Studies in Avian Biology No. 6:230–234, 1981.

PROBLEMS OF UNEQUAL OBSERVABILITY

JAN EKMAN[1]

ABSTRACT.—Generally the assumption of equal observability (probability of capture) forms the basis for the application of census models: (a) observability must remain stable if results from direct counts are to be compared and give a true picture of changes in number; and (b) capture-recapture models require equal capture probability of individuals (homogeneity) to yield unbiased estimates.

Errors arising when the conditions under (a) are not met could be seasonally changing behavior. Line transect indices of Willow Tits increased in late winter to early spring, although the population was apparently stable and no immigration occurred, as shown by detailed capture-recapture estimates from individually color-banded birds.

Conditions under (b) may not be met when behavior differs between individual categories. In groups of color-banded Willow Tits using the same area, some individuals were consistently less liable to observation and identification than others. These differences in observability were traced to height separation while foraging in trees.

Census models usually require that capture probability (observability) does not vary in time, between habitats (line transect census), or between individuals (capture-recapture models). The application of capture-recapture models may produce serious negative biases when the capture probability differs between individuals (heterogeneity) (Gilbert 1973). As biologists, however, we acknowledge individual variation in morphological traits and behavior as the very basis of natural selection and evolution. On this basis we should hardly expect the "equal catchability" assumption to be met in reality, and our confidence in census data has to be founded upon how robust models are to violations of their underlying assumptions. Gilbert (1973) and Carothers (1973) could, for instance, demonstrate that some capture-recapture models produce accurate estimates under certain conditions even when the individuals differ in their capture probabilities.

Unequal capture probability is not itself the only cause for the lack of accuracy in census estimates. The magnitude of the difference in individual capture probability, as well as the average capture probability for the entire population, are further critical attributes (Gilbert 1973). Successful application of census models will therefore not only require information on the heterogeneity itself, but also information on the distribution of capture probabilities, and preferably its behavioral bases.

Methods to identify and quantify heterogeneity are, however, poorly developed. Tests to reveal heterogeneity (Leslie 1958, Keith and Meslow 1968, Carothers 1971) are insensitive and their statistical justification has been queried (Roff 1973), but only recently have advances been made to quantify the effects of unequal catchability (Carothers 1979).

This work centers on attempts to identify and quantify heterogeneity in the capture probability of Willow Tits (*Parus montanus*) during census work in a population study. The approach developed here was to independently study the behavior of known individuals, rather than to start from the census data themselves. Further, different census techniques were operated simultaneously, and their results analyzed for temporal heterogeneity in capture probability.

METHODS

Willow Tits were studied in mature (about 70 years old) coniferous forest some 40 km east of Göteborg (Gothenburg), SW Sweden. The populations were censused either by line transect counts or by capture-recapture.

LINE TRANSECT

Willow Tits were censused along a 12.3 km trail. Between November 1968 and November 1975 this trail was censused by up to eight observers at the end of each month. From November 1975 to November 1978 censuses were conducted only every third month, but the trail was censused more times on each occasion. During census periods at the end of the months, only two observers censused on the same day. They usually started at the same place and time, moving in opposite directions. The census trail formed a loop, and both observers censused the entire trail. Censuses started within one hour after dawn, and, depending on weather conditions, took between four and six hours to walk (roughly 35 to 50 m/min). At the halfway point there was a lunch and resting break for 15 to 30 minutes. During periods with thick snowlayer the censuses were conducted on skis. As only two observers censused on the same day, each census period consisted of several days. When weather conditions allowed, census days were consecutive. However, no censuses were conducted when wind velocities exceeded 10 m/sec. We had no lower temperature limit where census work ceased. In practice the lowest temperatures encountered were around −15°C. All birds heard or seen were noted, regardless of their distance from the trail.

[1] Department of Zoology, University of Göteborg (Gothenburg), Box 250 59, S - 400 31 Göteborg, Sweden.

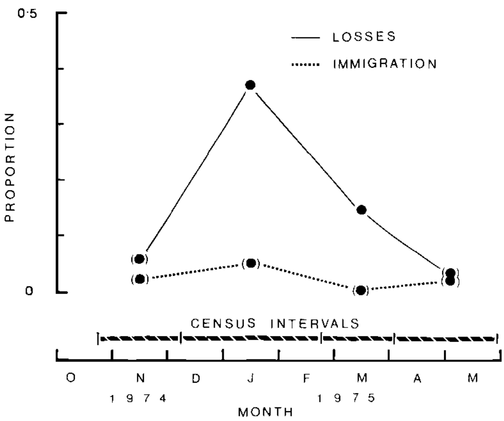

FIGURE 1. Disappearance and immigration rates of Willow Tits in winter (Jolly-Seber estimates. Brackets mark estimates not significantly larger than zero).

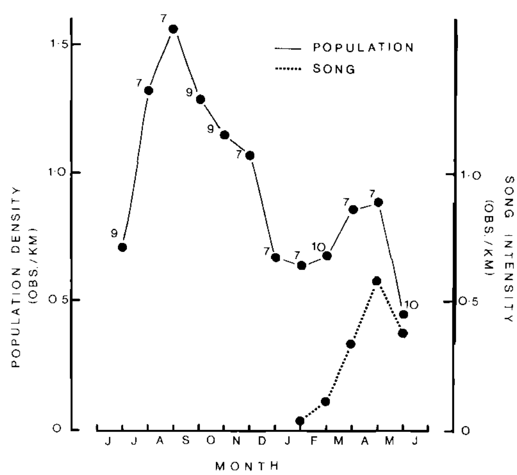

FIGURE 3. Pattern of seasonal change in line transect data for the Willow Tit. (Mean values for ten years census data. The figures show how many years data each point is based on.)

Willow Tits in the census trail area were never captured (or banded), so their observability was unaffected by human handling.

Capture-Recapture

Beginning in 1974, all Willow Tits were continuously banded in an 8 km^2 area whose southern border was roughly 400 m to the north of the census trail loop. Each Willow Tit received a unique combination of colorbands. This population was visited at bimonthly intervals, and samples of recaptures were collected by remote identification of their color combinations. Individuals could be, and usually were, observed and identified several times during each sampling occasion. However, no account was made for such repeated identifications in the capture-recapture sequence; it was only noted if an individual had or had not been identified. From the re-observation sequences, population size, survival and immigration was estimated, together with their variances, by the Jolly–Seber capture-recapture model (Jolly 1965, Seber 1965) using a computer and a slightly modified FORTRAN program (Davies 1971).

Winter survival in Willow Tits was age-specific (Ekman et al. in press) to an extent that necessitated separate calculations of adult and juvenile estimates (Manly 1970). Therefore no overall estimates of variance are available for population parameters as a whole, and tests for population changes had to be performed separately for adults and juveniles. Significant changes refer to t-tests with $P < 0.05$ within any of these groups.

Adult Willow Tits are highly sedentary (Ekman 1979), and all immigrants were therefore treated as juveniles. At the age of one year Willow Tits were considered recruited to the adult cohort.

The study area was provided with nest boxes; these were checked and natural nests searched for. The total number of nests found each year provided an independent control of our capture-recapture estimates of the breeding population.

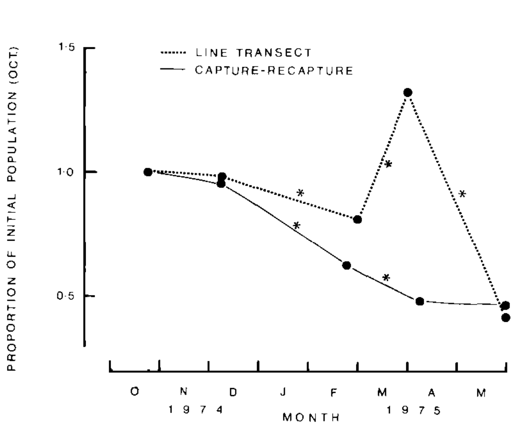

FIGURE 2. Seasonal changes in Willow Tit numbers measured by line transect census and capture-recapture. (Significant change between two samples indicated by asterisk by the line.)

Behavioral Studies of Observability

Willow Tits organize into small groups in winter. The groups contain a stable set of individuals and use restricted bordering, but non-overlapping, winter ranges (Ekman 1979). These groups, due to their sedentariness and stability in composition, provide excellent opportunities for studies of individual behaviors and mutual relationships. Registrations of observability were collected from such groups with known composition of individuals by keeping them under continuous surveillance for some hours while collecting identifications at random. After the identification of a banded bird, no further registrations were collected until the last identified bird had left the tree where it had originally been spotted. I then deliberately looked away so as not be see where it landed.

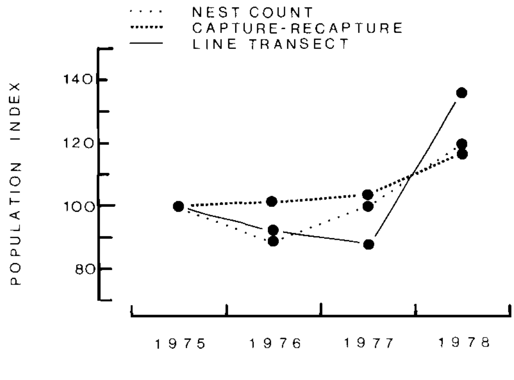

FIGURE 4. Annual fluctuations in Willow Tit numbers measured by three different methods.

I then resumed searching and the very first flock member spotted was identified. Identifications were registered on a tape recorder, which also gave the sequence in which group members were identified.

Simultaneously with the identification, the height at which the birds were first spotted was recorded, with height described as one of five height classes of the tree in which a bird was found.

RESULTS

SEASONAL DIFFERENCES IN OBSERVABILITY

Jolly-Seber estimates from the color-banded Willow Tit population could not verify any immigration in winter (Fig. 1). Hence, the population censused by capture-recapture steadily declined as the winter progressed, and the rate of this decline was determined solely by losses in the study population (Figs. 1 and 2). Presumably these losses were caused by natural mortality, as several tits were recovered dead (Ekman et al. in press), but no emigrants were found in neighboring areas.

Line transect censuses in the immediately adjacent forest simultaneously yielded increasing indices for Willow Tits in late winter and early spring (Fig. 2). Conceivably, the increasing line transect indices were caused by enhanced observability among local survivors, as we have no evidence for immigration to the banded population during the same period. Neither had there been any emigration from the banded population to the north into the census trail area, as no banded Willow Tits were ever found along the census trail. Hence, there are good reasons to believe that we have a bias in the line transect count censuses with seasonally changing observability. The conclusion that this is a consistent bias is reinforced by the fact that the spring increase of Willow Tit line transect indices is a regular phenomenon clearly reflected in the mean of ten years of census data (Fig. 3).

TABLE 1
OBSERVABILITY DIFFERENCES BETWEEN WILLOW TIT GROUP MEMBERS (ONE SELECTED EXAMPLE)

Individual category	Relative capture probability	Number of identifications[a]
Female, adult	0.183	37
Male, adult	0.208	42
Female, juvenile	0.277	56
Male, juvenile	0.332	67

[a] Distribution of identifications differs significantly from a random expectation ($\chi^2 = 12.04$, $P < 0.001$).

The increased observability of Willow Tits in spring could conceivably be connected to the onset of sexual activities and singing. The increase is fairly nicely timed to the increase in singing activities (Fig. 3), but other factors may also be involved, since song intensity reaches its peak in April/May without any corresponding steep increase in the population index in April. As the arthropod populations start to build up after the winter, a more mobile hunting strategy could pay for the tits (Norberg 1977), which would make encounters with a censusing observer more likely. Further, territorial defense may also call for conspicuous movements between different sections of the borders.

Data from line transect censuses are susceptible to changes in observability. This apparently makes them less powerful a tool to follow seasonal changes of Willow Tits. This objection does not, however, detract from the usefulness of line transect data in reflecting annual changes. Data obtained at the same phase of a seasonal cycle should still give an accurate picture of annual fluctuations. This point can be illustrated by comparing annual fluctuations in breeding numbers (early May) of Willow Tits estimated by different methods; (a) line transect census, (b) capture-recapture, and (c) nest counting. These three methods provide concordant patterns of annual fluctuation (Fig. 4).

INDIDIVUAL DIFFERENCES IN OBSERVABILITY

Winter social groups of Willow Tits are fairly uniform in composition, usually consisting of an adult pair and a juvenile pair (Ekman 1979). Group members are together most of the time, but are not equally liable to identification (Table 1). The difference in observability between group members did not vary at random. Old birds, particularly females, were observed and identified less often than young birds. Old females yielded the fewest identifications in all six four-member groups studied ($P = 0.0009$, Randomization test), and adults consistently yielded fewer observations than younger birds ($P = 0.03$, Randomization test).

The lower number of identifications for old females is due to low observability. An alternative hypothesis—that she might stick less to the group—is rejected because the identification sequences of other group members between consecutive identifications of old females do not depart from a random distribution (goodness-of-fit test; variance/mean-ratio = 1.025), indicating that she was present during the entire observation period. Otherwise we would have expected a bimodal distribution of the observation sequences of other group members, with one peak representing observations when old females were around, and one representing observations when they were away.

Low female observability is also suggested by their tree use pattern. While juvenile group members usually forage in the lower parts, old females are more frequently found in the upper sections, where the birds are more difficult to spot (Table 2).

TABLE 2
OBSERVATION HEIGHT OF DIFFERENT WILLOW TIT GROUP MEMBERS

	Tree height pentile	Frequency distribution of identification height[a] Female, ad.	Male, ad.	Juveniles
Top	5	0	0	0
	4	0.15	0.01	0.02
	3	0.64	0.45	0.11
	2	0.16	0.44	0.48
Bottom	1	0.05	0.11	0.38
Total number of observations		61	85	89

[a] Differences in distribution of heights differ significantly from a random expectation ($\chi^2 = 80.99$, df = 6, $P < 0.001$).

DISCUSSION

Census models can produce misleading and seriously biased estimates when the underlying assumptions fail. In this study it is obvious that Willow Tit line transect data cannot be used for studies of population processes occurring within one season. For instance, it is not possible to apply k-factor analysis since the slope of the line where survivors are regressed against initial population will be altered from the true relationship if population estimates do not have the same bearing on numbers.

Other methods, like the capture-recapture model, will not break down entirely when the assumptions fail. Unequal catchability will introduce only insignificant bias in Jolly-Seber estimates when the capture probability exceeds 0.5 (Gilbert 1973). Proper adjustment of the sampling design, however, requires that the heterogeneity can be identified and, preferably, also quantified. In this context, the failure of traditional "equal catchability" tests (Roff 1973) is distressing. If heterogeneity can be identified and sampling design is adjusted properly, unbiased estimates can still be obtained for separate population strata. For Willow Tits, direct behavioral observations, instead of tests on census data, proved a covenient and powerful tool to demonstrate unequal observability. The reobservations from a population with strata differing in capture probability, but where capture probability follows a Poisson distribution within strata, will produce a compound Poisson distribution when pooled (Feller 1970). Roff (1973) stressed that "equal catchability tests" will fail to identify strata with differing capture probability for this reason. The tests performed on data collected from Willow Tit flocks, where members of known age, sex and rank are compared on an individual basis, evade this objection since the only variation included is between single individuals from different strata.

On basis of what is known about the biological basis of unequal catchability in Willow Tits it is possible to elaborate the sampling (here reobservation procedures) to reduce the effect of heterogeneity. Sampling Willow Tit observations involves two steps: (a) localizing the flock, and (b) identifying the flock members. As an overwhelming majority of flocks contain the strata treated here (adults, juveniles, males, females), and have largely the same composition, differences in catchability are encountered during the identification of flock members. Hence, it is the procedures used once a flock is found which determine how observability differences will affect the total recapture material. Assume that our sampling schedule allows for repeated identification of the same individual. From the data embodied in Table 1 on the relative difference in observability of flock members, it is possible to calculate the number of observations necessary in a flock of given (or guessed) size to yield observations of the least observable individuals with a given probability. For instance, if we want to observe adult female Willow Tits with a probability of 0.5 in flocks of four individuals, we can extract the relative adult female observability from Table 1 (0.183) and calculate the required number to approximately four by substituting these values into the expression for a Poisson distribution ($0.5 = 1 - e^{-0.183\,n}$, where

$e^{-0.183\ n}$ is the probability of finding no adult female (zero term of Poisson distribution) in n identifications). At capture, probabilities exceeding 0.5 estimates will be only marginally biased by unequal catchability (Gilbert 1973).

With calculations such as these, it is possible to elaborate sampling procedures and data processing, by e.g., increasing sampling intensity or treating subgroups separately, to alleviate the consequences of heterogeneity. The procedural rationale used here is to study the behavior of individuals in situations where environmental "noise" can be eliminated, because unequal catchability is an attribute of individuals, and has a behavioral basis. The power of this approach is demonstrated by the ease by which heterogeneity was demonstrated for the Willow Tit. With the biological basis of heterogeneity known, the logical corollary of quantifying the individual differences would simply entail a sampling of the kind of data presented in Table 1 for a number of groups. Further, identification of the behavioral basis of heterogeneity provides an opportunity to consider the behavioral plasticity of individuals. Therefore, this approach also has the potential of understanding the dynamics of heterogeneity. For instance, how would observability of Willow Tits alter if flocks broke up? In our census models we will have to represent the dynamics of heterogeneity and these dynamics cannot be represented realistically by simple correction factors based on comparisons of different census models applied on the same population. What we need to know is under which conditions behavior, and heterogeneity, change.

The approach of looking at behavior proved powerful for Willow Tits where several population strata (or presumably all, since flocks are usually of identical composition) are present in the same flock, and can be readily compared. The same method should apply to many social species where population strata can be found in company. Differences in behavior between individuals of species living solitarily are less tractable.

ACKNOWLEDGMENTS

My thanks are due to C. Askenmo, G. Cederholm, A. Enemar, and U. Unger at the Department of Zoology, University of Gothenburg, who all had a large share in planning and conducting these studies, and to all those ornithologists who in their spare time indulged in line transect census in blazing sunshine or in the cold of the winter. I owe a special thanks to C. Kepler for his patience in turning my writing into readable English.