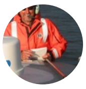
Line transect estimation of bird population density using a Fourier series

Article *in* Studies in Avian Biology · January 1981



Jeffrey L. Laake
Retired

Studies in Avian Biology No. 6:466–482, 1981.

LINE TRANSECT ESTIMATION OF BIRD POPULATION DENSITY USING A FOURIER SERIES

KENNETH P. BURNHAM,[1] DAVID R. ANDERSON[2]
AND JEFFREY L. LAAKE[3]

ABSTRACT.—A general approach to the estimation of bird population density from line transect data is discussed. This method is based on a nonparametric statistical analysis technique: the Fourier series method. The Fourier series estimator is both robust and efficient; i.e., it is not dependent on specific distributional assumptions about the detection probability of birds at various perpendicular distances from the transect line to provide relatively precise density estimates. The method is especially easy to compute for ungrouped, perpendicular distances, but can also be applied to grouped data commonly taken when sampling birds. A comprehensive computer program, TRANSECT, implements the Fourier series method, under a variety of options, by conducting hypothesis testing and point and interval estimation of population density. Examples of the Fourier series method based on nongame breeding bird transect data are provided. Because results will only be as reliable as the data collected, brief guidelines on field procedures and sample size are given. Finally, comments on other methods of analysis of line transect data are presented.

Line transect sampling to estimate the abundance of biological populations has been in use for over 40 years. However, only within about the last 10 years have there been substantial efforts to apply line transect sampling to the problem of estimating abundance of nongame birds. Similarly, it has only been within recent years that the statistical properties of this method have been intensively studied. Line transect sampling is now (almost) an established method for estimating densities of some species of nongame birds, especially breeding birds. Reliable results appear possible if good field practices are used to collect the data and robust, efficient analysis methods are used to analyze these data. The objective of this paper is to bring to the attention of ornithologists a general, robust, reliable data analysis method for use with line transect sampling data.

Line transect sampling embodies the explicit recognition of the fact that the probability of detecting birds decreases with increasing distance from the transect line. Because of this, distance data to birds detected are recorded. Estimation of bird abundance involves using these distance data to "correct" the sample size for the detectability of birds. This can be viewed as a refinement on strip transect sampling which requires the assumption that all birds are detected within a fixed perpendicular distance (i.e., within the strip) of the transect line.

Strip transect sampling predates line transects; it was used as early as 1906 (as reported in Forbes and Gross 1921). The use of distance data to "correct" for missed birds seems to have first been suggested in the 1930s (see Gates 1979). During the 1930s and 1940s, faltering attempts were made to put line transect sampling, i.e., estimation based on distance data, on a mathematical basis: see e.g., Leopold's (1933) reference to King's work in the late 1920s and early 1930s, Breckenridge (1935), Colquhoun (1940a, 1940b), Colquhoun and Morley (1941), Webb (1942), Kendeigh (1944), Southern (1944) and Kelker (1945). None of these papers presented any real theory of line transect sampling or estimation methods. A pioneering paper by Hayne (1949) was the first significant attempt to formulate an estimator of animal density based on line transect sampling data (Hayne's estimator has not actually been used much with nongame bird data).

Rigorous, general development of line transect theory did not really start until the late 1960s. Key papers by Gates et al. (1968) and Eberhardt (1968) laid the initial foundations of line transect theory. During the 1970s, work progressed and culminated in a good general understanding of, and theory for, line transect sampling and estimation of population abundance. The most comprehensive single reference is Burnham, Anderson and Laake (1980); however, other relevant literature during that decade is Anderson and Pospahala (1970), Seber (1973, 1979), Kovner and Patil (1974), Burnham and Anderson (1976), Hayes (1977), Schweder (1977), Anderson et al. (1978), Eberhardt (1978, 1979), Pollock (1978), Anderson et al. (1979), Burnham (1979), Crain et al. (1979), Gates (1979), Ramsey (1979), Ramsey and Scott (1979), Quinn (1977, 1979, 1980) and Patil et al. (1979). This represents a significant output of fundamental theory on line transect (and closely related) sampling; unfortunately, it has not yet been incorporated into ornithological practice.

[1] U.S. Fish and Wildlife Service, 2625 Redwing Road, Ft. Collins, Colorado 80526.
[2] Utah Cooperative Wildlife Research Unit, Utah State University, Logan, Utah 84322.
[3] Inter-American Tropical Tuna Commission, Scripps Institution of Oceanography, La Jolla, California 92037.

During this same period (approximately the past 20 years), ornithologists have been increasingly concerned with transect sampling to estimate bird abundance. However, there has been almost no basic transect sampling theory developed or presented in the ornithological literature (Yapp 1956, is an exception, but Royama 1960, concluded that Yapp's theory is not applicable in practice). Ornithologists have concentrated on conducting field studies to understand the numerous factors influencing the detection of birds (such as time of day, weather, habitat, bird species, and observer differences): see for example, Amman and Baldwin (1960), Bergerud and Mercer (1966), Brewer (1972), Fowler and McGinnes (1973), Järvinen and Väisänen (1975, 1976b), Franzreb (1976, 1977), Myrberget (1976), Tilghman (1977), Hickey and Mikol (1979). However, in a properly designed and conducted line transect study such factors can be safely ignored during data analysis if a suitably "robust" estimation method is used.

The best known line transect method used in ornithological studies is that of J. T. Emlen (1971, 1977a). The data collection aspects of Emlen's method can be improved, in principle, by more precise recording of distances. However, the estimation aspects of Emlen's method were developed with no theoretical basis and can be greatly improved. They should be replaced by rigorously developed, well founded estimation methods. We discuss one such method in this paper.

LINE TRANSECT SAMPLING

Background

A defined study area of known size, A, should be established before starting a transect study, especially if estimation of bird abundance at specific points in time is important (the alternative is to only compare changes in bird density over time). First, a set of transect lines must be established in the study area, along with a plan for sampling those lines. This constitutes an essential part of the study design and is of critical importance. Some comments on study design are presented in a different section of this paper. In general, one or more transect lines of fixed length are established in the area. Finally, the line is walked, at least once, and data on birds observed are recorded (replicates may, in fact, be different days of sampling the same line(s)).

In bird studies, it is common to establish a fixed distance, w, on either side of the line and only record birds observed within this distance. In strip transect sampling, all birds within the strip of length L and width $2w$ are assumed to have been observed (hence recorded). Thus, in strip transects, only the birds detected in the strip are counted. However, it is known that if w is large, detectability will decrease at increasing distances from the transect center line. Therefore, distance data on each bird, i.e., how far is it from the transect center line, must be recorded.

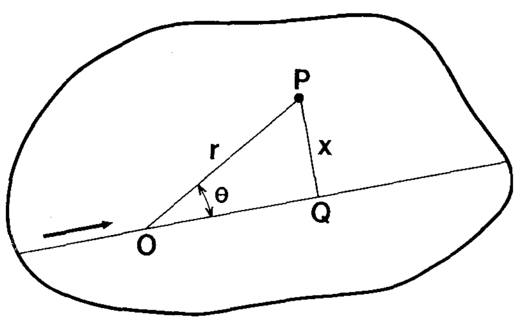


FIGURE 1. Diagram of the possible measurements that can be made for a detected object. The observer is at position O when an object is detected at position P, and Q is that point on the line perpendicular to the object. The sighting distance is r, the sighting angle is θ, and the perpendicular distance from the object to the center line is x. Note that the direction of the observer's travel, as shown by the arrow, is from O to Q.

Let there be R "replicate" lines, with lengths $l_1, l_2, \ldots, l_R$ and total length $L = l_1 + l_2 + \ldots + l_R$. We will sometimes treat the situation as if there were one overall line of length L. To facilitate further discussion, the following notation is defined: n_j = number of birds detected on line j, $j = 1, \ldots, R$; $n = n_1 + n_2 + \ldots + n_R$ = total count of birds for line length L; x = the recorded perpendicular distance from the transect (center) line to a detected bird—the total sample of such distances is x_i, $i = 1, \ldots, n$. Note that $x = r \cdot \sin(\theta)$; r = the sighting distance from the transect line to the bird; and θ = the sighting angle (see Fig. 1).

The basis for modeling line transect sampling is the concept that there is a decreasing probability of detection for birds at increasing distances from the transect line, and that this phenomenon can be represented by a "detection" function $g(x)$, where $g(x)$ = the probability of detecting a bird that is at perpendicular distance x from the transect line.

For strip transect sampling, it is assumed that $g(x) = 1$ for all distances less than w. However, for line transect sampling, the detection probability $g(x)$ decreases as perpendicular distance, x, increases.

In most ornithological field work, estimation of bird abundance has been based on the perpendicular distance data and we support that approach. However, these perpendicular distance data are often recorded by distance categories (i.e., as grouped data) rather than being recorded as exact measurements. The only justification for this is that the exact distances cannot always be determined; rather, the observer only knows that the bird, often heard rather than seen, is between some distance limits, such as 0 to 20 m or 20 to 50 m.

Assumptions

The goal of line transect sampling is to estimate the average density, D, of specified species in the study area. If N is the total number of birds in the area A, then $D = N/A$. A model is needed to relate the data to bird density, D, in order to derive a valid estimate of bird density. A model is just a set of assumptions; in their most concise form these are mathematical assumptions which, of course, have practical implications.

We recognize four basic assumptions in line transect sampling (in decreasing order of importance): (1) Birds directly on, or very near to, the line will always be detected; (2) There is no movement of birds in response to the observer and none are counted more than once during a given walking of the line; (3) All distance and angle data are recorded without measurement error; and (4) Sightings of different birds are statistically independent events.

In the abstract, line transect theory relates to sampling objects that do not move. This is reflected in assumption (2). However, movement that is random with respect to the location and movements of the observer causes no difficulty, provided the bird is counted only once during any one sampling of the transect and provided the distance to the transect line is accurately recorded when detection occurs. Assumption (2) will be violated by evasive movement, wherein birds move away from the transect line as the observer approaches, or by attraction of the birds toward the observer. Some degree of evasive movement is to be expected and can cause severe underestimation of bird density if it is extreme (e.g., even a moderate proportion of birds moving beyond the truncation distance, w).

Assumption (1) means that, if a bird is on, or very near, the line, the probability of detecting it is 1 (i.e., $g(0) = 1$). In practice, some birds will be missed during sampling. If they are well off the transect line, this causes no problems in estimating bird density. However, failure to detect birds that are on the line is a serious problem as regards density estimation. Birds that were on the line, but moved in response to the observer and then are missed, are also a problem. However, it is important to distinguish between these two situations as violating either assumption (1) or (2), respectively. Some degree of movement can be dealt with in the data analysis; i.e., moderate violation of assumption (2) may occur and transect sampling will still be useful. There is no way to deal with failure of assumption (1) from line transect data alone; failure to meet assumption (1) (all birds on the line are seen) directly and significantly biases *any* estimate of population density.

Assumption (3) is related to field techniques of distance measurement and the diligence of observers. If distances are to be recorded "exactly" (say to the nearest meter when $w = 100$ m), it is critical, for example, to avoid recording a distance between 1 and 10 m as zero (a not uncommon practice). It is necessary to have an objective way of measuring the distance (e.g., steel tape or pacing), otherwise the tendency is to record distances at convenient values like 5, 10, or 25 m. If the data are recorded by distance groups, then less rigor is needed because assumption (3) will be met if all birds detected are recorded in the correct distance category.

The primary way that assumption (4) is not met is if birds occur in distinct, small groups (large flocks of birds are not suitable for line transect sampling, anyway). We call this the case of birds (objects) occurring in "clusters." The proper treatment of clusters of birds is to regard the cluster itself as the object of interest and record only one distance per sighting, the distance to the cluster, and the cluster size. Standard line transect theory is then used to estimate the density of bird clusters and multiplying that estimate by the average cluster size in the total population gives the density of birds (see e.g., Burnham et al. 1980:192–194). Note, however, that the average cluster size observed from the actual line transect sampling may be a biased estimator of the true population average cluster size. In this case, estimation of average cluster size is not straightforward (see e.g., Burnham et al. 1980:192–194).

Approaches to Density Estimation

The total area sampled is $A = 2wL$. Let N be the total number of birds in this area. Given a properly designed study, an unbiased estimate of bird density is $\hat{D} = N/A$. For strip transect sampling, the total birds seen is $n = N$ (by assumption). But, for line transect sampling, N must be estimated as n/P, where P is the (average) probability of detecting a bird in the area sampled by the transect. This probability is related to the detection function; in fact,

$$P = \frac{1}{w}\int_0^w g(x)\,dx, \qquad (1)$$

which is the average value of $g(x)$ for $0 < x < w$ (J. T. Emlen 1971, 1977a, has called $1/P$ the "coefficient of detectability"). Define

$$a = \int_0^w g(x)\, dx \quad (2)$$

so that $P = a/w$. Then the estimate of D requires only an estimate of a (equivalent to estimating P because w is known):

$$\hat{D} = \hat{N}/A = \frac{n}{A\hat{P}} = \frac{n}{2wL\left(\frac{1}{w}\right)\hat{a}} = \frac{n}{2L\hat{a}},$$

$$\hat{D} = \frac{n}{2L\hat{a}}. \quad (3)$$

The estimate of this "correction factor" to account for birds missed that were off the transect center line depends on the recorded perpendicular distance data. The next step in this logical process is to derive the probability density function of the (random) variable x (=perpendicular distance). Seber (1973) has shown that

$$f(x) = \frac{g(x)}{a}, \quad (4)$$

where $f(x)$ represents the sampling distribution of x. Finally, by assumption (1), the probability of detecting a bird if it is on the transect center line is 1. Thus $g(0) = 1$, and from Eq. (4) we have

$$f(0) = \frac{1}{a} \quad (5)$$

Equation (5) provides a clear-cut relationship between the parameter a and the observed perpendicular distance data. Given any model for the detection function, or given a sampling model for perpendicular distances, there are many ways to estimate $f(0)$. Substituting Eq. (5) into Eq. (3) gives

$$\hat{D} = \frac{nf(0)}{2L}, \quad (6)$$

which is a general formula for estimating density, D.

The statistical estimation problem now is to specify a model, $f(x)$, of the sampling distribution of x (this is equivalent to modeling the detection function) and then derive an estimate of $f(0)$. A large variety of models for $f(x)$ have been used, for example, the negative exponential distribution (Gates et al. 1968), the half normal distribution (see, e.g., Gates 1979) and an exponential power series model which includes both these as special cases (Pollock 1978). Given the large variety of models for $f(x)$ and estimators for $f(0)$, reliable criteria are needed on which to base a choice of an estimator. It is not adequate or scientific to choose an estimator because one happens to like it or thinks it does well, or because one is comfortable with it. Finally, it is necessary to have an estimate of the sampling variance of the estimate of bird density. It is a major failing that most of the estimators in the biological literature have no associated estimates of precision.

CRITERIA FOR ROBUST ESTIMATION

The true detection function $g(x)$ is not known; moreover the work on line transect sampling in ornithology (and elsewhere) shows that the detection probability can vary due to numerous factors. Consequently, one cannot use a restrictive model for the detection probability and expect to get reliable estimates of density. A "robust" approach is needed; i.e., the estimator of bird density needs to be free of restrictive assumptions about the detection probability. The properties of $\hat{D}$ depend almost entirely on the estimator of $f(0)$, which depends, in turn, on the model chosen for the distribution of distances, and on the estimator used for $f(0)$. We have proposed several criteria that an estimator should satisfy in order to ensure reliable estimates of bird density from line transect sampling (Burnham et al. 1979).

Four criteria relate primarily to the properties of the assumed model for the sampling distribution of perpendicular distances. In order of importance these are: (1) *model robustness*; (2) *pooling robustness*; (3) *shape criterion*; and (4) *estimator efficiency*. Two additional criteria relate to promoting robustness of data analyses to common problems with transect distance data: (5) *data truncation*; and (6) *data grouping*. These last two criteria mean that the estimator of $f(0)$ should allow truncation of the data and should allow, or be developed to apply to, grouped data. Many line transect estimators in the literature are valid *only* for untruncated, ungrouped data.

Model robustness means that $f(x)$, the distribution of perpendicular distance data, is modeled with a general, flexible function, one that can take on a wide variety of shapes. Methods based on specific functional forms such as the negative exponential model are not *model robust* (see Burnham et al. 1980:162).

If an estimator is *pooling robust*, the fact that some birds off the line go undetected becomes totally irrelevant provided the basic assumptions are closely met. Data could be stratified by all possible factors likely to affect the detectability of birds and an estimate of density made for each strata. These separate estimates could then be combined into an estimate of total bird density;

this is a stratified estimator, $\hat{D}_s$. The alternative is to take the total set of data (n, $x_1, \ldots, x_n$ and L) and compute from this "pooled" data (pooled over replicate lines, observers and any other potential strata) an estimator, $\hat{D}_p$. An estimator of $f(0)$ is *pooling robust* if these two approaches produce the same estimate of density, i.e., $\hat{D}_s \equiv \hat{D}_p$. Thus, such an estimator is not affected by pooling the data over all the known and unknown factors that can effect the probability of detecting birds.

For line transect sampling of birds, it is very reasonable to assume that the detection function is 1 near the transect center line and hence has a "shoulder" near the line. This shoulder aspect of the shape of the detection function should be imposed on the model used to estimate $f(0)$. Mathematically, this is easy to do by specifying that the derivative of $f(x)$ at $x = 0$ is zero; hence $f'(0) = 0$ is the *shape criterion*. It means that the assumed detection function falls off very slowly near the transect line.

Criterion four, *estimator efficiency,* means that the estimator should have made the most use of the information in the distance data to estimate $f(0)$. An efficient estimator has a relatively small sampling variance. It is often easy to suggest ad hoc estimators; such ad hoc estimators are rarely efficient and are often badly biased.

The only general class of models that satisfy these four criteria are ones linear in their parameters, such as the polynomial:

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 \quad (7)$$

(see, for example, Gates and Smith 1980). However, $f(x)$ of Eq. (7) does not satisfy the *shape criterion* unless the parameter $a_1 = 0$. We considered the polynomial method, but found a better method for estimation of bird density. That method, the Fourier (pronounced *Foureay*) series estimator, is described in the next section.

Transect data in ornithology are generally taken with a finite truncation point, w. In other applications, w is often effectively infinite. Typically, it then will be necessary to delete a few "outliers" at extreme distances (which is why we presented criterion five). This is done by establishing a truncation point, w^*, and ignoring all data beyond distance w^*. This sort of data truncation may also be necessary with some species of birds. We distinguish between w^* and w because w^* is established after data collection while the transect width, w, is established before sampling. Such data truncation leads to more robust estimates of density (see e.g., Burnham et al. 1980:108–111). It is necessary to establish a truncation value w^* to apply the Fourier series estimator; however, it is entirely possible to take $w^* \equiv w$.

THE FOURIER SERIES ESTIMATOR

UNGROUPED DATA

The general estimator of density is

$$\hat{D} = \frac{n\hat{f}(0)}{2L}. \quad (8)$$

The estimator of $f(0)$ based on the Fourier series expansion is

$$\hat{f}(0) = \frac{1}{w^*} + \sum_{k=1}^{m} \hat{a}_k. \quad (9)$$

(The line length, L, and perpendicular distances must all be expressed in the same units.) The estimated Fourier coefficients $\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_m$ are computed from the ungrouped distance data $x_1, x_2, \ldots, x_n$ using the formula

$$\hat{a}_k = \frac{2}{nw^*}\left[\sum_{i=1}^{n} \cos\left(\frac{k\pi x_i}{w^*}\right)\right]. \quad (10)$$

Consider the second coefficient, $\hat{a}_2$ (i.e., $k = 2$), for a survey where 45 ($=n$) birds were detected within a 100 m wide (on each side) line transect ($w = w^* = 100$), then

$$\hat{a}_2 = \frac{2}{45 \times 100}\left[\sum_{i=1}^{45} \cos\left(\frac{2 \times 3.1416x_i}{45}\right)\right].$$

(If the computations are to be done on a small calculator, be certain that the cosine function allows the argument to be in radians, not degrees). After simplification,

$$\hat{a}_2 = 0.000444\left[\sum_{i=1}^{45} \cos(0.1396x_i)\right].$$

The number of Fourier coefficients computed to estimate $f(0)$ is determined by choosing the first value of m such that

$$\frac{1}{w^*}\left[\frac{2}{n+1}\right]^{\frac{1}{2}} \geq |\hat{a}_{m+1}|, \quad (11)$$

where $|\hat{a}_{m+1}|$ is the absolute value of $\hat{a}_{m+1}$. Equation (11) is called a "stopping rule." This rule for selecting the number of terms in the Fourier series represents a tradeoff between achieving small bias and always having a large number of terms (m), thereby getting a large sampling variance, or between always having a small m and having a possibly biased estimator. Typically, m is only 1, 2, or 3 and rarely needs to be as large as 5 or 6. In fact, if the above rule indicates $m \geq 6$ in nongame bird applications, something is probably wrong with the perpendicular distance data x_i (such as large rounding errors or mistakes in recording the field measurement). If the data, through Eq. (11), indicate, say, $m = 2$, then $f(0)$ is estimated as

$$\hat{f}(0) = \frac{1}{w^*} + \hat{a}_1 + \hat{a}_2.$$

An estimate of the sampling variance of $\hat{D}$ is given by

$$\hat{\text{v}}\text{ar}(\hat{D}) = \hat{D}^2\left[\frac{\hat{\text{v}}\text{ar}(n)}{n^2} + \frac{\hat{\text{v}}\text{ar}(\hat{f}(0))}{(\hat{f}(0))^2}\right]. \quad (12)$$

Estimates of the sampling variances var(n) and var($\hat{f}(0)$) are discussed below.

The estimate of var($\hat{f}(0)$) is based on the estimated Fourier coefficients, $\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_m$. The estimated variance-covariance matrix for these coefficients is

$$\hat{\text{c}}\text{ov}(\hat{a}_k, \hat{a}_j) = \frac{1}{(n-1)}\left[\frac{1}{w^*}(\hat{a}_{k+j} + \hat{a}_{k-j}) - (\hat{a}_k\hat{a}_j)\right], \quad (13)$$

For $k > j$, use $\hat{\text{c}}\text{ov}(\hat{a}_j, \hat{a}_k) \equiv \hat{\text{c}}\text{ov}(\hat{a}_k, \hat{a}_j)$ and, for $k = j$, use $\hat{a}_0 \equiv 2/w^*$. Of course, for $k = j$, $\hat{\text{c}}\text{ov}(\hat{a}_k, \hat{a}_j) = \hat{\text{v}}\text{ar}(\hat{a}_k)$.

Because the estimator of $f(0)$ (Eq. 9) is the sum of m Fourier coefficients (plus a constant term $1/w^*$), the sampling variance of $\hat{f}(0)$ is the sum of all the sampling variances and covariances of these m coefficients $\hat{a}_k$:

$$\hat{\text{v}}\text{ar}(\hat{f}(0)) = \sum_{j=1}^{m}\sum_{k=1}^{m} \hat{\text{c}}\text{ov}(\hat{a}_j, \hat{a}_k). \quad (14)$$

Equation (14) looks complex. However, consider the case $m = 2$; then Eq. (14) is the sum of the 4 elements in the 2×2 matrix

$$\begin{bmatrix} \hat{\text{c}}\text{ov}(\hat{a}_1, \hat{a}_1) & \hat{\text{c}}\text{ov}(\hat{a}_1, \hat{a}_2) \\ \hat{\text{c}}\text{ov}(\hat{a}_2, \hat{a}_1) & \hat{\text{c}}\text{ov}(\hat{a}_2, \hat{a}_2) \end{bmatrix},$$

or

$$\hat{\text{v}}\text{ar}(\hat{f}(0)) = \sum_{j=1}^{2}\sum_{k=1}^{2} \hat{\text{c}}\text{ov}(\hat{a}_j, \hat{a}_k). \quad (15)$$

The sampling variance of n is harder to estimate, but it can be approached in several ways. We will illustrate the special case where the survey is conducted on R replicate lines of equal length, then

$$\hat{\text{v}}\text{ar}(n) = \frac{R}{R-1}\sum_{j=1}^{R}(n_j - \bar{n})^2 \quad (16)$$

where $\bar{n} = \frac{1}{R}\sum_{j=1}^{R} n_j$. Other approaches are found in Burnham et al. (1980:51–55).

Finally, the estimated standard error of $\hat{D}$ (written $\hat{\text{s}}\text{e}(\hat{D})$) is the square root of the sampling variance of $\hat{D}$,

$$\hat{\text{s}}\text{e}(\hat{D}) = \sqrt{\hat{\text{v}}\text{ar}(\hat{D})}.$$

While we do not recommend doing these computations by hand, it is certainly possible and could be accomplished in one to two hours for a typical data set. To avoid rounding errors during calculations and to allow many additional features of the data to be explored, we recommend the use of a computer program to perform the arithmetic (see the section: PROGRAM TRANSECT).

Grouped Data

In many surveys, it is convenient to take the perpendicular distance data only by intervals (say, 0–20 m, 20–50 m, 50–100 m and 100–200 m), instead of measuring and recording the exact distance for each individual. The intervals need not be equal in size and, while there can be as few as two groups, it is preferable to have at least four, and 5–8 groups is more reasonable.

Taking the field data by intervals as grouped data should not be an excuse for inexact field procedures. Each observation should be properly recorded in the correct interval. Also, distance data should be taken ungrouped (i.e., distances precisely known), if possible.

Data can be analyzed as grouped, even when the original data were recorded in the field as ungrouped measurements. There are a variety of advantages in grouping the data for analysis, especially when the ungrouped data contain rounding errors, bias, or other anomalies. Estimates of density from the grouped data will then be more reliable than those based on the original ungrouped data.

Whereas, for ungrouped data, the perpendicular distance data are $x_1, x_2, \ldots, x_n$. For the grouped case, the data are the counts of birds seen in each interval, say $n_1, n_2 \ldots, n_k$, corresponding to the 1st, 2nd, . . . , kth interval. These counts can be used to estimate $f(0)$ with the Fourier series procedure. The proper computations are quite difficult and cannot be done without a computer.

Program "Transect"

We developed a comprehensive computer program, TRANSECT, to facilitate the analysis of line transect data. Program documentation is given by Laake et al. (1979). TRANSECT provides a convenient and thorough analysis tool which eliminates tedious calculations. It provides a variety of options for describing the basic data and includes several estimators of density and graphical and statistical goodness of fit tests.

The program consists of a main routine and 57 subroutines; there are approximately 7200 statements. Numerous comment statements document major features of the program. The program is written in ANSI Fortran IV. It is very portable and has been successfully run on CDC, Burroughs, IBM, and DEC computer systems. The program, example data, and output are available from the SHARE Program Library Agency, P.O. Box 12076, Research Triangle Park, N.C. 27709, at a cost of approximately $40.00. Specifications for the tape (e.g., 7 or 9 track, 800 or 1600 bpi, etc.) and the program No. 3600-05-003-007 should be given at the time of ordering.

EXAMPLE APPLICATIONS OF THE FOURIER SERIES ESTIMATOR TO NONGAME BIRD TRANSECT DATA

This section provides some examples of the Fourier series estimator. Hopefully, these examples will help make the method more fully understood and will help illustrate the points previously discussed. The data used in these examples is from a study done under a contract with the U.S. Fish and Wildlife Service to estimate breeding bird densities on coal lands (Hickey and Mikol 1979). We selected a subset of the data from two species: the Western Meadowlark (*Sturnella neglecta*) and the Lark Bunting (*Calamospiza melanocorys*). The Fourier series estimator is illustrated in both ungrouped and grouped formats. (Not all of the capabilities of program TRANSECT are illustrated here; for further examples see Burnham et al. 1980:90–120).

UNGROUPED DATA

The analysis of line transect data with the Fourier Series estimator can be thought of in terms of eight steps. These eight steps are the same regardless if the data are grouped or ungrouped (although different mathematical methods are employed at several steps). They are as follows: (1) estimate the Fourier coefficients a_i from the data; (2) determine the number of terms (m) to be used; (3) estimate $f(0)$; (4) estimate variances and covariances of the $\hat{a}_i$; (5) estimate the variance of $\hat{f}(0)$; (6) estimate D; (7) estimate the variance of $\hat{D}$; and (8) examine the goodness of fit.

These eight steps will be illustrated with the output of TRANSECT using ungrouped, replicated data on the Lark Bunting. For these data, there were 209 total observations recorded on five separate occasions. The length of the transect was 1000 meters (for each sampling occasion). For each observation, the perpendicular distance was recorded in meters and only birds within 100 meters ($w = 100$) were noted. A sample histogram of the perpendicular distances is illustrated in Fig. 2. This histogram shows that the detection of Lark Buntings decreased considerably at distances of 80–100 m.

For a first analysis, the data from the five replications were pooled to make one estimate of bird density. The estimates of the Fourier coefficients, $\hat{a}_i$, (step 1) were calculated by program TRANSECT using Eq. (10); more specifically for these data

$$\hat{a}_k = \frac{2}{100 \times 209}\left[\sum_{i=1}^{209} \cos\left(\frac{k(3.1416)x_i}{200}\right)\right]$$

where x_i are the individual measurements of perpendicular distance. The stopping rule value used to determine the number of terms (step 2) is calculated by the lefthand side of Eq. (11) and has the value of 0.000975 in this example. This results in ($m =$) five terms being selected; the estimates, $\hat{a}_1, \ldots, \hat{a}_5$ are shown in Table 1. The estimate of $f(0)$ (step 3), as calculated by Eq. (9), is

$$\hat{f}(0) = \frac{1}{w} + \sum_{k=1}^{5} \hat{a}_k = 0.006894.$$

The estimates of the variances and covariances of the $\hat{a}_i$ (step 4) require that $2m$ terms be computed. The estimates are computed from Eq. (13). The estimates of the standard errors are given with the point estimates in Table 1. The covariances, such as between $\hat{a}_1$, and $\hat{a}_2$, are not shown on the output; however, the related quantities, the correlation coefficients between $\hat{a}_i$ and $\hat{a}_j$, are printed by TRANSECT (see Table 1). The estimate of the variance of $\hat{f}(0)$ (step 5) is computed using Eq. (14); the result for this example is given in Table 1.

The estimate of density (step 6) only requires basic arithmetic and in this case is

$$\hat{D} = \frac{209 \times 0.006894}{2 \times 5000} = 0.0001441.$$

This estimate is in terms of numbers per square meter because both line length, L, and distances, x_i, were in meters. In order to get the result in Table 1 ($\hat{D}$ in numbers per hectare), the estimate must be multiplied by the number of square meters in a hectare (10,000), which gives 1.44 birds per ha.

The estimate of var($\hat{D}$) (step 7) can be accomplished in a variety of ways, for a discussion of this see Burnham et al. (1980:51–55). In this case, the estimated variance of $\hat{D}$ (the density estimate) was calculated using Eq. (12), and the variance of n was calculated empirically from the five replicates using Eq. (16). The estimated standard error $\hat{D}$ is given in Table 1.

The chi square goodness of fit test of the estimator of $f(x)$ (step 8) is shown in Table 2. From

```
FREQUENCY 23.  30.  35.  19.  25.  21.  22.  23.   6.   5.
----------------------------------------------------------------
    35              ****
    34              ****
    33              ****
    32              ****
    31              ****
    30         **** ****
    29         **** ****
    28         **** ****
    27         **** ****
    26         **** ****
    25         **** ****      ****
    24         **** ****      ****
    23    **** **** ****      ****           ****
    22    **** **** ****      ****      **** ****
    21    **** **** ****      **** **** **** ****
    20    **** **** ****      **** **** **** ****
    19    **** **** **** **** **** **** **** ****
    18    **** **** **** **** **** **** **** ****
    17    **** **** **** **** **** **** **** ****
    16    **** **** **** **** **** **** **** ****
    15    **** **** **** **** **** **** **** ****
    14    **** **** **** **** **** **** **** ****
    13    **** **** **** **** **** **** **** ****
    12    **** **** **** **** **** **** **** ****
    11    **** **** **** **** **** **** **** ****
    10    **** **** **** **** **** **** **** ****
     9    **** **** **** **** **** **** **** ****
     8    **** **** **** **** **** **** **** ****
     7    **** **** **** **** **** **** **** ****
     6    **** **** **** **** **** **** **** **** ****
     5    **** **** **** **** **** **** **** **** **** ****
     4    **** **** **** **** **** **** **** **** **** ****
     3    **** **** **** **** **** **** **** **** **** ****
     2    **** **** **** **** **** **** **** **** **** ****
     1    **** **** **** **** **** **** **** **** **** ****
---------+----+----+----+----+----+----+----+----+----+----+--
INTERVAL       10.       30.       50.       70.       90.
END-POINTS 0.00     20.       40.       60.       80.     0.10E+03
```

FIGURE 2. Histogram of perpendicular distances (in meters) for the Lark Bunting data when each bird is treated as a separate, independent sighting.

Table 2, the Fourier series model does not provide a very good fit. In fact, whenever the Fourier series requires more than three terms it reflects, from our experience, some anomalies in the data.

In this example, it appears there is "heaping" of distances at convenient values. This heaping resulted from two problems: (1) the distances were recorded at convenient values (they were not really measured); and (2) the birds (territorial males) were sometimes observed in temporary "groups" of two or three, during territorial interactions, and yet each bird sighted was treated as a single independent observation. The first problem can be corrected by more accurate measurements or can be made less severe by grouping the data. The second problem occurred because the above analysis is improper and it was done here only to show the effect of ignoring clustering (i.e., violating assumption 4) and heaping.

To properly analyze these data, any cluster of birds must be treated as a single entity and the density of birds estimated in two stages. First, a density of clusters is estimated by the Fourier series method, then that cluster density is multiplied by the average size of Lark Bunting clusters. This reduces the sample size from 209 birds observed to 166 clusters of birds observed; in most cases the "cluster" is only one bird. A

```
FREQUENCY    22.  22.  26.  16.  22.  15.  14.  18.   6.   5.
---------------------------------------------------------------
      26              ****
      25              ****
      24              ****
      23              ****
      22    **** **** ****      ****
      21    **** **** ****      ****
      20    **** **** ****      ****
      19    **** **** ****      ****
      18    **** **** ****      ****           ****
      17    **** **** ****      ****           ****
      16    **** **** **** **** ****           ****
      15    **** **** **** **** **** ****      ****
      14    **** **** **** **** **** **** **** ****
      13    **** **** **** **** **** **** **** ****
      12    **** **** **** **** **** **** **** ****
      11    **** **** **** **** **** **** **** ****
      10    **** **** **** **** **** **** **** ****
       9    **** **** **** **** **** **** **** ****
       8    **** **** **** **** **** **** **** ****
       7    **** **** **** **** **** **** **** ****
       6    **** **** **** **** **** **** **** **** ****
       5    **** **** **** **** **** **** **** **** **** ****
       4    **** **** **** **** **** **** **** **** **** ****
       3    **** **** **** **** **** **** **** **** **** ****
       2    **** **** **** **** **** **** **** **** **** ****
       1    **** **** **** **** **** **** **** **** **** ****
-----------+----+----+----+----+----+----+----+----+----+----+--
INTERVAL       10.       30.       50.       70.       90.
END-POINTS 0.00     20.       40.       60.       80.      0.10E+03
```

FIGURE 3. Histogram of perpendicular distances (in meters) for the Lark Bunting data. In this analysis, the observations are of clusters of birds (each bird is not necessarily treated as a separate, independent sighting).

TABLE 1
SUMMARY OF THE DENSITY ESTIMATION FOR THE LARK BUNTING DATA WITH THE FOURIER SERIES ESTIMATOR (EACH BIRD IS TREATED AS A SEPARATE, INDEPENDENT SIGHTING)[a]

Parameter	Point estimate	SE	% C.V.	95% C.I.	
A(1)	0.3447E-02	0.8819E-03	25.6	0.1718E-02	0.5175E-02
A(2)	−0.2635E-02	0.9009E-03	34.2	−0.4401E-02	−0.8692E-03
A(3)	0.1183E-02	0.9840E-03	83.2	−0.7456E-03	0.3111E-02
A(4)	−0.2425E-02	0.9930E-03	40.9	−0.4372E-02	−0.4792E-03
A(5)	−0.2676E-02	0.9679E-03	36.2	−0.4573E-02	−0.7785E-03
F(0)	0.6894E-02	0.2162E-02	31.4	0.2656E-02	0.1113E-01
D	1.441	0.4531	31.4	0.1830	2.699

Sampling correlation of estimated parameters

	1	2	3	4	5
1	1.000	0.335	−0.303	−0.036	−0.069
2	0.335	1.000	0.059	−0.161	0.043
3	−0.303	0.059	1.000	0.199	−0.062
4	−0.036	−0.161	0.199	1.000	0.107
5	−0.069	0.043	−0.062	0.107	1.000

[a] Density (D) units are numbers/ha.

TABLE 2

CHI-SQUARE GOODNESS OF FIT TEST FOR THE FOURIER SERIES ESTIMATOR (FROM FIG. 2) FIT TO THE LARK BUNTING DATA (EACH BIRD TREATED AS A SEPARATE, INDEPENDENT SIGHTING)[a]

Cell I	Cut points		Observed values	Expected values	Chi-square values
1	0.0	10.0	23.	17.56	1.69
2	10.0	20.0	30.	29.51	0.825E-02
3	20.0	30.0	35.	32.59	0.178
4	30.0	40.0	19.	22.89	0.662
5	40.0	50.0	25.	18.70	2.13
6	50.0	60.0	21.	25.73	0.871
7	60.0	70.0	22.	28.20	1.36
8	70.0	80.0	23.	18.69	0.992
9	80.0	90.0	6.	8.86	0.921
10	90.0	100.0	5.	6.26	0.255

[a] Total chi-square value = 9.063; degrees of freedom = 4; probability of a greater chi-square value = 0.05953485.

TABLE 4

CHI-SQUARE GOODNESS OF FIT TEST FOR THE FOURIER SERIES ESTIMATOR OF THE LARK BUNTING DENSITY WHEN THE OBSERVATIONS ARE CLUSTERS OF BIRDS[a]

Cell I	Cut points		Observed values	Expected values	Chi-square values
1	0.0	10.0	22.	19.49	0.324
2	10.0	20.0	22.	20.07	0.186
3	20.0	30.0	26.	20.84	1.28
4	30.0	40.0	16.	21.20	1.28
5	40.0	50.0	22.	20.58	0.987E-01
6	50.0	60.0	15.	18.70	0.732
7	60.0	70.0	14.	15.75	0.195
8	70.0	80.0	18.	12.36	2.57
9	80.0	90.0	6.	9.38	1.22
10	90.0	100.0	5.	7.64	0.912

[a] Total chi-square value = 8.793; degrees of freedom = 7; probability of a greater chi-square value = 0.26788825.

sample histogram of perpendicular distances is illustrated in Fig. 3. The summary of the density estimation from TRANSECT is provided in Table 3. In this analysis, the estimator only required two terms and the goodness of fit is acceptable (Table 4). The point estimate of cluster density ($\hat{D}$) is 1.938 clusters per hectare. In this example, the average cluster size ($\bar{c}$) can be calculated as the arithmetic mean; it is 1.259 with a standard error of 0.0498. This is an unbiased estimate of the true cluster size if the detection probability is independent of cluster size (Burnham et al. 1980:192–194). The estimated density of Lark Buntings is then

$$\hat{D} = \hat{D}_c{*}\bar{c} = 1.938 \times 1.259 = 2.440 \text{ birds/hectare.}$$

The standard error of $\hat{D}$ is calculated from

$$\hat{se}(\hat{D}) = \hat{D}(cv^2(\bar{c}) + cv^2(\hat{D}_c))^{\frac{1}{2}} = 2.440\left[\left(\frac{0.0498}{1.259}\right)^2 + \left(\frac{0.2963}{1.938}\right)^2\right]^{\frac{1}{2}} = 0.3853.$$

The point estimate of density increased dramatically compared to the results in the previous, incorrect, analyses. More importantly, the coefficient of variation is reduced considerably. This results from the reduction in the number of terms in the model and the reduction in the "heaping."

An alternative method of estimating the variance of density is to make separate density estimates for the replicate transects and calculate the variance empirically. This approach will be illustrated using the same Lark Bunting data to estimate density of clusters. Data were collected on five separate dates by the same observer,

TABLE 3

SUMMARY OF DENSITY ESTIMATION OF CLUSTERS OF LARK BUNTINGS WITH THE FOURIER SERIES ESTIMATOR[a]

Parameter	Point estimate	SE	% C.V.	95% C.I.	
A(1)	0.3628E-02	0.1007E-02	27.8	0.1654E-02	0.5601E-02
A(2)	−0.1956E-02	0.1055E-02	54.0	−0.4024E-02	0.1128E-03
F(0)	0.1167E-01	0.1662E-02	14.2	0.8415E-02	0.1493E-01
D	1.938	0.2963	15.3	1.115	2.760

Sampling correlation of estimated parameters

	1	2
1	1.000	0.298
2	0.298	1.000

[a] Density(D) units are numbers/hectares.

```
FREQUENCY            6.   8.   3.   3.   3.   4.   2.   1.   2.   2.
----------------------------------------------------------------------
          8              ****
          7              ****
          6         **** ****
          5         **** ****
          4         **** ****                ****
          3         **** **** **** **** **** ****
          2         **** **** **** **** **** **** ****      **** ****
          1         **** **** **** **** **** **** **** **** **** ****
-------------------+----+----+----+----+----+----+----+----+----+----+--

FREQUENCY            4.   4.   7.   6.   5.   5.   3.   3.   1.   0.
----------------------------------------------------------------------
          7                   ****
          6                   **** ****
          5                   **** **** **** ****
          4         **** **** **** **** **** ****
          3         **** **** **** **** **** **** **** ****
          2         **** **** **** **** **** **** **** ****
          1         **** **** **** **** **** **** **** **** ****
-------------------+----+----+----+----+----+----+----+----+----+----+--

FREQUENCY            6.   5.   2.   4.   5.   3.   2.   6.   1.   2.
----------------------------------------------------------------------
          6         ****                                ****
          5         **** ****           ****            ****
          4         **** ****      **** ****            ****
          3         **** ****      **** **** ****       ****
          2         **** **** **** **** **** **** **** ****      ****
          1         **** **** **** **** **** **** **** **** **** ****
-------------------+----+----+----+----+----+----+----+----+----+----+--

FREQUENCY            1.   3.   7.   2.   6.   1.   3.   4.   0.   1.
----------------------------------------------------------------------
          7                   ****
          6                   ****      ****
          5                   ****      ****
          4                   ****      ****            ****
          3              **** ****      ****       **** ****
          2              **** **** **** ****       **** ****
          1         **** **** **** **** **** **** **** ****      ****
-------------------+----+----+----+----+----+----+----+----+----+----+--

FREQUENCY            5.   2.   7.   1.   3.   2.   4.   4.   2.   0.
----------------------------------------------------------------------
          7                   ****
          6                   ****
          5         ****      ****
          4         ****      ****                **** ****
          3         ****      ****      ****      **** ****
          2         **** **** ****      **** **** **** **** ****
          1         **** **** **** **** **** **** **** **** ****
-------------------+----+----+----+----+----+----+----+----+----+----+--
INTERVAL                10.       30.       50.       70.       90.
END-POINTS 0.00              20.       40.       60.       80.      0.10E+03
```

FIGURE 4. Histograms of the perpendicular distances (in meters) for the five replicate transects of the Lark Bunting data; observations are clusters of birds.

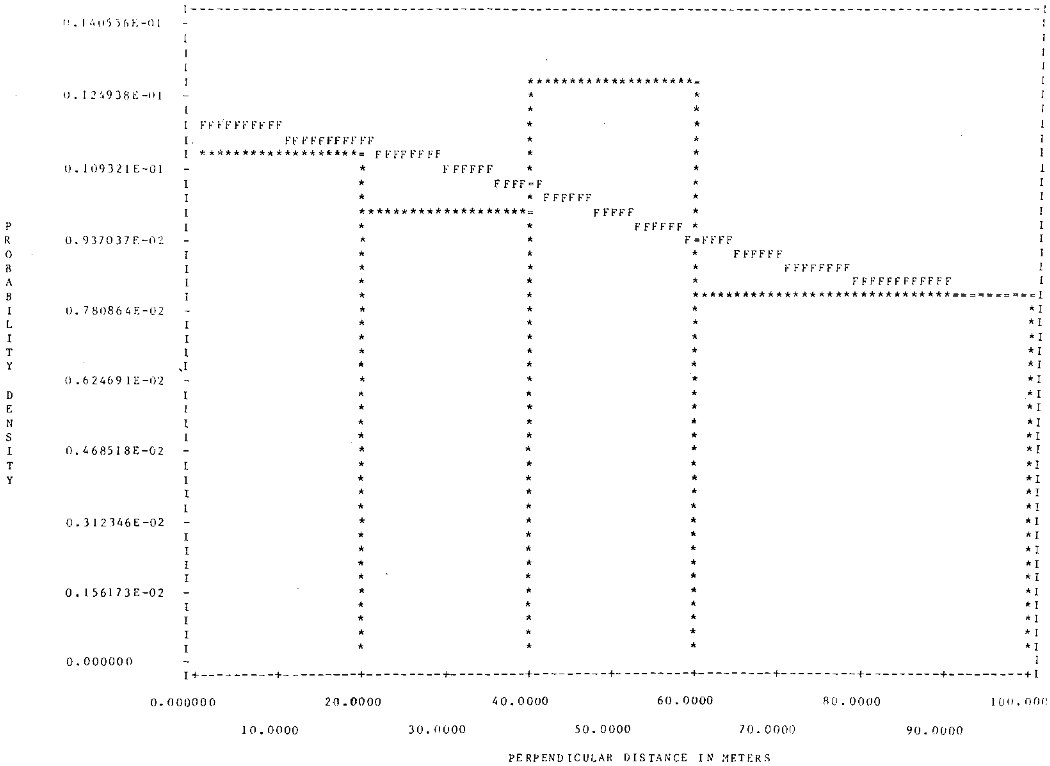


FIGURE 5.—Graphical representation of the goodness of fit of the Fourier series estimator (from Table 6) for the Western Meadowlark data.

over a three week period. The histograms of the perpendicular distances for these five replicates are illustrated in Fig. 4. These histograms reflect a considerable amount of variability. This further illustrates the need for a *model* and *pooling robust estimator*.

The Fourier series estimator was applied to each of the five replicates and the results summarized in Table 5. The average density estimate and its standard error are compared to the density estimate for the pooled data. If all of the estimates required the same number of terms, then the average density estimate would be exactly the same as the pooled density estimate. This is not the case and there is a considerable amount of variability in the shape of the histograms and the number of terms used. However, because the Fourier series is also very *model robust*, the two point estimates are very close. This method of calculating a variance for density is quite desirable but it requires a substantial sample size like this example.

Grouped Data

Often there are times in which the exact perpendicular distance cannot be measured, such as in aerial surveys, and the data are collected by distance intervals. The analysis of such grouped data with the Fourier series is not simple, rather it requires the use of a computer program. However, the same eight steps described previously can still be used but the estimates of the model coefficients a_i and their variances and covariances are calculated through numerical methods.

The analysis of grouped data will be illustrated with meadowlark data from the same transects used to collect the Lark Bunting data. There were 90 total observations of perpendicular distance on the five separate occasions. These perpendicular distance data have been grouped into four intervals.

The estimates of the parameters, a_i, are calculated by TRANSECT using maximum likelihood estimation. First, a model with one parameter is estimated, then two and successively more parameters are used in the model. The significance of adding each additional term is tested. For this example, the one-term Fourier series model is selected as the appropriate model for these meadowlark data. The estimates of $f(0)$, bird density, and their standard errors, are

TABLE 5
SUMMARY OF DENSITY ESTIMATES OF THE CLUSTERS OF LARK BUNTING FOR THE FIVE REPLICATE TRANSECTS[a]

Replicate	Sample size (n)	Number of terms in Fourier series (m)	Estimate of $f(0)$	Density estimate
1	30	1	0.0129	1.934
2	28	0	0.0100	1.400
3	36	0	0.0100	1.800
4	38	2	0.0100	1.959
5	34	1	0.0155	2.630
Averaged	166	—	0.0117	1.944 (0.198)
Pooled	166	2	0.0117	1.938 (0.296)

[a] Also shown are the weighted average of the five density estimates and the pooled density estimate (standard errors are in parentheses).

computed from the a_i just as in the ungrouped case; these results are shown in Table 6. The chi square goodness of fit test for this one-term Fourier series model indicated a good fit ($\chi^2 = 1.757$, 2 degrees of freedom, $P = 0.415$). Figure 5 shows the relative frequency histogram of the grouped data and the fitted one-term Fourier series model.

COMMENTS ON OTHER ANALYSIS METHODS

GENERAL COMMENTS

A variety of methods are available for the estimation of population size or density of biological populations. Here, we focus on variations of the line transect method, including the strip transect, in order to make some comparisons and suggestions.

Strip transects are merely very long, narrow quadrats and standard sampling theory applies. Strip transects do not require that distances be measured to estimate density. Line transect sampling offers two advantages over strip transects: (1) only animals on and near the centerline must be detected with certainty, and (2) the additional data taken at distances where the probability of detection is less than 1 can be used. This latter feature allows much more data to be used in the estimation of density. The ability to take the data as grouped greatly extends the applicability of the line transect procedure. In general, we recommend the use of line transect sampling over strip transect sampling in cases where both are appropriate.

The various approaches to density estimation using the sighting distance (r_i) and sighting angle (θ_i) are inferior to those based on perpendicular distances (x_i). Methods based on r_i and θ_i are quite sensitive to even small departures from the critical assumptions and these methods require additional assumptions as well as those required for estimation based on perpendicular distance data. The underlying models for the analysis of r_i and θ_i are very idealized and represent only crude approximations to the real situation. Finally, the estimators (e.g., Hayne's method, Hayne, 1949) are very sensitive to small sighting distances (i.e., the term $1/r_i$ will dominate the estimate if the i^{th} sighting distance is quite small). We do not recommend the estimation of density based on sighting distances and angles. If these data are all that are available, convert them to perpendicular distances and proceed on that basis.

Good methods for the estimation of bird density must be based on the following conditions: (1) sound theoretical development, (2) *model robustness*, (3) *pooling robustness*, (4) the *shape criterion*, and (5) high *estimator efficiency*. Consideration of these criteria leads us to recommend the Fourier series estimator as a useful, omnibus procedure. Further information on other analysis methods is given by Burnham et al. (1980: Part 4).

Finally, we caution against the use of the numerous ad hoc procedures that have been sug-

TABLE 6
SUMMARY OF DENSITY ESTIMATION FOR THE GROUPED WESTERN MEADOWLARK DATA USING THE FOURIER SERIES ESTIMATOR[a]

Parameter	Point estimate	SE	% C.V.[b]	95% C.I.[b]	
A(1)	0.1807E-02	0.1528E-02	84.6	−0.1188E-02	0.4803E-02
F(0)	0.1181E-01	0.1528E-02	12.9	0.8812E-02	0.1480E-01
D	1.063	0.1969	18.5	0.5159	1.609

[a] Density (D) units are numbers/ha.
[b] Notes on variance calculations and confidence intervals: the confidence intervals for the coefficients $A(I)$ and $F(0)$ were constructed by assuming asymptotic normality and using the Z-value 1.96; the variance of n was estimated using replicate line lengths (var(N) = 1.43); the confidence interval for density was constructed with a t distribution with the degrees of freedom equal to the number of line lengths − 1; and the t-value with 4 degrees of freedom is 2.776. Squared coefficient of variation for n = 0.1759E-01. Squared coefficient of variation for $F(0)$ = 0.1675E-01. Percent of the variation of density attributable to the sampling variance of n = 51.22.

TABLE 7
RESULTS OF 22 SURVEYS TO ESTIMATE THE DENSITY OF STAKES IN A 4 HA SAGEBRUSH-GRASS STUDY AREA IN UTAH, 1977–78

Observer	Number of stakes detected	Density estimate ($\hat{D}$)	$\hat{se}(\hat{D})$
1	81	35.4	4.72
2	72	33.6	4.86
3	54	25.4	4.66
4	56	29.1	4.25
5	57	26.2	4.11
6	68	38.8	6.12
7	48	29.2	6.65
8	49	28.0	5.32
9	51	23.0	3.77
10	68	36.9	4.82
11	84	36.9	4.49
12	48	34.6	8.59
13	75	30.2	4.44
14	61	35.9	6.53
15	60	31.0	8.51
16	100	27.2	4.27
17	55	33.2	6.41
18	61	34.4	7.60
19	46	24.7	5.21
20	41	33.6	8.51
21	54	34.1	5.74
22	72	34.5	6.20
Average	61.9	31.6	5.72

that the underlying detection function $g(x)$ differs greatly among observers. In field studies, the detection function would surely also vary among habitat types and species of birds. These factors, of course, affect n, making it an unreliable index.

In contrast to an index, estimates of density were computed for the 22 surveys using the Fourier series method. This is a better procedure as it allows both $g(x)$ and n to vary and still provide a valid *estimate* of density (not just a crude index). In these 22 surveys, the range in estimates of density is only 23.0 to 38.8 stakes per ha and the coefficient of variation for the density estimate (cv($\hat{D}$)) is about half that for the index (14.1%).

The use of a good estimation procedure allows estimates of density for various observers, surveying various habitat types for various species of birds. Estimates of precision are available as are tests of model fit.

An important assumption in line transect sampling is that all birds on, or very near, the centerline of the transect are detected. Mathematically, this is $g(0) = 1$; that is, the probability of detecting an animal at zero distance is one (or 100%). Note from Table 7 that the average density estimate is 31.6 stakes per ha while we know the true density to be 37.5. This bias is at least partially due to the failure of the assumption that $g(0) = 1$. Field procedures must focus on this assumption, or bias will be likely. The use of, for example, dogs, two observers, and all available cues will aid in meeting this important assumption. We can expect the failure of this assumption to be most severe with inanimate objects (e.g., stakes) rather than birds which often respond to the observer.

REASONS WHY BIRDS ARE UNDETECTED

The literature has many examples showing specific reasons why birds are not detected during a line transect survey. Limitations of the observer are often cited as a primary cause such as, inexperience, poor eyesight or hearing, lack of interest or training, or fatigue. The physical setting represents another broad class of reasons why birds that are present go undetected, including habitat type, sun angle, time of day and wind or other inclement weather. The species of bird being surveyed may preclude detection at the greater distances (e.g., small, drab colored birds that do not flush or vocalize readily). Many studies have looked at factors that are associated with incomplete detection and methods proposed to help lessen the proportion of birds that are not detected.

The two reasons for considering why birds are detected (or missed) are: (1) to design and conduct better studies; and (2) to improve distance estimation when the recorded distance depends on the detection cue(s) (rather than being a directly measured distance). The latter case is illustrated when detection depends entirely on hearing a bird and the distance *estimation* is also entirely based on this detection cue. In essence, studies on reasons why birds are detected (or missed) should be aimed at collecting improved *data*.

When, however, these efforts are directed at improved data analysis, they are largely misguided. A properly conducted line transect survey will provide valid estimates of density *even if a very substantial fraction of the birds go undetected*. In fact, the theory for line transect sampling *deliberately* allows birds to go undetected away from the centerline. Only in strip transect surveys is it necessary for all birds in the strip to be detected. As an example, in the 22 surveys summarized in Table 7, only 27–67% of the stakes present were detected. Still, relatively good estimates of density were obtained. The specific *reasons* why birds are not detected in either research or management programs. Such estimates are untrustworthy, not useful, and reflect poor procedure. Valid inferences or conclusions cannot be made without a good measure of the precision of the estimator. A variety of good procedures now exist for the careful and rigorous analysis of properly collected data. We see no excuse for using the many ad hoc approaches available in the literature.

ACKNOWLEDGMENTS

We thank S. A. Mikol and J. J. Hickey for letting us use some of their Lark Bunting and meadowlark data (from a much larger data set) to illustrate the Fourier series estimator. Gratitude is expressed to C. Snelling and M. Sieverin for assistance with manuscript preparation, and to C. Short for editorial assistance.