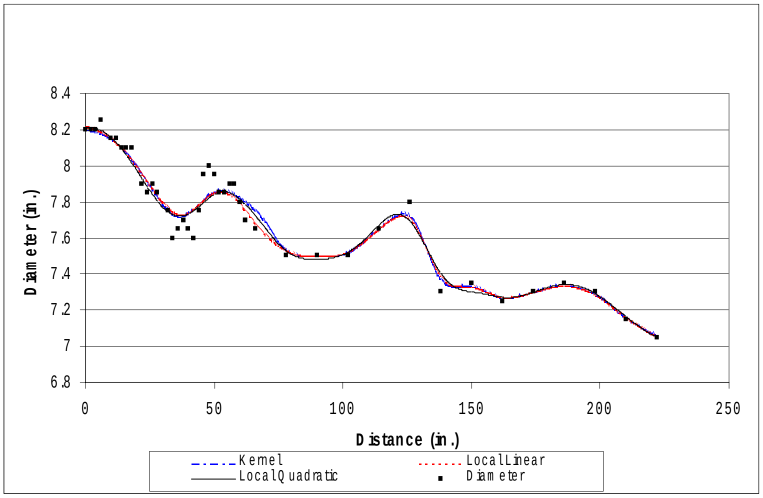



**Figure 4.** Nonparametric regression fits to one of the southern pine sawlogs.

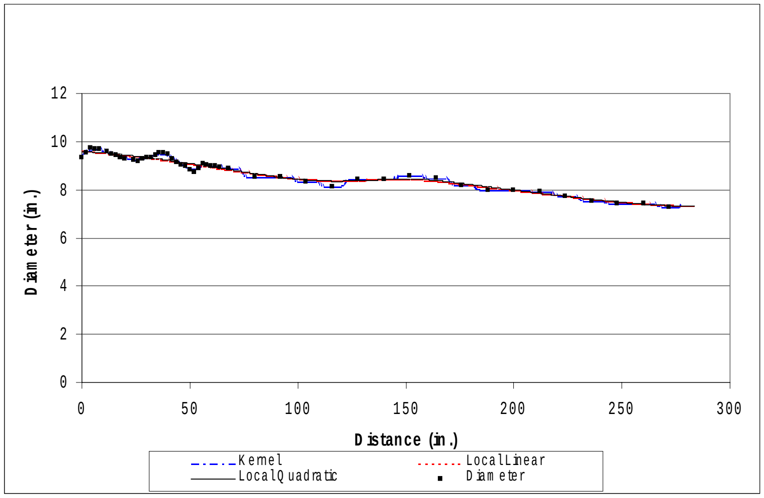



**Figure 5.** Nonparametric regression fits to one of the southern pine sawlogs.

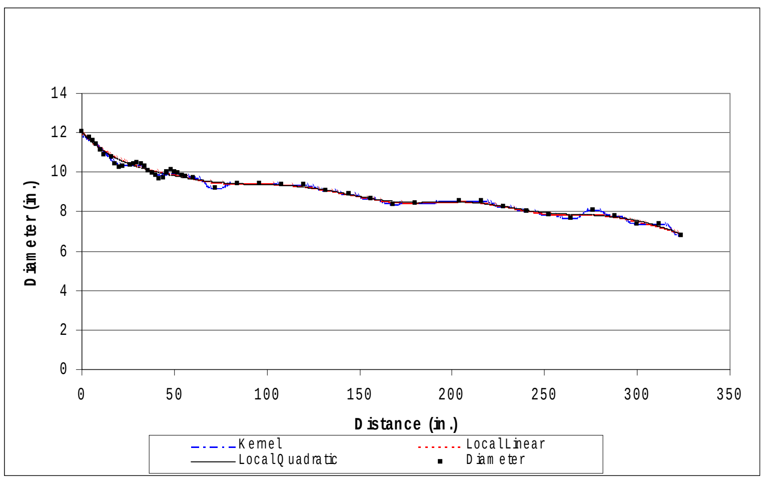



**Figure 6.** Nonparametric regression fits to one of the southern pine sawlogs.

A loose interpretation of the bandwidth is the proportion of the range in the X variable (distance from butt end of the log in this case) that carry weight in predicting Y (diameter in this case) at a given level of X. Since the kernel regression bandwidth was smallest of the three bandwidths in set of regression fits, it is not surprising that the $s^2$ of the kernel fits was smallest in all 5 fits shown.

Kernel regression, even when using PRESS* to select bandwidth, tended to overfit by assigning very little weight to points even slightly remote from the point of prediction. The kernel fits shown in Figures 5 and 6 show how the kernel fit just tends to connect the data points, and does not provide a smooth curve. More smoothing occurs with larger bandwidths, and the local linear and local quadratic regressions fit and shown herein possess this property.

## SUMMARY AND CONCLUSIONS

Nonparametric regression, at least with respect to kernel, local linear, and local quadratic regression, performed fairly well in describing taper of the irregular logs examined herein. Surprisingly, the kernel regression fits did not exhibit the boundary problems that typically accompany this form of nonparametric regression. However, kernel regression tended to result in model overfits, another undesirable quality. Based on the charts and fit statistics reported herein, local linear and local quadratic models provided adequate smoothing.

Unfortunately, closed form expressions of these taper curves do not exist (a feature of any form of nonparametric regression) so there is no taper curve to integrate from which volume can be determined. Though not undertaken in this paper, numerical integration techniques may be the solution to this situation. The authors intend to investigate this in another publication.

Future research using these data will include use of variable bandwidth nonparametric regression (see Müller and Stadtmüller 1987) as well as more recent advances in nonparametric regression such as locally nonparametric regression[3], locally non-parametric regression is the use of nonparametric regression only in those portions of a dataset where parametric regression performs poorly, use parametric regression everywhere else.

**ACKNOWLEDGMENTS**
The authors would like to thank the Arkansas Forest Resources Center for supporting this research Drs. Pelkki and Weih for their thoughtful review of this manuscript. This manuscript is approved for publication by the Director, Arkansas Agricultural Station.

**LITERATURE CITED**

Allen, D. 1974. The relationship between variable selection and data augmentation and a method for prediction. Technometrics 16:125-127.

Cleveland, W.S. 1979. Robust locally weighted regression and smoothing scatter plots. J. Am. Stat. Assoc. 74:829-836.

Doruska, P.F. and J.E. Mays. 1998. Crown profile modeling of loblolly pine by nonparametric regression analysis. For. Sci. 44:445-453.

Einsporn, R. 1987. HATLINK: A link between least squares regression and nonparametric curve estimation. Ph.D. Diss., Virginia Poly. Inst. and State University, Blacksburg, VA.

Härdle, W. 1990. Applied nonparametric regression. Econ. Soc. Monogr. No. 19. Cambridge University Press, New York. 333 p.

Hastie, T. and C. Loader. 1993. Local regression: automatic kernel carpentry (with discussion). Stat. Sci. 8:120-143.

Müller, H.-G. and U. Stadtmüller. 1987. Variable bandwidth kernel estimates of regression functions. Ann. Stat. 15:182-201.

Nadaraya, E.A. 1964. On estimating regressions. Theory Prob. Appl. 9:141-142.

Nadaraya, E.A. 1965. On non-parametric estimates of density functions and regression curves. Theory Prob. Appl. 10:186-190.

Rudemo, M. 1982. Empirical choice of histograms and kernel density estimates. Scand. J. Statist. 9:65-78.

Watson, G.S. 1964. Smooth regression analysis. Sankhya, Series A. 26:359-37.

[3] Conversation with Dr. James Mays, Virginia Commonwealth University.

# Fuzzy Set Classification of Old-Growth Southern Pine

by
**Don C. Bragg**[1]

**Abstract**

I propose the development of a fuzzy set ordination (FSO) approach to old-growth classification of southern pines. A fuzzy systems approach differs from traditional old-growth classification in that it does not require a "crisp" classification where a stand is either "old-growth" or "not old-growth", but allows for fractional membership in the set of old-growth stands. FSO produces a score ranging from 0.0 (highly different from old-growth) to 1.0 (completely residing in the set of old-growth stands). This value can also be interpreted as "apparent" age, or an approximation of stand age based on measured variables other than time. A FSO old-growth classification is less subjective than current regression or indexing procedures, most of which assign an arbitrary value to each classification variable. In this example highlighting southern pine, five characteristic features of old-growth (q-factor, maximum tree DBH, stand basal area, percent red heart infection, and large woody debris volume) are expressed as response functions to help differentiate between stands synthesized from historical descriptions of virgin timber in southern Arkansas.

## INTRODUCTION

Southern forests have been dramatically impacted by centuries of human influence. Logging, agriculture, settlement, introduced pests and pathogens, pollution, and other unnatural disturbances have notably altered natural ecosystems. Old-growth forests have been the most dramatically impacted: Davis (1996) estimated that less than 1% of original primary forests remain.

This rapid disappearance has placed modern-day public resource managers in a bind. Though expected to simultaneously protect threatened and endangered species (many of which are old-growth dependent), preserve and enhance existing resources, and maintain recreation and commodity production from a restricted land base, their options to address these issues are limited. New strategies like managing-for-old-growth have been proposed and are being implemented in small-scale field studies (e.g., Morton et al. 1991, Vora 1994), but the effectiveness of this approach has yet to be documented.

Management is further hindered by the lack of agreement on what constitutes old-growth (Hunter and White 1997). While the inherent differences between vegetation types precludes the development of a universal old-growth definition, there are also within-type issues. Different ecological thresholds such as minimum levels of woody debris or tree size are often considered, making it difficult to compare old-growth from one region (or study) to the next. Hunter and White (1997) noted that the arbitrariness of the current working definitions of old-growth did not improve management, as the thresholds used were often based on limited criteria poorly related to stand potential.

Several researchers have cautioned against the single feature classification strategy (e.g., Franklin and Spies 1991, Rusterholz 1996, Hunter and White 1997), preferring instead an index of "old-growthedness" in which multiple factors are scored to produce an old-growth evaluation system. For example, Franklin and Spies (1991) proposed a continuous scaling strategy, thus allowing for various degrees of old-growthedness. By assuming these characteristics fit a "U" or "S" shaped curve, they employed an arbitrary scale to reflect different stand developmental stages. The resulting values could then be summed to produce an index of old-growthedness. Rusterholz (1996) described a similar approach that applied criteria based on cover type. Each candidate stand was given a value for each of these criteria based on predetermined thresholds, and then the cumulative score (to a maximum of 65 points) was used to determine its status. For example, pine forests were evaluated using the following criteria: stand age, size, and context; degree of human intervention; pine regeneration; tree size class diversity; maximum tree size; and large woody debris volume. Stands with scores of $ 40 were recommended for old-growth protection.

Another old-growth classification scheme was described by Hale et al. (1999), who applied a logistic regression model to differentiate between managed mature hardwood forests and unmanaged old-growth. They evaluated seven parameters before settling on large woody debris (LWD) volume as the most

---

[1] Research Forester, USDA Forest Service, Southern Research Station, P.O. Box 3516 UAM, Monticello, AR 71656. Phone: (870) 367-3464, Fax: (870) 367-1164, E-mail: dbragg@fs.fed.us

*Citation for proceedings:* P.F. Doruska and D.C. Bragg (editors). 2002. Proceedings of the Southern Mensurationists' Conference, Chattanooga, TN, November 4-6, 2001. Arkansas Forest Resources Center, Monticello, AR. 43 p.

discriminating factor. While their fits were significant ($P < 0.05$), this procedure did not explain much of the variation in the data (adjusted $R^2 < 0.4$). This low explanatory power probably arose from the narrowness of old-growth defining criteria (LWD volume).

Part of the classification problem is the tendency to apply classical set theory when defining old-growth. In other words, a stand is either considered “old-growth” or “not old-growth,” usually with a singular threshold like stand age or average tree size. None of these systems have developed an objective approach in defining this critical threshold. Traditional set theory would require a threshold age (e.g., 200 years), for which anything older would be old-growth, and anything younger would not. But what do you do if a stand is 195 years old? A conceptually and mathematically rigorous multi-factor system to identify old-growth is needed. Even though numerical indexing (e.g., Franklin and Spies 1991, Rusterholz 1996) simultaneously incorporates multiple features in the identification of possible old-growth, it is prone to a high degree of subjectivity and statistical uncertainty. Logistic regression models are usually interpreted as all or nothing (even though their outcomes are fractional probabilities).

Recent developments in fuzzy set theory have provided a conceptual foundation that could address the problem of old-growth identification in a mathematically rigorous and ecologically sophisticated manner without burdening field managers with complex protocols and analysis procedures. Fuzzy sets, when generalized on dynamical systems theory (e.g., Roberts 1987*a*), produce a fuzzy systems theory which can then be used to determine the nature of the interaction between vegetative and environmental hyperspace (Roberts 1989). Thus, not only does fuzzy mathematics provide a more intuitive approach to many ecological questions, but it incorporates the dynamics involved between the key components of ecological systems.

A fuzzy set ordination (FSO) approach to old-growth classification represents an obvious departure from traditional approaches to old-growth delineation. The development of FSO old-growth classification was predicated on the following principles: 1) it must improve upon current classification procedures, and 2) it should be easily applied using traditional field measurements. This paper outlines conceptual and mathematical principles for a fuzzy set classification of old-growth southern pine.

## FUZZY SET ORDINATION

One of the major advantages of fuzzy sets is that they preserve the algebra of set theory, thus retaining the formal logic and mathematics of “crisp” (Euclidian) sets. Fuzzy mathematics can work on either continuous or discrete variables. FSO has considerable potential in ecological analysis because it does not depend on specific thresholds, but rather membership in fuzzy sets based on their degree of similarity to a reference set (Roberts 1989). Fuzzy set theory scores attributes based on their similarity to the largest (or upper) limit considered and assigns them a membership in the intended classification set. Thus, a stand that is 195 years old would receive a very high membership (probably > 0.95) in the fuzzy set of old stands. Traditional set theory, conversely, would reject this a stand as old-growth because it did not meet the minimum age threshold.

The following sections provide a brief synopsis of fuzzy set mathematics (more detailed reviews can be found in Roberts (1987*b*, 1989)). All parameters in this study were indexed to range from 0 to 1, with 1 representing the maximum value of that variable and 0 indicating the lowest value. As an example, a linear indexing would follow:

$$V = \frac{V_{actual} - V_{min}}{V_{max} - V_{min}} \tag{1}$$

where the indexed value ($V$) is a function of the actual ($V_{actual}$), minimum ($V_{min}$), and maximum ($V_{max}$) value of that parameter. By scaling variables between 0 and 1, the data became self-calibrated so that specific thresholds did not have to be predetermined. In traditional crisp sets, a stand would have membership in each of the defined sets as a 0 (absent) or 1 (present). Fuzzy sets allow for fractional membership in sets such that a stand could be anywhere in the range from 0 to 1. Each corresponding parameter set (generically referred to as $P$) are denoted by italicized capital letters. Thus,

$$P = \left\{ \left( x,\ m_P(x) \right) \right\} \tag{2}$$

where $x$ is an element of $P$ and $:_P(x)$ is the membership of $x$ in set $P$.

Because fuzzy sets retain the mathematics of traditional set theory, the following operators are defined for generic sets $P$ and $Q$:

union–

$$m_{P \cup Q}(x) = \max\left( m_P(x),\ m_Q(x) \right) \tag{3}$$

complement–

$$m_{\overline{P}}(x) = 1 - m_P(x) \tag{4}$$

difference–

$$m_{P-Q}(x) = \min\left( m_P(x),\ m_{\overline{Q}}(x) \right) \tag{5}$$

intersection–

$$m_{P \cap Q}(x) = \min\left( m_P(x),\ m_Q(x) \right) \tag{6}$$

Union, intersection, and complement can be considered “and,” “or,” and “not,” respectively

(Roberts 1986). Thus, in a multi-factor fuzzy set classification, one is first interested in how the sets orient themselves and then on defining factors. Roberts (1987*b*) adapted these operators to produce a new approach to sets: the anticommutative difference operator (ADO). The ADO (which can be considered as "while not") allows for contrasts between dissimilar sets:

$$m_{P\Gamma Q}(x) = \frac{\left[1 + \left(m_{\overline{Q}}(x)\right)^2 - \left(m_{\overline{P}}(x)\right)^2\right]}{2} \qquad (7)$$

The ADO allows for complex sets based on gradients to be developed. Since most environmental gradients tend to be complementary, stands can be considered "similar" to one end while not similar to the other. To arrive at this stage, plot variables must be associated with a similarity index. While many such indices abound, I applied Roberts' Index (Roberts 1986):

$$S_{xy} = \frac{\sum_{i=1}^{n}\left[\left(V_{xi} + V_{yi}\right)\left(\frac{\min\left(V_{xi}, V_{yi}\right)}{\max\left(V_{xi}, V_{yi}\right)}\right)\right]}{\sum_{i=1}^{n}\left(V_{xi} + V_{yi}\right)} \qquad (8)$$

where $S_{xy}$ is the similarity of stand $x$ to stand $y$, and $V$ is the indexed value calculated earlier for each of $n$ parameters. It is within the calculation of $S_{xy}$ that the parameters are combined to allow for a multi-factor old-growth classification. Since the primary consideration in determining old-growth is the age of the stand, I shall frame this in terms of age and "apparent" age. Age is directly measured for each stand, while apparent age is the fuzzy prediction deduced from their similarity to either young or old stands. To avoid circular conflicts, I did not use age as one of the indexed values. FSO is the juxtapositioning of apparent age with actual age. The set of stands *similar* to old stands (set $O$) was then:

$$m_O(x) = \frac{\sum_{y \neq x}\left[S_{xy}\left(m_A(y)\right)\right]}{\sum_{y \neq x}\left(m_A(y)\right)} \qquad (9)$$

where $\mu_O(x)$ is the membership of stand $x$ in set $O$ and $\mu_A(y)$ is the membership of stand $y$ in set $A$ (in this case, the set of old stands). Note that membership in $O$ is related to similarity to stands in set $A$ using equation (9). Likewise, the set of stands similar to young stands (set $Y$) is:

$$m_Y(x) = \frac{\sum_{y \neq x}\left[S_{xy}\left(m_{\overline{A}}(y)\right)\right]}{\sum_{y \neq x}\left(m_{\overline{A}}(y)\right)} \qquad (10)$$

where $\mu_Y(x)$ is the membership of stand $x$ in set $Y$ and $\mu_{\overline{A}}(y)$ is the membership of stand $y$ in set $\overline{A}$ (young stands, or the complement of the set of old stands). The newly defined sets ($\mu_O(x)$ and $\mu_Y(x)$) can then be placed in the ADO equation, re-standardized to range between 0 and 1, and then compared to measured stand age to indicate their position along the sere.

FSO results can be interpreted in several ways. Scores from individual stands can be ranked and evaluated. For example, it would be possible to use the apparent age gradient as a scaling for old-growthedness. Obviously, a score = 1 would indicate full membership in the set of old-growth stands, suggesting that all measured parameters were optimally met by this case. When scores fall between 0 and 1, then some condition(s) are less than maximum for a stand of a given age, which may or may not preclude the stand from further consideration as old-growth. Minimum levels of old-growthedness based on desired conditions could then be identified and managed for. For instance, old-growth reserve (i.e., no treatment) stands may have a value of 0.8 or greater, while those ranging from 0.6 to 0.8 could be considered as candidates for specialized treatment. Fuzzy set ordination scores could also be used to evaluate residual differences from the ordination graph and hence prove useful in identifying deficient or excessive conditions.

## METHODS

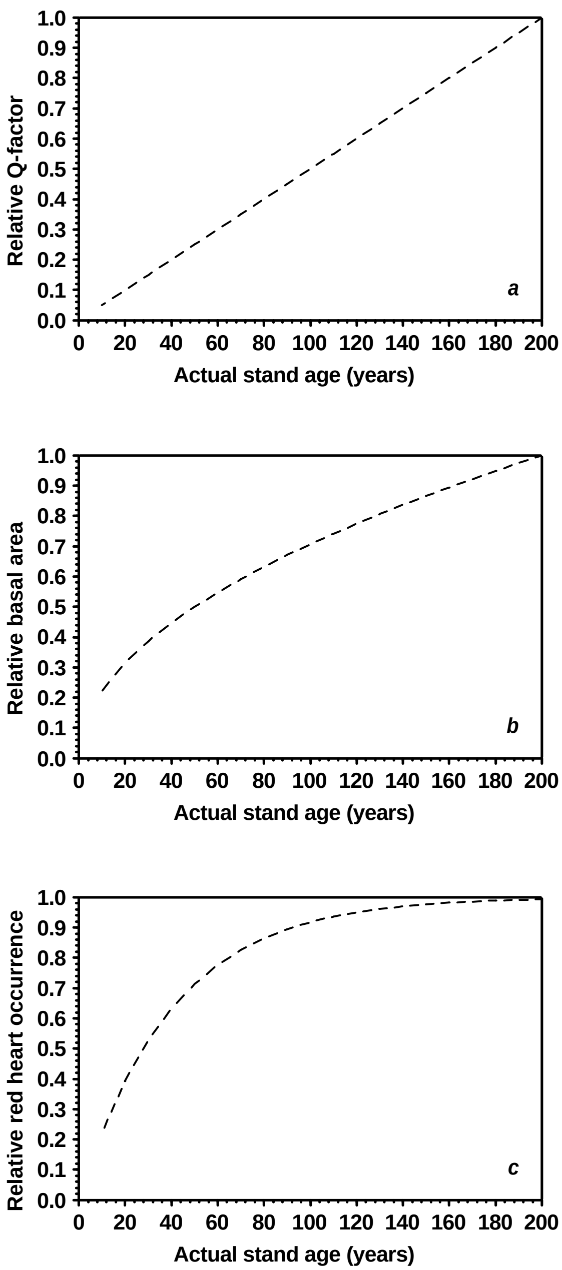
Using a set of derived gradients based on synthetic (but ecologically reasonable) trends for southern pine stands, a fuzzy set ordination was performed to anticipate stand age solely as a function of these parameters.

### Cover type selection and period delineation

The first step in any old-growth classification is the identification of the relevant cover type and time period. This is critical because one would not expect the parameters of interest for old-growth loblolly (*Pinus taeda* L.) and shortleaf pine (*Pinus echinata* Mill.) stands to be the same as those for baldcypress (*Taxodium distichum* (L.) Rich.) stands. The desired time period should also be identified, as conditions may also vary temporally. This effort considered factors for the virgin loblolly and shortleaf pine-dominated ecosystems of the Upper West Gulf Coastal Plain of Arkansas during the early 19th Century because 1) these forests were once common, but now are very limited; 2) they have an existing historical and contemporary literature base from which to parameterize; and 3) there is on-going research into managing-for-old-growth conditions, thus supporting the development of evaluative criteria.

**Parameter selection**
Any parameter with a functional relationship to stand age could be used (Roberts 1986). Conditions that were specifically quantifiable and unambiguous in the literature on old-growth pine were selected (rather than vague concepts like "absence of human disturbance"). Because the intention of this paper is to generally illustrate the FSO classification strategy, the values

**Figure 1.** Synthetic trends of parameters used to define response curves for fuzzy set ordination of old-growth southern pine (also see Appendix A).



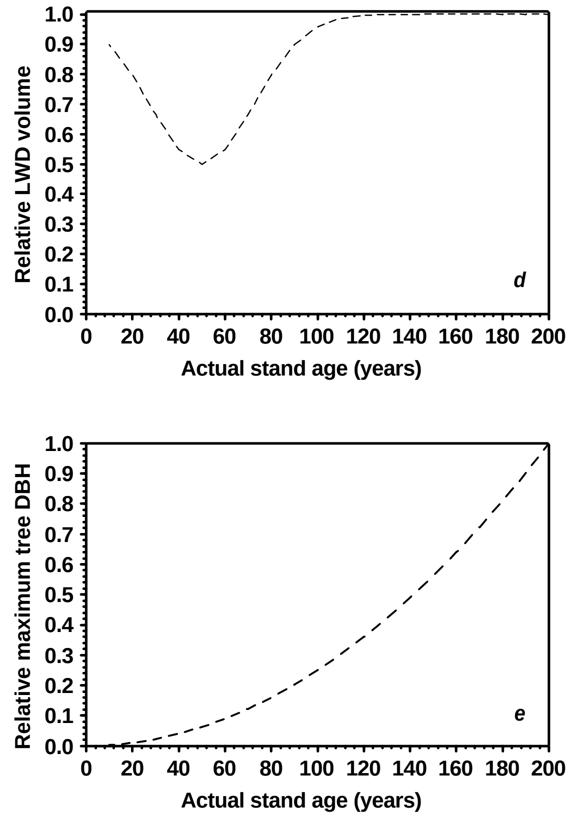




**Figure 1 (cont.).** Synthetic parameter trends.

presented were synthesized (without variance) from reasonable trends (Appendix A). Twenty stands were assembled from these synthetic values (Figure 1), combined using the Roberts similarity index, and then processed to produce an interpretable FSO.

The attributes used in this analysis included q factor, maximum tree diameter at breast height (DBH), stand basal area, red heart (*Phellinus pini* Ames) abundance, and LWD volume. These features are primarily structural, but should be well correlated with other less tangible old-growth attributes. Q factor is an abstraction of the relationship between stocking and diameter class, with higher numbers indicating a steeper trend (more small trees, few large ones) and a lower number suggesting fewer small trees and more big ones (typical of old-growth) (Smith 1986). Maximum tree DBH indicates the upper end of the structural condition of the forest, while stand basal area integrates size and stocking to suggest developmental stage. Red heart is a fungal heart rot that increases markedly as pine ages (Mattoon 1915). Dead wood volume is also strongly suggestive of development stage: old-growth usually contains substantial quantities of large LWD, while managed stands do not (e.g., Gore and Patterson 1986, Goodburn and Lorimer 1998, Hale et al. 1999).

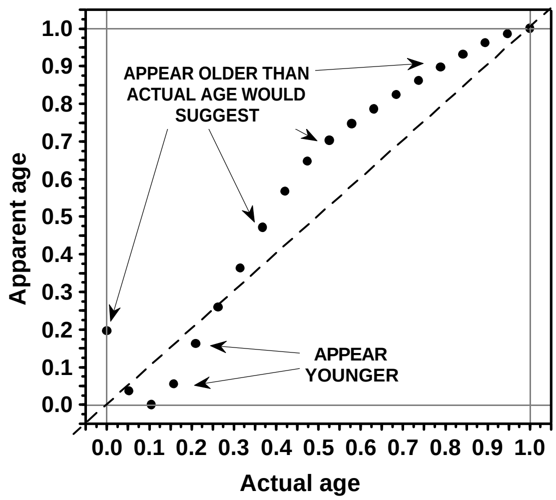



**Figure 2.** Apparent age (predicted from the fuzzy set ordination) compared to actual age (200 years when actual age = 1.0).

**RESULTS**

Fuzzy set ordination did a good job of predicting stand age from the variables it was provided (Figure 2). In general, the younger stands had attributes less like old-growth, while old stands were quite similar.

Residual differences are the deviations from the equivalence (dashed) line in Figure 2, and can be either positive or negative. Some stands appeared older than their chronological age would otherwise indicate, while others appeared younger than expected. The obvious departures from the 1:1 line in Figure 2 can be best understood by considering the features most responsible for this behavior. The bowing of the ordination results in Figure 2 is associated primarily with red heart abundance (Figure 1*c*). With the assumption of this study, the stands are noticeably overstocked with heart rot from a 1:1 expectation. The deviation apparent in young stands arose from the higher-than-expected volume of LWD present in these stands (Figure 1*d*). Large quantities of LWD are not unusual in young stands, especially those arising after catastrophic natural disturbances or timber harvesting (Sturtevant et al. 1997).

**DISCUSSION**

**FSO versus numerical indexing**

While the Franklin and Spies (1991) and Rusterholz (1996) procedures are more holistic than simple thresholds, they contain considerable subjectivity in their determination of old-growth point values. Since there is no mathematical basis to the values assigned, it could be argued that other sets of features or different emphasis on the criteria may result in a dramatically dissimilar outcome. FSO ordination avoids this issue because the measurements are scaled to those found in stands indisputably considered old-growth.

**FSO versus logistic regression**

A fuzzy set approach to old-growth classification is also an improvement over logistic regression analysis. Perhaps the biggest problem with a logistic approach is that it is inherently circular: to fit the regression, a stand must be classified *a priori* as “old-growth” or “not old-growth,” and then the coefficients are determined. Thus, using the resulting probability to categorize old-growth would not yield independent predictions. FSO does not require a defining variable like actual age to predict apparent age, and therefore avoids the problem of circularity. Additionally, the fitted nature of multivariate regression limits the interpretability of the residuals, and thus provide less utility in using that system to adaptively manage old-growth.

**Potential applications**

The interpretation and management directions suggested by residual analysis are some of the prime advantages to FSO. Identifying the factors leading to these discrepancies could be directly used to manage particular areas considered old-growth. Perhaps a stand appears younger than expected because of unusually low levels of LWD. This deficiency could be accommodated by the creation of new snags and/or downed logs. Individual parameters could be tested for their relative importance on the fuzzy old-growth classification by simple correlation analysis. Noticeable patterns may arise over part or all of the age gradient, which in turn can lead to further management emphasis on those components most sensitive to the correlation analysis.

The flexibility permitted by not having to define old-growth criteria *a priori* should also allow better customization of the process. This method also lacks the subjectivity of previous indexing methods as each variable used in the final analysis has been self-calibrated (as opposed to arbitrarily scored). The ability to combine multiple factors in an objective process will also improve classification from systems that key upon a single factor.

**Limitations and pitfalls of the method**

The success of a fuzzy approach to old-growth classification depends on our ability to identify clear patterns between stand age and parameters assumed to be indicative of old-growth-like conditions. Since old-growth stands are notoriously variable, only poor trends may appear, resulting in a weakly correlated classification outcome. FSO, however, is surprisingly robust to noise (Roberts 1998), so weak trends (noisy data) are not as detrimental to FSO as with other statistical approaches. It is also vital to sample a reasonably long temporal developmental gradient to help identify the key factors for classification because disturbances may cloud some of the relationships between stand structure and age (e.g., storm-related LWD accumulation).

**CONCLUSIONS**

Fuzzy set ordination appears to have considerable promise for old-growth classification. Even with a limited amount of structural parameters, it was possible to recover most of the structure of a synthetic gradient of different aged stands without specifically using age to organize the stands. FSO permits the direct interpretation of deviations from expected values in a manner rarely available for most old-growth classification strategies. This in turn suggests that management activities could be planned from the outcome of the ordination to optimize the value of existing stands for future action.

**ACKNOWLEDGMENTS**

I would like to thank Dave Roberts (Utah State University) for his inspiration and support of this effort, especially during its early stages. Thoughtful review comments by Mike Shelton (USDA Forest Service) and Hope Bragg (University of Arkansas-Monticello) improved the quality of this paper.

**LITERATURE CITED**

Davis, M.B. 1996. Extent and location. *In* Eastern Old-Growth Forests: Prospects for Rediscovery and Recovery, M.B. Davis, ed. Island Press, Washington, DC. p. 18-32.

Franklin, J.F. and T.A. Spies. 1991. Ecological definitions of old-growth Douglas-fir forests. *In* Wildlife and Vegetation of Unmanaged Douglas-fir Forests, L.F. Ruggiero, K.B. Aubry, A.B. Carey, and M.H. Huff, tech. coord. USDA Forest Service General Technical Report PNW-285. p. 61-69.

Goodburn, J.M. and C.G. Lorimer. 1998. Cavity trees and LWD in old-growth and managed northern hardwood forests in Wisconsin and Michigan. Canadian Journal of Forest Research 28:427-438.

Gore, J.A. and W.A. Patterson. 1986. Mass of downed wood in northern hardwood forests in New Hampshire: potential effects of forest management. Canadian Journal of Forest Research 16:335-339.

Hale, C.M., J. Pastor, and K.A. Rusterholz. 1999. Comparison of structural and compositional characteristics in old-growth and mature, managed hardwood forests of Minnesota, U.S.A. Canadian Journal of Forest Research 29:1479-1489.

Hunter, M.L. and A.S. White. 1997. Ecological thresholds and the definition of old-growth forest stands. Natural Areas Journal 17:292-296.

Mattoon, W.R. 1915. Life history of shortleaf pine. USDA Bulletin 244. 46 p.

Morton, D. and the New Perspectives Team. 1991. Maintaining old growth in northern hardwood ecosystems. *In* Proceedings of the 1991 Society of American Foresters National Convention, SAF Publication Number 91-05. p. 527-528.

Roberts, D.W. 1986. Ordination on the basis of fuzzy set theory. Vegetatio 66:123-131.

Roberts, D.W. 1987*a*. A dynamical systems perspective on vegetation theory. Vegetatio 69:27-33.

Roberts, D.W. 1987*b*. An anticommutative difference operator for fuzzy sets and relations. Fuzzy Sets and Systems 21:35-42.

Roberts, D.W. 1989. Fuzzy systems vegetation theory. Vegetatio 83:71-80.

Roberts, D.W. 1998. Fuzzy set ordinations are resistant to noise and inappropriate transformations of environmental data. Oral Presentation at the 83rd Annual Meeting of the Ecological Society of America. Ecological Society of America Abstracts, p. 112.

Rusterholz, K.A. 1996. Identification and protection of old growth on state-owned land in Minnesota. *In* Eastern Old-Growth Forests: Prospects for Rediscovery and Recovery, M.B. Davis, ed. Island Press, Washington, DC. p. 233-244.

Smith, D.M. 1986. The practice of silviculture. 8th ed. John Wiley & Sons, New York, NY. 527 p.

Sturtevant, B.R., J.A. Bissonette, J.N. Long, and D.W. Roberts. 1997. Coarse woody debris as a function of age, stand structure, and disturbance in boreal Newfoundland. Ecological Applications 7:702-712.

Vora, R.S. 1994. Integrating old-growth forest into managed landscapes: a northern Great Lakes perspective. Natural Areas Journal 14:113-123.

**Appendix A.** Realization of synthetic[a] trends assumed in Figure 1, including both actual and indexed (Rel.) values.

| Stand Number | Stand Age (yrs.) | Q Factor (unitless) | Rel. Q Factor | Max. DBH (cm) | Rel. Max. DBH | Basal Area ($m^2$/ha) | Rel. Basal Area | Red Heart (%) | Rel. Red Heart | LWD Volume ($m^3$/ha) | Rel. LWD Volume |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 200 | 1.1 | 1.0000 | 140.0 | 1.0000 | 20.0 | 1.0000 | 19.9 | 0.9933 | 40.0 | 1.0000 |
| 2 | 190 | 1.2 | 0.9500 | 127.3 | 0.9025 | 20.5 | 0.9747 | 19.8 | 0.9913 | 40.0 | 1.0000 |
| 3 | 180 | 1.3 | 0.9000 | 115.3 | 0.8100 | 21.0 | 0.9487 | 19.8 | 0.9889 | 40.0 | 1.0000 |
| 4 | 170 | 1.4 | 0.8500 | 103.9 | 0.7225 | 21.6 | 0.9220 | 19.7 | 0.9857 | 40.0 | 1.0000 |
| 5 | 160 | 1.5 | 0.8000 | 93.2 | 0.6400 | 22.1 | 0.8944 | 19.6 | 0.9817 | 40.0 | 1.0000 |
| 6 | 150 | 1.6 | 0.7500 | 83.1 | 0.5625 | 22.7 | 0.8660 | 19.5 | 0.9765 | 40.0 | 1.0000 |
| 7 | 140 | 1.7 | 0.7000 | 73.7 | 0.4900 | 23.3 | 0.8367 | 19.4 | 0.9698 | 40.0 | 0.9998 |
| 8 | 130 | 1.8 | 0.6500 | 64.9 | 0.4225 | 23.9 | 0.8062 | 19.2 | 0.9612 | 40.0 | 0.9992 |
| 9 | 120 | 1.9 | 0.6000 | 56.8 | 0.3600 | 24.5 | 0.7746 | 19.0 | 0.9502 | 39.9 | 0.9963 |
| 10 | 110 | 2.0 | 0.5500 | 49.3 | 0.3025 | 25.2 | 0.7416 | 18.7 | 0.9361 | 39.5 | 0.9864 |
| 11 | 100 | 2.1 | 0.5000 | 42.5 | 0.2500 | 25.9 | 0.7071 | 18.4 | 0.9179 | 38.4 | 0.9590 |
| 12 | 90 | 2.2 | 0.4500 | 36.3 | 0.2025 | 26.6 | 0.6708 | 17.9 | 0.8946 | 36.0 | 0.8991 |
| 13 | 80 | 2.3 | 0.4000 | 30.8 | 0.1600 | 27.4 | 0.6325 | 17.3 | 0.8647 | 31.9 | 0.7968 |
| 14 | 70 | 2.4 | 0.3500 | 25.9 | 0.1225 | 28.2 | 0.5916 | 16.5 | 0.8262 | 26.6 | 0.6649 |
| 15 | 60 | 2.5 | 0.3000 | 21.7 | 0.0900 | 29.0 | 0.5477 | 15.5 | 0.7769 | 21.9 | 0.5476 |
| 16 | 50 | 2.6 | 0.2500 | 18.1 | 0.0625 | 30.0 | 0.5000 | 14.3 | 0.7135 | 20.0 | 0.5000 |
| 17 | 40 | 2.7 | 0.2000 | 15.2 | 0.0400 | 31.1 | 0.4472 | 12.6 | 0.6321 | 21.9 | 0.5476 |
| 18 | 30 | 2.8 | 0.1500 | 12.9 | 0.0225 | 32.3 | 0.3873 | 10.6 | 0.5276 | 26.6 | 0.6649 |
| 19 | 20 | 2.9 | 0.1000 | 11.3 | 0.0100 | 33.7 | 0.3162 | 7.9 | 0.3935 | 31.9 | 0.7968 |
| 20 | 10 | 3.0 | 0.0500 | 10.3 | 0.0025 | 35.5 | 0.2236 | 4.4 | 0.2212 | 36.0 | 0.8991 |

[a] Trends are “synthetic” in that they reflect reasonable estimates of a parameter at the given age of the stand, but do not represent field-measured values. Due to rounding, some actual values many not precisely correspond to indexed ones.