12-2000

Sensitivity of Species Habitat-Relationship Model Performance to Factors of Scale

J. W. Karl

P. J. Heglund

E. O. Garton

J. M. Scott

N. M. Wright

See next page for additional authors

Authors

J. W. Karl, P. J. Heglund, E. O. Garton, J. M. Scott, N. M. Wright, and Richard L. Hutto

SENSITIVITY OF SPECIES HABITAT-RELATIONSHIP MODEL PERFORMANCE TO FACTORS OF SCALE

J. W. KARL,[1,6] P. J. HEGLUND,[2,7] E. O. GARTON,[3] J. M. SCOTT,[4] N. M. WRIGHT,[1,8] AND R. L. HUTTO[5]

[1]*Idaho Fish and Wildlife Cooperative Research Unit, University of Idaho, Moscow, Idaho 83844-1141 USA*
[2]*Department of Biological Sciences, University of Idaho, Moscow, Idaho 83844-3051 USA*
[3]*Department of Fish and Wildlife, University of Idaho, Moscow, Idaho 83844-1136 USA*
[4]*Biological Resources Division/USGS, Idaho Fish and Wildlife Cooperative Research Unit,
University of Idaho, Moscow, Idaho 83844-1141 USA*
[5]*Division of Biological Sciences, University of Montana, Missoula, Montana 59812-1002 USA*

Abstract. Researchers have come to different conclusions about the usefulness of habitat-relationship models for predicting species presence or absence. This difference frequently stems from a failure to recognize the effects of spatial scales at which the models are applied. We examined the effects of model complexity, spatial data resolution, and scale of application on the performance of bird habitat relationship (BHR) models on the Craig Mountain Wildlife Management Area and on the Idaho portion of the U.S. Forest Service's Northern Region. We constructed and tested BHR models for 60 bird species detected on the study areas. The models varied by three levels of complexity (amount of habitat information) and three spatial data resolutions (0.09 ha, 4 ha, 10 ha). We tested these models at two levels of analysis: the site level (a homogeneous area <0.5 ha) and cover-type level (an aggregation of many similar sites of a similar land-cover type), using correspondence between model predictions and species detections to calculate kappa coefficients of agreement. Model performance initially increased as models became more complex until a point was reached where omission errors increased at a rate greater than the rate at which commission errors were decreasing. Heterogeneity of the study areas appeared to influence the effect of model complexity. Changes in model complexity resulted in a greater decrease in commission error than increase in omission error. The effect of spatial data resolution on the performance of BHR models was influenced by the variability of the study area. BHR models performed better at cover-type levels of analysis than at the site level for both study areas. Correct-presence estimates (1 − minus percentage omission error) decreased slightly as number of species detections increased on each study area. Correct-absence estimates (1 − percentage commission error) increased as number of species detections increased on each study area. This suggests that a large number of detections may be necessary to achieve reliable estimates of model accuracy.

Key words: avian habitat; bird counts; GIS; Idaho; species habitat-relationship models.

Researchers have come to different conclusions about the validity and usefulness of habitat-relationship models (see Short and Hestbeck 1995, Davis 1996, Edwards 1996, Scott et al. 1996). Scott et al. (1993) and Edwards et al. (1996) reported high levels of performance from Gap Analysis wildlife habitat-relationship models in Idaho and Utah, respectively, when applied to managed areas generally >10 000 ha. However, Dedon et al. (1986) and Block et al. (1994) reported California habitat-relationship model performance low enough to question its use as a management tool. The key difference between these studies was the scale of analysis used by the researchers.

Block et al. (1994) found moderate error rates when investigating the ability of habitat-relationship models to predict species occurrence between study areas ranging from 1700 to 8000 ha in size. However, when they compared the ability of habitat-relationship models to predict species occurrence between major cover types within their study areas, their error rates rose. This has led some researchers to conclude that habitat-relationship models were only appropriate for regional, coarse-scale analyses (Block et al. 1994, Short and Hestbeck 1995).

The performance of a habitat-relationship model should be related to the scale of analysis. Csuti (1996) stated that the accuracy of model predictions generally improved at coarse map scales because very large areas were considered. Hollander et al. (1994) matched different data resolutions and sources of habitat information to create the most accurate models for different scales. Stoms (1992) reported that model predictions were sensitive to the resolution of the habitat data used. Additionally, assessment of model performance must

be conducted with an analysis unit that matches the target application of the model (Krohn 1996, Boone and Krohn 1999).

Csuti (1996:135) stated that "information used to create maps of animal distributional areas remains the same across the range of map scales." However, if this were the case, we would expect a properly constructed habitat-relationship model to perform at the same level across all scales. But, because species respond to ecological processes operating simultaneously over a wide range of scales, different factors may limit species distributions at different scales (Wiens 1989). Thus the wide range of reported values for the performance of habitat-relationship models at different scales suggests that information needs for accurately predicting suitable habitat for a species differ with level of application (Hollander et al. 1994).

These information needs involve matching appropriate sources of habitat information with the best spatial data resolution and an appropriate analysis unit for the management or research question being asked. Questions about the distribution of a species across its range or the best location for new conservation reserves are coarse-scale modeling applications. These questions require different habitat and spatial data than fine-scale questions. Landscape-scale questions, for which modeling would be useful, might regard the effect of management actions on vertebrate populations within a managed area or development of a sampling scheme for avian nest searches.

However, due to time and budget constraints, it is not always possible to create separate habitat-relationship models for each scale of application. Some programs (e.g., Gap Analysis) attempt to use a multiple-scale approach to modeling species habitats where a biogeographic range is first modeled and used as the modeling extent for regional habitat modeling (Scott et al. 1993, Butterfield et al. 1994, Csuti 1996). Resource managers and land use planners have shown considerable interest in applying widely available regional habitat-relationship models (e.g., Gap Analysis models) to local situations for which the models were not designed (e.g., Wright 1997).

We lack knowledge of how the extent of habitat information, resolution of input spatial data, and the level of analysis interact to affect the performance of habitat-relationship models at different spatial and temporal scales and the validity of using the models at these scales. Our objectives were to construct a series of bird habitat-relationship (BHR) models for a set of northern Idaho birds and use them to determine how accuracy of the models changes at different spatial levels of analysis, with different resolutions of geographic information system (GIS) input data, and with different levels of model complexity. We considered this objective for the Craig Mountain Wildlife Management Area, Idaho (Craig Mountain) and the Idaho portion of North-
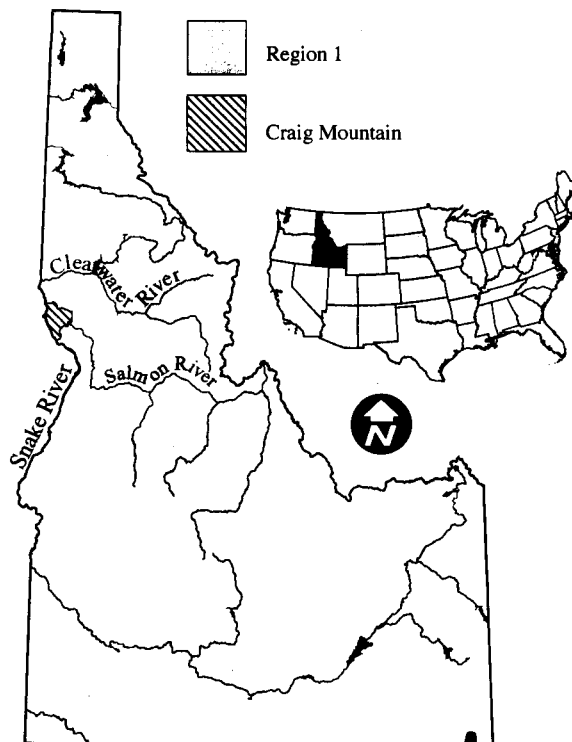

FIG. 1.   The Craig Mountain and Region 1 study areas in northern Idaho.

ern Rocky Mountain Region (Region 1, U.S. Forest Service).

STUDY AREAS

The Craig Mountain area extends from ~40 km south of Lewiston, Idaho southward to the confluence of the Salmon River and the Snake River (Fig. 1). The total Craig Mountain area comprises ~60 000 ha. The majority of this land consists of the Craig Mountain Wildlife Management Area, administered by the Idaho Department of Fish and Game. The Bureau of Land Management, the Nature Conservancy, Idaho Department of Lands, and the Nez Perce Tribe also own significant parcels of land and actively engage in management practices on the Craig Mountain area. The vegetation of Craig Mountain varies widely with elevation and aspect from native and exotic grasslands and riparian areas at low elevations to ponderosa pine (*Pinus ponderosa*), Douglas-fir (*Pseudotsuga menziezii*), grand fir (*Abies grandis*), and lodgepole pine (*Pinus contorta*) at higher elevations.

The second area, Region 1, is comprised mostly of the Idaho portion of U.S. Forest Service Northern Region (the Idaho Panhandle, Clearwater, Saint Joe, and Nez Perce National Forests) as well as land owned by the Potlatch Corporation (Fig. 1). This area (2.75 × 10[6] ha) begins just north of the Clearwater River, extending northward to the tip of the Idaho panhandle, but excluding the dry grasslands of the Snake River

TABLE 1. Common and scientific names, and number of detections for bird species detected in either Craig Mountain or Region 1, northern Idaho.

Common name	Scientific name	Region 1	Craig Mountain
Mallard	*Anas platyrhynchos*	2	
Northern Pintail	*Anas acuta*		1
Common Merganser	*Mergus merganser*	2	
Sharp-shinned Hawk	*Accipiter striatus*	2	2
Cooper's Hawk	*Accipiter cooperii*		2
Northern Goshawk	*Accipiter gentilis*	4	
Red-tailed Hawk	*Buteo jamaicensis*	5	4
American Kestrel	*Falco sparverius*	5	14
Gray Partridge	*Perdix perdix*		3
Chukar	*Alectoris chukar*		5
Spruce Grouse	*Falcipennis canadensis*	1	
Blue Grouse	*Dendragapus obscurus*	2	4
Ruffed Grouse	*Bonasa umbellus*	66	22
Wild Turkey	*Meleagris gallopavo*	1	5
California Quail	*Callipepla californica*		4
Mountain Quail	*Oreortyx pictus*	1	
Spotted Sandpiper	*Actitis macularia*	2	3
Common Snipe	*Gallinago gallinago*	3	5
Mourning Dove	*Zenaida macroura*	2	3
Common Barn Owl	*Tyto alba*		1
Great Gray Owl	*Strix nebulosa*		1
Common Nighthawk	*Chordeiles niger*		1
Vaux's Swift	*Chaetura vauxi*		7
Calliope Hummingbird	*Stellula calliope*	12	5
Broad-tailed Hummingbird	*Selasphorus platycercus*	2	3
Rufous Hummingbird	*Selasphorus rufus*	41	18
Belted Kingfisher	*Ceryle alcyon*	5	3
Lewis' Woodpecker	*Melanerpes lewis*	1	
Williamson's Sapsucker	*Sphyrapicus thyroideus*	7	8
Red-naped Sapsucker	*Sphyrapicus nuchalis*	70	24
Downy Woodpecker	*Picoides pubescens*	2	14
Hairy Woodpecker	*Picoides villosus*	49	20
Three-toed Woodpecker	*Picoides tridactylus*	3	
Black-backed Woodpecker	*Picoides arcticus*	2	
Northern Flicker	*Colaptes auratus*	56	67
Pileated Woodpecker	*Dryocopus pileatus*	34	10
Olive-sided Flycatcher	*Contopus borealis*	42	14
Western Wood-pewee	*Contopus sordidulus*	4	27
Willow Flycatcher	*Empidonax traillii*	19	4
Hammond's Flycatcher	*Empidonax hammondii*	239	32
Dusky Flycatcher	*Empidonax oberholseri*	174	115
Cordilleran Flycatcher	*Empidonax occidentalis*	22	27
Say's Phoebe	*Sayornis saya*		1
Western Kingbird	*Tyrannus verticalis*		13
Eastern Kingbird	*Tyrannus tyrannus*		2
Horned Lark	*Eremophila alpestris*		6
Violet-green Swallow	*Tachycineta thalassina*		18
Northern Rough-winged Swallow	*Stelgidopteryx serripennis*		1
Cliff Swallow	*Petrochelidon pyrrhonota*		8
Barn Swallow	*Hirundo rustica*	1	2
Gray Jay	*Perisoreus canadensis*	61	21
Steller's Jay	*Cyanocitta stelleri*	56	25
Clark's Nutcracker	*Nucifraga columbiana*	3	3
Black-billed Magpie	*Pica pica*		6
American Crow	*Corvus brachyrhynchos*		19
Common Raven	*Corvus corax*	9	9
Black-capped Chickadee	*Parus atricapillus*	143	77
Mountain Chickadee	*Poecile gambeli*	150	130
Boreal Chickadee	*Poecile hudsonicus*	4	
Chestnut-backed Chickadee	*Poecile rufescens*	222	7
Red-breasted Nuthatch	*Sitta canadensis*	559	120
White-breasted Nuthatch	*Sitta carolinensis*	22	13
Pygmy Nuthatch	*Sitta pygmaea*	4	
Brown Creeper	*Certhia americana*	71	25
Rock Wren	*Salpinctes obsoletus*	1	34
Canyon Wren	*Catherpes mexicanus*	2	18
House Wren	*Troglodytes aedon*	23	48
Winter Wren	*Troglodytes troglodytes*	302	11

TABLE 1. Continued.

Common name	Scientific name	Region 1	Craig Mountain
American Dipper	*Cinclus mexicanus*	13	
Golden-crowned Kinglet	*Regulus satrapa*	678	88
Ruby-crowned Kinglet	*Regulus calendula*	118	90
Western Bluebird	*Sialia mexicana*		9
Mountain Bluebird	*Sialia currucoides*	6	7
Townsend's Solitaire	*Myadestes townsendii*	56	7
Veery	*Catharus fuscescens*		23
Swainson's Thrush	*Catharus ustulatus*	470	37
Hermit Thrush	*Catharus guttatus*	49	37
American Robin	*Turdus migratorius*	266	226
Varied Thrush	*Ixoreus naevius*	192	7
Gray Catbird	*Dumetella carolinensis*	4	
Cedar Waxwing	*Bombycilla cedrorum*	42	37
European Starling	*Sturnus vulgaris*	1	5
Solitary Vireo	*Vireo solitarius*	183	94
Warbling Vireo	*Vireo gilvus*	263	83
Red-eyed Vireo	*Vireo olivaceus*	11	38
Orange-crowned Warbler	*Vermivora celata*	146	118
Nashville Warbler	*Vermivora ruficapilla*	113	29
Yellow Warbler	*Dendroica petechia*	80	113
Yellow-rumped Warbler	*Dendroica coronata*	588	163
Townsend's Warbler	*Dendroica townsendi*	667	62
American Redstart	*Setophaga ruticilla*	17	
Northern Waterthrush	*Seiurus noveboracensis*	6	
MacGillivray's Warbler	*Oporornis tolmiei*	554	157
Common Yellowthroat	*Geothlypis trichas*	3	
Wilson's Warbler	*Wilsonia pusilla*	138	14
Yellow-breasted Chat	*Icteria virens*		47
Western Tanager	*Piranga ludoviciana*	321	102
Black-headed Grosbeak	*Pheucticus melanocephalus*	75	30
Lazuli Bunting	*Passerina amoena*	72	163
Spotted Towhee	*Pipilo maculatus*	59	155
Chipping Sparrow	*Spizella passerina*	191	169
Vesper Sparrow	*Pooecetes gramineus*		41
Lark Sparrow	*Chondestes grammacus*		8
Savannah Sparrow	*Passerculus sandwichensis*	3	
Fox Sparrow	*Passerella iliaca*	112	2
Song Sparrow	*Melospiza melodia*	97	68
Lincoln's Sparrow	*Melospiza lincolnii*	6	1
White-crowned Sparrow	*Zonotrichia leucophrys*	4	1
Dark-eyed Junco	*Junco hyemalis*	683	198
Red-winged Blackbird	*Agelaius phoeniceus*		7
Western Meadowlark	*Sturnella neglecta*	8	44
Brewer's Blackbird	*Euphagus cyanocephalus*		1
Brown-headed Cowbird	*Molothrus ater*	39	50
Bullock's Oriole	*Icterus bullockii*		23
Pine Grosbeak	*Pinicola enucleator*	5	1
Cassin's Finch	*Carpodacus cassinii*	27	20
Red Crossbill	*Loxia curvirostra*	72	23
White-winged Crossbill	*Loxia leucoptera*	10	
Pine Siskin	*Carduelis pinus*	213	37
American Goldfinch	*Carduelis tristis*	6	49
Evening Grosbeak	*Coccothraustes vespertinus*	38	6

Notes: We detected 121 bird species between the two study areas. This represents 50.6% of the 239 confirmed breeding birds in Idaho (Stephens and Sturts 1998). Birds with ≥15 detections were modeled.

Valley and the Palouse agriculture lands. Most of the Region 1 area is dominated by coniferous forest communities that are in various stages of timber management.

METHODS

BHR model construction

We followed the model building methods proposed by Scott et al. (1993; see also Csuti [1996] and Smith and Catanzaro [1996]). This method consisted of four major steps: establishing a species list, defining species range limits, collecting species habitat information and determining habitat relationships, and modeling the species habitat in a GIS using the information gathered.

We selected only bird species that were detected at ≥15 survey points out of 440 and 1628 sites on either Craig Mountain or Region 1, respectively. We chose

15 points of observation as the lower limit for species selection because we felt that this would maximize the number of species included while allowing sufficient sample sizes to estimate kappa coefficients of agreement (Cohen 1960). The 15-point limit also allowed us to examine whether increased species detections yielded better accuracy estimates. In addition, we limited our list to breeding bird species observed in both study areas, and since no birds reached their range limits in our study areas (Groves et al. 1997), we omitted the range definition step of Scott et al. (1993). At the 15-point limit, we modeled 52 species on Craig Mountain and 52 on Region 1, for a total of 65 species between the two study areas (Table 1). We did not test BHR models for an additional 56 species (46.3%) that were detected at ≤14 stations. Because the lower limit of 15 detections represents a higher percentage of survey points on Craig Mountain (3.4%) than Region 1 (0.9%), the possibility exists that we might include species on Region 1 that would not have enough detections for inclusion from the Craig Mountain data set. We found three species that might fit this description: Pileated Woodpecker (scientific names in Table 1), Olive-sided Flycatcher, and Evening Grosbeak. Six species (Willow Flycatcher, Chestnut-backed Chickadee, White-breasted Nuthatch, Winter Wren, Wilson's Warbler, and Fox Sparrow) were detected in large numbers on Region 1, but were very rare or not detected on Craig Mountain. This was due to Craig Mountain either not containing appropriate habitat, or being outside (or on the edge of) the range of the species. Therefore, we did not consider the fixed lower limit of 15 points to significantly effect our research.

We conducted a literature search to find relevant habitat information for each bird species. We assumed that *Life Histories of North American Birds* accounts (Poole et al. 1997) provided comprehensive reviews of the literature for a given species. For each species with a published life history account, we used that account as the primary reference for habitat information. In addition, we searched four databases for literature on the remaining species: Essential Ecology, Zoology and Plant Science Abstracts (NISC 1997*a*), Essential Ornithological Abstracts (NISC 1997*b*), Essential Wildlife and Conservation Biology Abstracts (NISC 1997*c*), and Wildlife Worldwide (NISC 1997*d*). We then reviewed the references returned by the literature search for pertinent habitat information and entered it into a database. For all species considered, we took habitat information from modeling efforts in Idaho (Groves et al. 1997), Oregon (Csuti et al. 1997), Washington (Smith et al. 1997), and Wyoming (Merrill et al. 1996). Specifically, we looked for information on use of vegetation cover types, elevation, canopy closures, and tree size classes. We chose these habitat variables for two reasons. First, previous research has indicated that species respond to these variables in habitat selection at the scales we considered (Short and Williamson 1986,

Johnson and Grier 1988). Second, these habitat variables can be accommodated in a GIS.

We coded the habitat information using the scheme developed by Redmond et al. (1996) (Table 2). Aspect, canopy closure, and tree and shrub size class were classified into discrete categories. We converted the coded habitat relationships into a database table where it could be accessed by the model construction programs.

We built a series of models differing in degree of model complexity (the amount of habitat information included) using a linear, additive process. This is a sequential method of modeling that uses the previous model as a template for the next model being created. We found sufficient habitat information to make models with two levels of complexity for every bird species considered. The first model for each species was based on cover type only. The second model also included elevation (53 species [82%]), or if elevation was not relevant based on lack of information in literature, then canopy cover (seven species [11%]). Forty-two species (65%) added canopy cover as the third level of model complexity. Nine species (14%), for which four levels of complexity were possible, were the only species for which size-class information was available; therefore, we omitted the fourth model due to such small number of species.

We used the Current Vegetation Map of Northern Idaho and Western Montana data of Redmond et al. (1996) for the dominant cover type and canopy cover values for northern Idaho. Patch size for forested cover types averaged 45 ± 1485 ha and 34 ± 574 ha for Region 1 and Craig Mountain, respectively (means ± 1 SD). Patch size for nonforested cover types was 2 ± 5 ha and 18 ± 423 ha for Region 1 and Craig Mountain, respectively. Redmond et al. (1996) reported producer's accuracy measures (probability of a pixel being correctly classified [Congalton 1991]) for cover type and canopy cover classifications on a scene-by-scene basis (Table 3). Redmond et al. (1996) included elevation, slope, and aspect measures from 1:24 000 digital elevation models (DEMs) in the coverage with the vegetation measures. Redmond et al.'s (1996) data had an original cell size of 0.09 ha, which we resampled to 4 ha and 10 ha. These are commonly used data resolutions. Because the Craig Mountain area is only ~60 000 ha, spatial data resolutions of >10 ha would not capture the variability of cover types on the area. Also, when the Craig Mountain data was resampled to resolutions >10 ha, the proportions of cover types begin to change. For this reason we did not include resolutions >10 ha in our analyses.

A series of GIS programs automated the model construction process. Vegetation cover type always formed the first model for each species. We added habitat information to the vegetation-only model in the order of elevation and canopy cover. If a habitat variable (e.g.,

TABLE 2. Vegetation codes and descriptions of land cover types used in modeling avian habitats in Craig Mountain and Region 1 in northern Idaho (from Redmond et al. 1996).

Code	Land cover type	Land cover description
1000	Urban	Towns and cities
2000	Agricultural	Row crops, dry farm crops, pastures, fallow fields
3101	Foothills grassland	Grass and forb co-dominated dry meadows and ridges associated with species such as: sagebrush (*Artemisia* spp.), snowberry (*Symphoricarpus* spp.), and willow (*Salix* spp.)
3102	Disturbed grassland	Grasslands with >30% cover of noxious weeds and other exotic species
3104	Montane parklands and subalpine meadows	Grassland ridges, forest openings, and meadows dominated by native perennial montane or subalpine grass species
3202	Warm mesic shrubland	Upland shrublands naturally occurring or initiated by fire or clearcutting
4203	Lodgepole pine	Coniferous forest with >66% lodgepole pine (*Pinus contorta*)
4206	Ponderosa pine	Coniferous forest with >66% ponderosa pine (*Pinus ponderosa*)
4207	Grand fir	Coniferous forest with >66% grand fir (*Abies grandis*)
4208	Subalpine fir	Coniferous forest with >66% subalpine fir (*Abies bifolia*)
4210	Western red cedar	Coniferous forest with >66% western red cedar (*Thuja plicata*)
4212	Douglas-fir	Coniferous forest with >66% Douglas-fir (*Pseudotsuga menziesii*)
4215	Western larch	Coniferous forest with >66% western larch (*Larix occidentalis*)
4220	Mixed subalpine forest	Coniferous forest with <66% subalpine fir, with >10% lodgepole pine, Douglas-fir, or Engelmann spruce
4221	Mixed mesic forest	Coniferous forest codominated by western red cedar, grand fir, western larch, or western hemlock
4222	Mixed xeric forest	Coniferous forest codominated by Douglas-fir and ponderosa pine
4223	Douglas-fir–lodgepole pine forest	Coniferous forest with >80% Douglas-fir and lodgepole pine
4225	Douglas-fir–grand fir forest	Coniferous forest with >80% Douglas-fir and grand fir
4226	Western red cedar–grand fir forest	Coniferous forest with >80% western red cedar and grand fir
4301	Mixed needleleaf/broadleaf forest	Forest with >25% and <75% cover of both coniferous and deciduous forest
6101	Needleleaf-dominated riparian forest	Coniferous forest in areas influenced by the presence of running or standing water
6102	Broadleaf-dominated riparian forest	Deciduous forest in areas influenced by the presence of running or standing water
6103	Needleleaf–broadleaf riparian forest	Woodlands codominated by coniferous and deciduous trees influenced by the presence of running or standing water
6104	Mixed riparian	Mixed forest and nonforest areas influenced by the presence of running or standing water
6201	Graminoid- and forb-dominated riparian	Areas dominated by graminoids or forbs influenced by the presence of running or standing water
6202	Shrub-dominated riparian	Areas influenced by the presence of running or standing water with >15% shrub cover and <15% tree cover
6203	Mixed nonforest riparian	Areas influenced by the presence of running or standing water codominated by shrubs and forbs
7300	Exposed rock	Cliffs, rock outcrops, talus slopes, scree

elevation) did not apply for a species, the program skipped it and proceeded on to the next variable.

We then applied the models, which varied in complexity, to 0.09-ha, 4-ha, and 10-ha resolution raster GIS data for each species. This yielded up to nine model combinations for each species.

Breeding bird survey design and data collection

We used breeding bird survey data collected on Craig Mountain in 1993, 1994 (Cassirer 1995), and 1997 (J. W. Karl, *personal observation*), and on Region 1 in 1994–1996 (R. L. Hutto, *personal observation*), to assess the performance of our BHR models. All of the studies used methods comparable to Hutto and Hoffland (1996; see also Reynolds et al. [1980], Ralph and Scott [1981], Ralph et al. [1995]). This was a variable-radius, circular-plot method in which observers record the species observed, distance, activity (flying or not flying), and location of the bird (within or outside of same cover type as the survey point). Each of the Craig

TABLE 3. Producer's accuracy measures for cover-type and canopy cover classifications by scene as reported by Redmond et al. (1996).

Scene	Producer's accuracy level (%)	
	Cover type	Canopy cover
P43/R26	83.60 ($n = 103$)	97.90 ($n = 146$)
P43/R27	63.45 ($n = 118$)	93.70 ($n = 146$)
P42/R26	63.77 ($n = 124$)	97.20 ($n = 198$)
P42/R27	71.23 ($n = 479$)	93.00 ($n = 382$)
P42/R28	53.35 ($n = 320$)	98.60 ($n = 198$)
P41/R27	53.89 ($n = 1182$)	95.40 ($n = 262$)
P41/R28	57.87 ($n = 972$)	95.20 ($n = 328$)

Note: Scenes were merged together, with the most accurate scene taking precedence over less accurate ones in areas of overlap.

Mountain points was surveyed two to four times per year for up to three years. Each of the Region 1 points was surveyed once per year for up to three years. For Region 1 and the 1994 and 1997 Craig Mountain field season, we recorded all birds detected aurally or visually within a 10-min count at each site. Cassirer (1995) used a 5-min point count in 1993. Rarer species were likely missed due to shorter sampling periods in 1993 (Dawson et al. 1995). However, because Craig Mountain sites were sampled more than once per year, and because we were interested only in the presence or absence of a species at each point and not in estimating density, we did not consider the discrepancy in survey time between 1993 and other years a problem. All birds detected outside of the cover type at the survey point were eliminated from the data set. We also eliminated all birds that were flying when detected, except for those birds whose detections are mostly restricted to aerial foraging (i.e., swallows, swifts, hawks).

We further truncated the data set to only those observations occurring within 50 m of the survey point for two reasons. First, the ability to accurately judge the distance of an observation and the cover type in which it occurred decreases with distance from the survey point (Hutto and Hoffland 1996; see also Scott et al. [1981]). Second, limiting the area of analysis around the survey point reduces the potential for variation in the values of the GIS data layers around the survey point.

We converted differentially corrected global positioning system (GPS) coordinates for each of the Craig Mountain survey sites into a GIS point coverage. We received GIS coordinates for the Region 1 survey points from the University of Montana's Landbird Monitoring Program. These coordinates were digitized from georegistered aerial photographs of the study area. We then converted the vector point coverages from each study area to raster grids with a 0.09-ha cell size.

Assessment of model performance

We defined site-level of analysis as occurring for an area of homogeneous vegetation <0.5 ha in size. This corresponds to the area that can be sampled with a single survey point. Thus, a given patch of vegetation could consist of one or many sites. The cover-type level of analysis is an aggregation of many sites of similar vegetation type (e.g., grassland, shrubland, coniferous forest). We assessed the performance of a BHR model at the site and cover-type levels by examining the correspondence between a species' detection at a survey point and the model's prediction of the species' occurrence there. Because a species could have been detected anywhere within a 50-m radius of the survey point, we used a three-cell by three-cell neighborhood, totaling 8100 m², around the survey-point cell (Fig. 2).

A series of programs in ARC/INFO GRID (Environmental Systems Research Institute, Redlands, California) tallied the correspondence between model predictions and species occurrence at the site level for the Craig Mountain and Region 1 areas. The program coded each model as 1 for predicted habitat and 0 for not predicted. At each survey point we used the ARC/INFO GRID FOCALMAX command to determine if there was predicted habitat anywhere within the nine-cell neighborhood (Fig. 2). If predicted habitat occurred within the analysis window, then the survey point cell received a value of 1. Otherwise it remained 0. Survey data were coded in a similar manner: 1 for detected, 0 for not detected. At each survey point, the program compared the binary values for species detections and
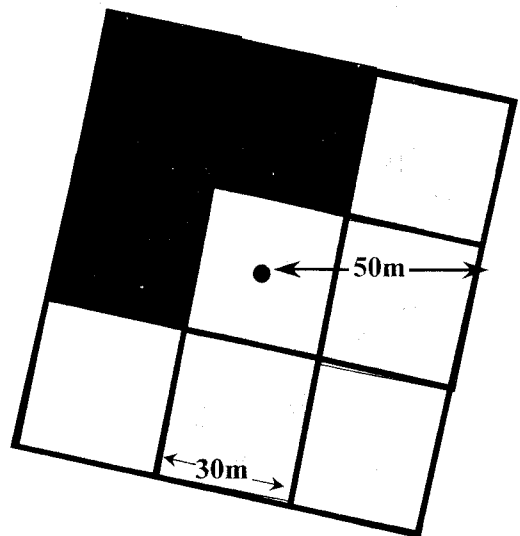

FIG. 2. The area associated with a variable-radius circular plot survey truncated to 50 m can be represented in a GIS with a 3 × 3 cell square of 0.09-ha cells. All birds detected within 50 m of the survey point were included in our study. The dark-shaded cells represent area predicted as suitable habitat by the species habitat relationship model. We used a focal function in the GIS to look for occurrence of suitable habitat within the 3 × 3 cell neighborhood.

TABLE 4. General vegetation types used for cover-type level assessment of model performance.

Vegetation type	Craig Mountain	Region 1
Agriculture		X
Grasslands	X	X
Mesic shrubland	X	X
Xeric shrubland		X
Deciduous forest		X
Coniferous forest	X	X
Mixed coniferous/deciduous forest		X
Riparian/wetland	X	X
Rock/barren	X	X

Note: These categories are generalizations of a hierarchical vegetation classification of Redmond et al. (1996) with plant species at its finest level (see Table 2).

model predictions. With this coding scheme there are four possible combinations: the species is detected and predicted, the species is detected but not predicted (omission error), the species is not detected but is predicted (commission error), or the species is not detected and not predicted. The program tallied the number of each of these four possibilities and output the result.

To define our cover-type level of analysis, we aggregated the vegetation data layer used in model construction (Redmond et al. 1996) to five and nine major cover types for Craig Mountain and Region 1, respectively (Table 4). For each study area we eliminated from consideration any cover type that did not have any survey points (i.e., agriculture, water). We then tallied the species predictions vs. observations for each of the cover types using methods similar to the site level calculations. For each study area, a species was considered present in a cover type if it was detected one or more times in that cover type.

We calculated the percentage omission (species detected but not predicted) and commission errors (species predicted but not detected) for each model. We calculated percentage correct-presence predictions as 100 minus the percentage omission error and percentage correct-absence predictions as 100 minus the percentage commission error for each model.

To examine the differences between models, we calculated Cohen's kappa coefficient of agreement (Cohen 1960, Rosenfield and Fitzpatrick-Lins 1986) for each of our 18 model combinations (three levels of complexity, three spatial data resolutions, and two levels of analysis). Kappa statistics are widely used in assessing the accuracy of remotely sensed land cover classifications (Congalton 1991) and represent the "... proportion of agreement (between a model and actual observation) after chance agreement is removed" (Rosenfield and Fitzpatrick-Lins 1986). High kappa values ($\gg 0$) indicate a greater than chance agreement between the model and the survey data. Kappa values close to zero suggest only chance agreement between the model and survey data. Negative kappa

values occur when there is less than chance agreement between the model and the field data (i.e., the model predictions are opposite of the field data).

The weighted kappa (Cohen 1968, Fleiss et al. 1969, Rosenfield and Fitzpatrick-Lins 1986), a derivation of the standard kappa, can be used when the "seriousness" of each model outcome (i.e., commission error, omission error, correct-occurrence prediction, or correct-absence prediction) is not the same.

Due to the typically high number of commission errors for BHR models (Krohn 1996), kappa estimates are generally quite low and occasionally negative. For the purposes of investigating model performance under varied conditions, we assigned weightings of 0.5 to commission errors and 1.0 to all other model outcomes (i.e., correct-presence, correct-absence, and omission error). Commission errors were given a lower weighting since not all of measured commissions are true errors (i.e., some species not observed, but predicted, will in fact occur in the study area). In most cases, this gave positive values for kappa.

Cohen (1960; see also Hudson and Ramm [1987]) provided the formula for calculating the approximate large sample variance of a kappa statistic and a test of significance between two kappa values using a Z test. For each species, we calculated Z for each model combination using the null hypothesis of

$$H_0: \hat{K}_m = 0 \qquad (1)$$

where $\hat{K}$ is the estimated weighted kappa, and $m = 1$, 2, . . . , 18 for each species. Under this null hypothesis, our Z test reduced to

$$Z = (\hat{K}_1)/[V(\hat{K}_1)]^{1/2} \qquad (2)$$

Because the estimates of kappa and the variance of kappa for the site level were calculated from 440 points on Craig Mountain and 1628 points on Region 1 areas, we considered them reliable measures. However, because we recognized only five and nine cover types for the Craig Mountain and Region 1 areas, respectively, estimates of kappa and the variance of kappa at those levels of analysis were unreliable. Thus, we calculated Z scores for only the site level of analysis.

Analysis of model performance

To analyze the effect of model complexity, data resolution, and level of analysis on model performance, we used a completely randomized factorial design with three levels of model complexity and three levels of data resolution as factors for the site level of analysis. Steinhorst (1979) described the use of factorial design and the analysis of variance tools to investigate the sensitivity of predictive models to changes in model parameters. To test for differences between models, we used the method of van Belle and Hughes (1984), in which sums of squares from the Z scores are partitioned using the completely randomized factorial design. We interpreted significant interactions with graphs of the

mean value of each model combination for one factor plotted against another.

We were also interested in whether the changes in model performance were due to change in correct-presence or correct-absence predictions. To determine this, we conducted a canonical variate analysis within our completely randomized factorial design with correct-presence and correct-absence as response variables (see Johnson and Wichern 1992). This analysis was conducted at the site and cover-type levels for both Craig Mountain and Region 1.

Because reliable Z scores could not be calculated for the cover-type level of analysis, we could not use the previously described methods to compare site and cover-type levels of analysis. To examine the effects of level of analysis on model performance, we compared kappa values from the site and cover-type levels. We also examined interaction graphs of kappa values at the cover-type level to look for patterns of model performance that were similar to the site level.

RESULTS

The chi-square analysis of Z scores for the site level of analysis indicated a significant interaction between model complexity and spatial data resolution for both Craig Mountain and Region 1 ($\alpha = 0.05$). This interaction precluded simple interpretation of the main effects of model complexity and spatial data resolution, but graphs of the two-way interaction provided insight as to the effects of these factors.

For both the Craig Mountain and Region 1 areas, we observed that fine spatial data resolutions (e.g., 0.09 ha) provided the highest weighted-kappa estimates as model complexity increased. For Craig Mountain, the interaction graph of model complexity and spatial data resolution at the site level indicated that changes in spatial data resolution at low levels of model complexity did not result in large changes in the performance of the BHR model (Fig. 3a). However, as model complexity increased, higher resolution data (i.e., 0.09 ha) yielded better results than the coarse resolution (i.e., 4 ha and 10 ha) spatial data ($\hat{K}_w = 0.282, 0.157,$ 0.160 for 0.09 ha, 4 ha, and 10 ha, respectively). Each spatial data resolution had its highest estimated kappa at the second level of model complexity. The canonical variate analysis for the model complexity × spatial data resolution interaction indicates that the significance of this interaction is driven both by changes in correct-presence and correct-absence predictions (Table 5).

For Region 1, the interaction graph of model complexity and spatial data resolution at the site level indicates that increasing model complexity resulted in higher estimated kappa values at all levels of spatial data resolution (Fig. 3b). At the highest level of model complexity, the 0.09-ha resolution gave the best results ($\hat{K}_w = 0.206$). For the 4-ha and 10-ha resolutions, the highest estimates of kappa were from the second level of model complexity ($\hat{K}_w = 0.199$ for both 4 ha and

10 ha). The canonical variate analysis for the (model complexity) × (spatial data resolution) interaction indicated that the significance of this interaction was driven both by changes in correct-presence and correct-absence predictions (Table 5).

The cover-type level of analysis had higher estimates of weighted kappas than the site level for all levels of model complexity on both study areas, although the kappa estimates tended to decrease with increasing model complexity for both levels of analysis. We used graphs of weighted-kappa estimates to examine the effect of level of analysis on the performance of BHR models (Fig. 3c, d). Each model complexity had higher estimates of kappa at the cover-type level across all model complexities for Craig Mountain ($\hat{K}_w = 0.458,$ 0.448, 0.415 for 0.09-ha, 4-ha, and 10-ha resolutions, respectively) and Region 1 ($\hat{K}_w = 0.540, 0.530, 0.460$ for 0.09-ha, 4-ha, and 10-ha resolutions, respectively). The canonical variate analysis showed that this difference was due to changes in correct-absence predictions across each area (Table 5).

We observed different effects of level of analysis for each spatial data resolution between study areas. For Craig Mountain, the graph of weighted kappa estimated by level of analysis for each spatial data resolution indicated little difference in model performance between the site- and cover-type levels of analysis at the 4-ha and 10-ha resolutions (Fig. 3e). At the 0.09-ha resolution, however, there was a large difference between model performance at the site and cover-type levels of analysis ($\hat{K}_w = 0.295$ and 0.866 for site- and cover-type levels, respectively). Results from canonical variate analysis suggests that this difference is due largely to changes in correct-absence predictions (Table 5).

For Region 1, however, the two-way graph of weighted-kappa estimates by level of analysis for each spatial data resolution indicated higher model performance at the cover-type level of analysis at all spatial data resolutions (Fig. 3f) ($\hat{K}_w = 0.531, 0.509,$ and 0.511 for 0.09-ha, 4-ha, and 10-ha resolutions, respectively). Results from canonical variate analysis indicated that this difference was due to changes in both model correct-presence and correct-absence predictions (Table 5).

Overall, for the Craig Mountain area, the BHR models yielded 20–100% correct-presence predictions across all model complexities, spatial data resolutions, and levels of analysis, and 0–100% correct-absence predictions. Weighted estimates of agreement ranged from $\hat{K}_w = 0.27$ to $\hat{K}_w = 0.58$. The overall means for predictions and agreement estimates on the Craig Mountain area were 96.4 ± 9.05% for correct-presence (mean ± 1 SD), 68.6 ± 25.05% for correct-absence, and $\hat{K}_w = 0.36$, respectively (Table 6). For each species we averaged the correct-presence and correct-absence measurement across all 18 model tests and plotted them against the total number of sites where that species was detected (Fig. 4). For Craig Mountain, we found cor-
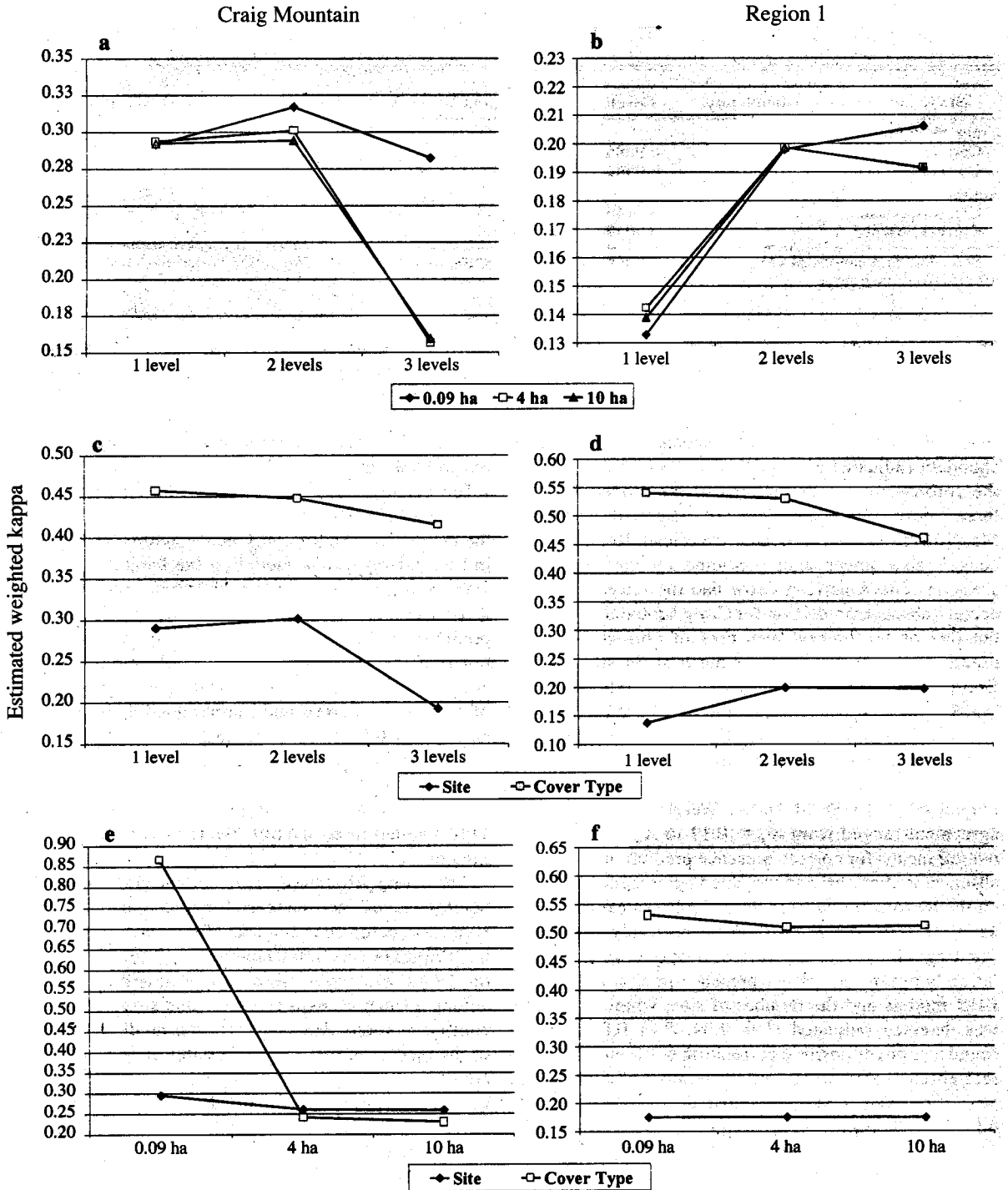
FIG. 3. Graphs of two-factor interactions for the Craig Mountain and Region 1 analysis. The graph of model complexity by spatial data resolution for the Craig Mountain area (a) indicates that, at low levels of model complexity, similar results are obtained from any of the three spatial data resolutions. However, at higher levels of model complexity, better results come at the 0.09-ha data resolution. (b) Two-way interaction graph for model complexity by spatial data resolution for Region 1 indicates that model performance generally increases as model complexity increases. At the highest level of model complexity, the 0.09-ha data resolution performs best. The cover-type level of analysis performs better at all levels of model complexity in (c) Craig Mountain and (d) Region 1. (e) A graph of spatial data resolution by level of analysis for Craig Mountain indicates that model performance is affected greatly by the level of analysis at the 0.09-ha resolution. At the 4-ha and 10-ha resolutions, there is little difference between the site and cover-type levels of analysis. (f) The Region 1 graph of spatial data resolution by level of analysis indicates higher model performance from the cover-type level of analysis at all levels of spatial data resolution.

TABLE 5. Within-design canonical structure for the first canonical variate for the BHR model performance from the Craig Mountain and Region 1 areas.

Study area	Complexity	Level	Resolution	Complexity by level	Complexity by resolution	Level by resolution
Craig Mountain						
Correct-presence	0.9574	0.9086	0.5474	−0.1327	0.9615	−0.0194
Correct-absence	−0.6190	0.0562	0.5785	0.9710	−0.6076	0.9376
Region 1						
Correct-presence	0.8057	0.2149	−0.2255	0.0386	0.9650	0.7559
Correct-absence	−0.7575	0.9039	1.0000	0.9653	−0.4714	−0.8071

Note: The first canonical variate for each main effect and interaction accounted for >97% of the variation between treatment groups from each area.

relations between number of species detections and BHR model performance. Correct-presence predictions of the models tended to decrease for the more abundant species (adjusted $r^2 = 0.42$, $P < 0.0001$). Correct-absence predictions increased as species became more abundant (adjusted $r^2 = 0.35$, $P < 0.0001$). However, the correct-absence predictions of the models for the three most abundant species (Spotted Towhee, MacGillivray's Warbler, and American Robin) were considerably lower than predicted by the linear regression. The American Crow had the lowest average correct-absence prediction for Craig Mountain (38.2%) but one of the highest measures of correct-presence predictions (99.3%). Most of the $n = 34$ species for Craig Mountain had correct-presence predictions >90% and correct-absence predictions >50% (Fig. 4a).

For the Region 1 area, the BHR models yielded 56–100% correct-presence predictions across all possible model combinations. Correct-absence predictions ranged from 0.17% to 100%. Weighted estimates of agreement ranged from $\hat{K}_w = 0.17$ to $\hat{K}_w = 0.52$. The overall means for correct-presence predictions, correct-absence predictions, and weighted agreement estimates on the Region 1 area were 96 ± 7.34% (mean ± 1 SD), 65 ± 27.34%, and $\hat{K}_w = 0.35$, respectively (Table 6). For Region 1 (Fig. 4b) we found only a weak correlation between the correct-presence predictions of the BHR models and the number of sites where a species was detected (adjusted $r^2 = 0.09$, $P = 0.0360$). We found a stronger positive correlation between BHR correct-absence predictions and the number of sites where a species was detected (adjusted $r^2 = 0.37$, $P < 0.0001$). All species on the Region 1 area with average correct-absence predictions >70% were detected at >70 survey sites. Three Region 1 species had average correct-absence predictions < 50%: Evening Grosbeak (48%), Hairy Woodpecker (46%), and Ruffed Grouse (44%). Each of these species had average correct-presence predictions of >95%.

Species that have larger home ranges (>10 ha), such as the American Crow, Evening Grosbeak, Northern Flicker, or Gray Jay, had low correct-absence predictions for both Craig Mountain and Region 1. Also, species with somewhat specific habitat requirements and smaller home ranges, such as the Brown Creeper,

Cedar Waxwing, or Canyon Wren, had average correct-absence predictions <70%.

DISCUSSION

The differences in effect of model complexity between study areas may be due to differences in landscape heterogeneity between the two. The vegetation of Region 1 was dominated by coniferous forests, and the majority of the Region 1 survey points were located in that habitat. When modeling the habitat on Region 1 for a species associated with coniferous forest, selection of that cover type in the GIS would result in prediction of large, continuous areas as suitable habitat. For rare species, this could result in a model that overpredicted suitable habitat, resulting in a large number of commission errors, and a corresponding low weighted-kappa estimate. By increasing the complexity of the model, thus reducing the area predicted as habitat, we decreased the commission errors at a greater rate than the rate at which the omission errors were increasing. This resulted in an overall rise in weighted-kappa estimates.

The Craig Mountain area, conversely, had a very heterogeneous landscape codominated by dry grasslands and coniferous forests. There was also a significant riparian vegetation component. The survey points on Craig Mountain were proportionally distributed among all major cover types. In this area, selection of coniferous forest types resulted in a smaller proportion of the landscape predicted as habitat. Increasing model complexity raised the rate of omission errors faster than the rate of commission errors decreased. This resulted in an overall decrease in weighted-kappa estimates as model complexity increased.

By adding layers of complexity to a BHR model, we attempted to reduce the amount of commission error for that model. By doing so, we increased the ability of that model to correctly reject unsuitable habitat. Theoretically this increased the ability of the model to correctly predict species absence. At the same time, however, it was difficult to maintain the desired level of correct-presence predictions. Block et al. (1994) described a trade-off between omission and commission errors. They reported that by varying BHR model parameters, they increased the omission error rate and

TABLE 6. BHR (bird habitat relationship) model performance measures for each of the independent variables for the Craig Mountain and Region 1 areas.

Independent variable	Craig Mountain						Region 1					
	Correct-presence (%)		Correct-absence (%)		Kappa		Correct-presence (%)		Correct-absence (%)		Kappa	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Model complexity												
1 layer	98.11	4.1	64.15	24.4	0.37	0.26	98.78	2.7	54.21	29.8	0.34	0.25
2 layers	97.83	4.3	67.08	24.6	0.38	0.26	96.23	5.9	65.54	25.1	0.37	0.24
3 layers	91.2	16.2	78.46	24.2	0.30	0.29	90.39	10.9	80.0	17.2	0.33	0.24
Spatial data resolution												
0.09 ha	97.95	4.5	74.5	23.9	0.58	0.34	95.84	7.2	64.01	27.9	0.35	0.25
4 ha	95.93	9.6	65.59	24.7	0.25	0.13	95.85	7.4	64.79	27.1	0.34	0.24
10 ha	95.39	9.6	65.55	25.5	0.24	0.12	95.82	7.4	64.93	27.2	0.34	0.24
Level of analysis												
Site	94.52	7.6	69.23	24.0	0.27	0.14	95.11	6.2	53.73	32.3	0.17	0.09
Cover type	98.32	10.0	67.87	26.1	0.44	0.33	96.57	8.3	75.42	14.8	0.52	0.23

decreased the commission error rate and vice versa. Our results support their findings to some degree, but we also observed that commission and omission error did not change at the same rate. Increasing model complexity resulted in a much greater decrease in commission errors than increase in omission errors, at least for Region 1, resulting in an overall increase in model performance.

The effect of spatial data resolution on the performance of BHR models was dependent on the other parameters of the model. The reason for this was again related to the difference in heterogeneity between the Craig Mountain and Region 1 landscapes. When the heterogeneous landscape of Craig Mountain was resampled to a coarser spatial data resolution, areas with high degrees of complexity that were smaller than the resampling unit were changed into the cover type dominant in that cell. Thus, resampling had a simplifying effect on heterogeneous landscapes. This could result in higher omission error rates due to small habitat patches being lost and higher commission error rates due to overrepresentation of major cover types. Because the Region 1 landscape was relatively homogeneous, large blocks of coniferous forests did not change shape or size much when resampled to a coarser unit (e.g., 4 ha or 10 ha). Therefore, the effect of spatial data resolution on the performance of BHR models was related to the heterogeneity of the landscape to which the models were applied.

The cover-type level of analysis produced higher weighted-kappa estimates than the site level in all applications except with the 4-ha and 10-ha spatial data resolutions on Craig Mountain. This again was likely due to the loss of detail from the heterogeneous Craig Mountain landscape when it was resampled to 4 ha and again to 10 ha.

We currently lack critical information on the ability of habitat-relationship models to predict the habitats of uncommon, wide-ranging species. Because we limited our species list to those with ≥15 sites of observation,

we were not considering the most uncommon species within each area. The high correct-presence predictions (mostly >90%) and low correct-absence predictions (mostly <70%) for species with <50 sites of observation may be a reflection of inadequate data for good estimates of model performance. In this case, the apparent error component of observed commission error may be contributing to the large difference between mean correct-presence and correct-absence estimates. It may take a large number of observations to achieve reliable estimates of model performance (Fig. 4). The number of points necessary is likely dependent on the size of the area modeled. Obtaining reliable estimates for less-common species is challenging because of the time and cost associated with collecting sufficient sample sizes. However, less-common species could potentially benefit the most from valid habitat-relationship models. Therefore, the expense of validating models for these species may be warranted.

Block et al. (1994) recommended at least 5 yr of field sampling to detect 98% of the species that would occur in an area and to gain reliable estimates of model performance. Krohn (1996) took a more conservative view, recommending 10 yr of surveying to ensure that most rare and incidental species are detected. Because of our large number of survey points, we feel confident that we detected most of the species occurring within our study areas after 3 yr of breeding bird surveys. The number of new species detected on Region 1 was <5% after 1 yr (P. J. Heglund, *unpublished manuscript*). However, we are highly confident in our measures of BHR model performance for only the most abundant species detected.

Theory of testing model performance

Bird species most likely use habitat in a continuous fashion (Wiens 1989), not as discretely as suggested by most habitat models. However, when predicting the presence or absence of a species at a given geographical location, there are only two possible types of measur-
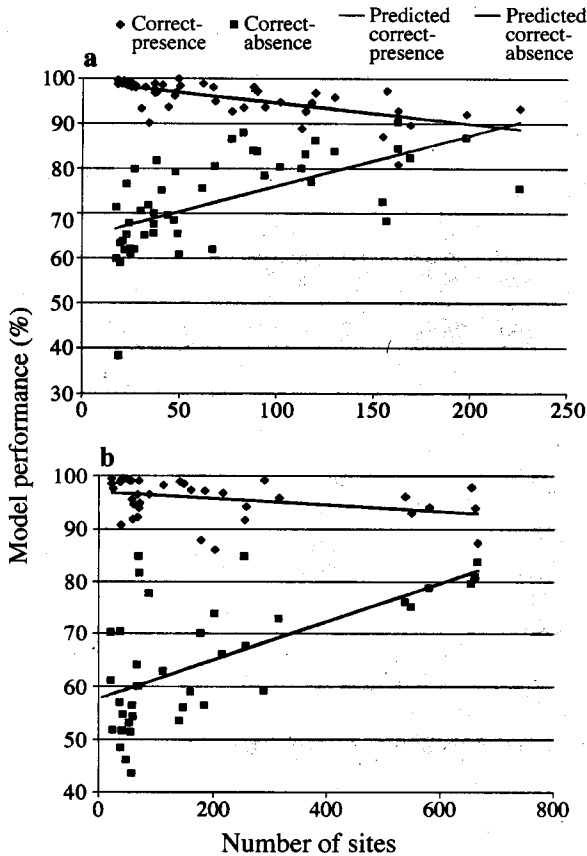
FIG. 4. Average model performance for bird species plotted against the number of sites where each species was detected. (a) For Craig Mountain, the least-squares regression lines indicate a decrease in correct-presence predictions as species became more common ($r^2 = -0.42$, $P = <0.0001$) and an increase in model correct-absence predictions as species became more common ($r^2 = 0.35$, $P = <0.0001$). (b) For Region 1, the least-squares regression lines suggest that correct-presence predictions do not change much as species become more abundant but that correct-absence predictions increase. This would lead to an overall increase in model performance for the more abundant species.

able model error, omission and commission error. A graphical representation of these different types of error rates is helpful (Fig. 5). Habitat selection for most species occurs nonlinearly along an environmental gradient (Whittaker 1967, Wiens 1989). The species is most abundant at some point along the continuum and decreases in abundance away from that point in either direction. For simplicity, we provide only one side of the habitat-selection curve. By modeling species occurrence, the modeler attempts to simulate the habitat selection of a species (represented as a curve) with discrete modeling methods (represented by the shaded box). With our method of modeling, which recognized only suitable or unsuitable habitat, a certain amount of error is inevitable. The modeler must decide the impact of omission and commission error to the model outcome and adjust the model accordingly. Difficulty in

measuring commission errors stems from differences between the actual habitat-use function for a species and that observed during field surveys. The difference between these two functions constitutes the apparent error component of commission error estimates.

Observations of a species give direct insight to the performance of a habitat-relationship model. If a species was observed in an area where it was predicted to occur, then the model performed correctly. Because these occurrence predictions represented the probability of correctly failing to reject our null hypothesis of habitat suitability, they contributed directly to the accuracy of the model. However, an omission error, observing a species where it was not predicted, is a definite model error, as long as the model was not used outside of its intended limits (season of use or geographic range) (Dedon et al. 1986). Because omissions represent the probability of rejecting a true null hypothesis (Type I error), they detracted directly from the accuracy of the model.

Theoretically, predictions of absence represent the probability of correctly rejecting a false null hypothesis of equal habitat suitability. However, measured commission errors and predictions of species absence are uncertain indicators of model performance because they consist of two components: true error in that the species does occur and should not be modeled, and apparent error in that the species does occur but was missed in the field surveys (Dedon et al. 1986, Scott
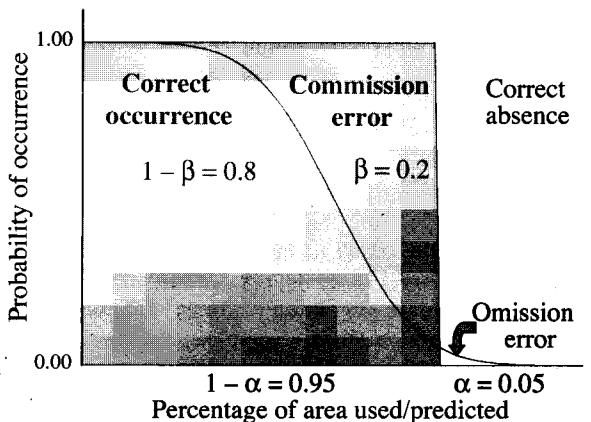

FIG. 5. A graphical representation of the theory of habitat-relationship model testing. Habitat selection in most species occurs along a continuum, represented in the graph by the curve. This method of modeling seeks to approximate this continuous event by discrete means, represented by the shaded portion of the graph. Omission errors (species is detected where not predicted) occur when the model does not include enough habitat types for a species. Commission errors (species is not detected where predicted) occur because of overprediction of suitable habitat. Overestimation of commission error rates occurs when the habitat use function is approximated incorrectly from field data. When using a discrete modeling method to simulate a continuous phenomenon, errors are unavoidable, and thus model users must make a priori decisions as to what error types are most critical to avoid.

et al. 1993, Block et al. 1994, Edwards et al. 1996, Krohn 1996). The latter component is not measurable in most field applications. Hence, measurements of commission error are usually an overestimate of true model error. A species may not have been detected because it did not occur or because it was missed due to inadequate sampling (Nichols et al. 1998). The inability to detect a species due to poor survey design, inadequate survey duration (too few years of surveying), or spatial and temporal variation in species distribution within its habitat can inflate the estimated commission error rate (Scott et al. 1993, Cassidy et al. 1994, Krohn 1996). Species that are rare, or have low population densities or low detection rates, may also inflate commission error rates (Krohn 1996, Boone and Krohn 1999). Block et al. (1994) and Krohn (1996) observed that commission errors tended to decrease as duration of the surveys increased. They both expected the rate of commission errors to asymptotically decrease toward the true model commission error. Our results (Fig. 4) support this.

Boone and Krohn (1999) reported that the likelihood of correctly predicting the presence of avian species could be estimated based on certain species attributes such as niche width, relative abundance, and home range size. Under this hypothesis, species that were common, occurred in a wide range of habitat types, and occupied small home ranges would be most likely to be modeled accurately (high correct-presence and correct-absence predictions). Our results for Craig Mountain and Region 1 support this. Our models from both study areas show that more-common species tend to have higher correct-absence predictions than the less-common ones (Fig. 4). However, the pattern we observed may be an artifact of inadequate sample sizes to reliably estimate accuracy of uncommon species.

Ideally, the investigator would select a minimum level of correct-presence predictions for the model (i.e., maximum omission error rate). This could be determined, in part, by the risks associated with making an omission (Type I; Dedon et al. 1986) vs. commission (Type II) error. Avery and VanRiper (1990) considered omission errors more serious modeling errors than commission errors. Dedon et al. (1986) concluded that extensive commission errors could also be the result of modelers trying to avoid omission errors. For example, if a manager were interested in determining the potential impacts of a management decision on a sensitive species, they would demand higher correct-presence predictions (lower omission error) than if they were setting up a sampling scheme to search for nest sites. The correct-absence predictions of the model could then be brought up to the desired level by restriction of model parameters and increased species detections. The critical assumption in this case is that the habitats that are used very infrequently (and thus become omission errors) are not essential habitats.

Many evaluations of habitat-relationship model performance (e.g., Block et al. 1994) have not considered the case of their model correctly predicting species absence. Given the purpose of habitat modeling in many cases is to exclude areas of unsuitable habitat from a greater region where the species is known to occur (Scott et al. 1993, Butterfield et al. 1994, Csuti 1996), this statistic should be of considerable interest. However, this statistic is difficult to measure for the same reasons that true commission errors are difficult to measure.

The potential exists that suitable habitat exists for a species but is not occupied at the time of survey (Fretwell and Lucas 1970). This situation would result in underestimation of correct-absence measures (overestimating commission error). Krohn (1996) believed that more complex models, incorporating multiple data layers, should be less sensitive to dynamic habitat use. This is because "if enough data are available to justify a presumably more complex habitat model, the density effects on habitat occupancy should be better understood than the simple species–vegetation associations" (Krohn 1996:150).

Conclusion

Overall, our results suggest that model complexity, spatial data resolution, and level of analysis play integral parts in determining the performance of habitat-relationship models. Changes in the ability of the model to predict species absence accounted for most of the difference in habitat-relationship model performance in our applications. Our results demonstrate that:

1) Species abundance affected the ability to test habitat-relationship models. The number of observations necessary to achieve reliable estimates of error likely depends on scale of model application.

2) Increasing model complexity improved accuracy on relatively homogeneous landscapes, whereas fine spatial data resolution improved model performance on heterogeneous landscapes.

3) Increasing model complexity resulted in a greater decrease in commission error than increase in omission error. This increased model accuracy overall.

4) The cover-type level of analysis performed better than the site level across all model complexities on each study area, suggesting that this method of modeling might perform better at coarser scales.

Survey requirements for testing habitat-relationship models should be based on a minimum number of detections for a species and not a specified number of years. Given proper study design, it should be possible to obtain sufficient data to test habitat-relationship models for certain abundant species with one season of surveys. For other less abundant species, it may take many years of very specialized surveys to gather enough observations to make reliable estimates of model performance. For rare species, obtaining sufficient detections could be prohibitively expensive. In this

case, adjustment of error rates based on species attributes (see Boone and Krohn 1999) may be necessary.

Ultimately, the level of model correct-presence predictions and correct-absence predictions that is acceptable will be determined by the application of the model. In some instances, the modeler may wish to sacrifice some correct-absence predictions and slightly overpredict the distribution of a species. For example, when examining the potential effects of a management decision on a threatened or endangered species, the manager may want to overpredict the distribution of the species, rather than risk jeopardizing essential habitat. Conversely, some situations may necessitate high correct-absence predictions at the cost of decreased model correct-presence predictions. Determining areas to sample for bird nests to characterize nesting habitat might be an example of this given that underestimating available habitat would not have dire consequences.

LITERATURE CITED

Avery, M. L., and C. VanRiper III. 1990. Evaluation of wildlife-habitat relationships database for predicting bird community composition in central California chaparral and blue oak woodlands. California Fish and Game 76:103–117.

Block, W. M., Morrison, M. L., Verner, J., and P. N. Manley. 1994. Assessing wildlife-habitat-relationships models: a case study with California oak woodlands. Wildlife Society Bulletin 22:549–561.

Boone, R. B., and W. B. Krohn. 1999. Modeling the occurrence of bird species: are the errors predictable? Ecological Applications 9:835–848.

Butterfield, B. R., Csuti, B., and J. M. Scott. 1994. Modeling vertebrate distributions for gap analysis. Pages 53–68 in R. L. Miller, editor. Mapping the diversity of nature. Chapman & Hall, London, UK.

Cassidy, K., E. O. Garton, W. B. Krohn, L. S. Mills, J. M. Scott, K. Williams, and B. Csuti. 1994. Assessing the predictive ability of gap analysis vertebrate distributions. Gap analysis handbook. Idaho Cooperative Fish and Wildlife Research Unit, Moscow, Idaho, USA.

Cassirer, E. F. 1995. Wildlife inventory, Craig Mountain, Idaho. U.S. Department of Energy Bonneville Power Administration. DOE/BP-62547-2.

Cohen, J. 1960. A coefficient of agreement for nominal scales. Educational and psychological measurement 20:37–46.

Cohen, J. 1968. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychological Bulletin 70:213–220.

Congalton, R. G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of the Environment 37:35–46.

Csuti, B. 1996. Mapping animal distribution areas for gap analysis. Pages 135–146 in J. M. Scott, T. H. Tear, and F. W. Davis, editors. Gap analysis: a landscape approach to biodiversity planning. American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland, USA.

Csuti, B., A. J. Kimerling, T. A. O'Neil, M. M. Shaughnessy, E. P. Gaines, and M. M. P. Huso. 1997. Atlas of Oregon wildlife: distribution, habitat and natural history. Oregon State University Press, Corvallis, Oregon, USA.

Davis, F. W. 1996. The nature of gap analysis. BioScience 46:74–75.

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 J. R. Ralph, J. R. Sauer, and S. Droege, editors. 1995. Monitoring bird populations by point counts. U.S. Forest Service General Technical Report PSW-GTP-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 115–119 in J. Verner, M. L. Morrison, and C. J. Ralph. Wildlife 2000. Modeling habitat relationships of terrestrial vertebrates. University of Wisconsin Press, Madison, Wisconsin, USA.

Edwards, T. C. 1996. Data defensibility and gap analysis. BioScience 46:75–76.

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.

Fleiss, J. L., J. Cohen, and B. S. Everitt. 1969. Large sample standard errors of kappa and weighted kappa. Psychological Bulletin 72:323–327.

Fretwell, S. D., and H. L. Lucas, Jr. 1970. On territorial behavior and other factors influencing habitat distribution in birds. I. Theoretical development. Acta Biotheoretica 19:16–36.

Groves, C. R., R. Butterfield, A. Lippincott, B. Csuti, and J. M. Scott. 1997. Atlas of Idaho's wildlife: integrating gap analysis and natural heritage information. Nongame and Endangered Wildlife Program, Idaho Department of Fish and Game, Boise, Idaho, USA.

Hollander, A. D., F. W. Davis, and D. M. Stoms. 1994. Hierarchical representations of species distributions using maps, images and sighting data. Pages 71–88 in R. I. Miller, editor. Mapping the diversity of nature. Chapman & Hall, London, UK.

Hudson, W. D., and C. W. Ramm. 1987. Correct formulation of the kappa coefficient of agreement. Photogrammetric Engineering and Remote Sensing 53:421–422.

Hutto, R. L., and J. Hoffland. 1996. USDA forest service northern region landbird monitoring project: field methods. U.S. Forest Service Report, Region 1, Division of Biological Sciences, University of Montana, Missoula, Montana, USA.

Johnson, D. H., and J. W. Grier. 1988. Determinants of breeding distributions of ducks. Wildlife Monographs Number 100:1–37.

Johnson, R. A., and D. W. Wichern. 1992. Applied multivariate statistical analysis. Third edition. Prentice-Hall, Upper Saddle River, New Jersey, USA.

Krohn, W. B. 1996. Predicted vertebrate distributions from gap analysis: considerations in the designs of statewide accuracy assessments. Pages 147–162 in J. M. Scott, T. H. Tear, and F. W. Davis, editors. Gap analysis: a landscape approach to biodiversity planning. American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland, USA.

Merrill, E. H., T. W. Kohley, M. E. Herdendorf, W. A. Reiners, K. L. Driese, R. W. Marrs, and S. H. Anderson. 1996. Wyoming gap analysis: a geographic analysis of biodiversity. Final Report. Wyoming Cooperative Fish and Wildlife Unit, University of Wyoming, Laramie, Wyoming, USA.

NISC (National Information Services Corporation). 1997a.

Essential ecology, zoology and plant science abstracts [CD-ROM]. Baltimore, Maryland. Literature search database.

NISC. 1997b. Essential ornithological abstracts [CD-ROM]. Baltimore, Maryland. Literature search database.

NISC. 1997c. Essential wildlife and conservation biology abstracts [CD-ROM]. Baltimore, Maryland. Literature search database.

NISC. 1997d. Wildlife worldwide [CD-ROM]. Baltimore, Maryland. Literature search database.

Nichols, J. D., T. Boulinier, J. E. Hines, K. H. Pollock, and J. R. Sauer. 1998. Inference methods for spatial variation in species richness and community composition when not all species are detected. Conservation Biology 12:1390–1398.

Poole, A., P. Stettenheim, and F. Gills, editors. 1997. The birds of North America. American Ornithologists' Union. Washington, D.C., USA.

Ralph, J. R., J. R. Sauer, and S. Droege, editors. 1995. Monitoring bird populations by point counts. U.S. Forest Service General Technical Report PSW-GTP-149.

Ralph, J. R., and J. M. Scott, editors. 1981. Estimating numbers of terrestrial birds. Studies in Avian Biology No. 6. Allen Press, Lawrence, Kansas, USA.

Redmond, R. L., Z. Ma, T. P. Tady, J. C. Winne, J. Schumacher, J. Troutwine, and S. W. Holloway. 1996. Mapping existing vegetation and land cover across western Montana and northern Idaho. Wildlife Spatial Analysis Laboratory, Missoula, Montana, USA.

Reynolds, R. T., J. M. Scott, and R. Nussbaum. 1980. A variable circular plot for censusing bird numbers. Condor 82:309–313.

Rosenfield, G. H., and K. Fitzpatrick-Lins. 1986. A coefficient of agreement as a measure of thematic classification accuracy. Photogrammetric Engineering and Remote Sensing 52:223–227.

Scott, J. M., F. Davis, B. Csuti, R. Noss, B. Butterfield, C. Groves, H. Anderson, S. Caicco, F. D'erchia, T. C. Edwards, J. Ulliman, and G. Wright. 1993. Gap analysis: a geographic approach to protection of biological diversity. Wildlife Monographs Number 123.

Scott, J. M., M. Jennings, R. G. Wright, and B. Csuti. 1996. Landscape approaches to mapping biodiversity. BioScience 46:77–78.

Scott, J. M., F. L. Ramsey, and C. B. Kepler. 1981. Distance estimation as a variable in estimating bird numbers. Pages 334–340 in C. J. Ralph and J. M. Scott, editors. Estimating

numbers of terrestrial birds: studies in avian biology. Number six. Cooper Ornithological Society, Lawrence, Kansas, USA.

Short, H. L., and J. B. Hestbeck. 1995. National biotic resource inventories and gap analysis. BioScience 45:535–539.

Short, H. L., and S. C. Williamson. 1986. Evaluating the structure of habitat for wildlife. Pages 97–104 in J. Verner, M. L. Morrison, and C. J. Ralph editors. Wildlife 2000: modeling habitat relationships of terrestrial vertebrates. University of Wisconsin Press, Madison, Wisconsin, USA.

Smith, K. G., and D. G. Catanzaro. 1996. Predicting vertebrate distributions for gap analysis: potential problems in constructing the models. Pages 163–169 in J. M. Scott, T. H. Tear, and F. W. Davis, editors. Gap analysis: a landscape approach to biodiversity planning. American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland, USA.

Smith, M. R., P. W. Mattocks, and K. M. Cassidy. 1997. Breeding birds of Washington State. Volume 4 in K. M. Cassidy, C. E. Grue, M. R. Smith, and K. M. Dvornich, editors. Washington State gap analysis - Final Report. Seattle Audubon Society Publications in Zoology Number 1, Seattle, Washington, USA.

Steinhorst, R. K. 1979. Parameter identifiability, validation, and sensitivity analysis of large system models. Pages 33–58 in G. S. Innis and R. V. O'Neil, editors. Systems analysis of ecosystems. International Cooperative Publishing House, Fairland, Maryland, USA.

Stephens, D. A., and S. H. Sturts. 1998. Idaho bird distribution. Idaho Museum of Natural History, Pocatello, Idaho. Special Publication Number 13.

Stoms, D. M. 1992. Effects of habitat map generalization in biodiversity assessment. Photogrammetric Engineering and Remote Sensing 58:1587–1591.

van Belle, G., and J. P. Hughes. 1984. Nonparametric tests for trend in water quality. Water Resources Research 20:127–136.

Whittaker, R. H. 1967. Gradient analysis of vegetation. Biological Review of the Cambridge Philosophical Society 49:207–264.

Wiens, J. A. 1989. The ecology of bird communities. Cambridge University Press, Cambridge, UK.

Wright, N. 1997. Wildlife habitat relationship models for terrestrial vertebrate species on Craig Mountain, Idaho. Idaho Cooperative Fish and Wildlife Research Unit, University of Idaho, Moscow, Idaho, USA.