Use of Regional-scale Exploratory Studies to Determine Bird-habitat Relationships

Jock S. Young and Richard L. Hutto

The wide geographic extent of regional bird monitoring programs usually makes them nonexperimental and exploratory in nature. In such studies, variables are often chosen by expedience, and sources of variation are not controlled. Even so, there is great potential to learn something meaningful about bird-habitat relationships when bird distribution or abundance information is linked with additional information about vegetation characteristics at the sites (Wiens and Rotenberry 1981a; Ralph et al. 1995a).

One goal of bird-habitat relationship studies is to identify environmental conditions that presumably control the distribution and abundance of a bird species. Knowledge of the biologically important variables would help us make more informed management decisions and more accurate predictions of bird occurrence in new, unsurveyed sites. Because we cannot measure all biologically important variables, however, the resulting models are heavily influenced by the choice of variables and the methods used for exploratory analysis. Consequently, there is a critical need to discuss the best approaches for getting the most out of such observational data sets.

In this chapter, we present analyses of data from the U.S. Forest Service (USFS) Northern Region Landbird Monitoring Program (Hutto and Young 1999). Full-scale monitoring began under this program in 1994. One aim was to conduct long-term monitoring, so a series of permanently marked transects were located randomly within a geographic stratification scheme. This is not the most efficient design for discerning habitat associations, because a preponderance of points lie in common cover types (Austin and Meyers 1996; Heglund, Chapter 1; Austin, Chapter 5). Nonetheless, we collected data on local vegetation characteristics in the area surrounding each point to obtain as much basic habitat-relationship information as possible to supplement the population monitoring data and to help explain the possible reasons for any population declines that may be detected later.

The distributions of various landbird species across cover types in the northern Rocky Mountains have already been published (Hutto and Young 1999). However, the eighteen cover types used in those analyses pooled together a diverse assemblage of vegetation structures. No bird species was detected at all points within any given cover type. Some of those absences may have been due to sampling error, population fluctuations, or chance, but we assume that there were also absences due to finer-resolution variation in habitat characteristics among the points within a single cover type. We need to know what these features are if we are to manage bird species successfully.

In this chapter, we discuss methods for this second step: building regression models to expose finer-resolution patterns of occurrence due to the continuously variable nature of vegetation features. We apply
our proposed model-building methods to a single example species, the Swainson’s thrush (Catharus ustulatus). We then examine several subsets of the data to determine the consistency and accuracy of our results under different conditions. Even though this is an exploratory exercise, and even though we make use of uncontrolled data not designed primarily to expose habitat relationships, the resulting descriptive models should at least suggest ecological relationships and help focus future research.

Methods

Sample points were distributed across twelve national forests of the USFS Northern Region, in northern Idaho and western Montana. Transects were geographically stratified by 7.5-minute topographic quadrangle maps throughout nonwilderness Forest Service lands and part of the Potlatch Corporation lands in central Idaho (Fig. 8.1). Potential transect start points were located by positioning a random point within each quarter of the quadrangle maps and then finding the nearest point on an unpaved, secondary or tertiary, open or closed road or trail (Hutto and Young 1999). Transects were selected randomly from this potential set, subject to logistical constraints. The ten points along each transect were placed at least 250 meters (straight-line distance) apart. Each point was sampled once during the breeding season in each of three consecutive years (1994 to 1996). There were some changes in transect locations between years. In this chapter, we used the 428 transects that were visited in all three years, although point selection criteria discussed below resulted in only 292 transects being represented in the data set used for analyses.

Bird Survey Methods

Our field technique followed recommendations discussed by Ralph et al. (1995a) and methods described by Hutto et al. (1986). All observers participated in a one-week training session. Points were visited once each breeding season between mid-May and mid-July. All birds seen or heard during a ten-minute count period were recorded, and the distance to each was estimated (Hutto and Young 1999). Field observers generally began counts about 15 minutes after sunrise (after the predawn chorus), and generally completed counts within four hours. Counts were not conducted on days with continuous rain or strong winds.

Vegetation Survey Methods

Field observers first determined whether a 100-meter-radius circle around a survey point could be considered a homogeneous cover type. If so, further vegetation measurements were then taken on relatively continuous variables representing vegetation physiognomy and floristics (Hutto and Young 1999).

The selection of variables to measure was based largely on our understanding of avian ecology from the literature and personal observation. Vegetation variables involved structural characteristics of the vegetation at different layers, as well as tree species compositions. Emphasis was placed on variables that were of potential biological importance for one or more of the bird species analyzed and could be collected quickly in the field by trained workers.

We estimated the tree species composition of the canopy layer because of the evidence that floristics may be important in habitat relationships (e.g., Mac Nally 1990b). Different tree species have different architecture and different invertebrate assemblages (e.g., Recher et al. 1991). Bird species forage non-randomly among plant species (Airola and Barrett 1985; Rottenberry 1985), and nests are often placed in some tree species preferentially over others (e.g., Martin 1992). General surveys of cover types (e.g., Hutto and Young 1999) have shown that many bird species are non-

Figure 8.1. The distribution of permanent landbird monitoring transects in northern Idaho and western Montana.
randomly distributed across stands of different tree species.

Because it had been determined qualitatively that the vegetation cover was homogeneous out to 100 meters, we assumed that quantification of vegetation variables within a 30-meter-radius circle (excluding the road corridor) was sufficient to represent the entire area. Therefore, all vegetation variables used in this report were estimated to 30 meters, except for counts of large-dbh (diameter at breast height more than 40 centimeters) trees (LGTREE), which were based on an 11.3-meter-radius circle (excluding the road corridor). Ocular estimates of the following variables were conducted within the 30-meter-radius circle (Hutto and Hoffland 1996; Hutto and Young 1999): HEIGHT—the typical height of the tree canopy layer; CANOPY—the percent cover of canopy trees (larger than saplings); SAPLING—the percent cover of sapling trees (between 5 and 10 centimeters dbh); SHRUB—the percent cover of tall shrubs (multistemmed woody plants greater than 1 meter tall); BUSH—the percent cover of low shrubs (less than 1 meter tall); and GROUND—the percent cover of grasses and forbs.

For tree species composition, we estimated the percentage of the total canopy cover made up by each of the tree species indicated below (some associated species were lumped together): PIPO—percentage of canopy made up by ponderosa pine (Pinus ponderosa); PSME—percentage of canopy made up by Douglas-fir (Pseudotsuga menziesii); LAOC—percentage of canopy made up by western larch (Larix occidentalis); PICO—percentage of canopy made up by lodgepole pine (Pinus contorta); SPRFIR—percentage of canopy made up by spruce/fir (Picea engelmannii, Abies lasiocarpa); MESIC—percentage of total canopy cover made up by western red cedar (Thuja plicata), western hemlock (Tsuga heterophylla), and grand fir (Abies grandis); and DECID—percentage of total canopy cover made up by deciduous trees (Betula papyrifera, Populus tremuloides, and P. trichocarpa).

Each year, new field observers independently estimated the values of vegetation variables in conjunction with the collection of bird data (although tree species composition was estimated in 1994 only). Except when noted, the data from the three separate years were averaged for each point.

### Analysis Methods

Points with more than one cover type within 100 meters were excluded from the analyses, so the estimates of vegetation structure could reasonably be expected to represent an average of a relatively homogeneous area around the point. The bird data were also limited to 100 meters so that both the bird and vegetation data represented samples of the 3.14-hectare (100-meter-radius) area surrounding the point. Some authors recommend the use of a 50-meter radius for bird data, but the effect of restricting count radius is not uniform across species (Wolf et al. 1995). Wide-ranging species with loud calls, such as the common raven (Corvus corax) and pileated woodpecker (Dryocopus pileatus), were detected within 50 meters only 5–15 percent of the time. Small birds with soft or high-pitched songs, such as the brown creeper (Certhia americana) and golden-crowned kinglet (Regulus satrapa), were detected within 50 meters 85–90 percent of the time. Some other species whose songs are unmistakable, such as the olive-sided flycatcher (Contopus cooperi) and varied thrush (Ixoreus naevius), were also identified at greater distances, perhaps due to observer confidence. Thus, it may be best to vary the cutoff radius for different species. Because between-species comparisons are not recommended for point counts (Wolf et al. 1995), the different radii should not be a concern. In the case of the Swainson’s thrush, only 35 percent of detections were within 50 meters, whereas 80 percent were within 100 meters. The song carries well and is easily identified, so we used all detections within 100 meters. To test the effect of this decision, however, we also constructed a model based on a 50-meter-radius plot.

To reduce the confounding effects of very different cover types, some of which we know this species would not occur in, we modeled the habitat associations of the Swainson’s thrush within the subset of conifer forest cover types. We included all points with conifer trees ranging from 5 to 35 meters tall and from 1 to 80 percent canopy coverage. By restricting the data set, we changed the question from the distribution of a bird species across a wide array of cover types to a more refined distribution within a subset of cover types.

Although different points on a transect can be in
different cover types and can be argued to be independent choices in habitat selection made by bird species with territory sizes of a couple hectares or less, the relative health of the local population, local meteorological conditions, and so forth, will always produce some dependence in the data within each transect. Nonetheless, we used individual points as sample units because (1) combining data from all points on a transect would create meaningless sample units with respect to vegetation variables, given that transects run through a series of different cover types; (2) given a mixture of cover types on each transect, and the elimination of points near edges, we included, on average, only 3.8 points per transect; and (3) our emphasis on the relative importance of variables rather than strict rules for inclusion of variables in the model made the sample size a less pressing issue, although it was still important relative to the ability of the data to support a model with many parameters. We also present models based on one point per transect.

For the main model of our example species (Swainson’s thrush), we pooled the three years of data by averaging the vegetation estimates over the three years and by counting the presence of the species in any of the three years as a presence for that point. This method allowed inclusion of sites where the species was simply missed in some of the three years. Alternative approaches are addressed below.

The statistical importance of a variable in any modeling procedure depends on how close the mathematical form of the model is to the form of the true relationship. The simplicity of linear regression has led to its adoption as the typical method in wildlife habitat studies (Young 1996). However, a linear model assumes that a unit-unit relationship holds true for the entire range of an environmental attribute (Meents et al. 1983) so that if more is better, then a lot more must be much better (Johnson 1981b). However, simple niche theory assumes that organisms respond to most important resource gradients in a unimodal fashion. This has been standard procedure among plant ecologists for decades (e.g., Whittaker 1967; Austin 1976; ter Braak and Prentice 1988), but animal ecologists have been much slower to embrace it (Young 1996; Heglund, Chapter 1; but see Meents et al. 1983; Heglund et al. 1994, etc.). There is no particular rea-

son why the relationship must be Gaussian or even symmetrical (Austin 1976); the mode may not even be the optimum (Austin 1980). We do not really know what the shape is likely to be in any particular case, so we used the simplest method possible to pick up at least some of any unimodal signal, while adding only one parameter, which was the addition of a quadratic term (ter Braak and Prentice 1988).

A significant quadratic term can result from nonlinear but monotonic relationships, such as asymptotes, as well as from unimodal relationships. Such response curves would be expected if there were a threshold in the response or if we did not have the complete gradient for the variable. In such cases, however, the linear regression would also show a strong relationship, so the quadratic term may not be necessary to determine the importance of a variable (although it may help improve predictions). The inclusion of the quadratic term is even more important in cases where linear regression indicates no relationship.

If a bird species is associated with one tree species, it is likely to show a quantitative response to others because the proportions are interdependent. One way around this might be an ordination procedure, although this may result in gradients due to overall productivity and/or forest structure rather than direct effects of tree species composition. We opted to use the direct variables to more easily interpret direct effects of tree species. We did not consider quadratic terms for the tree species variables because sparse data (many zeros) would make the additional parameter less supportable, and because such relationships would have little logical interpretation.

We visited point counts only once per year. Because single visits do not commonly produce multiple detections of any one bird species, we used logistic regression (Hosmer and Lemeshow 1989) to analyze the effects of vegetation variables on the presence or absence of each bird species at the points. In addition, biases due to detectability and observer variability should be less pronounced in presence/absence data than in abundance data.

To begin the model-building process, we discarded variables that had exceptionally high p-values in simple regressions and variables that made no biological sense for the particular species (in this case the Swainson’s
thrust). We then used the Akaike Information Criterion (AIC) for model selection (Akaike 1974; McQuarrie and Tsai 1998). AIC incorporates the tradeoff between bias and variance as variables are added to a model, and it provides a straightforward comparison between models that does not depend on a hypothesis-testing framework (Burnham and Anderson 1998). It moves the emphasis away from p-values and arbitrary cutoff criteria (Johnson 1999a), and extracts more information from the data regarding the relative strength of evidence for each variable (and model).

When choosing among models using statistical inference, it is best to work with only a few select models chosen a priori on biological grounds (Burnham and Anderson 1998). However, our data set was both correlative and unstructured, and we knew little about the expected relationships for many species, so we considered this an exploratory analysis. The best method of variable selection for modeling such a data set has been the subject of much discussion. Stepwise selection methods do not necessarily identify the “best” model even from a statistical perspective (James and McCulloch 1990). On the other hand, all-possible-subsets procedures will inevitably lead to overfitting of the data, because the model thus chosen will be highly specific to the data at hand (Burnham and Anderson 1998). In fact, no method can produce the “true” biological model from correlative data, and some overfitting is perhaps inevitable. We chose the most influential variables by forward selection, with AIC as the selection criterion for each step. As inclusion of variables became more uncertain, we modified the procedures to more closely resemble an all-possible-subsets methodology. This allowed the comparison of many likely models and embraced the idea of alternative models and model-selection uncertainty. Although we report the model thus chosen, the goal was not to produce a single final model but rather to determine the strength of evidence for the inclusion of each variable (Burnham and Anderson 1998: 202).

Accuracy Issues

We did not have independent data for testing the accuracy of our models, but we have analyzed the data in several different ways to get an indication of the robustness of the models in terms of the variables included. We performed a cross-validation procedure by splitting the full database in half. We sorted transects by latitude and longitude and selected every other transect for each subset of the data. We then built a logistic regression model for each subset and compared the classification accuracy of each model when predicting the observed data for the other (test) subset relative to the training set. For a classification accuracy assessment that was independent of the cut-point threshold, we used receiver operating characteristic (ROC) plots (Swets 1988; Fielding and Bell 1997; Pearce et al., Chapter 32).

We also checked the consistency of results by building a model for each of the three years separately. Examining each year separately is not an independent validation, of course, because we sampled the same points and, in some cases, the same individual birds returning the next year (or their philopatric offspring). However, the results from three consecutive years should be consistent if we expect the models to perform well on independent data. We used the same methods as above to build multiple logistic regression models for 1994, 1995, and 1996 separately, and compared the classification accuracy to that of the original three-year model using ROC plots. We used the same data set for the vegetation variables, with averages across all three years, so that any year-to-year variability was due to the bird data only, whether from sampling error or from actual changes in occupation of sites.

Sample Unit Considerations

We tested the sensitivity of our results to pseudoreplication issues by redoing the analyses with one randomly selected point from each transect. We had the luxury of doing this because of the large sample size in our regional program. We selected two subsamples, each with one randomly selected point from each transect. The second subsample was selected without making the points from the first set available for selection (i.e., sampling without replacement). The fifty-eight transects with only one available point were randomly divided between the two subsamples. This produced two subsamples of 263 points, each with only one point per transect and with no points in common. Multiple logistic regression models were built for these data sets using the same methods as above, and
the classification accuracy was compared to the original model using ROC plots.

**Poisson Regression**

When we pooled the bird data from three years at each point, we obtained considerable variation in abundance for some bird species, which was lost when we converted the data to presence or absence for logistic regression (Mac Nally 1990a). A point where a Swainson's thrush was detected in only one year (or was mistakenly identified) was given the same importance value as a point with many territorial thrushes singing every year. To better differentiate the relative use of the sites by this species, we reanalyzed the three-year data set using the summed abundances and Poisson regression (Jones et al., Chapter 35). Count data are more likely to follow a Poisson distribution than any other readily available distribution (but see White and Bennetts [1996] for a recommendation of the negative binomial distribution), and the method is fairly robust, requiring only that the variance in the data be proportional to the mean (McCullagh and Nelder 1989).

**Regional-scale Considerations**

In any study of habitat use, the set of "available" locations must be carefully chosen (Johnson 1980). If the species is not present in some areas for any reason other than the variables we have measured in the study, then it would be misleading to dilute the data with such absences, or "naughty noughts" (Austin and Meyers 1996), in potentially suitable habitat that is not occupied for other reasons (e.g., climate or landscape-scale factors, or current or historical dispersal barriers). If some measured vegetation variables also change across the same gradient, then it may look like those measured variables are controlling the distribution rather than the unmeasured factors that are truly limiting.

More than one scale is involved here. If data cover a large region, the geographic range of a species may not extend throughout the entire area. But even within a species' range, there may be suitable habitat that is not occupied due to landscape-scale factors. In a study of local-scale factors, both of these problems might be addressed by using only those transects along which a particular bird species was detected. Those occupied transects are ones in which the range, landscape, and season are apparently appropriate for the bird species' presence. If landscape-scale factors are to be included in the models, then we would want to analyze all transects within the occupied range.

Choosing the best approach to this problem may be a subjective exercise. For example, we detected the Swainson's thrush only rarely in south-central Montana (Fig. 8.2). However, because this area was well within the geographic range of the species (Montana Bird Distribution Committee 1996), we felt that it still could have been present in appropriate habitat. It therefore would be reasonable to use all occupied transects as our method to control for landscape for this species. However, because this abundant species was found on about 85 percent of the transects, this alternative procedure was not likely to produce different results. This method may thus be more useful for less-common species. We decided to restrict the data based on geographic area. Because we were interested in the effect of the geographic distribution of western larch on the importance of that variable in the habitat models, we restricted the data to the geographic range of the larch (all forests west of the Continental Divide except for the Bitterroot National Forest), which was also the area where Swainson's thrushes were most

---

**Figure 8.2.** The geographic distribution of the Swainson's thrush (*Catharus ustulatus*) across all 1,102 points used in the analyses. Closed circles indicate a presence of the Swainson's thrush in any of the three years and open circles indicate absences in all years. Points within transects are nearly congruent in this depiction.
common. We built a new logistic regression model using this subset of the data.

Another regional-scale consideration is a potential change in habitat use in different parts of a species' range. We did not pursue this question because, with this kind of correlative data, it would be very difficult to show that the inevitable differences between the models for two areas were due to actual biological differences in habitat selection rather than different competitive environments or simply sampling error.

**Results**

The final data set included a total of 1,102 points on transects visited in each of three years. We considered the area around a point to be relatively homogeneous (no edges) if only one (559) or none (543) of the three observers thought otherwise.

Almost any data set involving multiple vegetation variables will include a number of intercorrelations among the predictor variables. The highest correlations among the predictor variables in our data were among canopy cover, canopy height, and number of large trees, especially the latter two measures of tree size (Table 8.1). The proportion of mesic conifer species (western red cedar, western hemlock, and grand fir) in the canopy was also highly correlated with those three variables, especially canopy cover. Sites with more ponderosa pine had the lowest average canopy cover. The greatest understory development was under canopies of larch or, secondarily, spruce/fir. The proportions of ponderosa pine and Douglas-fir were negatively related to understory. Because we had already combined the most important species associations (spruce/fir and cedar/hemlock/grand fir), most of the correlations among conifer species variables were negative. The largest correlation coefficient ($r$) was less than 0.5 (Table 8.1), so with our sample size it should be possible to tease apart the effects of all variables, at least to some extent.

Swainson's thrushes were detected in at least one year at 555 of the 1,102 points. Therefore, the categories of presence and absence were nearly equal for the main analysis of the three-year data set using logistic regression. The most important variables in this main model (logistic regression of three-year averages; Table 8.2) appeared to be understory cover consisting of tall shrubs and conifer saplings, positive associations with larch and mesic tree species and, to a lesser degree, canopy cover.

The classification accuracy (at a cut-point of 0.5) of the main model was about 72 percent. When only the strongest variables were used (CANOPY, SAPLING, SHRUB, SHRUB$^2$, LAOC, and MESIC), then the classification accuracy was still 71 percent.

When the data set was split in half for cross-validation, each half had an internal classification accuracy of about 73 percent. When each of the resulting models was used to predict presence in the other half of the data, the ROC plots (Fig. 8.3) indicated a classification accuracy for this test set that was nearly as high as for the training set. In fact, the internal classification accuracy of the models for each half were similar to that for the main model (Fig. 8.4a).

When each year was analyzed separately, Swainson's thrushes were detected on 353 of the 1,102 points in 1994, 294 in 1995, and 281 in 1996. There were some differences in the apparent importance of the vegetation variables in models for the three different years (Table 8.2), most notably for canopy cover and tree species composition. The internal classification accuracy for these models was not quite as good as that for the main model (Fig. 8.4b).

To determine the sensitivity of our results to the use of points as sample units, we also randomly selected two subsets of the data that consisted of single points from each of 263 transects. Swainson's thrushes were detected on 128 and 134 of the 263 points in these two separate subsets. The models based on these two data sets were quite different from one another (Table 8.2), with more variables being included in the second model (including tree size). This second model was also the only model that did not include western larch as a tree species associate (although it would have if only positive tree associations were allowed). Although the internal classification accuracy for these models was better than that for the main model (Fig. 8.4c), validation of each model using the other subset as testing data gave relatively poor accuracy (Fig. 8.5).

The restriction of data to detections within a 50-meter radius did not change the core variables of the model (Table 8.2). The data supported fewer minor
## TABLE 8.1.

Nonparametric correlation coefficients (Kendall's tau-b) for bivariate comparisons of predictor variables.

<table>
<thead>
<tr>
<th></th>
<th>SAPLING</th>
<th>SHRUB</th>
<th>BUSH</th>
<th>GROUND</th>
<th>HEIGHT</th>
<th>LGTREE</th>
<th>PIPO</th>
<th>PSME</th>
<th>LAOC</th>
<th>PICO</th>
<th>SPRFIR</th>
<th>MESIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>CANOPY</td>
<td>+0.09</td>
<td>-0.04</td>
<td>-0.03</td>
<td>-0.12</td>
<td>+0.35</td>
<td>+0.31</td>
<td>-0.12</td>
<td>+0.02</td>
<td>-0.01</td>
<td>-0.04</td>
<td>-0.06</td>
<td>+0.24</td>
</tr>
<tr>
<td>SAPLING</td>
<td>+0.06</td>
<td>+0.02</td>
<td>-0.03</td>
<td>-0.11</td>
<td>-0.13</td>
<td>-0.18</td>
<td>-0.15</td>
<td>+0.15</td>
<td>+0.04</td>
<td>+0.10</td>
<td>+0.16</td>
<td></td>
</tr>
<tr>
<td>SHRUB</td>
<td>+0.47</td>
<td>+0.03</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>+0.01</td>
<td>&lt;0.02</td>
<td>+0.03</td>
<td>+0.25</td>
<td>-0.16</td>
<td>+0.08</td>
<td>+0.10</td>
<td></td>
</tr>
<tr>
<td>BUSH</td>
<td>+0.10</td>
<td>+0.10</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>+0.21</td>
<td>-0.07</td>
<td>+0.09</td>
<td>+0.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GROUND</td>
<td>-0.03</td>
<td>-0.03</td>
<td>&lt;0.01</td>
<td>+0.03</td>
<td>+0.08</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>-0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HEIGHT</td>
<td>+0.37</td>
<td>+0.04</td>
<td>+0.06</td>
<td>-0.04</td>
<td>-0.04</td>
<td>-0.15</td>
<td>-0.03</td>
<td>+0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGTREE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+0.10</td>
<td>+0.09</td>
<td>-0.04</td>
<td>-0.24</td>
<td>+0.02</td>
<td>+0.16</td>
</tr>
<tr>
<td>PIPO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+0.09</td>
<td>-0.13</td>
<td>-0.24</td>
<td>-0.22</td>
<td>-0.15</td>
<td></td>
</tr>
<tr>
<td>PSME</td>
<td>0.05</td>
<td>0.24</td>
<td>0.31</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAOC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+0.03</td>
<td>+0.07</td>
<td>-0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PICO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+0.07</td>
<td>-0.27</td>
<td>-0.12</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## TABLE 8.2.

Order of selection for variables included in multiple regression models of the habitat relationships of the Swainson's thrush (Catharus ustulatus), using AIC (Akaike information criterion).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Poisson</th>
<th>Logit 100m</th>
<th>Logit 50 m</th>
<th>1994</th>
<th>1995</th>
<th>1996</th>
<th>Tr A</th>
<th>Tr B</th>
<th>West</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N = 1102</td>
<td>N = 1102</td>
<td>N = 1102</td>
<td>N = 1102</td>
<td>N = 263</td>
<td>N = 263</td>
<td>N = 749</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- CANOPY + 7c 4c 6c 3c 4c 4c
- SAPLING + 8c 9c 5c 2c 3c 1c
- SHRUB + 1c 1c 1c 1c 1c 2c
- BUSH + 6c 3c 4c 6c 4c 7c
- GROUND + 9 9 8c 3c
- HEIGHT + 11 10 9c 3c 8 6c
- LGTREE + 12c 11
- PSME + 3c 3c 5c 5c 4c 2c 4c 2c 7c
- PICO + 10c 10 6 7c
- SPRFIR + 13 7c
- MESIC + 4c 4c 6c 3c 5c
- DECID

*All models used logistic regression except for the Poisson model.*

*All models used the same vegetation data, with the variables averaged over three years.*

*The first three models used accumulated data for abundance or presence of Swainson's thrush (Catharus ustulatus), over all three years.*

*The models designated by dates used presence data from each of the three years separately.*

*The Tr A and B models are based on two separate sets with one randomly chosen point per transect, and the last column is based on the subset of transects from west of the Continental Divide.*

*b*Indicates quadratic term for indicated variables; signs are for linear and quadratic term, respectively.

*c* Indicates variables included under traditional hypothesis-testing methods.
variables, with a shift to understory variables rather than tree size. The internal classification accuracy for this model was not quite as good as that for the main (100-meter) model (Fig. 8.4d).

The abundance of Swainson’s thrushes at the 555 points varied between one and eleven (sum of three visits; some high numbers brought the accuracy of the abundance data into question). More than one individual was detected at 348 points. Therefore, there was considerable variation in counts for use in a Poisson regression analysis. The resulting model was similar to that obtained by logistic regression (Table 8.2), although it was less likely to indicate nonlinear relationships and it did not include tree-size variables.

We detected the Swainson’s thrush on 493 of 749 points in the northwestern part of the region. The resulting logistic regression model was similar to that obtained for the full data set of 1,102 points (Table 8.2), although without canopy height, SPRFIR, and the quadratic term for canopy cover.

The AIC method usually indicated additional variables beyond those included by traditional hypothesis testing with alpha equal to 0.05. Because of the many models we examined in this exploratory analysis, it is likely that there was some overfitting of the data when the best model was chosen according to AIC.

**Discussion**

Two main goals of building regression models in habitat relationships are to identify biologically important variables and to predict the occurrence of bird species at previously unsampled sites. The first goal, identifying environmental conditions that a bird species needs to be present and successful, is of obvious scientific interest. In addition, it is only through understanding the true biological processes involved in a species’ distribution that we can hope to reach meaningful management recommendations. Determining the important variables can be difficult for a number of reasons. We know we have not measured many potentially important biological variables, such as food resource availability (Hutto 1990) or specific nest sites (Martin 1992). In addition, the biological importance of the measured variables cannot be directly confirmed from a correlative analysis. It is necessary to assume that the importance of the larger biological effects will be revealed by the statistical model, especially if several subsets of the data are examined, but there is no way to know how much of the observed effect is due to actual biological processes or to sampling error. This inherent model uncertainty should encourage us to emphasize the strength of evidence for each variable rather than try to decide which quantitative model is “correct.” The observed evidence of relative importance must then be used to form hypotheses for further investigation.

In this study, the various regression models of
Swainson’s thrush habitat relationships were fairly consistent in that they all included the same set of strongly influential variables (Table 8.2). The variables that appeared to be the weakest predictors of Swainson’s thrush occurrence in any one model were the same variables that were less consistently included in the other models. All of the variables chosen by AIC but not by hypothesis testing methods were in this category. In fact, a model with only the strongest and most consistent variables had nearly the same predictive ability as the full model. This suggests that the weaker variables were either biologically unimportant or inconsistently correlated with the true controlling variables. The increased resolution necessary to understand the possible effects of these variables would require a much larger sample size or a more intensive, controlled study.

This is a first attempt at getting a list of variables, more or less in order of statistical (but not necessarily biological) importance, within this data set. We can never be sure if the model reflects true biological relationships without confirming the results with inde-
cover and sapling cover were only weakly correlated ($r = +0.06$; Table 8.1), it seems that conifer saplings may provide an adequate substitute for this bird species as understory structure. Also, there appeared to be a threshold (asymptote) in the relationship of bird occurrence and understory cover (Fig. 8.6), indicating that 20–30 percent understory cover provided maximum benefit, as might be expected for a shrub-nesting species that forages more generally (Ehrlich et al. 1988). Because management practices tend to increase the amount of land with this level of understory cover, this species is not likely to be of management concern.

We do not know of any particular biological reason for larch to be important to the Swainson’s thrush, and this demonstrates the ambiguity of exploratory analyses. Larch cover was correlated with shrub cover ($r = 0.24$), but both variables were strongly significant in most multivariate models, so it is difficult to know whether this was a true biological relationship or an artifact of confounding variables. The fact that larch is restricted to west of the Continental Divide was our main reason for limiting the bird data to this western region for one model. In this way, we discovered that larch was still an important variable for the thrush within the tree’s geographic range, so the apparent association between the bird and tree species was not an artifact of geography. Further study would be necessary to determine if the retention of larch in the landscape is as important for this species as it is for many cavity-nesting birds (McClelland 1977), but this may be a good example of a relationship that was not apparent using simple cover type distributions (Hutto and Young 1999).

The negative association of Swainson’s thrush occurrence with ponderosa pine could be due simply to a positive association with other tree species, or perhaps the thrush does not do well in that type of tree architecture. Alternatively, it may have more to do with ponderosa pine stands typically having low canopy cover or minimal understory. Although the multivariate analyses may have been able to tease these apart, some residual effect probably remained.

We did not have an independent data set with which to test our models. However, we used a number of internal validation and classification accuracy...
procedures to explore the robustness of our results. We assumed that the most consistently included variables were more likely to have some biological foundation. This is not only of scientific interest, but should also increase the usefulness of the results for predicting the presence of the Swainson’s thrush at new sites, and for estimating its probable response to management decisions.

The cross-validation procedure seemed to show that we have created relatively robust and useful models. Models based on each of the two halves of the data set not only had classification accuracies nearly as large as the full model (Fig. 8.4a), but also the consistency of the models in predicting the other half of the data was encouraging. These results also suggest that doubling the sample size did not result in a greatly improved model.

The combination of three years of data appeared to improve the predictive power of the habitat model relative to the models based on single years. This is understandable given the sampling methods. A point count survey provides an incomplete sample of the birds in a given area. It is probably common for a species to be present but not detected. There is also true year-to-year variation in bird occupancy. It is important to design surveys to sample a representative cross-section of this variation (i.e., multiple years, places, etc.). In this respect, the Northern Region Landbird Monitoring Program may be unique. Samples were large in comparison with other controlled studies of habitat use, and data collection was repeated over several years. The results of these analyses suggest that improving the accuracy of data at each point may be more critical than increasing the total number of points. We are beginning to test these ideas by examining the accuracy of models for other bird species.

Vegetation variables are subject to measurement error and observer variability. There was considerable variation in the estimates by the three different observers at each point over the years. When the separate years were analyzed using the separate estimates of vegetation rather than the three-year average, there was a much greater difference between years than that shown in Table 8.2. This suggests that observer variability can be a serious problem, especially if vegetation is measured quickly by crews primarily trained to identify birds, or if only one year is available. We took the average of all three years for this reason, and it has prompted us to subsequently collect more vegetation data at the points, using experienced forestry crews and additional plots.

A plot radius of 50 meters is often recommended for comparison of bird abundance between different cover types because vegetation density can affect the detectability of individual birds. In addition, if both the bird and vegetation data were accurate, we would expect the 50-meter-radius model to have greater classification accuracy because the bird data would be more tightly associated with the vegetation near the point. In this study, however, the 50-meter-radius model was slightly less accurate then the main model using a 100-meter radius. This suggests that a 50-meter radius may have been insufficient to accurately represent occupancy in the stand. This is even more likely to be the case for less-common species. We conclude that the 100-meter-radius cutoff not only resulted in an adequate model, but also it may be preferable in studies with only one or a few visits to a point, where we are most likely to have an incomplete inventory.

Most bird-habitat relationship models explain only a small proportion of the variance in bird presence or
abundance (Maurer 1986; Morrison et al. 1987). This is due to a variety of factors that have often been mentioned (e.g., Rotenberry 1986; Wiens 1989b,c), and most of these are probably exacerbated by the nature of large exploratory studies. It is important for us to both realize the limitations of the method and to design surveys and analyses to decrease the effects of these problems as much as possible.

In spite of the numerous reasons that regional-scale monitoring data might not be conducive to rigorous habitat analyses, we determined a suite of vegetation characteristics that were strongly correlated with the presence of Swainson’s thrush in forest stands (Table 8.2). We think it is very important that such data are used as fully as possible, as long as the results are not overinterpreted. Managers must be made aware of model uncertainties so that potential problems are not overlooked when final decisions are made (Conroy and Moore Chapter 16).

We may also wish to use these bird-habitat relationship models for the second main goal of building regression models—predicting the likelihood of a particular bird species being present at new, unsurveyed sites. Such predictions would be more robust and less location-specific if the predictor variables were more relevant to biological processes (Austin and Meyers 1996), but this is not absolutely necessary for a useful model as long as the new sites requiring prediction have the same correlational linkages between the measured surrogate variables and the true variables that influence bird occurrence. In any case, the expense of measuring predictor variables over wide regions may be prohibitive unless remotely sensed data can be used. There is little reason to develop models for region-wide prediction until we know what variables are likely to be available to managers over all target areas. We can then determine if models based on those variables are adequately robust for management needs.