ESTIMATING WINTER SPECIES RICHNESS WITH UNLIMITED-DISTANCE POINT COUNTS

KEVIN J. GUTZWILLER

Department of Biology and Institute of Environmental Studies, Baylor University, Waco, Texas 76798-7388 USA

ABSTRACT.—During February in 1987 and January and February in 1988, I conducted 167 unlimited-distance point counts in central Texas to assess how count duration, daily period (time of day), and site type (woodland perimeter, woodland interior) were associated with winter richness estimates. Annual and seasonal influences on richness were controlled for statistically. Based on comparisons of means, levels of count duration and daily period may be useful in various combinations to obtain estimates with comparable magnitude and precision throughout the day. Such combinations might be used in sampling schedules to increase point-count sample size. Point counts in woodland interiors produced higher estimates than those at adjacent woodland perimeters. When estimates from before and after mid-February were compared, I observed declines in richness at woodland perimeters but not at woodland interiors. The standard error of richness increased exponentially during the day; this increase was most evident for 10- and 15-min counts. To avoid declines in precision, one should sample earlier in the day, perhaps from 0700 to 1345. With some preliminary data, investigators also can apply standard formulas to allocate counts to different daily periods, count durations, or both, to obtain minimum-variance estimates of richness for an area. Through statistical power effects, increases in the number and precision of unlimited-distance point counts should improve the accuracy of ecological inferences. *Received 4 September 1990, accepted 20 March 1991.*

WITH UNLIMITED-DISTANCE point counts (Verner 1985), separate estimates of bird species richness can be obtained efficiently, especially in patchy habitats (Blondel et al. 1981) where transects are inappropriate. The method's time effectiveness (Verner 1988) and the ease with which it is used in rugged terrain (D. K. Dawson 1981) have contributed to its broad appeal. Recent studies of its application (Hutto et al. 1986, Verner and Ritter 1986, Verner and Milne 1989) have provided new information for sampling designs that will increase its utility. Yet there are several important issues about its use during winter that have not been studied. First, the length of unlimited-distance point counts per station during winter, which will maximize the number of counts per day, the magnitude of richness estimates, and the precision of richness estimates, is unknown. A clear advantage to maximizing the number of counts is that the power of statistical tests increases as sample size increases (Sokal and Rohlf 1987). There are potential trade-offs, however, between the increased number of counts that can be completed (by decreasing count duration), and the quality of the resulting estimates. Two-minute counts may yield inaccurate estimates, or result in estimates imprecise enough to offset the statistical

power gained from an increase in sample size. Although 20-min counts may produce more accurate and precise estimates, they also may reduce statistical power in associated analyses because fewer counts are possible (see Scott and Ramsey 1981, Fuller and Langslow 1984). Alternatively, a single person might increase point-count sample size in the winter by sampling during more hours of the day (Rollfinke and Yahner 1990), not just during the early- or mid-morning hours as is typical in breeding-season studies.

A second problem is to determine when counts should be made during the day to maximize the magnitude and precision of winter richness values. A winter study in western Mexico (Hutto et al. 1986) demonstrated that most of the morning hours were equally productive for using unlimited-distance point counts to estimate richness. But in the western Sierra Nevada of California, winter richness based on point counts varied significantly among hourly periods of the morning (Verner and Ritter 1986). Researchers who employed other censusing techniques also found differences in winter bird detectability associated with time of morning (Robbins 1971, Shields 1977) and between morning and afternoon periods (Robbins 1972).

Compared with breeding birds, winter birds during midday and afternoon periods may be more detectable because of increased feeding activity, lower resource densities, or both. On the other hand, richness estimates could be smaller or more variable at these times (e.g. Conner and Dickson 1980), and there would be little to gain by increasing sample size this way.

Finally, it has not been determined whether unlimited-distance point counts yield richness estimates that differ between a habitat's interior and its perimeter. In highly fragmented habitats, patch centers may be only tens of meters from their perimeters, yet there are likely to be subtle differences in richness between interior and perimeter sites. Because some species can be detected readily up to 150 m away, small-scale differences in richness may be difficult to detect with this technique. But if such differences are detectable, this method could be used to collect data for within-habitat comparisons in patchy environments. A disadvantage of this detectability would be that count position within a patch would be a source of variation in richness estimates that would have to be balanced or avoided via study design, or accounted for analytically.

The influences of count duration, daily period, and site type on point-count results have been examined separately (see references above), but it may be more fruitful to assess simultaneous effects. A more productive way to maximize sample size and time efficiency, for example, might be to determine how different levels of these variables could be combined in a sampling scheme to yield comparable estimates of richness (see Robbins 1981a, Hutto et al. 1986, Verner and Ritter 1986). This requires simultaneous analysis of main and interaction effects and subsequent comparisons of level means. Further, most research on unlimited-distance point counts has been conducted during the breeding season (but see Hutto et al. 1986, Verner and Ritter 1986), so additional winter research is essential. Many of the relations observed in the present analysis for richness are also evident in the probability of detection of individual species. My objectives are to demonstrate how count duration, daily period (time of day), and site type (woodland perimeter, woodland interior) are related simultaneously to richness estimates during the winter, and to describe how these relations can be used to maximize the number of counts and

the magnitude and precision of richness estimates. The results indicate how the modified use of unlimited-distance point counts can improve the accuracy and statistical power of winter studies.

METHODS

Study area.—I studied woodland birds in the Blackland Prairie and Grand Prairie regions of central Texas (Riskind and Diamond 1988), within 50 km of Waco (31°33'N, 97°10'W). Observations were made primarily throughout McLennan County and to a lesser extent in northern Bell County. Soils of the Blackland Prairie region are dark calcareous clays mixed with gray acid sandy loams; Grand Prairie soils are dark calcareous clays over limestone (Correll and Johnston 1979). Annual precipitation averages about 86 cm, and the climate is subtropical humid (Riskind and Diamond 1988). McLennan and Bell counties lie in a transition zone that grades from east to west through deciduous forest, tallgrass prairie, and evergreen woodland; areal percentages of these three plant formations average 10%, 51%, and 39%, respectively, in McLennan County, Bell County, and two comparable adjacent counties (F. R. Gehlbach in prep.). The dominant plants in the study area are *Acer negundo, Carya illinoinensis, Celtis laevigata, Juniperus ashei, Maclura pomifera, Melia azedarach, Populus deltoides, Prosopis glandulosa, Quercus virginiana, Salix nigra, Ulmus crassifolia, Baccharis neglecta, Forestiera pubescens, Ligustrum quihoui, Smilax bona-nox, Andropogon gerardi, Schizachyrium scoparium,* and *Sorghum halepense* (names from Correll and Johnston 1979).

Data collection.—Before collecting data each year, I spent approximately 40 h during a 3–5 week period identifying winter calls and songs in the field. From 3 February to 24 February 1987, and from 19 January to 25 February 1988, I recorded the number of bird species and individuals detected during a total of 167 20-min unlimited-distance point counts. Before January and after early March, species richness in the study area fluctuates as winter residents arrive or leave, and as migrants pass through (F. R. Gehlbach pers. comm.). To minimize confounding influences of seasonal changes (see Robbins 1972, Anderson et al. 1981, Rollfinke and Yahner 1990), I limited the study to January and February.

I recorded birds detected during four consecutive 5-min intervals within the 20-min period (Robbins 1981a). I identified 22 areas of woodland that were typical of those in central Texas in terms of plant species composition, areal extent, and successional stage. Random starting points defined initial count sites for each of these localities. Subsequent sites were established by pacing at least 200 m from the first (or previous) site and stopping at nearby woodland perimeters (borders of woodlands that adjoined pasture, fallow fields, or cropland) or woodland interiors, de-

pending on available habitat and the site type needed to obtain a balance of site types in each locality. Depending on the size of localities, each yielded from 2–15 count sites. All count sites were established at least 200 m apart to minimize dependencies in the data from consecutive sites (see Blondel et al. 1981, D. G. Dawson 1981, Hutto et al. 1986). The majority of sites within a given locale were approximately 300 m apart, and individual localities were separated by 3–30 km. I began recording data when I reached a point 25 m from the next point-count site (at least 175 m from the previous site). This enabled me to record species that were present at a site but that, on my approach, stopped vocalizing or flushed without returning (Hutto et al. 1986).

I studied species that used the perimeters or interiors of woodlands during the winter for shelter, feeding, or resting. Occasionally, I detected individuals of these species outside or high above woodlands; typically they flew <40 m above me from one woodland patch to another, or across fields to reach woodlands. I included these individuals in my analyses; I excluded individuals detected more than 50 m above the ground. In short, I analyzed data for individuals that actually used woodland perimeters or interiors, either at the time of detection or immediately thereafter. Exceptions to this involved several species that were often audible or visible from distances exceeding 200 m (waterfowl, wading birds, Turkey Vulture [*Cathartes aura*], Black Vulture [*Coragyps atratus*], Red-shouldered Hawk [*Buteo lineatus*], Red-tailed Hawk [*B. jamaicensis*], American Crow [*Corvus brachyrhynchos*]). To minimize the chance for statistical dependencies among point counts, I excluded these particular species from my analysis. I emphasized smaller, less conspicuous birds.

I did not census when wind speed was greater than 20 km per hour (Robbins 1981b), air temperature was lower than 0°C, or more than a light drizzle fell. Snow cover is rare in central Texas, and many species are noticeably less active when it occurs; I did not census when snow was on the ground. Immediately after each 20-min count, I (1) measured air temperature with a hand-held thermometer, relative humidity with a digital hygrometer, and windspeed with a hand-held anemometer; (2) recorded wind direction and whether sunshine occurred during the count; (3) estimated by eye (after practice with a measured distance of 50 m) whether visibility through vegetation exceeded 50 m; and (4) used the "plant cramming" technique (Hays et al. 1981) to estimate percent canopy coverage of *Quercus-Juniperus* vegetation within a circular area (radius = 25 m) centered on the count point. Counts were categorized according to their starting times into one of five daily periods: 0700–0915, 0916–1130, 1131–1345, 1346–1600, and 1601–1815. Point counts hampered by noises or activities of dogs, people, or large bird flocks were not used.

Habitat and physical conditions varied somewhat among the 167 count sites. I wanted to identify associations between richness and count duration, daily period, and site type that transcended such influences, so I did not remove analytically richness associated with these differences from the analysis. My intent was to provide censusing guidelines effective for the entire range of conditions I encountered. I sampled once at 167 distinct but comparable sites to ensure statistical independence among point counts and to provide information for a greater variety of environmental conditions than repeated censuses at fewer sites would have permitted.

Statistical analyses.—Although a 20-min count was conducted at each of the 167 sites, I randomly selected sites for analysis of their 5-, 10-, or 15-min results. Researchers usually use one of these three durations because shorter durations (e.g. 2 min) may yield less accurate estimates (see Verner 1988) and longer durations (e.g. 20 min) can be inefficient or cause successive counts to be statistically dependent (Fuller and Langslow 1984, Verner 1988). I applied a square-root transformation to richness estimates (counts) to normalize error terms and stabilize error-term variances (Neter and Wasserman 1974: 507, equation 15.11). I computed type III sums of squares with Proc GLM (SAS Institute 1985) to test for differences in mean richness associated with count duration, daily period, site type, and their interactions. Type III sums of squares enable one to assess more clearly the influence of a given effect because all other effects in the model are accounted for first. A runs test (Draper and Smith 1967: 97) on the residuals (ordered in time and space) from the initial model revealed serial dependencies. Residual plots (Neter and Wasserman 1974: 99) indicated that year (1987 or 1988) and winter date (before 14 February, after 14 February) were necessary in the model to eliminate these dependencies.

Following the inclusion of year and winter date, I tested for two-way interactions involving these variables and count duration, daily period, and site type. I was interested in whether count duration × daily period interaction effects varied with year, site type, or winter date, and therefore I tested for the appropriate three-way interactions. Often, main effects are not interpreted when associated interactions are significant. But there are circumstances in which knowledge about main effects in the presence of interactions is valuable (Sokal and Rohlf 1987: 198). Consider, for example, a researcher who does not intend to vary count duration or the habitat type to be studied, but who wants to choose optimal daily periods for censusing. An understanding of the main effects of daily period would be helpful here, even though there may be associated interactions. I present main-effect data for this reason and because the patterns exhibited by levels of the main effects are also evident in the interactions. My research recommendations, however, emphasize interaction effects. Many of the two-way and all of the three-way interactions were non-

significant ($P > 0.05$). Because statisticians are unsure about how the deletion of nonsignificant interactions affects the power and significance of tests on remaining terms, I retained all of the interactions in the analysis (Neter and Wasserman 1974: 582, 655). Residual plots and a runs test confirmed the statistical validity of the final model.

I used Tukey pairwise comparisons to test for differences in mean richness among levels of effects that were significant ($P < 0.05$) in the general linear model (Zar 1984: 186, 226, 251). Because these effects were identified as significant after controlling for other effects in the model, I applied the Tukey comparisons to level means that were adjusted for other effects in the general linear model. I used harmonic interpolation to compute critical values for Tukey tests (Zar 1984: 477, 537).

I did not sample repeatedly at the same count points because this would have prevented me from using the statistical analyses that were needed to meet my objectives. When data are repeated measurements on the same experimental units, only repeated measures analysis of variance is appropriate to assess main and interaction effects simultaneously. This method assumes that the influences of different treatments on the same experimental unit are independent. That is, it is assumed that the unit returns to a normal or pretreatment state sometime after one treatment is applied and before the next treatment is administered (Ott 1988: 793). In the context of the present study, the bird assemblage around a given count point was the experimental unit. To assess the effects of, say, daily period on richness estimates with this technique, one would assume, for example, that the influences of the 0700–0915 period on species' detectabilities were independent of the influences of the 0916–1130 period on species' detectabilities. This is unrealistic. For instance, an individual's ability to secure food during 0700–0915 will determine in part its foraging activity (hence detectability) during 0916–1130. Time-related dependencies originating from other factors also undoubtedly influence species' detectabilities. Such dependencies invalidate repeated measures analysis of variance. An analysis of repeated measures data with conventional analysis of variance techniques would incur the problems of pseudoreplication in time (Hurlbert 1984) because some or many of the same individuals around a given count point would be detected repeatedly (during the daily cycle, during different times of winter, and perhaps even during different years). The approach I took avoided these problems.

Values of the standard error (SE) of richness were estimated from the square-root transformed values of the single counts at separate sites using the conventional formula (Zar 1984: 87). Ideally, these SEs could have been estimated by first computing the means of repeated counts at each of many sites and then calculating the standard deviation (SD) of these means (i.e. the SE of the mean [$SE_{\bar{x}}$]). This approach would have limited the number of sites I could have visited, and the results would have applied to a much narrower range of physical and biological conditions. I used nonlinear regression to determine whether the SE of richness varied significantly with count duration and daily period. In these analyses, I did not want to exclude variation associated with site type, year, and winter date because one encounters these influences on estimate precision in practice. Thus, I did not adjust the SEs of richness used in these nonlinear regression analyses for other effects in the general linear model. Differences among regression slopes were assessed using pairwise comparisons (Zar 1984: 302) and interaction terms in regression models. Statistical power computations followed equations in Zar (1984: 173, 312).

RESULTS

Biotic and abiotic conditions.—Forty-six species were detected in 2 yr (Table 1). The number of bird species detected varied from 1 to 15, with a mean ($\pm$SD) of 7.3 ± 3.4 for individual counts. The numbers of count sites with *Quercus-Juniperus* canopy coverage of 0%, 1–33%, 34–67%, and >67% within 25 m of the count point were 59, 74, 23, and 11, respectively. Visibility through vegetation in all directions exceeded 50 m at 55 count sites and was less than 50 m at 112 sites. The mean ($\pm$SD) air temperature was $16.0 \pm 6.2°C$ (range: 1.1–28.6°C). Mean ($\pm$SD) relative humidity was $45.9 \pm 19.9\%$ (range: 14.0–83.9%). Sunshine was constant or intermittent during 99 counts but was blocked by clouds or haze during 68 counts. Winds came from the north (270–90°) during 84 counts and south (91–269°) during 83 counts. Wind speeds were <3.2 km per hour at 104 sites, 3.2–9.7 km/h at 45 sites, and 9.8–16.1 km/h at 18 sites.

Statistical analyses.—A runs test for the final general linear model indicated residuals were independent ($z = 0.461$, $P = 0.322$). All main effects and several interaction effects were significantly associated with richness (Table 2). Pairwise comparisons revealed differences among level means for main effects (Table 3) and among level means for significant interaction effects (Table 4).

The standard error of richness increased significantly with the square of the number of 2.25-h periods after 0700 (PERIODS^2) (Fig. 1). Increases were most evident for 10- and 15-min counts (Fig. 2). In a regression model, the interaction between count duration and PERIODS^2 (DUR $\times$ PERIODS^2) was associated significantly

TABLE 1. Proportion of 20-min unlimited-distance point counts during which bird species were detected in 1987 ($n = 66$) and 1988 ($n = 101$). The total number of different individuals detected during counts is listed in parentheses.

	Year	
Species	1987	1988
Falco sparverius	0.02 (1)	0.02 (2)
Meleagris gallopavo	0.03 (2)	0.01 (1)
Colinus virginianus	0.03 (10)	0.02 (2)
Zenaida macroura	0.09 (8)	0.02 (2)
Geococcyx californianus	0.00 (0)	0.01 (1)
Bubo virginianus	0.00 (0)	0.03 (3)
Strix varia	0.03 (3)	0.02 (2)
Melanerpes carolinus	0.68 (46)	0.20 (21)
Sphyrapicus varius	0.03 (2)	0.02 (2)
Picoides scalaris	0.12 (8)	0.12 (12)
P. pubescens	0.32 (21)	0.05 (5)
Colaptes auratus	0.44 (29)	0.32 (34)
Sayornis phoebe	0.03 (2)	0.03 (3)
Cyanocitta cristata	0.47 (33)	0.38 (39)
Parus carolinensis	0.91 (64)	0.49 (53)
P. bicolor	0.52 (35)	0.17 (17)
Certhia americana	0.00 (0)	0.01 (2)
Thryothorus ludovicianus	0.64 (45)	0.28 (29)
Thryomanes bewickii	0.00 (0)	0.15 (15)
Troglodytes aedon	0.02 (1)	0.00 (0)
Regulus satrapa	0.12 (10)	0.01 (1)
R. calendula	0.48 (35)	0.11 (11)
Sialia sialis	0.20 (14)	0.32 (33)
Catharus guttatus	0.02 (1)	0.03 (3)
Turdus migratorius	0.74 (61)[a]	0.53 (72)
Mimus polyglottos	0.36 (26)	0.38 (39)
Toxostoma rufum	0.00 (0)	0.03 (3)
Bombycilla cedrorum	0.11 (39)	0.08 (19)
Lanius ludovicianus	0.02 (1)	0.05 (5)
Sturnus vulgaris	0.12 (8)	0.05 (10)
Dendroica coronata	0.67 (51)	0.51 (54)
Cardinalis cardinalis	0.74 (64)	0.67 (87)
Pipilo erythrophthalmus	0.06 (4)	0.09 (9)
Spizella passerina	0.02 (1)	0.01 (1)
S. pusilla	0.03 (3)	0.05 (8)
Pooecetes gramineus	0.02 (3)	0.05 (6)
Passerella iliaca	0.08 (9)	0.06 (8)
Melospiza melodia	0.00 (0)	0.01 (1)
Zonotrichia albicollis	0.05 (15)	0.02 (7)
Z. leucophrys	0.02 (5)	0.01 (1)
Z. querula	0.09 (9)	0.18 (31)
Junco hyemalis	0.05 (5)	0.06 (22)
Sturnella spp.	0.23 (15)	0.11 (30)
Molothrus ater	0.08 (5)	0.00 (0)
Carduelis tristis	0.32 (22)	0.39 (41)
Passer domesticus	0.05 (3)	0.01 (3)

[a] During six counts, I detected from tens to hundreds of flocking individuals in addition to those listed in parentheses.

TABLE 2. Statistics from the general linear model for winter species richness ($n = 167$).[a]

Effect	df	F	P
Site type	1	6.95	0.0097
Count duration	2	7.96	0.0006
Daily period	4	5.17	0.0008
Year	1	20.93	0.0001
Winter date	1	20.16	0.0001
Site type × Count duration	2	1.62	0.2036
Site type × Daily period	4	0.53	0.7140
Site type × Year	1	0.73	0.3949
Site type × Winter date	1	6.33	0.0134
Count duration × Daily period	8	3.54	0.0012
Count duration × Year	2	0.52	0.5965
Count duration × Winter date	2	1.26	0.2892
Daily period × Year	4	0.43	0.7875
Daily period × Winter date	4	2.83	0.0282
Year × Winter date	1	0.33	0.5670
Count duration × Daily period × Site type	8	1.78	0.0893
Count duration × Daily period × Year	8	1.08	0.3803
Count duration × Daily period × Winter date	8	1.11	0.3612

[a] F and P values are based on type III sums of squares (SAS Institute 1985).

and 0.014 for 5-, 10-, and 15-min counts, respectively) indicated no significant differences (experimentwise-error rate = 0.05). These latter tests were lower in power than the overall test for interaction (via regression, $n = 15$) because the groups for count duration had smaller sample sizes ($n = 5$). The SE of richness for combinations of count duration and daily period was related inversely to associated sample sizes ($r = -0.768$, df = 13, $P = 0.001$), which suggests the relations above were spurious consequences of sample size. But after I accounted for sample size with regression, the coefficient for PERIODS² and the coefficient for DUR × PERIODS² (from a separate model) were still significant ($t = 3.29$, df = 12, $P = 0.007$; $t = 2.77$, df = 12, $P = 0.017$, respectively). The relations between the SE of richness and daily period and count duration were therefore not artifacts of sample size.

DISCUSSION

The absence of serial dependency among the residuals for the final model indicates that spacing count points at least 200 m apart was adequate in my study area. Evidently, any dependency that may have arisen from detecting the same individuals during successive counts was

with the SE of richness ($t = 3.63$, df = 13, $P = 0.003$), which indicates differences among the three slopes (one for each count duration) that related the SE of richness to PERIODS². But pairwise comparisons of the slopes (0.001, 0.012,

TABLE 3. Differences and similarities in richness among levels of main effects.[a]

Main effect	$\bar{x} \pm$ SE (n)
Site type	
Woodland interior	4.95 ± 0.17 (88) A
Woodland perimeter	4.29 ± 0.17 (79) B
Count duration	
5 min	3.97 ± 0.21 (56) A
10 min	5.03 ± 0.20 (56) B
15 min	4.87 ± 0.18 (55) B
Daily period	
0700–0915	5.00 ± 0.27 (35) AC
0916–1130	5.35 ± 0.24 (38) A
1131–1345	4.34 ± 0.21 (37) B
1346–1600	4.36 ± 0.34 (29) BC
1601–1815	4.06 ± 0.22 (28) B
Year	
1987	5.19 ± 0.20 (66) A
1988	4.05 ± 0.13 (101) B
Winter date	
Before 14 February	5.16 ± 0.15 (107) A
After 14 February	4.08 ± 0.18 (60) B

[a] Statistics are for square-root transformed data; $\bar{x}$s and SEs listed here are adjusted for other effects in the general linear model. Results of the two-group comparisons are based on F tests from the general linear model (Table 2). Results of the three- and five-group comparisons are based on Tukey pairwise comparisons of means that were adjusted for the other effects in the general linear model; each of the latter two sets of comparisons had experimentwise-error rates = 0.05. Within each main effect, levels not marked with a common letter had adjusted means that differed significantly.

minimal. Although true richness at perimeter and interior sites can be expected to differ, it is not axiomatic that these differences will be detectable with unlimited-distance point counts. This is especially true in patchy or fragmented habitats in which these two site types sometimes may be only 20–30 m apart, as in the present study. From a given point (perimeter or interior) in such habitats, both perimeter and interior species will be detected during a single count. Nevertheless, point counts in woodland interiors and at woodland perimeters yielded significantly different richness estimates, which indicates that unlimited-distance point counts from different parts of even small patches are sensitive enough to detect differences in richness. These results also indicate that count positions within small patches should be accounted for in data analyses, or that counts should be made from comparable places.

The differences between these site-type means and other means (Tables 3 and 4) are small. These differences suggest that, although the

TABLE 4. Differences and similarities in richness among levels of significant interaction effects.[a]

Interaction effect	$\bar{x} \pm$ SE (n)
Site type × Winter date[b]	
Woodland interior, Before 14 Feb.	5.20 ± 0.20 (53) A
Woodland interior, After 14 Feb.	4.70 ± 0.24 (35) A
Woodland perimeter, Before 14 Feb.	5.13 ± 0.19 (54) A
Woodland perimeter, After 14 Feb.	3.46 ± 0.30 (25) B
Count duration × Daily period[c]	
5 min, 0700–0915	3.46 ± 0.63 (10) A
5 min, 0916–1130	5.56 ± 0.54 (9) BC
5 min, 1131–1345	3.54 ± 0.42 (9) AD
5 min, 1346–1600	3.72 ± 0.38 (16) AD
5 min, 1601–1815	3.55 ± 0.29 (12) AD
10 min, 0700–0915	5.35 ± 0.39 (9) BE
10 min, 0916–1130	4.82 ± 0.28 (17) BD
10 min, 1131–1345	4.79 ± 0.30 (16) CDEFG
10 min, 1346–1600	4.80 ± 0.74 (6) ABF
10 min, 1601–1815	5.37 ± 0.43 (8) BG
15 min, 0700–0915	6.19 ± 0.31 (16) B
15 min, 0916–1130	5.67 ± 0.36 (12) BF
15 min, 1131–1345	4.68 ± 0.35 (12) ACEFG
15 min, 1346–1600	4.56 ± 0.62 (7) ACEFG
15 min, 1601–1815	3.26 ± 0.40 (8) A
Daily period × Winter date[b]	
0700–0915, Before 14 Feb.	5.72 ± 0.21 (27) A
0916–1130, Before 14 Feb.	5.34 ± 0.21 (27) AC
1131–1345, Before 14 Feb.	4.59 ± 0.22 (25) BCDEF
1346–1600, Before 14 Feb.	5.14 ± 0.56 (16) AD
1601–1815, Before 14 Feb.	5.04 ± 0.39 (12) AE
0700–0915, After 14 Feb.	4.28 ± 0.47 (8) BCDEG
0916–1130, After 14 Feb.	5.36 ± 0.41 (11) AB
1131–1345, After 14 Feb.	4.08 ± 0.38 (12) BDEG
1346–1600, After 14 Feb.	3.58 ± 0.36 (13) FG
1601–1815, After 14 Feb.	3.09 ± 0.30 (16) G

[a] Statistics are for square-root transformed data; $\bar{x}$s and SEs listed here are adjusted for other effects in the general linear model. Level differences are based on Tukey pairwise comparisons of means that were adjusted for the other effects in the general linear model. Within each interaction effect, levels not marked with a common letter had adjusted means that differed significantly.
[b] Experimentwise-error rate = 0.05.
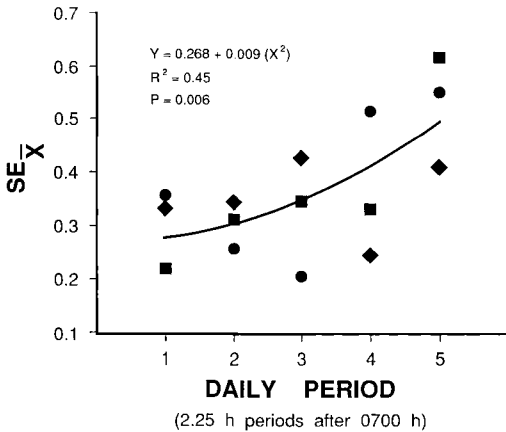[c] Experimentwise-error rate = 0.10.

Fig. 1. Relation between standard error of species richness (SE_x) (transformed data) and the number of 2.25-h periods after 0700 (daily period). Standard errors for 5-min (◆), 10-min (●), and 15-min (■) counts are shown.
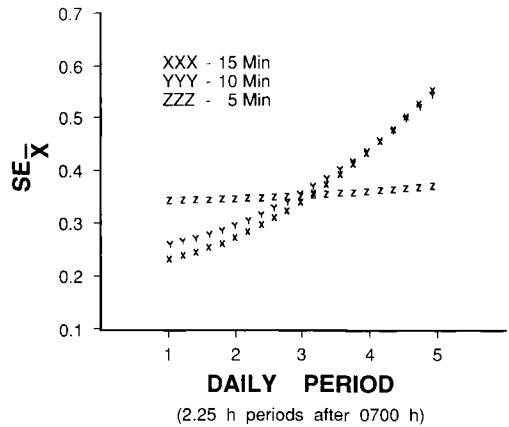

Fig. 2. Separate relations for each count duration between standard error of species richness (SE_x) (transformed data) and the number of 2.25-h periods after 0700 (daily period).

means were statistically significant, the differences were not substantial in the field. But after these transformed values are converted back to original units (species), it becomes evident that some means found to be significantly different differed only by one or two species, while others involved greater than fourfold differences in richness. For example, the untransformed means for richness were 4.12 at perimeter sites and 5.63 at interior sites (cf. Table 3), whereas they were 9.10 for 15-min, 0700–0915 counts and 2.18 for 15-min, 1601–1815 counts (cf. Table 4). Thus, among the levels of some variables, statistically significant differences reflected substantial field differences in richness.

Ten-minute counts produced estimates with the highest magnitude for the least amount of expended time. Fifteen-minute counts yielded magnitudes that were not significantly different from those for 10-min counts (Table 3); estimate precision was comparable for 10- and 15-min counts. Because 15-min counts require more time, if count duration alone is considered, 10-min counts are most appropriate for simultaneously maximizing estimate magnitude, precision, and sample size (cf. Fuller and Langslow 1984, Verner 1988).

Under the conditions of my study, the first two daily periods appear optimal for estimating richness (cf. Rollfinke and Yahner 1990). The third and fifth periods yielded significantly lower estimates than the first two. The fourth period produced the least-precise estimates (see

Conner and Dickson 1980). I observed important yearly differences in the proportion of counts during which particular bird species were detected (Table 1). Thus, during many counts the chance of detecting a number of species declined in 1988. These differences may have been responsible for the lower richness estimates in 1988 (Table 3) (cf. Smith 1984). I controlled for these differences by including year in the general linear model. Consequently, yearly effects on richness, associated with regional population changes, differences in sampled habitats, or perhaps other variables, did not significantly influence assessments of the other main and interaction effects. Some investigators (e.g. Kricher 1975, Rollfinke and Yahner 1990) have noted late-winter declines in species richness. I also detected fewer species after 14 February than before 14 February (Table 3). Overall, if just the main effects are considered, 10-min counts between 0700–1130 before 14 February seem most appropriate.

If the need or opportunity exists in one's sampling design to consider interaction effects, then additional criteria should be examined to obtain comparable estimates. For example, comparing estimates from before and after 14 February revealed declines in richness at woodland perimeters but not at woodland interiors (Table 4). Comparisons of level means for the count duration × daily period interaction indicate how levels of count duration and daily period can be used in combination to yield comparable

KEVIN J. GUTZWILLER

richness estimates for a variety of sampling plans. Five-minute counts during periods 1, 3, 4, and 5, for instance, produced similar estimates (Table 4) and therefore could be used to get a large sample of reasonable estimates. Alternatively, 10-min counts during the first and last daily periods and 15-min counts during the second, third, and fourth periods would produce richness values with higher and comparable magnitudes. If an investigator wanted to sample until the end of February, comparable estimates could be obtained by using the last three daily periods before 14 February and the first two periods after 14 February, or the last two daily periods before 14 February and the first three periods after 14 February. Thus, many combinations of the levels of these variables could be used to obtain like estimates. The object would be to choose combinations that yield comparable richness values and that maximize the number of counts and the magnitude and precision of estimates. The results on which these suggested combinations are based are statistically sound, biologically reasonable, and consistent with results from other studies (see above). Nevertheless, additional work is needed in different areas and with higher levels of replication to substantiate the results and test the general appropriateness of the combinations.

At dawn, many species and individuals are usually visible and audible, and it can be difficult to record all detections accurately (Bystrak 1981, Bart and Schoultz 1984). Conceivably, one could miss detecting uncommon species altogether. Although this difficulty was not a problem in my study, winter researchers who encounter this situation could use longer counts (e.g. 15 min) to ameliorate it. If this problem was severe, investigators might want to skip this period of the day and shift their sampling effort to later periods (D. G. Dawson 1981, Bart and Schoultz 1984). For example, 10-min counts from 0916–1815 would provide estimates that were comparable to those based on counts from 0700–1600 (Table 4), with no reduction in sample size.

None of the three-way interactions involving count duration and daily period were significant, which indicates the interaction between these two variables did not vary with site types, winter dates, or years. This demonstrates that, even when site-type, winter-date, or yearly changes are encountered, it is possible for var-

ious combinations of count duration and daily period to yield comparable estimates. To the extent that the biotic and abiotic conditions in other study areas match those in this study (see Results and Table 1), my findings may be applicable. It is probably more appropriate, though, for researchers to conduct a 2- or 3-yr study in their own area (Verner and Ritter 1986).

The increase in the standard error of richness after 1345 was higher for 10- and 15-min counts than for 5-min counts (Fig. 2). After 1345, the 10- and 15-min counts were long enough to result in both large and small estimates. Although 5-min counts yielded small afternoon estimates, they were typically too short for a large number of species to be detected. Five-minute counts did not produce both large and small estimates as often as the longer counts did, and 5-min counts were more precise as a consequence. In terms of precision, 10- and 15-min counts were more susceptible than 5-min counts to the sporadic bird activity that occurred after 1345. Before this hour, when birds were detectable more consistently, the precision of 10- and 15-min counts was slightly better than that of 5-min counts (Fig. 2). The larger SEs toward the end of the day (Figs. 1 and 2) imply daytime increases in the variation in activities that influence detection (including vocalization, foraging, resting, and cover seeking).

For 10- and 15-min counts, less-precise estimates of richness were obtained during afternoon periods, especially 1601–1815 (Fig. 1) (cf. Conner and Dickson 1980). The problem can be avoided by working before 1345 (Fig. 1). Another strategy would be to apply an optimal allocation formula (Cochran 1977: 98, eq. 5.26) to determine how the maximum number of possible counts should be allocated to daily periods, given a fixed sampling cost (say, 15 min total for each consecutive 10-min count). The allocation would yield a minimum-variance richness estimate for an area based on the set of daily periods used for counts. This approach requires estimates of the standard deviation of richness for a given count duration and specific daily periods, which could be obtained from some preliminary sampling. Such values could be approximated from the standard errors and sample sizes (n) in Table 4, although data for the area and conditions of interest would be preferable. The formula will allocate more counts to those daily periods that yield less pre-
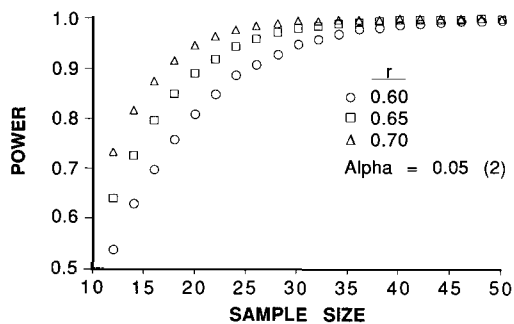
Fig. 3.   Relations among statistical power, sample size, and correlation coefficients for two-tailed tests involving Pearson's *r*, α = 0.05.
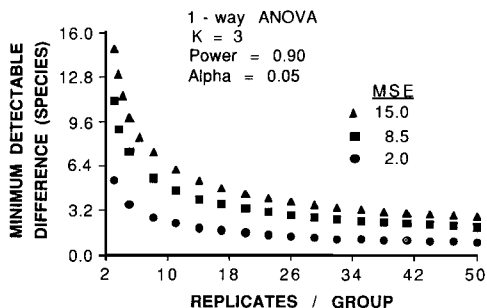

Fig. 4.   Relations among the minimum detectable difference in number of species, replicates per group, and mean square error (MSE) for a three-group, one-way analysis of variance. Power = 0.90, α = 0.05.

cise richness estimates. Alternatively, given a fixed time budget (e.g. total sampling time available each day), 5-, 10-, and 15-min count durations could be viewed as the strata in an optimal-allocation formula that accommodates unequal within-stratum sampling costs (Cochran 1977: 98, eq. 5.23). The computations would indicate how the number of counts should be allocated to different count durations to produce minimum-variance estimates of richness under a fixed time-budget constraint.

One also can obtain minimum-variance estimates for a fixed total cost by allocating sampling effort to primary sampling units and their subunits (Cochran 1977: 280). For instance, given a fixed amount of sampling time, this approach could be used to decide how many forest stands (primary units) and how many separate points within each stand (subunits) should be sampled to achieve a minimum-variance richness estimate for an area. This method is not appropriate if count points or count periods (e.g. 10 min) are considered to be the primary units and individual minutes within a period are considered to be the subunits. There are two reasons for this. First, although the primary units (10-min counts) could be arranged objectively so that they would be spatially and temporally independent, the subunits (individual minutes) would probably not exhibit such independence. This holds because counts made from the same place, and during several 1-min intervals within a single 10-min period, would likely be based on detections of the same individuals. Secondly, it is assumed that a simple random sample or a systematic sample of subunits (individual minutes) would be used to generate the estimates. But if one were to sam-

ple randomly minutes 2, 7, and 10, for example, the remaining 7 min would not be spent sampling. Thus, point-count sampling with such an allocation would not be time-efficient either.

Increased sample sizes and precision will improve statistical power. Consequently, significant differences and relations are more easily detected, and clearer interpretations of results are possible. Power computations for the statistical conditions (*n*, *r*, α, β, mean square error (MSE), minimum detectable difference) I used or observed, and that field ornithologists are likely to use or encounter, illustrate this. For example, at *n* < 30, the power of Pearson's *r* is higher for less variable relations (Fig. 3). As *n* increases to 35, so does the power of this test. Similarly, one can detect smaller differences in richness with analysis of variance when the number of replicates per group is higher and when the MSE is lower (Fig. 4). Through power effects, increases in point-count sample size and estimate precision can improve the accuracy of ecological inferences. When daily sample sizes can be increased, advantages also accrue for monitoring programs. Specifically, more areas can be sampled, a fixed set of sites can be sampled during fewer days, or both.

tifying plants; and D. Anderson and R. Cochran for statistical advice. Numerous central Texas landowners kindly gave me permission to work on their property. Funding was provided by the Baylor University Research Committee.

LITERATURE CITED

ANDERSON, B. W., R. D. OHMART, & J. RICE. 1981. Seasonal changes in avian densities and diversities. Stud. Avian Biol. 6: 262–264.

BART, J., & J. D. SCHOULTZ. 1984. Reliability of singing bird surveys: changes in observer efficiency with avian density. Auk 101: 307–318.

BLONDEL, J., C. FERRY, & B. FROCHOT. 1981. Point counts with unlimited distance. Stud. Avian Biol. 6: 414–420.

BYSTRAK, D. 1981. The North American breeding bird survey. Stud. Avian Biol. 6: 34–41.

COCHRAN, W. G. 1977. Sampling techniques. New York, John Wiley and Sons.

CONNER, R. N., & J. G. DICKSON. 1980. Strip transect sampling and analysis for avian habitat studies. Wildl. Soc. Bull. 8: 4–10.

CORRELL, D. S., & M. C. JOHNSTON. 1979. Manual of the vascular plants of Texas. Richardson, Texas, Univ. Texas–Dallas.

DAWSON, D. G. 1981. Counting birds for a relative measure (index) of density. Stud. Avian Biol. 6: 12–16.

DAWSON, D. K. 1981. Sampling in rugged terrain. Stud. Avian Biol. 6: 311–315.

DRAPER, N. R., & H. SMITH. 1967. Applied regression analysis. New York, John Wiley and Sons.

FULLER, R. J., & D. R. LANGSLOW. 1984. Estimating numbers of birds by point counts: How long should counts last? Bird Study 31: 195–202.

HAYS, R. L., C. SUMMERS, & W. SEITZ. 1981. Estimating wildlife habitat variables. FWS/OBS-81/47. Washington, D.C., U.S. Fish and Wildl. Serv.

HURLBERT, S. H. 1984. Pseudoreplication and the design of ecological field experiments. Ecol. Monogr. 54: 187–211.

HUTTO, R. L., S. M. PLETSCHET, & P. HENDRICKS. 1986. A fixed-radius point count method for nonbreeding and breeding season use. Auk 103: 593–602.

KRICHER, J. C. 1975. Diversity in two wintering bird communities: possible weather effects. Auk 92: 766–777.

NETER, J., & W. WASSERMAN. 1974. Applied linear statistical models. Homewood, Illinois, Richard D. Irwin.

OTT, L. 1988. An introduction to statistical methods and data analysis. Boston, Massachusetts, PWS-Kent Publ. Co.

RISKIND, D. H., & D. D. DIAMOND. 1988. An introduction to environments and vegetation. Pp. 1–15 in Edwards Plateau vegetation: plant ecological studies in central Texas (B. B. Amos and F. R. Gehlbach, Eds.). Waco, Texas, Baylor Univ. Press.

ROBBINS, C. S. 1971. Winter bird survey of central Maryland. Maryland Birdlife 27: 31–38.

———. 1972. An appraisal of the winter bird-population study technique. Am. Birds 26: 688–692.

———. 1981a. Effect of time of day on bird activity. Stud. Avian Biol. 6: 275–286.

———. 1981b. Bird activity levels related to weather. Stud. Avian Biol. 6: 301–310.

ROLLFINKE, B. F., & R. H. YAHNER. 1990. Effects of time of day and season on winter bird counts. Condor 92: 215–219.

SAS INSTITUTE, INC. 1985. SAS user's guide: statistics. Version 5 edition. Cary, North Carolina, SAS Institute, Inc.

SCOTT, J. M., & F. L. RAMSEY. 1981. Length of count period as a possible source of bias in estimating bird densities. Stud. Avian Biol. 6: 409–413.

SHIELDS, W. M. 1977. The effect of time of day on avian census results. Auk 94: 380–383.

SMITH, P. G. R. 1984. Observer and annual variation in winter bird population studies. Wilson Bull. 96: 561–574.

SOKAL, R. R., & F. J. ROHLF. 1987. Introduction to biostatistics. New York, W. H. Freeman.

VERNER, J. 1985. Assessment of counting techniques. Pp. 247–302 in Current ornithology, vol. 2 (R. F. Johnston, Ed.). New York, Plenum Press.

———. 1988. Optimizing the duration of point counts for monitoring trends in bird populations. Res. Note PSW-395. Berkeley, California, Pacific Southwest For. Range Exp. Sta.

———, & K. A. MILNE. 1989. Coping with sources of variability when monitoring population trends. Ann. Zool. Fennici 26: 191–199.

———, & L. V. RITTER. 1986. Hourly variation in morning point counts of birds. Auk 103: 117–124.

ZAR, J. H. 1984. Biostatistical analysis. Englewood Cliffs, New Jersey, Prentice-Hall.