Studies in Avian Biology No. 6:502–508, 1981.

THE EFFECT OF GROUP SIZE ON LINE TRANSECT ESTIMATORS OF ABUNDANCE

TERRANCE J. QUINN II[1]

ABSTRACT.—Line transect methodology is appropriate for transect experiments where some measure of distance is made to the animal that is sighted or flushed. This methodology is extended to populations where animals are sighted in groups (schools, flocks, etc.). Thus, the probability of sighting increases as a function of the size of the group. The first method presented for such sighting data pools the data over group size and uses a line transect model that is appropriate to fit the data. The estimate of the number of groups is approximately unbiased provided a flexible estimator is chosen. The estimator of the number of individuals is the product of the estimated number of groups and the estimated average group-size. The estimated average group size must be weighted to account for the increased probability of sighting larger groups. The second method presented post-stratifies the sighting data by group size and then proceeds as in the first method. The two methods are evaluated theoretically and by computer simulation. The method of post-stratification produces estimates that are closer to the true value but have larger variances than the method of pooling.

Transect methods for the estimation of animal abundance have been carried out for many years on a variety of species. These surveys have usually been designed as strip transect surveys, defined by a fixed width from the transect line wherein all animals were thought to be seen, or index surveys, where all animals sighted are counted and the results are interpreted as relative indices between years or regions. Population estimates can be obtained only for the strip transect surveys and are calculated intuitively from extrapolating the number sighted in the strip to the entire population area. Although distances to sightings have been measured occasionally, they usually have been used for checking that all animals are sighted in the strip. Some heuristic estimators using distances have been developed (Amman and Baldwin 1960, J. T. Emlen 1971, see Gates 1979 for others) but lack of statistical formulation has prevented assessment of an estimator's properties.

Incorporation of measured distances into the experimental design of transect experiments forms a powerful technique for estimating abundance called a line transect experiment. The roots of its methodology are contained in statistical models for sampling theory and recent advances in non-parametric density estimation and robust estimation, as well-described in recent reviews (Eberhardt 1978, Gates 1979, Quinn and Gallucci 1980, Burnham et al. 1980). The focus of the methodology is to construct a sighting model from the measured distances to correct for animals that are overlooked. The sighting model $g(y)$ is the non-increasing probability of a sighting at perpendicular distance y from the transect line, and animals on the transect line are assumed to be sighted with probability 1 (i.e., $g(0) = 1$). The strip transect method is a special case of the more general line transect sampling methodology (Seber 1973).

When a population is made up of groups (i.e., schools, flocks, herds) of varying sizes, line transect methodology is still appropriate, but with some modification. The purpose of this paper is to describe and compare two methods to analyze data from populations where sightings are made in groups. The key concept in the methodology is that the probability of sighting is likely to be an increasing function of group size that need not be linear. Empirical experiments on porpoise populations support this assertion (R. Holt and J. Powers, in prep.).

The general estimation framework for line transect methodology is briefly reviewed below. Three sighting models are described which represent common classes of estimators for line transect sampling. The two methods of analyzing transect data from populations of groups are also discussed below. The first method is to pool the transect data over groups of all size classes in order to estimate the total number of groups. The estimator from this method is robust, because the pooled sighting model is self-weighted by the true relative abundance of each group-size class in the population (Quinn 1979, Burnham et al. 1980). In the second method, the total sample of n sightings of groups is partitioned into t group-size classes. This method is referred to as post-stratification, because the total sample is partitioned after the completion of the survey, rather than taking an independent sample of each group-size class. The number of groups in each class is estimated and the estimates are summed to get the total number of groups in the population. The salient estimation formulae for both methods of the total number of groups, the total number of individuals and their variability are presented from Quinn (1980). I will compare the two methods with computer simulation and theoretically. In the last section below, the results are discussed in terms of specific recom-

[1] International Pacific Halibut Commission, P.O. Box 5009, University Station, Seattle, Washington 98105.

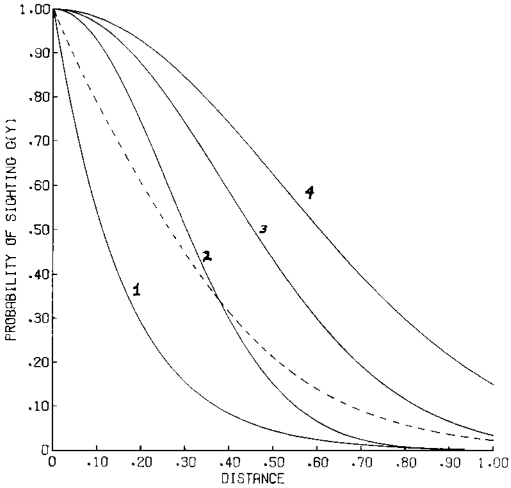

FIGURE 1. Plot of individual sighting models (solid lines) and the resultant pooled sighting model (dashed line) for the text example.

mendations for the planning and analysis of transect data from populations of groups.

ESTIMATION

For a line transect experiment on a population of individuals, the estimator of the number of individuals N in the population is given by

$$\hat{N} = \frac{An}{2L}\hat{c}^{-1} = \frac{An}{2L}\hat{f}(0), \quad (1)$$

where A is the population area, L is the transect length, c is the effective half-width sampled [defined by the integral of $g(y)$], $f(0)$ is the probability density function of sightings evaluated at the origin, and a caret ($\hat{\ }$) indicates an estimate (Quinn and Gallucci 1980, Burnham et al. 1980). The number of sightings is extrapolated to the total number by the ratio of the population area and the effective area sampled $2L\hat{c}$ (Quinn and Gallucci 1980).

The estimated variance of (1) is

$$\hat{\mathrm{Var}}(\hat{N}) = \hat{N}^2[\hat{\mathrm{c.v.}}^2(n) + \hat{\mathrm{c.v.}}^2(\hat{c}^{-1}|n)]. \quad (2)$$

The term $\hat{\mathrm{c.v.}}^2(n)$ is the estimated squared coefficient of variation [i.e., $\hat{\mathrm{Var}}(n)/n^2$] of the number of sightings obtained from subsampling or jackknifing. The term $\hat{\mathrm{c.v.}}^2(\hat{c}^{-1}|n)$ is the estimated squared coefficient of variation of the inverse of the estimated effective half-width, which is a derived formula from the sighting model (Quinn and Gallucci 1980, Burnham et al. 1980).

Three sighting models are used in this comparative study which are representative of available models. They are:

(1) The exponential model (EM)—a one-parameter model which postulates a sharp spiked decrease in sighting probability as distance from the transect increases; thought useful for flushing birds (Gates 1979, Eberhardt 1978).

TABLE 1
PARAMETERS AND MODELS USED IN THE SIMULATION STUDY[a]

Group-size class i	1	2	3	4
S_i	5	25	125	625
N_i/N	.4	.3	.2	.1
$c_i(\alpha \ln S_i)$	.161	.323	.484	.646
Sighting model $g_i(y)$	EM	HNM	HNM	HNM
$E(n_i/n)$	.2	.3	.3	.2

[a] $A = L = 1$; $N = 154.9$; $n = 50$.

(2) Fourier model (FOUR—a non-parametric approach from a Fourier series expansion of the probability density function $f(y)$; with the ability to assume a variety of non-spiked sighting curves (Burnham et al. 1980).

(3) Kelker model (KELK)—a version of the strip transect model as named after one of its earliest progenitors (see Gates 1979); a nonparametric approach, because no parameter of the model is estimated.

The mathematical representation of these sighting models and corresponding estimators is given in Quinn (1979). A fourth type of estimator, a generalized parametric approach, produces results similar to the Fourier series (Burnham et al. 1980).

METHODS FOR ANALYSIS OF GROUP SIGHTINGS

Let there be t classes of groups in the population where S_i is the number of individuals per group (group-size) for the ith size class. Let N_i be the true number of groups in the ith class and let $\Sigma N_i = N$. Suppose that a transect experiment is carried out and n total sightings occur with n_i sightings in class i. Each class has an associated sighting model $g_i(y)$ and effective half-width c_i. The group size S_i associated with each sighting is assumed to be determined without error.

The following example of such a population illustrates the important parameters of the experiment and forms the basis of the later computer simulation exercise. Four classes of groups in the population are constructed with true relative abundances N_i/N of 0.4, 0.3, 0.2, and 0.1 and represent group-sizes of 5, 25, 125, and 625 individuals. The exponential model (EM) is chosen for the underlying sighting model for the first class to represent the situation where sightings of small groups fall off rapidly at short distances from the transect line. The half-normal model (HNM) is chosen for the underlying sighting model for the three larger classes to represent the situation where sightings are fairly uniform at distances near the transect line and fall off smoothly at larger distances depending on the size of the group. The effective half-width c_i is chosen to be a logarithmic function of group-size in order to specify the scale of each individual curve (Fig. 1). The parameters which determine these relationships are shown in Table 1.

In general, the sighting model for groups of all sizes during the transect experiment is formed by weighting each individual sighting model by its relative abundance in the population (Quinn 1979). Applying this principle to the example produces the pooled sighting model for all sightings as shown in Figure 1. This pooled sighting model exhibits the heavier weighting of the first, more abundant class and has a shape that is functionally different from its component parts.

Method of Pooling

The first method of estimation for transect data pools the data over group-size classes. First, the estimated number of groups and its variance are calculated from (1) and (2), where N is redefined as the number of groups rather than individuals. A reliable estimator of N is obtained when the sighting model used for the pooled data approximates the unknown pooled sighting model.

Secondly, the estimated number of groups $\hat{N}_P$ is multiplied by an estimate of the average group-size $\bar{S} = \Sigma N_i S_i / N$ to estimate the total number of individuals in the population, i.e.,

$$\hat{T}_P = \hat{N}_P \hat{\bar{S}}. \tag{3}$$

The variance of $\hat{T}_P$ in (3) is the variance of a product (Seber 1973:7–9). The sample average group-size is not an unbiased estimate of $\bar{S}$ if there is a relationship between group-size and probability of sighting (or equivalently effective half-width), because larger groups are more likely to be in the sample than their presence in the population indicates. The estimate of $\bar{S}$ using (1) is

$$\hat{\bar{S}}_1 = \sum \hat{N}_i S_i \Big/ \sum \hat{N}_i = \sum n_i S_i \hat{c}_i^{-1} \Big/ \sum n_i \hat{c}_i^{-1}, \tag{4}$$

which is approximately unbiased when the n_i are near their expectations.

If the data are pooled over group-size, then (4) is not estimable, because each c_i is not estimated. One method of alleviating this problem is to: assume a functional relationship between c_i and S_i, i.e., $c_i = h(S_i)$(which may include a constant term); assume the mean sighting distance $\bar{y}_i$ is proportional to c_i for all classes; regress $\bar{y}_i$ against S_i to establish $\hat{h}$; and finally replace $\hat{c}_i^{-1}$ by $[\hat{h}(S_i)]^{-1}$ in (4). In particular, if c_i is proportional to the logarithm of group-size $\ln S_i$, then

$$\hat{\bar{S}}_2 = \sum \frac{n_i S_i}{\ln S_i} \Big/ \sum \frac{n_i}{\ln S_i} \tag{5}$$

results. An alternative estimator of $\bar{S}$, called $\hat{\bar{S}}_3$, is the average group size from sightings in a small interval about the transect line where groups of all sizes are likely to be seen. The estimators $\hat{\bar{S}}_2$ and two versions of $\hat{\bar{S}}_3$ using different intervals are evaluated by computer simulation below in the section dealing with comparison of the pooled and post-stratified methods.

Method of Post-Stratification

The second method of analysis is to partition the data by group-size. This method requires a sufficient number of sightings in each group-size class, say 25. The estimated number of groups in each class N_i is obtained from (1) using only the sightings from that class. Since the total sample of n sightings is stratified after the experiment is completed, this method is called post-stratification.

The intuitive post-stratified estimator of the total number of groups is

$$\hat{N}_s = \sum_{i=1}^{t} \hat{N}_i, \tag{6}$$

with estimated variance

$$\hat{\mathrm{Var}}(\hat{N}_s) = \sum_{i=1}^{t} \hat{\mathrm{Var}}(\hat{N}_i) + 2 \sum_{i<j} \hat{\mathrm{Cov}}(\hat{N}_i, \hat{N}_j). \tag{7}$$

The covariance terms are necessary because the n_i come from a multinomial distribution with parameters n and p_i^*, $i = 1, \ldots, t$, where p_i^* is the expected proportion of sightings $E(n_i)/E(n)$. Using results of conditional variance and covariance derived by Quinn (1980), the estimated variance of $\hat{N}_i$ is

$$\begin{aligned} \hat{\mathrm{Var}}(\hat{N}_i) = \hat{N}_i^2[&\hat{\mathrm{c.v.}}^2(n_i \mid n) + \hat{\mathrm{c.v.}}^2(n) \\ &+ \hat{\mathrm{c.v.}}^2(\hat{c}_i^{-1} \mid n_i)], \end{aligned} \tag{8}$$

where

$$\hat{\mathrm{c.v.}}^2(n_i \mid n) = (n - n_i)/nn_i,$$

and the estimated covariance between $\hat{N}_i$ and $\hat{N}_j$ is

$$\hat{\mathrm{Cov}}(\hat{N}_i, \hat{N}_j) = \hat{N}_i \hat{N}_j [\hat{\mathrm{c.v.}}^2(n) - 1/n] \tag{9}$$

Methods of estimating c.v.$^2(n)$ are given by Quinn and Gallucci (1980). The estimates (8) and (9) are substituted into (7) for the estimated variance of $\hat{N}_s$.

Finally, the post-stratified estimate $\hat{T}_s$ of the total number of individuals is obtained by multiplying the estimated number of groups $\hat{N}_i$ for class i by its group-size S_i and adding up over groups, so that

$$\hat{T}_s = \sum \hat{N}_i S_i. \tag{10}$$

Its estimated variance is

$$\begin{aligned} \hat{\mathrm{Var}}(\hat{T}_s) = &\sum S_i^2 \hat{\mathrm{Var}}(\hat{N}_i) \\ &+ 2 \sum_{i<j} S_i S_j \hat{\mathrm{Cov}}(\hat{N}_i, \hat{N}_j). \end{aligned} \tag{11}$$

Thus, the post-stratified estimator $\hat{T}_s$ does not require estimation of the average group-size $\bar{S}$ in contrast to $\hat{T}_p$. However, if S_i refers to a range of group-sizes, then this source of variability should be incorporated into (11) using (4), although its effect is likely to be minor compared to the variability of group-sizes over the entire population.

COMPARISON OF THE POOLED AND POST-STRATIFIED METHODS

In order to quantitatively compare the two methods, a computer simulation study was conducted using population parameters from the previous example, which are summarized in Table 1. The total number of sightings was fixed at 50, and the term c.v.$^2(n)$ was thus set to 0. This

TABLE 2
SIMULATION ESTIMATES OF THE NUMBER OF SCHOOLS

Estimator	Simulation average	Standard error (c.v.)		Root mean squared error $\sqrt{MSE}$[a]	
		Theoretical	Empirical	Theoretical	Empirical
Method of pooling					
EM	185.1	3.8 (.021)	3.6 (.019)	40.2	39.3
FOURIER	140.9	3.7 (.026)	4.3 (.031)	29.4	33.2
KELK	134.2	3.8 (.028)	4.4 (.033)	33.7	37.1
Method of post-stratification					
EM	200.7	5.1 (.025)	3.8 (.019)	57.6	52.5
FOURIER	154.3	5.6 (.036)	6.9 (.045)	39.2	48.3
KELK	146.1	4.8 (.033)	4.5 (.031)	34.7	32.7
True value	154.9				

[a] $\sqrt{MSE} = \sqrt{\hat{E}(\hat{N} - u)^2} = \sqrt{(n_r - 1)s^2 + (\bar{x} - u)^2}$
where n_r = number of replications; s = empirical or theoretical standard error; $\bar{x}$ = simulation average, and u = true parameter.

approach produces a smaller variance than a normal transect study where n is itself a random variable. However, the comparison of the two methods of analysis is still valid, because the term c.v.$^2(n)$ occurs equally in the variance expressions for both methods [Equations (2) and (8)]. The simulation was replicated 50 times to provide empirical means and standard errors for comparison with known or theoretical values. Further details concerning the mechanics of the simulation are found in Quinn (1979, 1980).

ESTIMATION OF THE NUMBER OF GROUPS

The simulation estimates of the number of groups N in the population are shown in Table 2 for the two methods. For the method of pooling, the EM estimator is positively biased, and the Fourier and Kelker estimators are negatively biased. The Fourier estimator is the least biased of the three. The spiked nature of the pooled sighting model (Fig. 1) causes underestimates to occur for estimators that assume a rounded shape near the origin (Crain et al. 1978, Quinn 1977). The EM estimate has the lowest coefficient of variation, followed by the Fourier and then the Kelker estimators. The root mean squared error, a convenient statistic incorporating the effects of variance and bias, favors the Fourier and then the Kelker estimator.

For the method of post-stratification, the Fourier estimator is the only estimator that produces an unbiased estimator. The Kelker and EM estimators produce under- and over-estimates, respectively. By examining the results of each group-size class (Table 3), the explanation for the bias is apparent. The EM overestimates the last three classes and correctly estimates the first class, producing an overall overestimate, in accord with the sighting models used for each group-size class (Table 1). The Fourier simulation average is unbiased for all classes, which produces an overall unbiased estimate. The Kelker estimate is negatively biased only for the first class, since the Kelker estimator performs poorly for spiked sighting models but reasonably well for rounded models (provided the truncation width is chosen small enough). The coefficients of variation show the same trends as for the method of pooling. The root mean squared error favors again the Fourier and then the Kelker estimator.

These results, which form a subset of a larger simulation study (Quinn 1980), suggest two general results. First, the simulation averages of a reasonable estimator such as the Fourier or the Kelker for the method of post-stratification generally are closer to the true parameter than for the method of pooling. This result is not unexpected, because the pooled sighting model generally has a more complicated shape and often a wider range of distances than the individual sighting models. Secondly, the theoretical and, generally, the empirical coefficients of variation for the method of post-stratification are larger than for the method of pooling (Table 2). This result is expected, because the number of sightings for each group-size class is substantially smaller than the total number of sightings, and the variance of an estimator of effective half-width is generally proportional to the inverse of the number of sightings (Quinn 1980, Burnham et al. 1980). The root expected mean squared error (Table 2) favors the method of pooling for the EM and Fourier, and either method for the Kelker.

These two generalizations from the simulation study have roots in theoretical relationships between sighting models and estimators. If an estimator has a functional form that is additive (Quinn 1980), then the method of pooling and the method of post-stratification produce identical estimators. This condition of additivity is satisfied by the Fourier and Kelker estimators, but only if their prespecified parameters are assumed constant for all classes. However, since the functional form of the sighting model for each class was different (e.g., Fig. 1), these prespecified parameters were not constant. In the simulation, these parameters were allowed to vary, so that each estimator could better estimate the number of groups in each class. Hence, simulation estimates for the post-stratified method were closer to the true parameter than for the pooling method.

The second generalization concerning the increased precision of the pooled estimator can also be verified theoretically by assuming that the precision of an estimator is proportional to the number of sightings, i.e.,

$$\text{c.v.}^2(\hat{c}^{-1} \mid n) = \sigma^2/n,$$

where $\sigma_c{}^2$ is an asymptotic constant dependent on the sighting model. This assumption appears to be reasonable by examination of the form of the variance estimator although non-parametric estimators are slightly less precise (Quinn 1980, Burnham et al. 1980, Eberhardt 1978). By substituting this relationship into (2) and (8) and using the Cauchy-Schwarz inequality, it can be shown that the theoretical coefficient of variation for the method of pooling is always less than or equal to that for the method of post-stratification (Quinn 1980, theorem 4). The only situation where the two are equal is when there is no relationship between the effective half-width and group-size.

The impact of the above results concerning transect estimation for grouped populations involves the trade-offs in accuracy (closeness to the true value) versus precision (as measured by the inverse of the coefficient of variation of the estimates). By post-stratifying the data, it is often possible to estimate each class accurately and, hence, the total number of groups N is estimated accurately. However, a single incorrect choice of a sighting model for a class leads to a biased estimate of N, and may not be detected by goodness-of-fit tests if there is a small number of sightings in the class. By pooling the data, the resultant sighting model may have a shape that is difficult to approximate by common sighting models, especially if the effective half-widths are substantially different. When a flexible model such as the Fourier is applied to both methods, the method of post-stratification is usually more accurate. On the other hand, the precision of the method of post-stratification, as compared to the method of pooling, becomes increasingly poor as greater differences occur in the effective half-widths.

The method of pooling is recommended for estimating the number of groups as long as an estimator derived from a flexible sighting model is chosen. In general, the bias of the estimator with the method of pooling is not large, and both the coefficient of variation and the mean-squared error are likely to be smaller than with the method of post-stratification.

TABLE 3
SIMULATION ESTIMATES OF N_i FOR EACH GROUP-SIZE CLASS[a]

Estimator	Group-size class (i) 1	2	3	4
EM	62.7	68.1	46.9	23.0
	4.1	3.5	2.3	1.5
	3.6	3.2	2.0	1.2
FOURIER	59.5	44.2	33.3	17.3
	4.8	3.0	2.1	1.5
	5.7	3.6	2.5	1.7
KELK	53.2	46.6	29.1	17.3
	3.9	3.0	1.9	1.3
	3.5	2.8	1.7	1.3
True value	62.0	46.5	31.0	15.5

[a] Reported for each group and estimator are the simulation average, theoretical standard error, and empirical standard error.

ESTIMATES OF AVERAGE GROUP-SIZE AND THE TOTAL NUMBER OF INDIVIDUALS

In order to estimate the total number of individuals T for the method of pooling, the average group size $\bar{S}$ in the population must be estimated as an intermediate step as shown in the section above on methods for analysis of group sightings. The data from the simulation are used to compare estimators of $\bar{S}$ and to illustrate the magnitude of bias of the sample average group-size $\bar{s}$. The true mean group-size, $\bar{S}$, and expected sample average group-size, $E(\bar{s})$, are computed from the values in Table 1. Four estimates are computed from the simulation replications: $\bar{s}$, $\hat{\bar{S}}_2$ (the log-weighted estimator), and two estimates $\hat{\bar{S}}_3$ the first uses sightings in the interval $[0,\Delta]$, the second in the interval $[0,2\Delta]$, where 2Δ includes no more than 75% of the total sightings).

The results are straightforward (Table 4). The true and expected sample average group-sizes are radically different (97 as compared to 171). The simulation average group-size is close to its theoretical value. By correcting $\bar{s}$ for increasing

TABLE 4
ESTIMATES OF AVERAGE SCHOOL SIZE FOR THE SIMULATION STUDY

n	n_r	$\bar{S}$	$E(\bar{s})$	$\bar{s}$	$\hat{S}_2$	$\hat{S}_3$ Interval $[0,2\Delta]$[a]	$\hat{S}_3$ Interval $[0,\Delta]$[a]
50	50	97	171	170.3 ± 4.9	98.1 ± 3.1	127.0 ± 5.0	119.4 ± 6.6

[a] Δ is chosen so that the interval [0,2Δ] encompasses no more than 75% of the observations.

probability of sighting by deleting more and more sightings at large distances, the simulation average becomes closer to the true group-size. However, even $\hat{S}_3$ in the interval $[0,\Delta]$ is positively biased and has the highest coefficient of variation of the four estimates. The log-weighted estimator, $\hat{S}_2$, has the lowest coefficient of variation and no bias. Thus, when effective widths are proportional to the logarithm of group-size, $\hat{S}_2$ is the best estimator.

However, additional studies have shown that the estimator $\hat{S}_2$ is not robust to the relationship between effective width and group-size (Quinn 1980). The estimator $\hat{S}_3$ is fairly robust but usually biased upward. Thus, there appears to be a need for more efficient and robust approaches to the estimation of a weighted average group-size.

The final comparison between the two methods involves estimates of the total number of individuals from (3) and (10). As shown in Table 5, three estimation models are considered: the method of pooling using the efficient estimator $\hat{S}_2$, the method of pooling using the robust estimator $\hat{S}_3$, and the method of post-stratification. The estimates in Table 5 are calculated directly from the values in Tables 2, 3, and 4.

The same trends for estimating the total number of individuals are found as for estimating the number of groups. One interesting difference is that the differences in the coefficients of variation of $\hat{T}$ for the pooled estimators are not as large as for the coefficients of $\hat{N}$. The contribution to the coefficients of the average group-size predominates, because the range in group-size is two orders of magnitude.

The best estimator for T appears to be the precise estimator $\hat{N}_P\hat{S}_2$ using the Fourier series estimator. If a more robust estimator is desired, then $\hat{T}_s$ with the Fourier estimator is fairly robust and reasonably precise. However, it may not be possible to compute $\hat{T}_s$ when the number of sightings is small. Another robust estimator that may be used as a last resort is the pooled estimator $\hat{N}_P\hat{S}_3$ with the Kelker estimator, although this estimator is the least precise.

TABLE 5
POOLED AND POST-STRATIFIED ESTIMATORS OF THE TOTAL NUMBER OF INDIVIDUALS T[a]

Estimator	Pooled estimator $\hat{N}_p\hat{S}_3$	Pooled estimator $\hat{N}_p\hat{S}_2$	Post-stratified estimator $\hat{T}_s$
EM	22,092	18,160	22,254
	1312 (.059)	684 (.038)	932 (.042)
	1304 (.059)	674 (.037)	729 (.033)
FOURIER	16,816	13,824	16,378
	1036 (.062)	568 (.041)	949 (.058)
	1068 (.064)	607 (.044)	1087 (.066)
KELK	16,017	13,166	15,881
	1001 (.062)	558 (.042)	821 (.052)
	1035 (.065)	599 (.046)	814 (.051)

[a] Reported for each estimator are its estimate, theoretical and empirical standard errors, and coefficients of variation. True parameter is 15025.

DISCUSSION

Based on simulation results and theoretical principles found here and in Burnham et al. (1980) and Quinn (1980), the following recommendations are given for a line transect sampling experiment of populations made up of groups.

1. The experiment should assure that a minimum of 50 groups are sighted for the method of pooling or 25 sightings per class for the method of post-stratification. Otherwise, criteria of accuracy and precision may not be met. If possible, a pilot study should be carried out to provide preliminary estimates of transect parameters. The preliminary parameters are necessary to calculate formulae for determining the amount of effort needed to be expended in the main experiment to achieve a certain level of precision (Quinn 1980, Burnham et al. 1980).

2. Critical assumptions of the line transect method are that all groups on the transect line are sighted and that there is no directional movement toward or away from the transect line. If possible, experimental design should include a test of these assumptions and ancillary experiments to develop correction factors if the assumptions are not met. Other assumptions of the line transect method are found in Burnham et al. (1980).

3. It is necessary to measure distances accurately to the geometric center of the group. Otherwise, the estimated number of groups is unreliable. The size of each group must also be

counted accurately. Otherwise, the estimated total number of individuals and average group-size are unreliable.

4. The method of pooling is recommended over the method of post-stratification for estimating the total number of groups, because of its increased precision, lower mean-squared error, larger number of sightings for hypothesis tests, and lack of arbitrary determinations about number of classes and sighting models for each class. Some stratification may be necessary, however, if the pooled sighting model has a complicated shape that is not well-represented by common sighting models in usage.

5. The recommended sighting model is the non-parametric Fourier series model based on these results and other studies (Burnham et al. 1980, Quinn 1980). The Kelker estimator should be used only as a last resort. One-parameter models such as the exponential model should not be used unless an extremely good fit to the data is produced. Generalized parametric estimators and other non-parametric estimators (Quinn and Gallucci 1980, Burnham et al. 1980) are often an acceptable alternative to the Fourier series.

6. The average group-size in the population must be estimated using a weighting procedure based on the relationship between the effective half-width and group-size. Current weighting procedures are largely empirical and have limitations of robustness. Better weighting procedures are needed.