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', 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 IO-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

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°F in January to 72°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 homeblende 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 ≥ I -in dbh. Plot data consisted of four environmental variables: elevation (nearest 10 feet), aspect (classes of 45" 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 IO-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 mockemut (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 2 10 plots is an index (I,) 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 l in dbh at the beginning of the period, expressed as

$$I_{,} = 0.1736(\sum_{N}(d_f^2 - d_i^2)/\sum_{N} d_i),$$

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 multinornial 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 IO-yr productivity ranged from 0.134 ft^2/ac/in to 1.2 18 ft^2/ac/in on the 2 10 sample plots and averaged 0.405 ft^2/ac/in (-10.188 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 2 10 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 2 10 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 $\geq$10 plots and were retained for the analysis (Table 1). The hickory, birch, and serviceberry tree groups each consisted of more than one species. The overall (27 rows x 3 columns) contingency analysis indicated that species frequency on plots was not independent of classification by productivity (P = 0.0001). Chi-square tests of hypothesized ratios of species occurrence of 1:2:1 for sites of low, medium, and high productivity respectively indicated that chestnut oak, scarlet oak, and hickory were associated ($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 2 10 FIA plots in western North Carolina

Species	Productivity class			Chi-square probability	Total Plots	Percent of total
	Low	Medium	High			
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 pseucioacacia*)	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	b.202	40	18.8
Northern red oak (*Quercus rubra*)	15	37	10	0.299	62	29.1
Virginia pine (*Pinus virginiana*)	2	10	7	0.261	19	8.9
Cucumber-tree (*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	0.93 1	21	9.8
Shortleaf pine (*Pinus echinata*)	4	7	3	0.93 1	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.

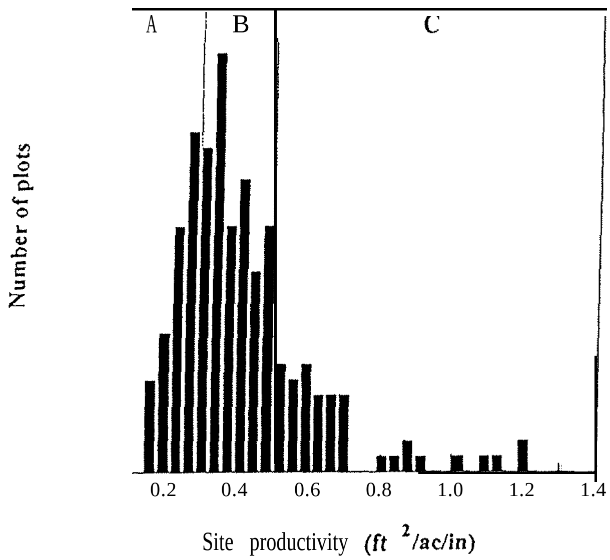

Figure 1. Distribution of the 2 10 sample plots by low (A), medium (B), and high (C) productivity classes. The low (<0.280 ft^2/ac/in) and high (>0.489 ft^2/ac/in) classes each consist of the *lower* and upper 25 percentile of plots ranked by productivity; the medium (LO.280 to ≤0.489 ft^2/ac/in) class consists of the remaining SO 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	0.02 1
	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	0.02s
	Chestnut oak	-1.17s	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 Iower 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 defming 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 2 10 FIA plots in western North Carolina.

Logistic model[a]	Independent variable	Regression Coefficient	P > |z|
Low	Constant	-1.460	0.100
	Elevation	0.0003	0.25 1
	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	0.00 1

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 2 10 FIA plots in western North Carolina.

Observed class	Predicted class: Low	Medium	High	Totals
Low	7	39	8	54
Medium	S	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 (Bums 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, Bums 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 (Bums 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 (Bums 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 (7 1 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. 198 1. Yellow-poplar: characteristics and management. Agricultural Handbook 583. U.S. Department of Agriculture. Forest Service. Washington, DC. 91 p.

Bums, 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.

Cat-mean, 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):9 15-9 18.

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

Golden, MS. 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-15 1. 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. 53 1 p.

Lexen, B. 1943. Bole area as an expression of growing stock. Journal of Forestry 4 1:883-885.

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.

Spun, 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.