Effects of Plot Size on Forest-Type Algorithm Accuracy

James A. Westfall[1]

Abstract: *The Forest Inventory and Analysis (FIA) program utilizes an algorithm to consistently determine the forest type for forested conditions on sample plots. Forest type is determined from tree size and species information. Thus, the accuracy of results is often dependent on the number of trees present, which is highly correlated with plot area. This research examines the sensitivity of a forest-type algorithm to changes in amounts and types of input data that result from altering the sample plot area. Logistic regression was used to determine which plot metrics have the most influence on algorithm output. Relationships between plot area and key variables such as number of species, number of trees, and total basal area were established and applied to the regression models. The results allow for assessment of algorithm accuracy over a range of plot sizes. The algorithm was generally robust to changes in area for loblolly/shortleaf, oak/hickory, and oak/gum/cypress type groups. Algorithm accuracy was mediocre for other type groups, with oak/pine having the poorest performance. A comparison between field-observed forest type and algorithm output showed average agreement rates of near 90 percent when computed types were conifer. However, agreement rates were lower for hardwood groups, especially when the computed type was aspen/birch. Better alignment between the field- and algorithm-based determinations may be achieved by providing real-time algorithm output to field crews.*

Keywords: forest inventory, logistic regression, species diversity, classification

[1] Research Forester, USDA Forest Service, Northern Research Station, Newtown Square, PA.

In: McWilliams, Will; Moisen, Gretchen; Czaplewski, Ray, comps. 2009. 2008 Forest Inventory and Analysis (FIA) Symposium; October 21-23, 2008: Park City, UT. Proc. RMRS-P-56CD. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 1 CD.

Introduction

Eyre (1980) describes forest type as a “descriptive classification of forestland based on present occupancy of an area by tree species”. The contributions to site occupancy are often determined via the numbers and sizes (e.g., diameter at breast height [dbh]) of trees for each species (Hansen and Hahn 1992). The relative occupancies among species (or groups of species) are used to establish the forest-type classification. Due to the relatively large number of described forest types and pronounced similarities among a number of types, forest-type groups are often created. This allows a number of related forest types to be classified under a single designation, which is often useful for broader analytical summarizations.

In many forest inventories, the forest type may be assessed by the field crew at the time the sample data are collected, determined at a later time by applying a computer algorithm to the sample plot condition data, or both. The Forest Inventory and Analysis (FIA) program of the U.S. Forest Service uses both field- (USDA 2007) and algorithm-based (Arner et al. 2003) forest-type information. Generally, the algorithm-based forest type is used in estimation. However, if a forested condition is less than one subplot in area (~0.0415 ac) the field-based forest type is used. It is assumed the algorithm cannot accurately determine the forest type when the area is relatively small, because often few trees are present on which to make a determination. As area and numbers of trees are highly correlated, the question that arises is what affect does sampled-area size have on an algorithm-based determination of forest-type. An understanding of the accuracy of algorithm-determined forest type in relation to area sampled will allow forest managers to make informed decisions regarding the appropriate method of forest-type classification for particular forest inventory designs.

Data

Evaluation of algorithm classifications at various sampled-area sizes was accomplished using FIA data from Indiana (1999-2003), South Carolina (2002-2006), and Maine (1999-2003). The states were chosen so that many of the forest types encountered in the eastern United States. would be represented. The data were collected under the annual inventory design outlined by Bechtold and Patterson (2005). Sample plots are composed of four subplots, each having a 24-ft radius. Within each subplot is a microplot having 6.8-ft radius. Trees having 5.0 in. or larger dbh were tallied on the subplots. Sapling (1.0-4.9 in. dbh) and seedling (< 1.0 in. dbh with minimum height criteria) data were recorded on the microplots. To facilitate the analysis, only single-condition plots were retained. In order to have a large number of possible plot combinations for repeated simulations, only forest-type groups having more than 100 plots were evaluated. There were 3,712 plots in the study data representing 55 forest types within eight forest-type groups. Table 1 provides a summary of the data by forest-type group and forest type.

Methods

The forest-type algorithm used in this study is described by Arner et al. (2003). This algorithm uses relative stocking to assess the site occupancy of sample trees. Individual-tree stocking values are computed from species-specific equations using tree dbh. Further adjustments (e.g., weighting) may be made based on tree-size classification and social position. The individual-tree values are aggregated into initial type assemblages and the stocking totals of these initial groups are evaluated via decision rules to determine the final forest type. Forest types are hierarchically assigned to a more generic forest-type group, so forest-type group determination is straightforward once the forest type is established.

The accuracy of the algorithm-based classifications was examined for forest-type groups, which are assemblages of similar forest types. The analysis consisted of two phases: 1) combining a number of plots with the same forest-type group and then systematically reducing the area of the combined plots and re-evaluating the forest-type group to see if the classification changes; and 2) using the results of (1), perform logistic regression to evaluate which plot attributes are correlated with the classification changes and predict probabilities of correct classification.

In the first phase, a Monte-Carlo simulation (Metropolis and Ulam 1949) was performed by combining 30 randomly selected plots (without replacement) having identical forest-type group classification into a ‘population’ of 5 acres in size (30 x 1/6 ac = 5 ac). Forest-type group was determined for this combination of plots. The area was then reduced by 1/24 ac by removing a randomly selected subplot and the forest-type group was re-evaluated. This area reduction method was carried out until only a single subplot remained (1/24 ac). This allowed for evaluation of potential forest inventory plot sizes ranging from 1/24 ac to 5 ac. The resultant output for the 120 different plot sizes included a binary variable that indicated whether the classification had changed from the original type and also summary variables such as numbers of species and numbers of stems for seedlings, saplings (1.0-4.9 in. dbh), and trees (5.0+ in. dbh), and basal area for saplings and trees. This process was repeated 500 times for each forest-type group; results were quite stable after 300 iterations.

These data were then used in a logistic regression analysis where the binary response variable was whether or not the type classification had changed at any given reduced area. Independent model variables considered were the summary variables described above (with two-way and three-way interactions). A stepwise variable-selection procedure was used to identify variables having significant (α = 0.10) predictive ability. The α level of 0.10 was chosen to promote inclusion of more variables that may help explain the classification changes. These logistic regression models provided the basis for predicting the probability that forest-type group would be correctly identified at a specified plot size.

Regression models relating the summary variables to plot area were developed to describe average plot attributes at the various plot sizes. The relationships in the data suggested linear relationships between plot area and numbers of stems as well as plot area and basal area. Nonlinear relationships existed between area and numbers of species. The model forms were:

$$S_{jk} = \beta_{1jk} A_j + \varepsilon_{jk} \quad [1]$$

$$SPP_{jk} = \beta_{2jk} \times A_j^{\beta_{3jk}} + \varepsilon_{jk} \quad [2]$$

$$BA_{jk} = \beta_{4jk} A_j + \varepsilon_{jk} \quad [3]$$

where: j = tree size class (seedling, sapling, and tree)
k = forest-type group
S_{jk} = number of stems tallied for tree size class *j*, forest type *k*
SPP_{jk} = number of species tallied for tree size class *j*, forest type *k*
BA_{jk} = basal area of stems tallied for tree size class *j*, forest type *k*
A_j = sampled plot area (ac) for tree size class *j*
ε_{jk} = random error component for tree size class *j*, forest type *k*
$\beta_{1jk} - \beta_{4jk}$ = estimated coefficients for tree size class *j*, forest type *k*

The estimated coefficients are presented in Table 2. The predicted values from models [1] through [3] were used as inputs into the logistic regression model to predict the probability of misclassification for a given plot area. This analytical approach was carried out separately for each forest-type group.

Results

The logistic regression analyses were conducted for the eight forest-type groups. The general form of the model was:

$$P_k(\text{Correct}) = f\left(S_{jk}, BA_{jk}, SPP_{jk}, \times 2, \times 3\right) + \varepsilon_k \quad [4]$$

where: P_k(Correct) = Probability of correct classification for forest-type group *k*
×2 = all two-way interactions of the predictor variables
×3 = all three-way interactions of the predictor variables
ε_k = random error component for forest-type group *k*
all others as defined above

The variables chosen by the stepwise selection procedure varied considerably among the groups. Across all eight type groups analyzed, there were 34 different significant predictor variables related to the probability of correct classification of forest-type group (the detailed information is not provided here due to size limits). The models fit the data reasonably well with R^2 values ranging from 0.43 to 0.64 (Table 3). The AIC (Akaike 1974) statistics also showed that the addition of covariates to the model substantially improved the prediction when compared to an intercept-only model.

The probability of correct classification of the white/red/jack pine and spruce/fir groups was influenced primarily by numbers of stems, numbers of different species, and basal area for saplings and trees. The classification accuracy of the loblolly/shortleaf pine group was affected mostly by numbers of stems, numbers of different species, and basal area for trees only. Conversely, the hardwood-type groups were more complex due to increased numbers of significant predictor variables, such numbers of stems and numbers of species for seedlings and various two-way interactions between these variables and the sapling and tree covariates. The oak/pine group had the most intricate model, with numerous three-way interactions being significant explanatory variables.

Inputs into the logistic regression model for each forest-type group were generated using models [1] through [3] for plot sizes ranging from 1/24 to 5 ac. The sensitivity of the algorithm to changes in forest parameters due to sample plot size was dependent upon the type group of interest. Within conifer types, the loblolly/shortleaf pine forest-type group was the most robust, as the probability of classification error was only 0.15 for 1/24 ac plot size (Figure 1c). The white/red/jack pine and spruce/fir type groups were more sensitive to area sampled, with the probability of misclassification being 0.3 - 0.4 at a plot area of only 1/24 ac (Figure 1a, 1b). A sampled area of roughly 0.2 ac. was needed to attain a nearly zero misclassification probability for loblolly/shortleaf, while the other two conifer types required about 0.5 ac.

For hardwood forest-type groups, the most stable classifications across the various plot sizes were in the oak/hickory and oak/gum/cypress groups (Figure 1e, 1f). For these groups, the probability of misclassification was near 0.1 at the smallest plot size evaluated (1/24 ac). Near-zero probabilities were achieved at a plot size of roughly 0.25 ac for oak/gum/cypress and nearly 0.5 ac for oak/hickory. The oak/pine group required plot sizes of over 2.5 ac to attain near-zero misclassification rates (Figure 1d). At a 1/24 ac plot size, the oak/pine group had misclassification probability of 0.62 and was 0.24 for the maple/beech/birch group. The maple/beech/birch group required a plot size of about 0.45 ac to obtain a misclassification probability less than 0.001 (Figure 1g). For the aspen/birch group, the maximum misclassification probability was near 0.29 (at 1/24th ac plot size) and near-zero probabilities occurred at about 0.9 ac (Figure 1h).

The forest-type algorithm always provides the same forest-type group for a given set of input data from the sample plot. However, the field crews have the advantage of viewing the entire area – their determination is not limited to only trees within the sample plot. Also, a certain amount of subjectivity is introduced based on the field crew's perception of the area. These factors can result in differing outcomes between the field-based and algorithm-based forest-type group. Table 4 quantifies the agreement/disagreement proportions for the forest-type groups analyzed in this study. Agreement was relatively high for softwoods, with red/white/jack pine having ~81 percent agreement and both spruce/fir and loblolly/shortleaf having agreement rates exceeding 90 percent. The conformity for hardwoods was poorer, as both aspen/birch and oak/pine had agreement rates less than 50 percent. When the algorithm determined the type was aspen/birch, the field call was spruce/fir for nearly 40 percent of the plots. The best agreement between algorithm and field hardwood type groups was for oak/gum/cypress, which had identical results for roughly 88 percent of the plots. Overall, agreement between field crew and algorithm occurred for ~ 75 percent of the plots.

Discussion/Conclusion

For the red/white/jack pine, spruce/fir, and maple/beech/birch groups, the algorithm classification accuracies decreased relatively quickly at plot sizes below 1/4 ac. This outcome is a reflection of the algorithm threshold for information needed to accurately classify these type groups. A review of the description for each type group indicates a wide range of species occur within these type groups (Eyre 1980). For example, spruce and fir species occur in areas where aspen, birch, and maple are also present. As plot size is reduced below 1/4 ac, the dominance of the spruce/fir species becomes more ambiguous, and the decision rules employed in the algorithm may produce a classification outside the spruce/fir group. The most common classification error for both red/white/jack pine and spruce/fir groups was maple/beech/birch. Similarly, a common misclassification of maple/beech/birch was spruce/fir type.

In contrast, there should be much less concern regarding misclassification of the loblolly/shortleaf pine group. These plots often come from planted areas where other species (primarily hardwoods) occur in the understory, which makes the preeminence of the primary species more apparent for smaller plots. In cases where loblolly/shortleaf was misclassified, oak/pine was by far the most common outcome.

A notable characteristic for the oak/pine and (to a lesser extent) aspen/birch groups was a relatively slow improvement in classification accuracy as plot sizes increased. For oak/pine, numbers of species, numbers of stems, and basal area among the three tree size classes all contributed to the misclassification rate. The confusion within aspen/birch was due primarily to species, stems, and basal area of trees having dbh 5.0 in. or larger. The oak/pine group required over 2.5 ac plot size to attain near-zero misclassification probabilities, while the aspen/birch group needed slightly less than 1 ac. In addition, the oak/pine group had the worst classification accuracy of all groups evaluated, with a probability of misclassification exceeding 0.6 when plot size was 1/24 ac. This gives further support to the argument given above related to species mixes. On plots where there is a wide range of species, it is difficult to determine the dominant type and relatively small shifts in the tree list can sway the classification in a different direction. Common misclassifications of oak/pine were loblolly/shortleaf and oak/hickory groups. The aspen/birch group was most often mistaken with spruce/fir and while maple/beech/birch, owing to the primary species of this group often being replaced by more shade-tolerant species, resulting in relatively high numbers of species and differing tree sizes.

The relationships between area sampled and misclassification probability for the oak/hickory and oak/gum/cypress groups were similar to those for loblolly/shortleaf pine. This is presumably attributable to the tendency for these species groups to be fairly well defined, such that the dominant species are likely to survive and flourish relative to species that are primary to other type groups. The oak/gum/cypress sites also tend to be undisturbed and have large diameter trees. These large trees provide high stocking values that are very influential in the computations, especially at the smaller plot sizes. Misclassifications were due primarily to confusion with the oak/pine and either maple/beech/birch or elm/ash/red maple groups.

There are two primary differences between field observation and algorithm-based forest-type group determination. The field crews have the advantage of viewing the broader area, not just the area within the plot. However, there is also an element of subjectivity such that different crews may resolve different forest types when assessing the same area. A feature of the algorithm is that the same forest type will be computed for a given tree list, removing any subjectivity. The drawback of the algorithm is that performance is suspect when there are not many trees. These differences can result in conflicting determinations of forest-type group. It is shown in Table 4 that when a computed type group is either oak/hickory or oak/pine, a wide range of different types are recorded by the field crew. It is also shown that a computed aspen/birch type is seen as spruce/fir for almost 40 percent of the plots and is judged to be maple/beech/birch for 14 percent of the plots. This suggests that 1) the tree species and size composition over the broader area differs somewhat from that within the sample plot area only; and/or 2) the relative importance afforded to the various tree sizes and species differ between the field crew and the algorithm.

This leads to another point regarding species composition. One would expect that increases in species diversity occur in transition zones near the edges of stands of differing type groups and more generally near the indistinct boundaries of natural ranges of type groups. In these zones, the increased diversity may lead to higher levels of classification error, as well as additional disparity between the field determination and algorithm output. Such analyses are beyond the scope of this paper, but the concept is worth highlighting as a future research topic.

A dilemma for analysts is whether to use an algorithm or the field-observed forest-type group. This choice could result in large shifts in estimated area for certain forest-type groups. There is a need to better align the field forest-type group with that computed by the algorithm. Given that crews collect data with electronic data recorders, improved consistency may be obtained by having the algorithm provide real-time feedback on the computed forest-type group. This would allow the field crews to see when there is disagreement. This may 1) allow the field crews to calibrate their observations to be more consistent with algorithm output; and 2) provide feedback that sheds light on needed modifications to improve algorithm accuracy.

In summary, the algorithm was generally robust to changes in plot size for loblolly/shortleaf, oak/hickory, and oak/gum/cypress groups. For classification of other forest-type groups, the recommended plot size should reflect the relative proportions of occurring type groups and be consistent with levels of misclassification that are considered tolerable. For example, if the area is composed primarily of aspen/birch then a larger plot size should be considered than if the area is mostly oak/gum/cypress. Ultimately, it would be desirable to refine the algorithm such that all forest-type groups had similar (small) misclassification probabilities. This paper provides an analytical framework for evaluating whether changes to the algorithm provide improved classification consistency.

Literature Cited

Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 19(6): 716–723.

Arner, S.L.; Woudenberg, S.; Waters, S.; Vissage, J.; MacLean, C.; Thompson, M.; Hansen, M. 2003. National algorithms for stocking class, stand size class, and forest type. Available: http://www.fia.fs.fed.us/library/field-guides-methods-proc/docs/National%20algorithms.doc

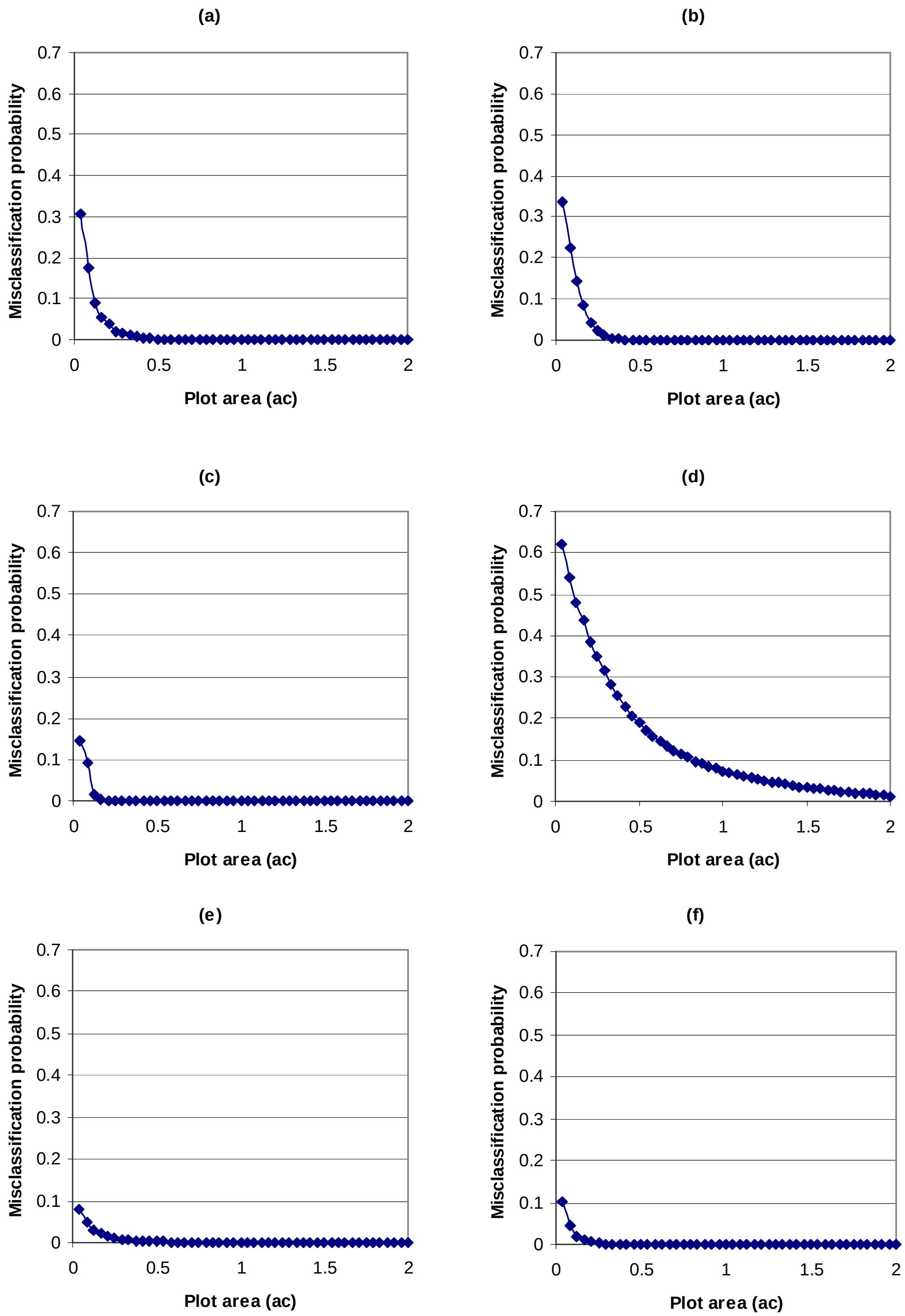
Bechtold, W.A.; Patterson, P.L., eds. 2005. The enhanced forest inventory and analysis program – national sampling design and estimation procedures. Gen. Tech. Rep. SRS-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 85 p.

Eyre, F.H., ed. 1980. Forest cover types of the United States and Canada. Washington, DC: Society of American Foresters. 148 p.

Hansen, M.H.; Hahn, J.T. 1992. Determining stocking, forest type, and stand-size class from forest inventory data. Northern Journal of Applied Forestry. 9(3): 82-89.

Metropolis, N.; Ulam, S. 1949. The Monte Carlo method. Journal of American Statistical Association. 44: 335-341.

U.S. Department of Agriculture. 2007. Forest inventory and analysis national core field guide: volume I: field data collection procedures for phase 2 plots, version 4.0. Washington, DC: U.S. Department of Agriculture, Forest Service, Forest Inventory and Analysis. Available: http://www.fia.fs.fed.us/library/field-guides-methods-proc/docs/core_ver_4-0_10_2007_p2.pdf


(a)
Misclassification probability
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.5
1
1.5
2
Plot area (ac)
(b)
Misclassification probability
Plot area (ac)
(c)
Misclassification probability
Plot area (ac)
(d)
Misclassification probability
Plot area (ac)
(e)
Misclassification probability
Plot area (ac)
(f)
Misclassification probability
Plot area (ac)

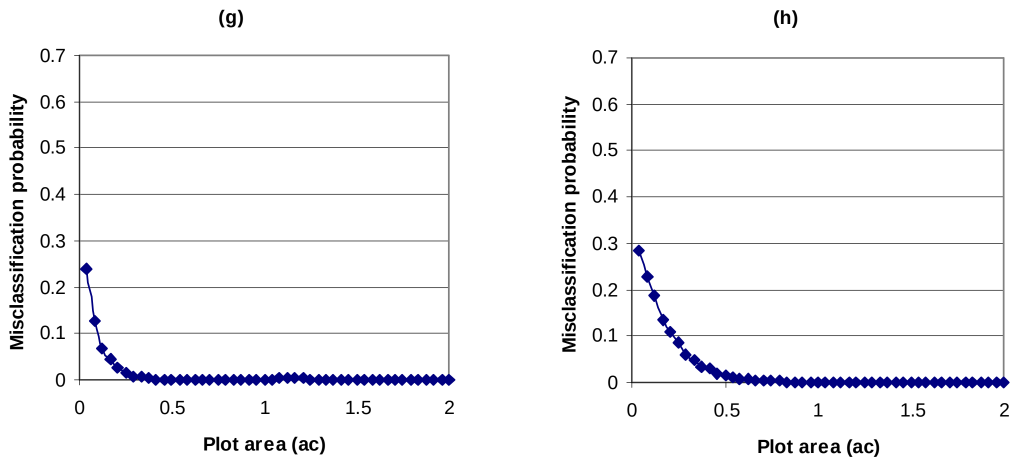


Figure 1: Misclassification probability vs. plot area for a) red/white/jack pine; b) spruce/fir; c) loblolly/shortleaf; d) oak/pine; e) oak/hickory; f) oak/gum/cypress; g) maple/beech/birch; and h) aspen/birch type groups.

Table 1: Data summary statistics by forest type and forest-type group.

Forest type group	Forest type	# plots	No. stems/plot[a] Min.	No. stems/plot[a] Mean	No. stems/plot[a] Max.	No. species/plot Min.	No. species/plot Mean	No. species/plot Max.
White/red/jack pine	Jack pine	1	55	55	55	5	5	5
White/red/jack pine	Red pine	7	41	86	113	6	10	14
White/red/jack pine	Eastern white pine	58	30	74	154	3	8	15
White/red/jack pine	White pine/hemlock	28	36	74	119	6	8	14
White/red/jack pine	Eastern hemlock	30	43	93	135	4	9	13
		124	**30**	**79**	**154**	**3**	**9**	**15**
Spruce/fir	Balsam fir	269	7	95	194	2	8	15
Spruce/fir	White spruce	14	12	53	103	1	5	9
Spruce/fir	Red spruce	151	32	104	187	1	8	13
Spruce/fir	Red spruce/balsam fir	137	25	97	177	3	8	13
Spruce/fir	Black spruce	73	4	73	141	1	5	13
Spruce/fir	Tamarack	5	31	72	122	5	7	10
Spruce/fir	Northern white-cedar	119	38	116	177	3	9	16
		768	**4**	**97**	**194**	**1**	**8**	**16**
Loblolly/shortleaf pine	Loblolly pine	512	4	65	308	1	7	17
Loblolly/shortleaf pine	Shortleaf pine	7	3	89	167	1	13	19
Loblolly/shortleaf pine	Virginia pine	9	32	76	114	10	14	19
Loblolly/shortleaf pine	Pond pine	9	33	59	106	2	7	10
		537	**3**	**66**	**308**	**1**	**7**	**19**
Oak/pine	White pine/red oak/white ash	37	12	66	103	3	9	14
Oak/pine	Eastern redcedar/hardwood	14	7	68	124	6	13	20
Oak/pine	Longleaf pine/oak	13	20	41	70	4	7	12
Oak/pine	Shortleaf pine/oak	9	31	72	97	6	11	18
Oak/pine	Virginia pine/southern red oak	6	28	60	92	3	13	22
Oak/pine	Loblolly pine/hardwood	80	12	62	147	3	9	19
Oak/pine	Slash pine/hardwood	4	31	44	60	9	10	12
Oak/pine	Other pine/hardwood	6	42	65	92	3	6	8
		169	**7**	**62**	**147**	**3**	**10**	**22**
Oak/hickory	Post oak/blackjack oak	12	20	68	116	8	12	18
Oak/hickory	Chestnut oak	12	24	55	115	4	8	15
Oak/hickory	White oak/red oak/hickory	190	15	63	283	5	12	24
Oak/hickory	White oak	30	20	66	138	4	12	18
Oak/hickory	Northern red oak	19	34	65	84	5	8	14
Oak/hickory	Yellow-poplar/white oak/red oak	35	21	67	158	6	14	26
Oak/hickory	Sassafras/persimmon	19	1	54	101	1	9	18
Oak/hickory	Sweetgum/yellow-poplar	50	27	59	122	3	10	18
Oak/hickory	Bur oak	1	24	24	24	3	3	3
Oak/hickory	Scarlet oak	3	57	63	72	8	11	14
Oak/hickory	Yellow-poplar	9	33	71	152	9	12	16
Oak/hickory	Black walnut	2	27	33	38	7	8	9
Oak/hickory	Black locust	1	68	68	68	10	10	10
Oak/hickory	Southern scrub oak	10	19	35	73	1	6	11
Oak/hickory	Chestnut oak/black oak/scarlet oak	15	18	47	87	3	11	21
Oak/hickory	Red maple/oak	11	2	59	143	1	8	12
Oak/hickory	Mixed upland hardwoods	103	2	62	380	2	10	21
		522	**1**	**62**	**380**	**1**	**11**	**26**
Oak/gum/cypress	Swamp chestnut oak/cherrybark oak	7	33	46	64	8	11	15
Oak/gum/cypress	Sweetgum/Nuttall oak/willow oak	84	3	47	97	1	9	17
Oak/gum/cypress	Overcup oak/water hickory	4	13	27	46	6	10	14
Oak/gum/cypress	Baldcypress/water tupelo	33	25	54	92	1	7	15
Oak/gum/cypress	Sweetbay/swamp tupelo/red maple	77	19	56	145	3	8	21
		205	**3**	**51**	**145**	**1**	**8**	**21**
Maple/beech/birch	Sugar maple/beech/yellow birch	978	20	92	172	4	9	19
Maple/beech/birch	Black cherry	2	17	32	47	4	6	7
Maple/beech/birch	Cherry/ash/yellow-poplar	33	22	72	156	4	10	20
Maple/beech/birch	Hard maple/basswood	9	33	60	119	5	10	14
Maple/beech/birch	Elm/ash/locust	1	22	22	22	6	6	6
Maple/beech/birch	Red maple/upland	66	10	87	147	2	9	14
		1089	**10**	**91**	**172**	**2**	**9**	**20**
Aspen/birch	Aspen	114	3	90	151	1	9	17
Aspen/birch	Paper birch	173	4	89	185	1	8	17
Aspen/birch	Balsam poplar	11	57	97	162	6	10	15
		298	**3**	**90**	**185**	**1**	**9**	**17**

[a] Includes all tallied seedlings, saplings, and trees.

Table 2: Estimated coefficients for numbers of stems, numbers of species, and basal area models [1] through [3] by tree size class.

		Number stems [1]		Number spp [2]				Basal area [3]	
Forest type group	**Tree size class**	**β_1**	**Pr > |t|**	**β_2**	**Pr > |t|**	**β_3**	**Pr > |t|**	**β_4**	**Pr > |t|**
Red/white/jack pine	Seedling	2718.5473	<.0001	47.4815	<.0001	0.3425	<.0001	-	-
Spruce/fir	Seedling	3725.9365	<.0001	35.6997	<.0001	0.2836	<.0001	-	-
Loblolly/shortleaf	Seedling	1618.1861	<.0001	60.6337	<.0001	0.4410	<.0001	-	-
Oak/pine	Seedling	2172.1193	<.0001	101.0000	<.0001	0.4567	<.0001	-	-
Oak/hickory	Seedling	2504.3036	<.0001	92.2299	<.0001	0.4079	<.0001	-	-
Oak/gum/cypress	Seedling	996.9864	<.0001	62.8282	<.0001	0.5128	<.0001	-	-
Maple/beech/birch	Seedling	3993.0822	<.0001	50.4331	<.0001	0.3389	<.0001	-	-
Aspen/birch	Seedling	3694.8156	<.0001	41.9937	<.0001	0.2634	<.0001	-	-
Red/white/jack pine	Sapling	548.4302	<.0001	35.0012	<.0001	0.4799	<.0001	16.8095	<.0001
Spruce/fir	Sapling	1468.3205	<.0001	29.8807	<.0001	0.3948	<.0001	36.6782	<.0001
Loblolly/shortleaf	Sapling	623.1041	<.0001	45.3810	<.0001	0.5509	<.0001	21.1109	<.0001
Oak/pine	Sapling	648.6880	<.0001	83.9266	<.0001	0.6005	<.0001	19.4474	<.0001
Oak/hickory	Sapling	517.2282	<.0001	76.0099	<.0001	0.5895	<.0001	15.9016	<.0001
Oak/gum/cypress	Sapling	647.3272	<.0001	51.7741	<.0001	0.5139	<.0001	19.9905	<.0001
Maple/beech/birch	Sapling	857.7573	<.0001	36.6037	<.0001	0.4043	<.0001	23.9603	<.0001
Aspen/birch	Sapling	1277.4473	<.0001	35.1636	<.0001	0.3859	<.0001	32.3657	<.0001
Red/white/jack pine	Tree	215.0946	<.0001	18.1165	<.0001	0.3987	<.0001	112.2513	<.0001
Spruce/fir	Tree	175.9560	<.0001	12.8520	<.0001	0.2973	<.0001	69.5593	<.0001
Loblolly/shortleaf	Tree	220.1037	<.0001	13.0503	<.0001	0.5233	<.0001	84.2581	<.0001
Oak/pine	Tree	147.5297	<.0001	29.4719	<.0001	0.4790	<.0001	67.6585	<.0001
Oak/hickory	Tree	132.9210	<.0001	31.7232	<.0001	0.4219	<.0001	79.1111	<.0001
Oak/gum/cypress	Tree	177.5039	<.0001	25.3210	<.0001	0.4436	<.0001	113.7109	<.0001
Maple/beech/birch	Tree	161.3880	<.0001	18.1601	<.0001	0.3663	<.0001	74.5506	<.0001
Aspen/birch	Tree	152.9498	<.0001	15.5264	<.0001	0.3395	<.0001	54.5049	<.0001

Table 3: Fit statistics by forest-type group for model [4].

Forest type group	$R^{2\,a}$	AIC Intercept only	AIC Intercept + covariates	AIC % reduction (covariates)
White/red/jack pine	0.57	4419.4	1975.7	55.3%
Spruce/fir	0.50	2952.0	1525.5	48.3%
Loblolly/shortleaf pine	0.64	1389.9	519.6	62.6%
Oak/pine	0.46	25862.0	15441.7	40.3%
Oak/hickory	0.43	2657.6	1536.3	42.2%
Oak/gum/cypress	0.51	2025.0	1010.3	50.1%
Maple/beech/birch	0.58	4854.6	2090.3	56.9%
Aspen/birch	0.49	7717.3	4054.7	47.5%

[a] Max. rescaled R^2

Table 4: Frequency of agreement between field forest-type group and computed forest-type group for 3,712 FIA plots.

Field Forest Type Group[a]

Algorithm Forest Type Group

Frequency Row Pct Col Pct	AB	EAR	ES	LLS	LS	MBB	OGC	OH	OP	PJ	WRJ	SF	Total
AB	133 44.63 74.72	2 0.67 3.70	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	42 14.09 5.17	0 0.00 0.00	1 0.34 0.24	2 0.67 0.90	0 0.00 0.00	1 0.34 0.55	117 39.26 10.75	298
EAR	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0
ES	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0
LLS	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0
LS	0 0.00 0.00	1 0.19 1.85	0 0.00 0.00	6 1.12 40.00	492 91.62 90.61	0 0.00 0.00	0 0.00 0.00	4 0.74 0.95	34 6.33 15.25	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	537
MBB	34 3.12 19.10	15 1.38 27.78	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	693 63.64 85.34	0 0.00 0.00	43 3.95 10.17	25 2.30 11.21	0 0.00 0.00	54 4.96 29.67	225 20.66 20.68	1089
OGC	0 0.00 0.00	19 9.27 35.19	0 0.00 0.00	0 0.00 0.00	1 0.49 0.18	0 0.00 0.00	180 87.80 93.75	4 1.95 0.95	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	1 0.49 0.09	205
OH	2 0.38 1.12	12 2.30 22.22	1 0.19 50.00	3 0.57 20.00	16 3.07 2.95	55 10.54 6.77	4 0.77 2.08	355 68.01 83.92	70 13.41 31.39	1 0.19 25.00	2 0.38 1.10	1 0.19 0.09	522
OP	1 0.59 0.56	2 1.18 3.70	0 0.00 0.00	5 2.96 33.33	34 20.12 6.26	1 0.59 0.12	8 4.73 4.17	16 9.47 3.78	82 48.52 36.77	1 0.59 25.00	17 10.06 9.34	2 1.18 0.18	169
PJ	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0
WRJ	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	1 0.81 0.12	0 0.00 0.00	0 0.00 0.00	10 8.06 4.48	0 0.00 0.00	101 81.45 55.49	12 9.68 1.10	124
SF	8 1.04 4.49	2 0.26 3.70	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	20 2.60 2.46	0 0.00 0.00	0 0.00 0.00	0 0.00 0.00	1 0.13 25.00	7 0.91 3.85	730 95.05 67.10	768
Total	178	53	1	14	543	812	192	423	223	3	182	1088	3712

[a] AB = Aspen/Birch, EAR = Elm/Ash/Red Maple, ES = Exotic Softwood, LLS = Longleaf/Slash Pine, LS = Loblolly/Shortleaf Pine, MBB = Maple/Beech/Birch, OGC = Oak/Gum/Cypress, OH = Oak/Hickory, OP = Oak/Pine, PJ = Pinyon/Juniper, WRJ = White/Red/Jack Pine, SF = Spruce/fir.