THE EFFECTS OF VISIBILITY BIAS ON TIME-BUDGET ESTIMATES OF NICHE BREADTH AND OVERLAP

DAVID W. BRADLEY
Data Processing Department, California State University, Long Beach, California 90840 USA

ABSTRACT.—Four time-budget estimation strategies are compared with respect to their sensitivity to two components of visibility bias in the observation process: discovery bias and loss bias. Monte Carlo simulations and a brief field study both indicate that visibility bias (particularly discovery bias) can substantially affect the results of time-budget studies. Estimators designed to curtail these biases performed best. Counting only initial contacts was least satisfactory. Bootstrap confidence intervals for niche overlap from the field study were so broad that overlap estimates seem nearly useless with very small sample sizes, such as the 93 observation series with 1,065 data points obtained here. Investigators who measure time or energy budgets in the field should take care to minimize sample biases, obtain adequate sample sizes, select analysis techniques appropriate for their sampling scheme, and confine inference to a scope compatible with the temporal and spatial scale of their study. *Received 13 June 1984, accepted 27 November 1984.*

STUDIES of niche characteristics and energy budgets often depend on estimates of the proportion of time birds spend in arbitrarily defined activity states. These estimates may be biased by differences in observability of distinct activity states. Bias in the time-budget profiles also may adversely affect summary statistics such as niche breadth and overlap, or energy expenditure. Conclusions drawn from these data then may be undermined, since selection of different sampling techniques may induce substantially different results (Wiens 1983).

Altmann (1974), Dunbar (1976), and Tyler (1979) compared many time-sampling techniques and concluded that the most generally applicable methods for estimating time budgets are continuous time sampling and instantaneous (scan) sampling. Continuous time sampling requires precise recording of the transition times between activity states (using a stopwatch). Scan sampling records the state at fixed time points (e.g. every 20 s), which can be conveniently defined in the field by an electronic metronome (Wiens et al. 1970). Altmann (1974) discussed continuous sampling in the context of focal-animal sampling, where a single individual is followed throughout the observation period. In this paper, two data-gathering techniques are compared: (1) the "count all observations" method, where a focal-animal approach is used in combination with scan sampling; and (2) the "count first observation" strategy, where only a single, instantaneous record is made of the bird's activity state.

Wiens (1969) showed that the effects of visibility bias can cause count-first and count-all observation sampling to differ in estimated activity profiles. Easily observed activities, such as singing on wires and fence posts, were overrepresented in his count-first samples when compared with the count-all samples. Wagner (1981) pointed out that visibility bias may influence these two sampling techniques in the reverse direction as well. This reflects varying contributions from the two components that induce visibility bias. Activities may differ either in the ease with which a bird can be found (discovery bias) or the chances that it will be lost during the bout (loss bias). For example, during flight birds may be easily discovered but also frequently lost. Sleep fits the opposite pattern (difficult to find but seldom lost once found), while some activities, such as canopy foraging, may suffer downward bias from both components (rarely discovered and readily lost). Some of the most biased activities (e.g. flight and sleep) may not be relevant to studies of niche parameters but still would be important for energy expenditure analyses.

I investigated the effects of both of these aspects of visibility bias on time-budget estimates of activity profiles and niche breadth and overlap. Monte Carlo simulation is used to assess the accuracy of the two methods discussed above (count-first and count-all sampling) along

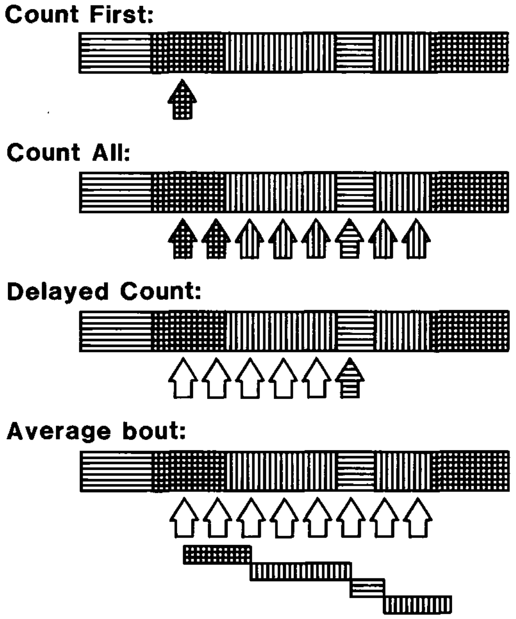

Fig. 1. Time-budget estimators used in this study. The separately hatched bars represent different activity states occupied by a bird through a short time period. With the count-first approach, only the state when the bird is initially found is recorded. Count all (scan sampling) records the state at regular intervals until loss of contact. The delayed count records the state at only a single time point, but after a fixed delay from the initial contact. The average-bout method records the bout lengths (periods with no change in state) and the transitions between states. For the first three methods, the time budget is estimated as the proportion of observations per state. The average-bout estimator is more complex (see text).

with two other methods (described below) that attempt to adjust these to reduce the effects of visibility bias. A small field study provides additional comparisons among the estimators.

METHODS

Time-budget estimators.—Four time-budget estimators are compared throughout this study (Fig. 1). Two of these, the delayed-count and average-bout methods, adjust the familiar count-first and count-all methods to reduce the impact of visibility bias. The delayed-count approach reduces the vulnerability of the count-first method to discovery bias by recording the activity state at a single point in time only after an arbitrarily selected delay from the time of initial contact. This study used a 2-min delay. The average-bout method is more complex.

If complete activity histories for all birds under study were available, the time budget could be derived directly from the sum of bout lengths for each activity state. This also could be computed for each state by multiplying its relative number of bouts by its average bout length. The average-bout method separately estimates these two components with methods designed to minimize bias and multiplies them together to obtain the time budget. The matrix of transitions between bouts is used to estimate the first component: relative frequency of bouts for each activity state. This is done by assuming a Markov process and computing the long-run distribution from the transition matrix (Isaacson and Madsen 1976). Because some bouts are truncated due to loss of contact with the bird, special care must be taken to estimate the second component (mean bout length) with minimal effect of loss bias. Use of the product-limit survival curve estimator for censored data achieves this objective without assuming any particular parametric distribution for the bout lengths (Kaplan and Meier 1958, Kalbfleisch and Prentice 1980).

Continuous time sampling would provide ideal data for the average-bout method. However, the definitions for activity states used in the field study (described below) require finite time intervals to identify a bird's current state. Consequently, instantaneous scan samples (taken at 20-s intervals) were recorded instead of continuous samples. The same data records provided time-budget estimates for all four methods, although the two single-count approaches used only part of each observation sequence.

Monte Carlo trials.—Unfortunately, for many species, such as arboreal birds, there is no way to obtain an exact measure of the time budget to serve as a criterion for evaluating the comparative accuracy of the estimators. As an alternative, computer simulation can provide data where the true time budget is exactly known. The observation process (including bias effects) also can be simulated, providing a basis for accuracy comparisons.

Time budgets in this study were simulated with a stochastic process that yielded a random series of bouts in arbitrarily defined activity states. Transitions between bouts were controlled by a Markov process from randomly generated transition matrices. Lengths of bouts were obtained from survival curves following a Weibull distribution, with separate values for the scale and shape parameters for each activity state. Use of Weibull distributions allowed the simulated states to exhibit differences with respect to both the mean and variance of bout lengths. The true time budget then was obtained from the product of the means for these distributions and the long-run distribution from the transition matrices.

The next step simulated the observation process by modeling an observer discovering the bird, following it as long as possible, and losing contact either during a bout or at the point where the bird flew away. Eight main activity states were used, with another added for flight. Loss of contact during a bout implemented loss bias when loss rates were different for separate activity states. Loss at transition to flight was considered an unbiased loss of contact (uncensored observations). Discovery of the bird was simulated from a random start by using exponential distributions for lag times between the start of a bout and discovery. If the lag exceeded the bout length, then the next state was examined. Discovery bias was implemented by using different values for the mean lag time per state.

Simulations were performed in a 2 × 2 design with 250 replications. The main effects were presence and absence of discovery bias and loss bias. Each replication consisted of a trial with 1,000 observation series (each running from discovery to loss of contact). The bias level was constant over replications. Consequently, these simulations indicate the effects at the chosen bias levels without comparison to possible alternatives of higher or lower levels of bias. Time-budget estimates were produced for each trial with each of the four methods and compared with true time budgets for the trial. The error rate of the estimates was measured as the root mean squared deviation between the estimated and true profiles. The results are presented as box plots (Tukey 1977), with vertical bars for the range, a box covering the quartiles, and a horizontal bar marking the median.

Additional simulation trials were run to investigate the effects of bias on estimates of niche breadth and overlap. Each replicate included generation of two actual time budgets (one per simulated species). Both biases were included in these simulations, which incorporated 100 trials of 1,000 observation series per species. Niche overlap was computed with Pianka's (1975) familiar method as the cosine of the angle between the resource-utilization profiles (Petraitis 1981). Inspired by the work of Petraitis (1981) and Smith (1982), niche breadth was computed relative to simulated profiles of available resources by the cosine of the angle between the two vectors, providing a generalization of the well-known reciprocal sum of squares measure to adjust for inequalities in state availability.

Field study.—The basic plan for the field study was to gather data reminiscent of that used in foraging-behavior studies but with activity categories designed to emphasize effects of bias. Three species, Yellow-rumped Warbler (*Dendroica coronata*), Ruby-crowned Kinglet (*Regulus calendula*), and Townsend's Warbler (*D. townsendi*), were observed in a 2.5-ha plot in Recreation Park, Long Beach, California (33°47′N, 118°8′W) between 10 January and 13 March 1982. Observations took place on 12 days between 0700 and 0900, averaging 1 h/day. A total of 93 series of bouts with individual birds incorporating 1,065 time-sampling points were recorded, with an average contact time of 4 min/bird.

Measurements were made with a 20-s metronome counter to provide timing signals and a Hewlett Packard 41C calculator attached to the binocular to record the data. Four attributes of the bird's state were recorded: two related to activity level and two substrate measures. Activity level was coded as high (movement rate exceeding 6 m/min, indicated by total movement of at least 2 m during the 20-s interval), low (maintaining a single perch), and medium (in between). In addition, presence or absence of vocalizations during the interval was recorded. The substrate was partitioned into gross categories of high or low, with a cut point of 6.7 m above the ground, using 6.7-m lampposts for reference. In addition, a coarse taxonomy of tree types was used: *Eucalyptus*, deciduous (without leaves during the study), broadleaf evergreen, and coniferous.

Vegetation at the study site consisted of a wide variety of ornamental trees. To make estimates of niche breadth relative to available resources, the total volume of foliage in each substrate category was estimated. Each of the 139 trees in the plot was measured, with maximal crown radius (in meters) estimated by pacing. The percentage of 6.7-m high cylinders of this radius occupied by foliage was estimated separately for low and high height classes by visual assessment. Percentages of the total foliage volume in the eight available categories were computed from these data and showed *Eucalyptus* with 6% (low) and 40% (high), deciduous with 11% each for low and high, broadleaf with 10% (low) and 8% (high), and coniferous with 6% (low) and 8% (high).

Confidence intervals for niche breadth and overlap for these species were obtained with bootstrap resampling using the percentile method generally described by Diaconis and Efron (1983), with additional details in Efron (1983) and an extensive technical treatment in Efron (1982). This is essentially the strategy suggested by Ricklefs and Lau (1980). It proceeds by drawing random samples with replacement from the observed data to simulate the sampling variation that would arise by repeated sampling from the true (infinite) population under study. Estimated sampling distributions for niche breadth and overlap were obtained in this way, with 95% confidence intervals determined from the 2.5 and 97.5 percentile points on these distributions. The bootstrap has the advantage that the sampling distribution is derived empirically without assuming any particular parametric form for the true distribution of the data. A second, more important advantage is that it can be readily adapted to complex sampling situations, taking into account dependencies among successive observations. This was achieved here by considering each entire series of observations from the initial sighting to loss of contact as a single point for resampling, rather than resampling on the basis of individual time points.

RESULTS

The accuracy of the four methods for estimating time-budget profiles is compared under four different simulated bias conditions in Fig.

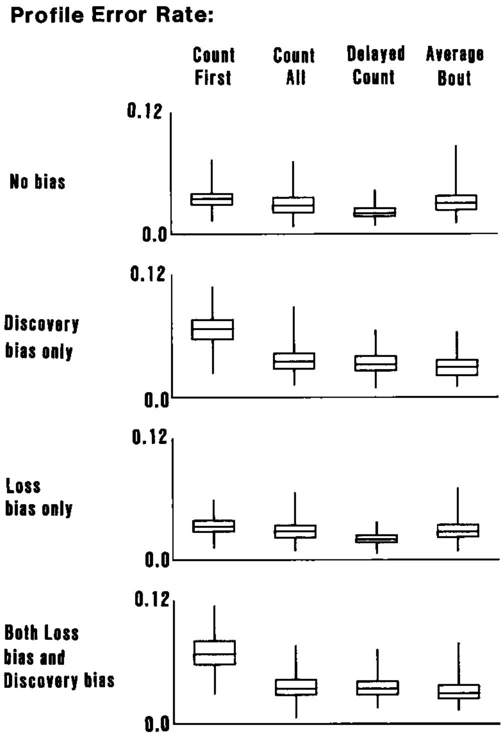

Fig. 2. Profile error rates for Monte Carlo trials with four combinations of simulated discovery and loss bias. The boxes span the quartiles of the distributions of error rates obtained from 250 trials, with a bar at the median. The vertical lines cover the full observed range of error rates. The four estimators employ the methods described in Fig. 1, with error rates calculated as the root mean squared deviation of the estimated profile from the true profile for the trial (exact proportions of time the simulated birds would spend in each activity state).

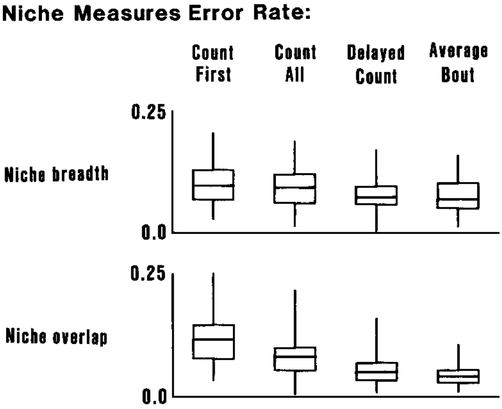

Fig. 3. Niche measures error rates for simulations including both loss and discovery bias. The error rates and box plot structure correspond to Fig. 2, although only 100 Monte Carlo trials were used here. Niche breadth was measured for a single simulated species relative to a simulated profile of available resources. Niche overlaps were based on two simulated species. Error rates were calculated from the (exactly known) activity profiles used to control the simulations.

2. The results seem paired, with similar patterns arising for no bias and loss bias, while the discovery bias results closely matched those with both biases. Evidently, loss bias had little effect in these simulations. Discovery bias was more prominent. In its absence, all four methods performed about equally well. However, presence of discovery bias substantially increased the error rate for the count-first method but had little deleterious effect on the other methods.

The methods are compared with respect to niche breadth and overlap in Fig. 3. Results are displayed for the simulations that included both biases. Differences in accuracy for niche breadth were small, with the delayed-count and average-bout methods performing the best. Similar but more pronounced differences arose for niche overlap. In both cases, the count-first method had the highest error rates. These results confirm that effects of bias on the raw profiles also can affect the estimates of niche breadth and overlap.

Field estimates of time budgets for both activity and substrate also were computed with all four methods. The methods did show differences in profiles, which are best illustrated with data from the most numerous species, the Yellow-rumped Warbler. Profiles for the activity patterns using six states obtained from all combinations of calling/silent with slow/medium/fast movement rates are given in Table 1. The results are presented with the average-bout method as a baseline and the other three methods as deviations from the average-bout results. The most frequently employed activities were calling with medium and fast movement. The average-bout method showed the highest of the four estimates for slow activities.

The greatest differences from the average-bout results occurred with the count-first method. It showed particularly high values for calling/fast and relatively low values for slow activities. Most dramatic is the entire lack of observations with slow, silent movement for

TABLE 1. Activity profile for Yellow-rumped Warbler.

	Deviations from average-bout estimates			Actual profile
	Count first	Count all	Delayed count	Average bout
Slow				
Silent	−18	−5	−8	18%
Calling	−3	−1	−3	7%
Medium				
Silent	+5	+2	+11	9%
Calling	+5	+1	−1	33%
Fast				
Silent	+1	0	+1	5%
Calling	+10	+3	0	28%

the count-first method. This pattern follows precisely what would be expected from the effects of discovery bias.

Estimates of niche breadth and overlap for all four methods are shown in Fig. 4. Estimates from the four methods seem similar, although there are some differences in value that are sufficient to alter even the rank orderings. The greatest interspecies differences occurred in substrate, with high values for Yellow-rumped Warbler in breadth and for kinglet-Townsend's overlap. However, the most striking aspect of Fig. 4 is the broad range covered by many of the confidence intervals. In some cases the intervals cover nearly the entire range of possible values, indicating that the data are inadequate to precisely estimate the true niche overlaps.

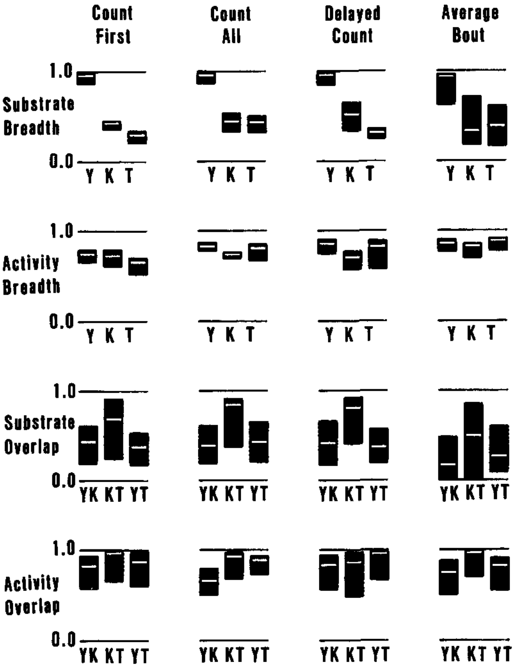

Fig. 4. Niche breadths and overlaps for the field study as estimated with four sampling strategies. Separate panels appear for activity category and for substrate (with niche breadth for substrate considered relative to available resources). The boxes represent 95% bootstrap confidence intervals for the estimates. The white lines inside the boxes mark the actual estimates obtained with the field data. Y = Yellow-rumped Warbler, K = Ruby-crowned Kinglet, and T = Townsend's Warbler. Overlaps are pairwise (e.g. YK = Yellow-rumped Warbler overlap with Ruby-crowned Kinglet).

DISCUSSION

Clearly, visibility bias can degrade estimates both of time-budget profiles and of summary statistics such as niche breadth and overlap. Yet differential visibility is not the only factor that could affect time-budget estimates. As in avian censusing (Scott and Ramsey 1981), bias caused by attraction to or avoidance of the observer by the bird also could skew the profiles. Other factors (e.g. season, time of day, habitat, and weather) may influence bird visibility (Anderson et al. 1981; Best 1981; Ekman 1981; Robbins 1981a, b). However, in the simulations and limited field study reported here, the effects of discovery bias seem predominant. In contrast, loss bias appeared paramount in Wagner's (1981) study of small passerines.

High sensitivity to discovery bias caused the count-first method to perform poorly in the simulations. The delayed-count approach fared much better. Discarding observation sequences that failed to reach the delay time may induce loss bias similar to the problems with the count-all estimates in Wagner's (1981) study. However, inclusion of short series could introduce discovery bias. It may not be possible to design a single-count strategy immune to both biases simultaneously.

The count-all strategy avoids discovery bias only by overwhelming the potentially biased early observations with far more numerous subsequent observations in each sequence. Additional protection against discovery bias could be obtained by discarding the first few observations in each sequence (Wiens 1983). This approach was used by Wagner (1981) and East (1982). Conceivably, this may actually increase the overall visibility bias by discarding observation series that disproportionately fail to reach the minimum required time due to difficulty in maintaining the contact (Wagner 1981). This would occur if increased loss bias compensated for reduction in discovery bias.

The average-bout method adjusts for loss bias by using techniques specifically developed for right-censored data (where some bouts cannot be followed to the end). Discovery bias is reduced by discarding left-censored bouts (those encountered after having begun). This method performed best in the simulations, as would be expected since the model employed in the simulation perfectly fits the requirements of the average-bout approach (that the bout length and next transition depend only on the current activity state and are independent of previous bout history). Violations of these assumptions in actual field situations could reduce its effectiveness. Heavy loss bias also could cause problems with downward bias of bout lengths based on the product-limit estimator (Dixon 1983). Perhaps the most serious problem with the average-bout method is the deleterious effect of the discrete time scale used with scan sampling. The lack of precision in time measurements might bias the bout length estimates (Tyler 1979), and very short bouts could be missed altogether. This method should properly be used with exact measurement of the beginning and end of each bout.

Selection of the best estimation technique depends upon the situation under study. In open areas where long records may be easily obtained but birds are sparse, single-point sampling would be unproductive. Scan sampling with the count-all method or continuous time sampling with the average-bout method would be more effective. In dense forests, however, long series may be difficult to maintain and loss bias may be intense. Here a single-count strategy, such as the delayed-count method, may be the best choice. The single-point approach has the advantage that it is amenable to use with standard statistical analysis techniques, because separate observations are independent. The count-all and average-bout methods require special methods such as the bootstrap (Efron 1982, 1983; Diaconis and Efron 1983) or the randomization test (Sokal and Rohlf 1969, Edgington 1980).

One problem that is readily apparent with the field data used for this study is the large uncertainty in estimates of niche breadth and overlap indicated by the broad confidence intervals in Fig. 4. Actually, uncertainty in substrate niche breadth may be underestimated, because it does not take into account sampling variation in the measures of resource availability. The enormous confidence intervals for niche overlap defy attempts to infer patterns from these data. Certainly, much larger sample sizes are needed before useful generalization from the data is possible.

Even with more data, however, the nature of this field study would limit the scope of reliable generalization. By restricting observations to a single site, season, and time of day, the data remain largely homogeneous, precluding estimation of spatial or temporal variance. Thus, the confidence intervals in Fig. 4 indicate how the results might have differed had observations been taken at comparable times (e.g. weekdays instead of weekends) during the same year and season at the same site. They do not indicate the range of possible responses to different conditions, which presumably would be even larger. Replication at additional sites and times is a necessary prerequisite to generalization beyond this one set of study conditions. As Wiens (1981) has shown, temporal and spatial scale are vitally important in the interpretation of survey results.

Thus, there are at least four prerequisites to reliable inference from avian survey data. First, the scale of the study (in both space and time) must be carefully designed and explicitly stated. Second, observation techniques should be selected that minimize the impact of sampling biases. Third, adequate sample sizes must be obtained (with respect to the number of sites and seasons studied, as well as the number of data gathered per site). Finally, analysis methods must be used that are appropriate for the sampling scheme, taking into account complications such as dependence among successive observations in a single sequence.

ACKNOWLEDGMENTS

I wish to thank Richard Bradley, John Wiens, and other members of the Friday afternoon seminar at the University of New Mexico for helpful suggestions in the early stages of this project. In particular, I thank Bob Pietruszka for extensive discussions of time-sampling techniques and suggestions for improving the paper. I also thank two anonymous reviewers for helpful comments on the manuscript. Computer time was provided by the data processing department at California State University, Long Beach.

LITERATURE CITED

ALTMANN, J. 1974. Observational study of behavior: sampling methods. Behaviour 49: 227-267.

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

BEST, L. B. 1981. Seasonal changes in detection of individual bird species. Stud. Avian Biol. 6: 252-261.

DIACONIS, P., & B. EFRON. 1983. Computer-intensive methods in statistics. Sci. Amer. 248: 116-130.

DIXON, W. J. 1983. BMDP statistical software. Berkeley, California, Univ. California Press.

DUNBAR, R. I. M. 1976. Some aspects of research design and their implications in the observational study of behaviour. Behaviour 58: 78-98.

EAST, M. 1982. Time-budgeting by European Robins *Erithacus rubecula*: inter- and intrasexual comparisons during autumn, winter, and early spring. Ornis Scandinavica 13: 85-93.

EDGINGTON, E. S. 1980. Randomization tests. New York, Marcel Dekker, Inc.

EFRON, B. 1982. The jackknife, bootstrap, and other resampling plans. SIAM, Monogr. No. 38, CBMS-NSF.

———. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. Amer. Stat. 37: 36-48.

EKMAN, J. 1981. Problems of unequal observability. Stud. Avian Biol. 6: 230-234.

ISAACSON, D. L., & R. W. MADSEN. 1976. Markov chains, theory and applications. New York, John Wiley & Sons.

KALBFLEISCH, J. D., & R. L. PRENTICE. 1980. The statistical analysis of failure time data. New York, John Wiley & Sons.

KAPLAN, E. L., & P. MEIER. 1958. Nonparametric estimation from incomplete observations. J. Amer. Stat. Assoc. 53: 457-481.

PETRAITIS, P. S. 1981. Algebraic and graphical relationships among niche breadth measures. Ecology 62: 545-548.

PIANKA, E. R. 1975. Niche relations of desert lizards. Pp. 292-314 *in* Ecology and evolution of communities (M. L. Cody and J. M. Diamond, Eds.). Cambridge, Massachusetts, Belknap Press.

RICKLEFS, R. E., & M. LAU. 1980. Bias and dispersion of overlap indices: results for some Monte Carlo simulations. Ecology 61: 1019-1024.

ROBBINS, C. S. 1981a. Effect of time and day on bird activity. Stud. Avian Biol. 6: 275-286.

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

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.

SMITH, E. P. 1982. Niche breadth, resource availability, and inference. Ecology 63: 1675-1681.

SOKAL, R. R., & F. J. ROHLF. 1969. Biometry. San Francisco, W. H. Freeman and Co.

TUKEY, J. W. 1977. Exploratory data analysis. Reading, Massachusetts, Addison-Wesley.

TYLER, S. 1979. Time-sampling: a matter of convention. Anim. Behav. 27: 801-810.

WAGNER, J. L. 1981. Visibility bias in avian foraging data. Condor 83: 263-264.

WIENS, J. A. 1969. An approach to the study of ecological relationships among grassland birds. Ornithol. Monogr. 8: 1-93.

———. 1981. Scale problems in avian censusing. Stud. Avian Biol. 6: 513-521.

———. 1983. Avian community ecology: an iconoclastic view. Pp. 355-403 *in* Perspectives in ornithology (A. H. Brush and G. A. Clark, Jr., Eds.). Cambridge, England, Cambridge Univ. Press.

———, S. G. MARTIN, W. R. HOLTHAUS, & F. A. IWEN. 1970. Metronome timing in behavioral ecology studies. Ecology 51: 350-352.