J. Field Ornithol., 55(3):330–335

INFLUENCE OF SAMPLE SIZE ON DISCRIMINANT FUNCTION ANALYSIS OF HABITAT USE BY BIRDS

By Michael L. Morrison

Characterization of habitat use by birds has long intrigued biologists (see reviews by Thorpe 1945, Hilden 1965). A popular means of describing habitat use by birds has been to make measurements within territories of nesting birds, using a perch-site or nest-site as the center of a circular plot in which habitat is measured (see review by Holmes 1981). A multivariate method, often discriminant function analysis, is then used to search for differences in habitat use among species (e.g., James 1971, Martinka 1972, Anderson and Shugart 1974, Whitmore 1975, Titterington et al. 1979). There are, however, problems associated with this sampling method and the subsequent analytical procedures (Collins 1981, Noon 1981). The objective of this study was to assess the influence of sample size on discriminant function analysis of bird habitat use (based on perch sites).

METHODS

Data for this study were gathered on 2 early-growth clearcuts in the Oregon Coast Range, Siuslaw National Forest, Alsea Ranger District (Lane and Lincoln counties). The 2 sites (21 and 36 ha in size) were clearcut about 5 years prior to this study. Physiognomic characteristics of the sites were similar and were described earlier (Morrison 1982: Appendix).

Species analyzed were the Song Sparrow (*Melospiza melodia*), MacGillivray's Warbler (*Oporornis tolmiei*), and Wilson's Warbler (*Wilsonia pusilla*); no other species had adequate sample sizes for analysis. Habitat use was assessed during 1979 and 1980 by sampling vegetation in 5-m-radius plots centered at the perch of singing males (James 1971). I assumed that the vegetation around a perch gave an indication of habitat within the territory of the bird. Visual estimates of height (nearest .1 m) and % cover of shrubs, and density, height, and cover of conifers and deciduous trees were made in each plot; Morrison (1981, 1982) described specific methodology. Vegetation available to birds was sampled in 1979 by placing fifty 5-m-radius plots on each site following a stratified random sampling scheme; the data recorded were the same as those for bird habitats.

Two-group discriminant function analysis (DFA) with direct inclusion of variables (i.e., all variables included simultaneously) was performed using the Statistical Package for the Social Sciences (Klecka 1975). Groups were (1) the measures of vegetation made at the perch sites of the birds, and (2) measurements taken in the random plots. The number of random plots (n = 50 plots/site) was held constant for all DFA's. Subsamples

TABLE 1. Habitat used by (based on territory plots), and available to (random plots), the 3 bird species analyzed in this study (based on maximum sample size).

Sample size	Shrub cover (%)		Shrub height (m)		Conifer cover (%)		Conifer height (m)		Deciduous tree cover (%)	
	x̄	SD	x̄	SD	x̄	SD	x̄	SD	x̄	SD
					Song sparrow					
77	46.7	23.7	1.0	0.4	4.9	4.3	1.5	0.5	5.0	10.9
					MacGillivray's warbler					
59	54.4	25.1	1.1	0.3	4.7	3.7	1.6	0.6	7.6	11.2
					Wilson's warbler					
43	51.5	26.6	1.1	0.4	2.5	2.2	1.6	0.7	36.9	25.5
					Availability					
100	46.4	28.1	0.9	0.4	4.9	4.7	1.3	0.5	16.2	23.5

of varying numbers of bird plots were used to assess effects of sample size on the 2-group DFA's; 5 replicates (DFA's) were run for each subsample. A random number table was used to select bird plots that comprised each replicate. Standardized discriminant function coefficients were used to assess the relative importance of variables included in DFA. Interpretation of these coefficients is often difficult if highly intercorrelated variables are included in the analysis (Williams 1981). Only variables with an $r < .5$ with any other variable were included in DFA. Tests were considered significant at $P < .05$. Transformations of the original data did not significantly improve the covariance matrices; analyses were run on the original data (see Morrison 1982).

RESULTS

The study sites were typified by low and ubiquitous shrub cover, small conifers, and scattered patches of deciduous trees. The frequency of use of shrubs by the 3 bird species was similar to the frequency of shrubs available (Table 1). Song Sparrows and MacGillivray's Warblers also used conifers in proportion to availability. Wilson's Warblers, however, used areas with less conifer cover than was available. The most distinct pattern of habitat use was seen for deciduous tree cover: Song Sparrows and MacGillivray's Warblers used little deciduous cover, while the use of deciduous trees by Wilson's Warblers was over twice their availability.

The significance of the separation of the 2 groups (song perch and random plots) as analyzed by discriminant function analysis (DFA) increased with increasing sample size for all 3 species of birds (Table 2). The discriminant function was significant at n = 30 for the Song Sparrow and MacGillivray's Warbler, and at n = 10 for the Wilson's Warbler. The significance of the test of equality of the variance-covariance ma-

TABLE 2. Two-group discriminant function analysis of habitat use by birds (territory versus random plots). Analyses were replicated 5 times for each sample size; values are $\bar{x}$ (SD).

Sample size	Standardized discriminant function coefficients: Shrub cover (%)	Shrub height (m)	Conifer cover (%)	Conifer height (m)	Deciduous tree cover (%)	Variance-covariance matrices: Box's M	P	Significance of discriminant function: χ^2	P
					Song sparrow				
5	0.70 (0.22)	0.24 (0.34)	0.51 (0.29)	−0.54 (0.79)	0.45 (0.42)	NC[a]	NC	8.78 (2.29)	0.14 (0.09)
10	0.02 (0.31)	0.02 (0.42)	0.45 (0.35)	−0.30 (0.69)	0.49 (0.53)	45.78 (22.21)	0.09 (0.19)	9.82 (5.16)	0.22 (0.31)
20	0.24 (0.17)	−0.14 (0.24)	0.58 (0.17)	−0.74 (0.19)	0.64 (0.08)	45.74 (25.54)	0.07 (0.13)	13.28 (9.38)	0.12 (0.13)
30	0.23 (0.22)	0.07 (0.14)	0.53 (0.14)	−0.79 (0.08)	0.73 (0.11)	61.58 (27.92)	0.01 (0.02)	19.07 (6.55)	0.03 (0.06)
40	0.20 (0.09)	0.01 (0.24)	0.41 (0.11)	−0.74 (0.08)	0.73 (0.15)	48.56 (18.07)	0.01 (0.01)	18.02 (2.07)	0.01 (0.01)
60	0.24 (0.05)	−0.03 (0.10)	0.51 (0.07)	−0.74 (0.01)	0.74 (0.07)	50.76 (16.71)	0.01 (0.01)	21.60 (2.07)	0.01 (0.01)
Total (n = 77)	0.22	−0.02	0.49	−0.71	0.77	57.37	<0.01	25.36	<0.01
					MacGillivray's warbler				
5	0.08 (0.36)	−0.06 (0.24)	0.19 (0.44)	−0.60 (0.83)	0.05 (0.43)	NC	NC	8.92 (5.04)	0.20 (0.11)
10	0.04 (0.15)	−0.07 (0.42)	0.18 (0.53)	−0.73 (0.10)	0.20 (0.45)	30.60 (8.17)	0.12 (0.12)	8.56 (4.50)	0.22 (0.20)
20	−0.06 (0.15)	−0.34 (0.19)	0.17 (0.31)	−0.47 (0.70)	0.04 (0.48)	31.80 (10.36)	0.05 (0.05)	10.42 (6.16)	0.18 (0.20)
30	−0.28 (0.09)	−0.19 (0.28)	0.09 (0.41)	−0.30 (0.72)	0.27 (0.40)	37.31 (2.20)	0.01 (0.01)	18.50 (8.53)	0.04 (0.08)
40	−0.11 (0.06)	−0.37 (0.07)	0.33 (0.08)	−0.70 (0.08)	0.43 (0.06)	42.52 (7.54)	0.01 (0.01)	18.54 (6.37)	0.01 (0.01)
50	−0.07 (0.10)	−0.35 (0.10)	0.33 (0.11)	−0.73 (0.07)	0.43 (0.08)	43.20 (2.13)	0.01 (0.01)	20.10 (2.60)	0.01 (0.01)
Total (n = 59)	−0.13	−0.32	0.34	−0.73	0.45	48.43	<0.01	23.29	<0.01
					Wilson's Warbler				
5	−0.07 (0.11)	−0.40 (0.38)	0.26 (0.25)	−0.34 (0.46)	−0.25 (0.70)	NC	NC	15.38 (5.90)	0.07 (0.15)
10	0.15 (0.20)	−0.50 (0.20)	0.49 (0.12)	−0.51 (0.28)	−0.52 (0.21)	31.88 (5.90)	0.08 (0.08)	20.84 (7.94)	0.02 (0.04)
20	0.13 (0.13)	−0.52 (0.12)	0.37 (0.18)	−0.51 (0.02)	−0.72 (0.11)	28.12 (6.53)	0.09 (0.11)	34.76 (9.27)	0.01 (0.01)
30	0.03 (0.08)	−0.48 (0.06)	0.37 (0.08)	−0.47 (0.08)	−0.77 (0.05)	43.06 (7.30)	0.01 (0.01)	39.18 (7.19)	0.01 (0.01)
37	0.05 (0.03)	−0.53 (0.04)	0.39 (0.02)	−0.50 (0.03)	−0.71 (0.03)	43.84 (2.76)	0.01 (0.01)	45.94 (2.44)	0.01 (0.01)
Total (n = 43)	0.05	−0.50	0.40	−0.47	−0.74	50.52	<0.01	48.73	<0.01

[a] Could not be calculated because of low sample size.

trices (Box's M) also increased with sample size, and was significant for all 3 species at n = 30 (Table 2).

Deciduous tree cover and conifer height were the dominant variables that separated song perch and random plots for the Song Sparrow at n = 40; the standard deviation (SD) of the estimates was small by this n (Table 2). Deciduous tree cover and conifer height also were associated with separation at n < 40, but could not be clearly removed from the contribution of the cover of conifers.

Conifer height was the dominant variable separating song perch and random plots for the MacGillivray's Warbler at all n (Table 2). The SD, however, did not become small until n = 40.

Deciduous tree cover separated song perch and random plots at all n > 5 for the Wilson's Warbler (Table 2). The importance of deciduous tree cover could not, however, be clearly separated from shrub height and conifer cover until n = 30.

DISCUSSION

Only subtle differences existed between the use and availability of shrubs and conifers for the 3 bird species on the study sites. There was, however, a distinct pattern of use of deciduous trees by the birds. Thus, it was not surprising that deciduous tree cover was identified as one of the most important variables separating vegetation used from that available for each bird species. The few studies specifically addressing the habitats used by these birds also suggest a preference for various combinations of shrubs and deciduous trees in their territories (e.g., see reviews by Gabrielson and Jewett 1940, Bent 1953, 1968).

Results suggest that about 35 plots were needed to obtain stable values using discriminant function analysis. Johnson (1981b) hypothesized that the number of plots needed in such studies could be approximated by starting with a minimum of 20 plots, plus 3 to 5 additional plots for each variable in the analysis. Since there were 5 variables in my study, 35 to 45 plots would have been adequate to describe habitats used by the birds; my results parallel Johnson's approximation. Many authors, however, have accepted low sample sizes (e.g., fewer than 20 plots/species) as the minimum criteria for multivariate analyses (e.g., James 1971, Whitmore 1975, 1977, Conner and Adkisson 1977, Morrison 1982). My results raise questions concerning the validity of conclusions based on such low sample sizes. Simulation analyses have also indicated that a sample of 40 is the minimum necessary for multivariate analyses (see Carnes and Slade 1982).

A relationship between the equality of the variance-covariance matrices (Box's M) and the discriminant functions was shown; that is, Box's M was usually significant when the discriminant function was significant. Further, as the number of plots increased, both Box's M and the discriminant function became significant. These results raise an interesting paradox: a large sample size was needed to identify that the habitat used by the birds was different from the vegetation available on the study sites. But because the variance of the distributions of song perch and random plots was dissimilar, the usual test of the null hypothesis of no difference in group centroids was invalid.

Standardized discriminant function coefficients were used to interpret the discriminant functions. The robustness of these coefficients under violation of the assumptions of multivariate analysis is open to question (Johnson 1981a). My results suggest that stability of coefficients must be assessed with respect to sample size regardless of the results of Box's M. That is, a nonsignificant Box's M is no guarantee that the associated coefficients are reliable. Therefore, an additional method of interpreting discriminant functions, such as correlating the original variables with the canonical variates, could be used to strengthen results (Williams 1981).

The ability to predict the effects of a management practice on animals is critical for the resource manager. Results of this study indicate, however, that the application of management guidelines based on low sample sizes to broader geographic areas is tenuous.

ACKNOWLEDGMENTS

Dr. H. Johnson is thanked for assistance during the design and analysis of this study; D. M. Leslie, Jr., and B. R. Noon reviewed the earlier drafts. L. Mauer and L. Merkle prepared several drafts. This study was conducted under the auspices of the Oregon Cooperative Wildlife Research Unit: Oregon State University, Oregon Department of Fish and Wildlife, U.S. Fish and Wildlife Service, and the Wildlife Management Institute cooperating. Partial funding was provided by the U.S. Fish and Wildlife Service, Patuxent Wildlife Research Center, Cooperative Agreement 14-16-0009-78-977. Oregon State University Agricultural Experiment Station Technical Paper 6414.

LITERATURE CITED

ANDERSON, S. H., AND H. H. SHUGART, JR. 1974. Habitat selection of breeding birds in an east Tennessee deciduous forest. Ecology 55:828–837.

BENT, A. C. 1953. Life histories of North American wood warblers. U.S. Natl. Mus. Bull. 203.

———. 1968. Life histories of North American cardinals, grosbeaks, buntings, towhees, finches, sparrows, and allies. U.S. Natl. Mus. Bull. 237.

CARNES, B. A., AND N. A. SLADE. 1982. Some comments on niche analysis in canonical space. Ecology 63:888–893.

COLLINS, S. L. 1981. A comparison of nest-site and perch-site vegetation structure for seven species of warblers. Wilson Bull. 93:542–547.

CONNER, R. N., AND C. S. ADKISSON. 1977. Discriminant function analysis: a possible aid in determining the impact of forest management on woodpecker nesting habitat. Forest Sci. 22:122–127.

GABRIELSON, I. N., AND S. G. JEWETT. 1940. Birds of Oregon. Oregon State College, Corvallis.

HILDEN, O. 1965. Habitat selection in birds. Ann. Zool. Fenn. 2:53–75.

HOLMES, R. T. 1981. Theoretical aspects of habitat use by birds. Pp. 33–37, *in* The use of multivariate statistics in studies of wildlife habitat, D. E. Capen, ed., USDA For. Serv. Gen. Tech. Rep. RM-87.

JAMES, F. C. 1971. Ordinations of habitat relationships among breeding birds. Wilson Bull. 83:215–236.

JOHNSON, D. H. 1981a. The use and misuse of statistics in wildlife habitat studies. Pp. 11–19, *in* The use of multivariate statistics in studies of wildlife habitat. D. E. Capen, ed., USDA For. Serv. Gen. Tech. Rep. RM-87.

———. 1981b. How to measure habitat—a statistical perspective. Pp. 53–57 The use of multivariate statistics in studies of wildlife habitat, *in* D. E. Capen ed., USDA For. Serv. Gen. Tech. Rep. RM-87.

KLECKA, W. R. 1975. Discriminant analysis. Pp. 434–467, *in* Statistical package for the social sciences, N. H. Nie, C. H. Hull, J. G. Jenkins, K. Steinbrenner, and D. H. Bent, eds., McGraw-Hill, New York.

MARTINKA, R. R. 1972. Structural characteristics of Blue Grouse territories in southwestern Montana. J. Wildl. Manage. 36:498–510.

MORRISON, M. L. 1981. The structure of western warbler assemblages: analysis of foraging behavior and habitat selection in Oregon. Auk 98:578–588.

———. 1982. Response of avian communities to herbicide-induced vegetation changes, western Oregon. Ph.D. thesis, Oregon State Univ., Corvallis. 77 pp.

NOON, B. R. 1981. Techniques for sampling avian habitats. Pp. 42–52, *in* The use of multivariate statistics in studies of wildlife habitat, D. E. Capen, ed., USDA For. Serv. Gen. Tech. Rep. RM-87.

THORPE, W. H. 1945. The evolutionary significance of habitat selection. J. Anim. Ecol. 14:67–70.

TITTERINGTON, R. W., H. S. CRAWFORD, AND B. N. BURGASON. 1979. Songbird responses to commercial clear-cutting in Maine spruce-fir forests. J. Wildl. Manage. 43:602–609.

WHITMORE, R. C. 1975. Habitat ordination of passerine birds of the Virgin River Valley, southwestern Utah. Wilson Bull. 87:65–74.

———. 1977. Habitat partitioning in a community of passerine songbirds. Wilson Bull. 89:253–265.

WILLIAMS, B. K. 1981. Discriminant analysis in wildlife research: theory and applications. Pp. 59–71, *in* The use of multivariate statistics in studies of wildlife habitat, D. E. Capen, ed., USDA For. Serv. Gen. Tech. Rep. RM-87.

Cooperative Wildlife Research Unit, Oregon State University, Corvallis, Oregon 97331. (Present address: *Department of Forestry and Resource Management, University of California, Berkeley, California 94720*). Received 10 May 1983; accepted 25 Feb. 1984.