A test of an expert-based bird-habitat relationship model in South Carolina

John C. Kilgo, David L. Gartner, Brian R. Chapman, John B. Dunning, Jr., Kathleen E. Franzreb, Sidney A. Gauthreaux, Cathryn H. Greenberg, Douglas J. Levey, Karl V. Miller, and Scott F. Pearson

Abstract Wildlife-habitat relationships models are used widely by land managers to provide information on which species are likely to occur in an area of interest and may be impacted by a proposed management activity. Few such models have been tested. We used recent avian census data from the Savannah River Site, South Carolina to validate BIRDHAB, a geographic information system (GIS) model developed by United States Forest Service resource managers to predict relative habitat quality for birds at the stand level on national forests in the southeastern United States. BIRDHAB is based on the species-habitat matrices presented by Hamel (1992). Species-specific accuracy rates for BIRDHAB predictions (the percentage of all stands in which a species was predicted correctly as present or absent) ranged from 33.6–93.0%, with a mean of 67.4±17.3% (SD, n=46 species). Accuracy was >90% for 5 species, but <50% for 9 species. BIRDHAB performed well (P <0.05) in predicting presence-absence of 32 species. Generally, the model was more accurate in predicting presence-absence for habitat specialists than for generalists. Habitat-specific accuracy rates (the percentage of species for which a habitat's prediction was correct) ranged from 52.7–92.7%, with a mean of 71.8±(SD) 9.8% (*n-26* habitat types). BIRDHAB was a useful tool for many of the species that we tested, but it had no predictive ability for many others. Such species-specific variation in accuracy probably is common among wildlife-habitat relationships models, reinforcing the need for thorough testing before these models are used in land-use planning.

Key words BIRDHAB, bird habitat model, GIS, habitat, model accuracy, model validation, Savannah River Site, South Carolina, species-habitat matrix, wildlife-habitat relationships

Land managers frequently require site-specific information on species occurrence to predict potential impacts of proposed management. Because it is cost-prohibitive to conduct comprehensive surveys at every proposed impact site, predictive models of species occurrence are needed. Models with predictions that are based on various habitat components are known as wildlife-habitat

Address for John C. Kilgo: United States Forest Service, Southern Research St&ion, P. 0. Box 700, New Ellenton, SC 29809, USA; e-mail: jkilgo@fs.fed.us. Address for David L. Gartner, United States Forest Service, Southern Research Station, Center for Forested Wetlands Research, 2730 Savannah Highway, Charleston, SC 29434, USA; present address: United States Forest Service, Southern Research Station, 4700 Old Kingston Pike, Knoxville, TN 37919, USA. Address for Brian R. Chapman and Karl V. Miller: D. B. Warnell School of Forest Resources, University of Georgia, Athens, GA 30602, USA; present address for Chapman: College of Arts and Sciences, Sam Houston State University, Huntsville, TX 77341, USA. Address for John B. Dunning, Jr.: Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN 47907, USA. Address for Kathleen E. Franzreb: United States Forest Service, Southern Research Station, Clemson University, Clemson, SC 29634, USA; present address: United States Forest Service, Southern Research Station, University of Tennessee, Knoxville, TN 37919, USA. Address for Sidney A. Gauthreaux: Department of Biological Sciences, Clemson University, Clemson, SC 29634, USA. Address for Cathryn H. Greenberg, United States Forest Service, Southern Research Station, 1577 Brcvard Rd., Asheville, NC 28806, USA. Address for Douglas J. Levey and Scott F. Pearson: Department of Zoology, University of Florida, Gainesville, FL 32611, USA; present address for Pearson: Washington Department of Natural Resources, 1111 Washington St. SE, Mail Stop 47010, Olympia, WA 98504.

relationships (WHR) models (see Morrison et al. 1992 for review).

Species-habitat matrices are a type of WHR model that provides estimates of the relative quality of several habitats for each of many species (Morrison et al. 1992). The quality of each habitat usually is rated on a qualitative scale (e.g., unsuitable, marginal, suitable, or optimal). Although such scales generally are defined to reflect increasing frequency of occurrence or abundance (Hamel 1992), species-habitat matrices do not explicitly predict occurrence or abundance. They often are used in impact assessments to evaluate habitat quality for various species under alternative management scenarios. For example, a hypothetical management alternative might be predicted to change the relative quality of an impact site for a species of interest from optimal to unsuitable, or vice versa. Such information may affect the type of management action, its scale, or whether it is even implemented.

When WHR models are used in land management planning, they may have a dramatic effect both on species' habitat quality and availability, and on the management options available to the landowner. Thus, it is ecologically and economically important that models be as accurate as possible and that the level of accuracy be known. The importance of WHR model validation has long been acknowledged (Marcot et al. 1983). However, most models remain untested (Morrison et al. 1992, but see Dedon et al. 1986, Raphael and Marcot 1986, Timothy and Stauffer 1991, Block et al. 1994, Edwards et al. 1996).

We used field data collected during recent and ongoing research to test the predictions of BIRDHAB (United States Forest Service [USFS] 1994), a WHR model developed for national forests in the southeastern United States (USFS Region 8). BIRDHAB was created to assist in impact assessment and in the development of biological evaluations of proposed management actions. It is an ArcInfo geographic information system (GIS) analysis program that uses standard USFS stand inventory data and a species-habitat matrix (Namel 1992) to predict relative habitat quality (unsuitable, marginal, suitable, or optimal) at the stand level for 271 bird species that breed or winter in the southeastern United States. Hamel's (1992) matrix represented the expert opinion of a team of ornithologists, and included species-specific habitat quality predictions for 4 successional stages of 23 vegetation types.

Specifically, we determined model accuracy for 46 species and 26 habitats and tested BIRDHAB's ability to correctly predict species' presence-absence. Because the habitat needs of rare or sensitive species are often overriding considerations in management decisions, we focused on the model's performance for these species. More generally, our study provides an evaluation of how species-habitat matrix WHR models are constructed: can qualitative models, based on expert opinion, provide accurate estimates of species occurrence or should other, more quantitative approaches be encouraged in the construction of WHR models?

Study area

To test BIRDHAB, we used data collected on the United States Department of Energy's (USDOE) Savannah River Site (SRS), a 78,000-ha National Environmental Research Park located in Aiken, Barnwell, and Allendale counties, South Carolina, in the Upper Coastal Plain physiographic province. Approximately 95% of the land area was undeveloped and managed by the USFS-Savannah River according to USDOE policies. Of the forested acreage, 71% was managed pine (loblolly [*Pinus taeda*], longleaf [*P. palustris*], and slash [*P. elliottii*]), 3% was upland hardwood, 22% was bottomland hardwood, and 3% was mixed pine-hardwood (Workman and McLeod 1990). Pine forests were managed on 80- to 120-yr rotations. Clearcutting was conducted occasionally to convert off-site slash pine to loblolly or longleaf pine. Hardwood forests typically were not managed, although harvest occurred occasionally.

Methods

The BIRDHAB model

BIRDHAB provides users with 2 options: querying a list of predicted habitat qualities for all bird species in a single stand or querying habitat quality for a single bird species within an area of several stands (e.g., a timber compartment). Output for the latter query is displayed as a habitat map.

BIRDHAB's predictions are based on Hamel's (1992) species-habitat matrix. The only difference between BIRDHAB and Hamel's (1992) matrix relates to the assumptions made in the derivation of the 2 input variables (vegetation type and successional stage); once vegetation type and successional stage are obtained, predicted habitat quality for all species that occur on the user's site (as determined from Hamel's [1992] range maps, digitized for BIRDHAB) is taken directly from Hamel's (1992) matrix. BIRDHAB uses information from the USFS corporate Continuous Inventory of Stand Condition database (CISC) to determine the vegetation type and successional stage of the queried stand(s). The CISC database contains USFS forest type codes (different from Hamel's [1992] vegetation types), stand-condition class codes (general descriptors of timber stocking rates and merchantability), and stand age. Hamel's (1992) vegetation type was derived from the USFS forest type with a relational cross-walk table taken directly from an appendix provided by Hamel (1992). However, successional stage is derived using rules that combine condition class and stand age (Appendix 1), and these rules require several assumptions that are not outlined by Hamel (1992). These assumptions may lead to erroneous determination of successional stage for a stand, and therefore to inappropriate predictions, the responsibility for which should not be borne by Hamel's (1992) matrix. Therefore, our results are not intended to be viewed as direct validation of Hamel's (1992) matrix. Nevertheless, we believe that our testing represented a reasonable approximation of a validation of Hamel's (1992) matrix for the SRS.

Test data

Because many WHR models do not predict occurrence (i.e., presence-absence) or relative abundance but rather relative habitat quality, it is difficult to validate them using quantitative count data. Further complicating the issue is the fact that habitat quality and species density are not necessarily related (Van Horne 1983, Vickery et al. 1993). For example, the quality of a forest stand with apparently excellent habitat structure for a given species may be compromised by landscape-level factors so that it actually functions as a sink habitat (Pulliam 1988). Such difficulties notwithstanding, we elected to use observed occupancy rate as an index of habitat quality. In doing so we assumed that within any given classification scheme (e.g., habitat types, predicted quality levels, etc.), the percent of stands that were occupied represented a better reflection of the quality of an individual class than did mean density of birds in stands of that class. The intended objective of BIRDHAB is not to explicitly predict species occurrence, and our tests therefore did not constitute a true validation (i.e., we were not testing how well BIRDHAB achieved its intended objective).

We used data from 6 avian research projects conducted at the Savannah River Site from 1991-1998 (Table 1) to validate the breeding season component of the BIRDHAB model for selected species (see below). All data sets except J. B. Dunning's unpublished data were comparable in that standard point-count techniques (Ralph et al. 1995) were used. Counts were conducted from sunrise to 3.5-hr post-sunrise between May 5 and June 30, except during rain or high winds. Birds flying over the plot were recorded separately. K. E. Franzreb (unpublished data) used 50-m radius 10-min point counts to sample all stands within a 1-km buffer zone surrounding active red-cockaded woodpecker (*Picoides borealis*) clusters in 1995. Likewise, S. A. Gauthreaux (unpublished data) used 50-m radius 10-min point counts to sample at permanent vegetation monitoring plots of the USFS Forest Inventory and Analysis program from 1993-1995. Kilgo et al. (1997, 1998) and Buffington et al. (1997) used 50-m radius 5-min point counts, also recording species detected in the stand ±3 min of the count period, to sample stands of upland and bottomland

Table 1. Habitats and sample sizes of data sets collected from 1991-1998 used to validate BIRDHAB for the Savannah River Site, South Carolina. With the exception of the Kilgo et al. (1997, 1998) and Buffington et al. (1997) data sets, in which a total of 62 individual stands were sampled in multiple years, all stands in the database were sampled in only one year.

Data set	No. of stands	Habitats sampled
Buffington et al. (1997)	6[a]	Bottomland hardwood (early, mid, late succession)
J. B. Dunning, Jr. (unpublished data)	131	Longleaf, loblolly pine (early, late succession)
K. E. Franzreb (unpublished data)	99	All[b]
S. A. Gauthreaux (unpublished data)	158	All
Kilgo et al. (1997)	76"	Upland hardwood, pine (late succession)
Kilgo et al. (1998)	45"	Bottomland hardwood (late succession)
D. J. Levey et al. (unpublished data)	56	All
Total	571	

[a] Number represents stand-years.

[b] Included early, mid, and late succession upland and bottomland hardwood, loblolly and longleaf pine, and mixed-pine hardwood.

hardwood and mature pine forests from 1993-1995. D. J. Levey et al. (unpublished data) used 50-m radius 5-min point counts, followed by a 5-min period during which an audio recording of a screech owl (*Otus asio*) was played, to sample stands of several habitats in 1998 in which fleshy fruit availability and use were measured. All 5 point-count studies separately recorded birds detected beyond 50 m. The studies sampled at various intensities, both in terms of numbers of points per stand and number of visits per point. Because 2 studies sampled at only 1 point per stand and visited that point only once, we randomly selected only 1 visit to 1 point per stand from the studies with >1 visit or point. Collectively, these studies comprised a database of point counts from 440 stands that represented all major habitats occurring on the SRS in rough proportion to their acreage.

The point-count data were used to determine presence or absence of species only. We defined a species as present in a stand if it was recorded in any way except flying over the stand. Thus, in general we determined presence by unlimited distance point counts of 10 min, a duration determined sufficient for estimation of bird-habitat model performance (Dettmers et al. 1999).

Unpublished data from Dunning's surveys (hereafter Dunning counts), which were more time-intensive than the point counts, were used to assess the ability of point counts to detect presence of each species (see below). These surveys were described in Dunning and Watts (1990). Briefly, 2- to 7-yr-old pine regeneration stands and mature pine stands were surveyed for 6 min from each of 6 points, distributed evenly throughout the stand. Birds were recorded during a 3-min listening period, a 1.5-min period during which a tape-recorded Bachman's sparrow (*Aimophila aestivalis*) song was played (densities of this species were of particular interest in this study), and a final 1.5-min listening period. Sampling was conducted during 1991 and 1993-1994. Stands were visited twice per year, and data were recorded as total number of individuals of each species detected across all 6 points per visit. Thus each stand was surveyed for 72 min, which likely yielded more accurate presence-absence information than the more conventional point counts.

We assessed BIRDHAB's predictions for 51 species. Species detected in <5 stands and species for which point counts did not adequately assess presence were not considered. This included wetland birds, nocturnal birds, birds most commonly recorded as fly-avers, and birds that rarely vocalize. Accordingly, the following taxa were excluded: waterfowl (Anatidae), wading birds (Ardeidae), raptors (Falconiformes, Strigiformes), goatsuckers (Caprimulgidae), aerial insectivores (Apodidae, Hirundinidae), hummingbirds (Trochilidae), kingfishers (Alcedinidae), and corvids (Corvidae).

Testing

Using BIRDHAB, we obtained species-specific predictions of habitat quality for each of the 571 (440 point counts+ 13 1 Dunning counts) stands for which we had validation data. We defined a prediction of "present" to include predictions of suitable or optimal habitat quality and a prediction of "absent" to include predictions of unsuitable or marginal habitat quality. We calculated the accuracy of the model in predicting presence-absence (assuming the relationship defined above between habitat quality and presence-absence) of each species as the percentage of the total number of stands in which presence or absence was predicted correctly ([(the number of stands in which the species was predicted to be present and was found + number in which it was predicted to be absent and was not found)/total number of stands] x 100). Error of commission rate was calculated for each species as the number of stands in which the species was predicted to occur but did not, divided by the total number of stands. Error of omission rate was calculated as the number of stands in which the species was not predicted to occur but did, divided by the total number of stands. We tested the accuracy of the presence-absence predictions using Fischer's exact test for 2 x 2 contingency tables (PROC FREQ, SAS Institute Inc. 1988). We also calculated the Kappa statistic (Kappa= [error rate(random) -error rate(model)] / error rate(random); Landis and Koch 1977, Fielding and Bell 1997), which represented the reduction in error rate as compared to the error rate associated with a random allocation of birds to stands (1 - [(proportion of stands predicted present * proportion of stands occupied) + (proportion of stands predicted absent * proportion of stand unoccupied)]). The observed accuracy rates represented an average of the accuracy rates of the 4 years of data.

Because species often are missed on l-visit point counts of short duration (Dawson et al. 1995), the validation data we used may have falsely inflated errors of commission and underestimated errors of omission. We attempted to assess the ability of the point counts to detect species by comparing detection rates (proportion of stands in which a species was detected) of the 51 test species between the point-count data and Dunning's more time-intensive unpublished survey data using Fisher's exact test for 2 x 2 contingency tables. We had sufficient numbers of stands for both survey methods in longleaf pine-shrub-seedling (16 point counts, 45 Dunning counts) and loblolly pine-shrub-seedling (17 point counts, 59 Dunning counts), each of which was tested separately. We used the one-tailed alternative because the Dunning counts were expected ***a priori*** to yield higher detection rates.

Finally, we calculated accuracy rates by habitat types (defined as forest type-successional stage combinations). For each habitat type, the accuracy rate was the percentage of the 51 test species whose presence or absence was predicted correctly. Accuracy rate was calculated for each stand and averaged by habitat type. Error of commission rate and error of omission rate also were calculated.

Results

BIRDHAB predicted 95 species to have at least marginal habitat in at least 1 of our 440 test stands. Of the 51 species selected for analysis, only ovenbird (*Seiurus* ***aurocapillus)*** was not predicted to occur. Only in recent years has this species expanded its breeding range to include the SRS (Meyers and Odum 2000), and its range map in BIRDHAB should be modified accordingly. For 4 additional species (eastern kingbird [*Tyrannus tyrannus*], northern mockingbird [*Mimus polyglottos*], orchard oriole [*Icterus spurius*] , and American goldfinch [*Carduelis tristis*]), the greatest level of habitat quality predicted was marginal. Because we treated a prediction of "marginal" habitat as "absent" in our presence-absence analyses, the species were never predicted to occur for purposes of these analyses. As these species were relatively common at SRS (J.B. Dunning, unpublished data; Mayer et al. 1997), habitat conditions for them are presumably better than "marginal," so their BIRDHAB predictions should be viewed as inaccurate. However, accuracy could not be calculated for these 5 species, so BIRDHAB predictions were tested for only 46 species.

Accuracy rates for BIRDHAB predictions by species ranged from a low of 33.6% for Carolina chickadee (*Poecile* ***carolinensis)*** to a high 93.0% for field sparrow (*Spizella pusilla*), with a mean for all species of 67.4±17.3% (SD) (Table 2). Accuracy was >90% for 5 species, and <50% for 9 species. Generally, the model was more accurate in predicting presence-absence of habitat specialists (species with occurrence restricted to few habitats; e.g., field sparrow, Kentucky warbler [*Oporornis formosus*], prothonotary warbler ***[Protonotaria citrea]*** , Swainson's warbler [*Protonotaria* ***citrea])*** than of generalists (species that use many habitats; e.g., brown thrasher | *Toxostoma rufum*], mourning dove [*Zenaida macroura*], Carolina chickadee, red-bellied woodpecker ***[Melanerpes*** *carolinus*]). Mean error of omission rate was 6.8±6.8% (SD) and mean error of commission rate was 26.1± 17.5%. Error of commission rates exceeded error of omission rates for all but 10 species, but several species with relatively high accuracy had high error of omission rates (e.g., white-eyed vireo [Vireo griseus] , red-eyed vireo [*Vireo olivaceous*] , and indigo bunting ***[Passerinn*** *cyanea*]). Fischer's exact test indicated that BIRDHAB performed well ($P<0.05$) in predicting presence-absence for 32 of the 46 species. Observed frequency of occurrence in stands with the four levels of BIRDHAB prediction are presented for these 32 species (Table 3).

Our comparison of detection rates between point counts and Dunning's counts revealed that 14 of the 40 (35.0%) test species that occurred in longleaf pine-shrub-seedling stands and 24 of the 39 (61.5%) test species that occurred in loblolly pine-shrub-seedling stands had greater ($P< 0.05$) detection rates in the Dunning counts. Mean detection rates (i.e., the percentage of stands in which each species was detected) were 15.7% greater for Dunning counts (41.5%) than point counts (25.8%) in longleaf pine-shrub-seedling stands and 30.9% greater for Dunning counts (45.4%) than point counts (14.5%) in loblolly pine-shrub-seedling stands. **Thus,** as expected, point counts frequently underestimated species' habitat occupancy.

Accuracy rates for habitat types ranged from 52.7% for white oak-red oak-hickory-sawtimber to 92.7% for the same type in the sapling-poletimber stage (Table 4). By gross habitat types, predictions were worst for mixed pine-hardwood ($\bar{x}$= 66.4±7.4%), intermediate for hardwood ($\bar{x}$=72.2± 12.2%), and best for pine ($\bar{x}$=75.2±5.4%). Generally, hardwood and mixed pine-hardwood had high error of commission rates relative to error of omission rates (more species were predicted present

Table 2. Accuracy and error rates (%) for BIRDHAB's predictions of presence-absence for the Savannah River Site, South Carolina. Test data were collected from 1991-l 998. Species are listed in descending order of *Kappa* statistic.

Species	Accuracy	Error rate[a]		N[b]	*P*[c]	*Kappa*
		Omission	Commission			
Red-eyed vireo *(Vireo olivaceous)*	80.7	10.2	9.1	137	<0.001	0.556
Acadian flycatcher *(Empidonax virescens)*	81.1	7.0	11.8	111	<0.001	0.529
Northern parula *(Parula americana)*	81.6	5.7	12.7	66	<0.001	0.395
Blue-gray gnatcatcher *(Polioptila caerulea)*	73.0	2.3	14.8	121	<0.001	0.341
Kentucky warbler *(Oporornis formosus)*	78.0	3.0	19.1	50	<0.001	0.325
Indigo bunting *(Passerina cyanea)*	81.4	2.3	6.4	76	<0.001	0.247
Hooded warbler *(Wilsonia citrina)*	62.5	2.7	34.8	74	<0.001	0.238
Field sparrow *(Spire//a pusilla)*	93.0	0.2	6.8	5	<0.001	0.195
Wood thrush *(Hylocichla mustelina)*	58.0	4.1	38.0	69	<0.001	0.186
Yellow-breasted chat *(Icteria virens)*	91.4	5.2	3.4	28	0.006	0.165
Yellow-billed cuckoo *(Coccyzus americanus)*	65.7	14.1	20.2	105	0.004	0.163
Swainson's warbler *(Limnothlypis swainsonii)*	81.8	0.0	18.2	9	<0.001	0.150
Bachman's sparrow *(Aimophila aestivalis)*	68.2	3.2	28.6	40	<0.001	0.150
Brown-headed nuthatch *(Sitta pusilla)*	54.1	2.0	43.9	60	<0.001	0.150
Prothonotary warbler *(Protonotaria citrea)*	82.3	0.0	17.7	8	<0.001	0.144
Louisiana waterthrush *(Seiurus motacilla)*	77.7	1.1	21.1	17	<0.001	0.139
Carolina wren *(Thryothorus ludovicianus)*	56.8	16.8	26.4	215	0.005	0.139
Eastern towhee *(Pipilo erythrophthalmus)*	42.0	1.4	56.6	93	<0.001	0.132
Common yellowthroat *(Geothlypis trichas)*	91.1	5.2	3.6	27	0.028	0.119
White-eyed vireo *(Vireo griseus)*	85.9	13.6	0.5	6.5	0.001	0.113
Chipping sparrow *(Spizella passerina)*	78.0	6.4	15.7	42	0.014	0.113
Northern cardinal *(Cardinalis cardinalis)*	53.4	13.6	33.0	173	0.030	0.111
Yellow-throated vireo *(Vireo flavifrons)*	71.6	5.2	23.2	42	0.007	0.107
Downy woodpecker *(Picoides pubescens)*	52.3	3.6	44.1	49	0.036	0.107
Prairie warbler *(Dendroica discolor)*	92.0	4.5	3.4	23	0.061	0.101
Blue grosbeak *(Guiraca caerulea)*	91.8	4.8	3.4	24	0.068	0.099
American redstart *(Setophaga ruticilla)*	71.6	0.9	27.5	15	<0.001	0.096
Pine warbler *(Dendroica pinus)*	65.2	21.8	13.0	132	0.028	0.096
Eastern wood-pewee *(Contopus virens)*	43.4	2.7	53.9	a4	<0.001	0.094
Pileatcd woodpecker *(Dryocopus pileatus)*	68.4	6.8	24.8	5.3	0.019	0.092
Northern bobwhite *(Colinus virginianus)*	52.7	2.0	45.2	39	<0.001	0.085
Eastern bluebird *(Sialia sialis)*	69.5	3.2	27.3	29	0.012	0.081
Tufted titmouse *(Baeolophus bicolor)*	51.6	37.3	11.1	2.56	0.042	0.078
Yellow-throated warbler *(Dendroica dominica)*	62.3	0.0	37.7	6	0.003	0.043
Summer tanager *(Piranga rubra)*	49.5	2.5	38.0	135	0.220	0.036
Brown thrasher *(Toxostoma rufum)*	88.9	7.0	4.1	33	0.452	0.026
Mourning dove *(Zenaida macroura)*	36.6	5.5	58.0	84	0.793	0.025
Brown-headed cowbird *(Molothrus ater)*	42.7	5.2	52.0	44	0.884	0.005
Great crested flycatcher *(Myiarchus crinitus)*	48.2	1.1	40.7	203	0.513	0.004
Carolina chickadee *(Poecile carolinensis)*	33.6	4.8	61.6	1 05	0.640	-0.002
Red-bellied woodpecker *(Melanerpes carolinus)*	48.2	15.2	36.6	148	0.573	-0.004
Red-headed woodpecker *(M. erythrocephalus)*	59.5	4.8	35.7	33	0.665	-0.010
Hairy woodpecker *(Picoides villosus)*	44.5	1.3	54.1	10	0.901	-0.013
Gray catbird *(Dumatella carolinensis)*	81.1	2.0	16.8	10	0.849	-0.022
Black-and-white warbler *(Mniotilta varia)*	87.3	2.9	9.8	13	1.000	-0.050
Northern flicker *(Colaptes auratus)*	58.4	6.8	34.8	41	0.950	-0.05 1
Mean	67.4 ± 17.3	6.8 ± 6.8	26.1 ± 17.5			0.156

[a] Error of omission = not predicted but did occur; error of commission = predicted but did not occur.

[b] Number of stands in which species was detected out of 440.

[c] Fischer's exact test.

Table 3. Observed frequency of occurrence among BIRDHAB's predictions of relative habitat quality for species on the Savannah River Site, South Carolina, 1991-1998, for which tests of presence-absence predictions were significant (Table 2). Blank cells indicate that no stands were predicted in that category. Species are ordered as in Table 2.

	Relative Habitat Quality [a]			
Species	Unsuitable	Marginal	Suitable	Optimal
Red-eyed vireo	0.110	0.194	0.694	0.593
Acadian flycatcher	0.032	0.208	0.140	0.831
Northern parula	0.057	0.170	0.263	0.462
Blue-gray gnatcatcher	0.171	0.179	0.352	0.615
Kentucky warbler	0.030	0.182	0.000	0.474
Indigo bunting	0.183	0.120	0.440	
Hooded warbler	0.030	0.250	0.064	0.463
Field sparrow	0.002		0.188	0.056
Wood thrush	0.086	0.105	0.138	0.306
Yellow-breasted chat	0.048	0.111	0.250	
Yellow-billed cuckoo	0.194	0.333	0.200	0.393
Swainson's warbler	0.000		0.096	0.125
Bachman's sparrow	0.029	0.156	0.000	0.1 72
Brown-headed nuthatch	0.043	0.083	0.1 18	0.233
Prothonotary warbler	0.000	0.000	0.114	0.000
Louisiana waterthrush	0.006	0.111	0.114	
Carolina wren	0.500	0.407	0.522	0.551
Eastern towhee	0.250	0.045	0.230	0.385
Common yellowthroat	0.078	0.023	0.111	1.000
White-eyed vireo	0.139	0.133	0.714	
Chipping sparrow	0.009	0.110	0.169	
Northern cardinal	0.333	0.335	0.429	0.455
Yellow-throated vireo	0.068	0.182	0.085	0.226
Downy woodpecker	0.053	0.086	0.122	0.152
American redstart	0.014	0.000	0.043	0.129
Pine warbler	0.278	0.263		0.387
Eastern wood-pewee	0.286	0.055	0.016	0.287
Pileated woodpecker	0.019	0.113	0.182	0.174
Northern bohwhite	0.027	0.078	0.137	0.056
Eastern bluebird	0.009	0.067	0.111	
Tufted titmouse	0.475	0.563	0.735	0.529
Yellow-throated warbler	0.000	0.000	0.021	0.05 1

[a] Predicted values are not given for categories represented by <10 stands.

[b] No asterisk denotes a non-significant ($P > 0.05$) BIRDHAB test (BIRDHAB has little or no predictive ability).

* denotes significance for both the goodness-of-fit ($P < 0.15$) and the BIRDHAB test ($P < 0.05$) (BIRDHAB has moderate predictive ability).

** denotes significance only for BIRDHAB test (BIRDHAB has good predictive ability). See text for detailed explanation.

than were found) (Table 4). By successional stages, predictions were best for sapling-poletimber habitats (stage 3, $\bar{x}$=79.4±7.6%) and worst for sawtimber habitats (stage 4, $\bar{x}$= 66.2 ±7.7%). Generally, grass-forb habitats (stage 1) had relatively high error of omission rates (more species were found than were predicted present) and sawtimber habitats had high error of commission rates (Table 4).

Discussion

BIRDHAB significantly predicted presence-absence for 32 of the 46 species we tested. For 14 species, however, BIRDHAB had little to no predictive ability. BIRDHAB tended to perform better for habitat specialists than for habitat generalists. The occurrence of a species restricted to one or a few habitats was easier to predict than a species that occurred in several habitats, presumably because each additional habitat predicted added an additional level of uncertainty. A model is considered useful "if at least some predictions are empirically correct" (Morrison et al. 1992:256). Therefore, BIRDHAB was a useful model for most habitat specialists on our study site.

Generally, there was good agreement between BIRDHAB accuracy rates and results of the 2 x 2 contingency analyses. However, this was not always the case; BIRDHAB had significant Fischer's exact tests for a few species that had low accuracy rates (e.g., eastern towhee [*Pipilo erythrophthalmus*], eastern wood-pewee [*Contopus virens*]), and it had non-significant tests for a few species with high accuracy rates (e.g., brown thrasher, gray catbird [*Dumatella carolinensis*] , black-and-white warbler [*Mniotilta varia*]). The Kappa statistic helped explain these discrepancies. Species with high accuracy rates but poor predictive abilities generally bad low Kappa statistics, indicating that the high accuracy rates were likely due to random chance. High accuracy rates can occur when a species' occupancy rates are very low and it is predicted absent in most stands. Species with low accuracy rates but good predictive abilities generally had moderate (about 0.10) Kappa statistics. Kappa statistics in this range occurred when the stands with predictions of present had occupancy rates that were relatively low (<0.50) but still higher than the occupancy rates for stands with predictions of absent. We suggest that results of the Fischer's exact tests should be interpreted as the indication of BIRDHAB's ability to predict species presence-absence, and that accuracy rates should be used as supplemental information. Accordingly, BIRDHAB may be considered most appropriate for species with significant tests and high accuracy.

Table 4. Accuracy and error rates (%) of BIRDHAB predictions by habitat type on the Savannah River Site, South Carolina. Test data were collected from 1991–1998. Accuracy represents the percent of the 51 species analyzed whose presence or absence was predicted correctly for stands within the habitat type. Predictions were the same for all USFS forest types within a vegetation type of Hamel (1992).

Vegetation type –Hamel (1992)	USFS forest type (code)	Successional stage[a]	Accuracy	Error rate		No. of stands
				Omission	Commission	
Longleaf-slash pine	longleaf pine (21)		72.5	23.5	3.9	2
		2	78.9	14.8	6.3	16
		3	81.0	10.5	8.5	22
			71.4	6.3	22.3	93
	slash pine (22)		70.2	6.7	23.1	17
Loblolly-shortleaf pine	loblolly pine (31)		75.4	6.9	17.6	17
			83.6	6.2	10.2	29
		4	66.6	8.3	22.1	83
Mixed pine hardwood	shortleaf pine-oak (12)		80.4	19.6	0.0	
	loblolly pine-hardwood (13)	2	58.8	25.5	15.7	
		4	64.3	6.7	29.0	
	southern red oak-yellow pine (44)		64.7	0.0	35.3	
	white oak-black oak-yellow pine (47)		64.2	2.9	32.8	
Oak-hickory	white oak-red oak-hickory (53)		92.8	7.2	0.0	12
			53.7	1.4	45.0	43
Southern scrub oak	scrub oak (57)		86.3	11.8	2.0	
Elm-ash-cottonwood	sweetgum-yellow poplar (58)		71.2	9.8	19.0	
		4	63.3	2.6	34.1	62
Oak-gum-cypress	swamp chestnut oak-cherrybark oak (61)		76.5	23.5	0.0	
	sweetgum-Nuttall oak-willow (62)		73.2	10.5	16.3	
		4	64.7	4.9	30.4	
	laurel oak-willow oak (64)	4	56.9	9.8	33.3	
	sweet bay-swamp tupelo-red maple (68)	2	88.2	5.9	4.9	
			73.3	7.5	19.2	
			66.7	0.8	32.5	
	bottomland hardwood-yellow pine (46)	4	65.7	0.7	33.7	
Mean			71.4	9.1	19.5	

[a] Successional stages: 1 = grass-forb; 2 = shrub-seedling; 3 = sapling-poletimber; 4 = sawtimber.

BIRDHAB tended to err on the side of over-predicting habitat suitability rather than under-predicting it; error of commission rates generally were much higher than error of omission rates. There may be several components to this error. First, inaccuracies may have existed in the input data (i.e., the CISC database). If a stand's forest type or age was incorrect in CISC, the habitat type designation would have been incorrect and predictions would have been inappropriate. We corrected all known errors, but a few still may have existed. Users should be aware of the critical importance of accurate stand inventory data. Second, inaccuracy certainly existed in the validation data. Determination of presence-absence based on one ten-minute point count was problematic; species frequently were overlooked, as indicated by our comparison of the point count and Dunning survey data sets. Because BIRDHAB did tend to over-predict habitat quality, we feel the detection problem resulted in missing species where they were predicted present far more often than it resulted in missing species where they were predicted absent. Consequently, our results may be viewed as minimal estimates of BIRDHAB's accuracy. Had detection rates been greater for those species with high error of commission rates, accuracy undoubtedly would have improved. Third, many WHR models are intentionally designed conservatively so that if a species may possibly be present on a site, a manager will be prompted to consider it (Dedon et al. 1986). BIRDHAB is apparently no exception. However, conservative predictions could lead to habitat loss if relatively poor habitat was defined as suitable and was used in mitigation for destruction of optimal habitat. Finally, the model may simply be inaccurate. Regardless of the cause, errors of commission may be considered risk-aversive in conservation planning in that it is better to over-predict than under-predict (Edwards et al. 1996).

A few species had relatively high error of omission rates (e.g., white-eyed vireo, indigo bunting, pine warbler, tufted titmouse), indicating that they occurred with some regularity in habitat predicted to be unsuitable. The first, second, and last potential sources of error discussed above may apply to errors of omission as well. With regard to the second (the potential effect of errors in validation data), species detected in adjacent stands or those attracted from adjacent stands by tape playback would have caused errors of omission if the adjacent stand was of a different type. Similarly, detections of species foraging in habitats that were unsuitable to them for breeding would have caused errors of omission (e.g., many forest species occasionally forage in clearcuts and young pine plantations, Krementz and Christie 1999). Additionally, a portion of the error of omission rate likely was attributable to our treatment of "marginal" habitat predictions as predictions of "absent." Tbat is, the species actually occurred in habitat predicted to be marginal, an outcome that should not be considered model error. True model error of omission should be considered more serious as it leads to the exclusion of species from consideration in impact assessments. It was difficult to determine the proportion of the observed error (both commission and omission) that was due to true model error, but clearly it was less than the observed rate, and thus true accuracy likewise was greater than the observed rate.

BIRDHAB performed more poorly for some habitats than others, either because less was known about the bird communities of certain habitats at the time of model development, or because the communities in those habitats were inherently more difficult to predict. It consistently predicted too many species in sawtimber hardwood (both upland and bottomland), with error of commission rates >30% for 6 forest types, and it predicted too few species in grass-forb and shrub-seedling successional stages of several habitats (primarily longleaf pine) where error of omission rates were 23.4 and 14.8%, respectively. Conversely, in 12 stands of oak-hickory-sapling-poletimber, not a single error of commission was made and accuracy was 92%. Likewise, habitat-specific accuracy rates varied widely within and among other studies. Raphael and Marcot (1986) reported accuracy of 35-78% for breeding birds in 4 seral stages of mixed evergreen forests of northwestern California, although their figures were conservative because they did not include species not predicted and not found. Rice et al. (1986) reported accuracy rates of 88-96% for breeding birds in two habitats along the lower Colorado River, though they excluded several problematic species with intermediate predictions.

BIRDHAB's good predictive ability for several species indicated that qualitative, expert-based WHR models can be useful and accurate tools. Therefore, for some species, model performance may be improved by adjustment of the predicted suitability of certain habitats within the bird-habitat relationship matrix upon which BIRDHAB is based (Hamel 1992). Examination of species- and habitat-specific error rates may facilitate this refinement by highlighting where errors currently exist. However, for other species, incorporation of quantitative vegetation and landscape structural variables (e.g., understory density, canopy closure, stand size, surrounding habitat type) may be necessary to improve predictive ability (Bolger et al. 1997). Finally, for some extreme habitat generalists, accurate prediction may not be possible by either qualitative or quantitative models.

Management implications

From a land management and conservation perspective, a model should perform well for those species most likely to be the focus of impact assessments. Thus, the fact that BIRDHAB performed well for habitat specialists is fortuitous. Many species with narrow habitat associations tend to be relatively rare, particularly if the habitat with which they are associated is rare, and thus are often the focus of impact assessments. For example, BIRDHAB had at least moderate predictive ability for all of the tested species that were ranked by Partners in Flight as "very high" (4 species) or "extremely high" (I species) priority for conservation attention in the Atlantic coastal plain and for most (13 of 18 species) that were ranked as "high priority." Likewise, WHR models for species for which predictive ability is poor clearly should not be used in land management planning. If their use is restricted to appropriate species (those with significant predictive models), BIRDHAB and similar models should be useful tools for impact assessment by land managers.

BIRDHAB, as well as Hamel's (1992) matrix, may be useful tools at a regional scale, such as that at which gap analysis is frequently conducted. Accuracy of WHR models is dramatically affected by the scale of prediction, with most models performing better at scales larger than the individual stand

(Hamel et al. 1986, Raphael and Marcot 1986). For example, Hamel et al. (1986) reported that model success (i.e., the percentage of predicted species that were observed) increased from 41-63% from the stand to the regional scale for oak-hickory forests in Kansas, and Edwards et al. (1996) reported accuracy of 81-95% in predicting breeding bird occurrence at the large scale of 8 national parks in Utah. BIRDHAB's habitat-based accuracy rates fell within the range of those previously reported for breeding birds, even at larger scales. Scaling up to the landscape or regional level should yield even higher accuracy rates. As with any model, however, thorough testing should be conducted for each scale at which it is intended to be used, especially if such scales would require the use of input other than CISC data to determine habitat types, as in GAP analysis.

Acknowledgments. This research was supported by the United States Department of Energy-Savannah River Operations Office through the USFS-Savannah River under inter-agency agreement DE-IA09-76SR0056, and by the USFS Southern Research Station, Clemson University, the University of Florida, and the University of Georgia. The research benefited from discussions with J. I. Blake, M. J. Conroy, I? B. Hamel, W. Jarvis, E.T. LeMaster, and B. G. Marcot. We thank B. Parresol, P. B. Hamel, W. C. Hunter, and C. E. Moorman for reviewing the manuscript.

Literature cited

BLOCK, W. M., M. L. MORRISON, J. VERNER, AND I? N. MANLEY. 1994. Assessing wildlife-habitat relationships models: a case study with California oak woodlands. Wildlife Society Bulletin 22: 549-560.

BOLGER, D. T., T. A. SCOTT, AND J. T. ROTENBERRY. 1997. Breeding bird abundance in an urbanizing landscape in coastal southern California. Conservation Biology 11: 406-421.

BUFFINGTON, J. M., J. C. KILGO, R. A. SARGENT, K. V. MILLER, AND B. R. CHAPMAN. 1997. Comparison of breeding bird communities in bottomland hardwood forests of different successional stages. Wilson Bulletin 109: 314-319.

DAWSON, D. K., D. R. SMITH, AND C. S. ROBBINS. 1995. Point count length and detection of forest neotropical migrant birds. Pages 35-43 *in* C. J. Ralph, J. R. Sauer, and S. Droege, editors. Monitoring bird populations by point counts. United States Forest Service General Technical Report PSW-GTR-149.

DEDON, M. F., S. A. LAYMON, AND R. H. BARRETT. 1986. Evaluating models of wildlife-habitat relationships of birds in black oak and mixed-conifer habitats. Pages 115-120 *in* J. Verner, M. L. Morrison, and C. J. Ralph, editors. Wildlife 2000: modeling habitat relationships of terrestrial vertebrates. University of Wisconsin, Madison, USA.

DETTMERS, R., D. A. BUEHLER, J. G. BARTLETT, AND N. A. KLAUS. 1999. Influence of point count length and repeated visits on habitat model performance. Journal of Wildlife Management 63: 815-823.

DUNNING, J. B., AND B. D. WATTS. 1990. Regional differences in habitat occupancy by Bachman's Sparrow. Auk 107: 463-472.

EDWARDS, T. C., E. T. DESHLER, D. FOSTER, AND G. G. MOISEN. 1996. Adequacy of wildlife habitat relation models for estimating spatial distributions of terrestrial vertebrates. Conservation Biology 10: 263-270.

FIELDING, A. H., AND J. F. BELL. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24: 38-49.

HAMEL, I? B. 1992. The land manager's guide to the birds of the South. The Nature Conservancy, Chapel Hill, North Carolina, USA.

HAMEL, P. B., N. D. COST, AND R. M. SHEFFIELD. 1986. The consistent characteristics of habitats: a question of scale. Pages 121-128 *in* J. Verncr, M. L. Morrison, and C. J. Ralph, editors. Wildlife 2000: modeling habitat relationships of terrestrial vertebrates. University of Wisconsin, Madison, USA.

KILGO, J. C., R. A. SARGENT, B. R. CHAPMAN, AND K. V. MILLER. 1998. Effects of stand width and adjacent habitat on breeding bird communities in bottomland hardwoods. Journal of Wildlife Management 62: 72-83.

KILGO, J. C., R. A. SARGENT, K. V. MILLER, AND B. R. CHAPMAN. 1997. Landscape influences on breeding bird communities in hardwood fragments in South Carolina. Wildlife Society Bulletin 25: 878-885.

KREMENTZ, D. G., AND J. S. CHRISTIE. 1999. Scrub-successional bird community dynamics in young and mature longleaf pine-wiregrass savannahs. Journal of Wildlife Management 63: 803-814.

LANDIS, J. R., AND G. G. KOCH. 1977. The measurement of observer agreement for categorical data. Biometrics 33: 159-174.

MARCOT, B. G., M. G. RAPHAEL, AND K. H. BERRY. 1983. Monitoring wildlife habitat and validation of wildlife-habitat relationships models. Transactions of the North American Wildlife and Natural Resources Conference 48: 315-329.

MAYER, J. J., L. D. WIKE, J. C. KILGO, AND 1. L. BRISBIN. 1997. An annotated checklist of the birds of the Savannah River Site. Environmental Information Document WSRC-TR-98-0000, Westinghouse Savannah River Company, Aiken, South Carolina, USA.

MEYERS, J. M, AND E. I? ODUM. 2000. Early avian research at the Savannah River Site: historical highlights and possibilities for the future. Pages 18-31 in J. B. Dunning and J. C. Kilgo, editors. Avian research at the Savannah River Site: a model for integrating basic research and long-term management. Studies in Avian Biology 21.

MORRISON, M. L., B. G. MARCOT, AND R. W. MANNAN. 1992. Wildlife-habitat relationships: concepts and applications. University of Wisconsin, Madison, USA.

PULLIAM, H. R. 1988. Sources, sinks, and population regulation. American Naturalist 132: 653-661.

RALPH, C. J., J. R. SAUER, AND S. DROEGE. 1995. Monitoring bird populations by point counts. United States Forest Service General Technical Report PSW-GTR-149.

RAPHAEL, M. G., AND B. G. MARCOT. 1986. Validation of a wildlife-habitat relationships model: vertebrates in a Douglas-fir sere. Pages 129-138 *in* J. Verner, M. L. Morrison, and C. J. Ralph, editors. Wildlife 2000: modeling habitat relationships of terrestrial vertebrates. University of Wisconsin, Madison, USA.

RICE, J. C., R. D. OHMART, AND B. W. ANDERSON. 1986. Limits in a data-rich model: modeling experience with habitat management on the Colorado River. Pages 79-86 in J. Verner, M. L.

Morrison, and C. J. Ralph, editors. Wildlife 2000: modeling habitat relationships of terrestrial vertebrates. University of Wisconsin, Madison, USA.

SAS INSTITUTE, INC. 1988. SAS/STAT® user's guide. Release 6.03 edition. SAS Institute Inc., Cary, North Carolina, USA.

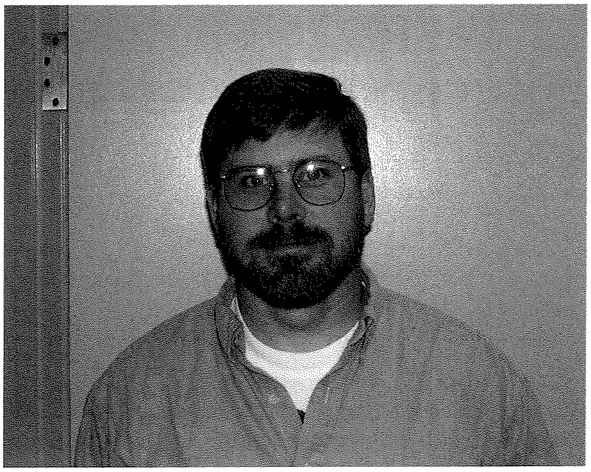
TIMOTHY, K. G., AND D. F. STAUFFER. 1991. Reliability of selected avifauna information in a computerized information system. Wildlife Society Bulletin 19: 80-88.

UNITED STATES FOREST SERVICE. 1994. BIRDHAB: GIS bird habitat evaluation for resource managers, user guide. Beta version 0.999. United States Forest Service, Atlanta Georgia, USA.

VAN HORNE, B. 1983. Density as a misleading indicator of habitat quality. Journal of Wildlife Management 47: 893-901.

VICKERY, P. D., M. L. HUNTER, AND J. V. WELLS. 1993. Is density an indicator of breeding success? Auk 109: 706-710.

WORKMAN, S. W., AND K. W. MCLEOD. 1990. Vegetation of the Savannah River Site: major community types. National Environmental Research Park Program Publication SRO-NERP-19. United States Department of Energy Savannah River Site.



John C. Kilgo (photo) is a research wildlife biologist with the U.S. Forest Service's (USFS) Southern Research Station at the Savannah River Site near Aiken, South Carolina. He received a B.S. in biology from Wofford College, an M.S. in wildlife ecology from the University of Florida, and a Ph.D. in wildlife ecology from the University of Georgia. His research interests include habitat relationships of forest birds, particularly in bottomland hardwood forests, and the ecology of white-tailed deer and wild turkeys. ***David L. (Dave) Gartner*** is a mathematical statistician with the USFS Southern Research Station in Knoxville, Tennessee. ***Brian R. Chapman*** is Dean of the College of Arts and Sciences at Sam Houston State University in Huntsville, Texas. ***John B. (Barney) Dunning, Jr.*** is associate professor of wildlife ecology in the Department of Forestry and Natural Resources at Purdue University. ***Kathleen E. (Kay) Franzreb*** is a research wildlife biologist in the USFS Southern Research Station at Knoxville. ***Sidney A. (Sid) Gauthreaux*** is professor of zoology in the Department of Biological Sciences at Clemson University. ***Cathryn (Katie) Greenberg*** is a research wildlife biologist in the USFS Southern Research Station at Bent Creek Experimental Forest in Asheville, North Carolina. ***Douglas J. (Doug) Levey*** is professor of zoology at the University of Florida in Gainesville. ***Karl V. Miller*** is professor of wildlife ecology in the Warnell School of Forest Resources at the University of Georgia in Athens. ***Scott F. Pearson*** is a biologist with the Washington Department of Natural Resources, Lands and Resources Division, in Olympia. All authors have studied or monitored forest bird populations on the Savannah River Site. Much of their work there was summarized in a recent edition of Studies in *Avian* Biology entitled "Avian research at the Savannah River Site: a model for integrating basic research and long-term management," edited by John Kilgo and Barney Dunning (Vol. 21).

Associate Editor: Pace

Appendix 1. Derivation of successional stages from condition class and stand age information by BIRDHAB.

Look up direction in BIRDHAB →		
IF USFS Condition class (code):	AND Stand age:	THEN Successional stage:
In regeneration (1)	≤2	Grass-forb
Damaged poletimber (2)	Any	Sapling-poletimber
Damaged sawtimber (3)	Any	Sawtimber
Forest pest infestation (4)	<30	Sapling-poletimber
Forest pest infestation (4)	≥30	Sawtimber
Sparse poletimber (5)	Any	Sapling-poletimber
Sparse sawtimber (6)	Any	Sawtimber
Low quality poletimber (7)	Any	Sapling-poletimber
Low quality sawtimber (8)	Any	Sawtimber
Mature poletimber (9)	Any	Sapling-poletimber
Mature sawtimber (10)	Any	Sawtimber
Immature poletimber (11)	Any	Sapling-poletimber
Immature sawtimber (12)	Any	Sawtimber
Seedling/sapling adequately stocked (13)	≤10	Shrub-seedling
Seedling/sapling adequately stocked (13)	≥11	Sapling-poletimber
Seedling/sapling inadequately stocked (14)	≤10	Shrub-seedling
Seedling/sapling inadequately stocked (14)	≥11	Sapling-poletimber
Non-stocked (15)	Any	Shrub-seedling
Group selection management (16)	Any	Sawtimber