L. DANIEL MAXIM
Everest Consulting Associates
Princeton Jct., NJ 08550
LEIGH HARRINGTON
Mathtech, Inc.
Arlington, VA 22209
MARY KENNEDY
Consultant
Great Falls, VA 22066

A Capture-Recapture Approach for Estimation of Detection Probabilities in Aerial Surveys

Employed when ground truth data are non-existent, the approach requires the assumptions that detections are independent and that no false positives occur.

ABSTRACT: *A simple approach for estimating detection probabilities from imagery when ground truth data are non-existent is presented. Based upon what are termed capture-recapture statistics, the method requires only an independent examination of the imagery by two or more observers. In its simplest form the approach requires the assumptions that detections are independent and that no false positives occur. When data from three or more observers are available, checks upon model assumptions can be performed and less restrictive models can be developed. The approach is illustrated with several numerical examples.*

THE PROBLEM SETTING

IN SOME remote sensing applications, (see Heller (1968) for example) the detection, classification, and identification of objects of interest is essentially a deterministic process. Though there may be some probabilistic elements (e.g., those relating to weather), in these cases (after development of the detection and interpretation logic) the problem is largely that of developing an efficient search and inventorying strategy. If a complete census of a geographic area is to be taken, as opposed to a statistical sample, the analysis process can fairly be characterized as a problem in accounting.

Often, however, circumstances are less fortunate (see Huddleston and Roberts (1968) or Green *et al.* (1977) for example) and detection itself is a probabilistic process. Even if a complete census is to be taken of a population, the problem is statistical in character and it is appropriate to speak in terms of the *probability* of detection and identification. The number of objects observed needs to be adjusted upwards to reflect those objects that were missed in the interpretation process. This results in an *estimate* rather than a count. The adjustment or 'scale-up' factor depends upon the true, but often unknown, probability of detection and identification. Thus, it becomes necessary to measure, determine, or otherwise estimate this quantity. Estimation of detection and identification probabilities is often an objective of *ground truth* collection efforts. Here a geographical area is surveyed to determine the density or number and location of objects of interest (see Benson *et al.* (1971) for a discussion of ground truth considerations). Frequently the ground truth data are divided into two sets: a training set used to determine interpretation keys and a measurement sample or calibration set used to estimate detection probabilities. Given the experiment, computation of detection and identification probabilities is often relatively straightforward in statistical terms. But what if such ground truth data are difficult, impossible, or very costly to obtain? Counts of icebergs in the North Atlantic, polar bears in the arctic, or farmer's fields in the Amazon serve as examples where ground truth collection would be difficult. Even in circumstances where data collection

PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING,
Vol. 47, No. 6, June 1981, pp. 779-788.

poses lesser challenges, the cost of ground truth collection can be high. For example, the purpose of a preliminary feasibility study might be to determine which of several survey methods would be a cost effective choice (see Maxim and Cullen (1977) for one illustration). Here a rough or "first-order" estimate of detection probabilities might suffice. If an aerial survey were selected, later ground truth experiments could be conducted for rigorous estimation and validation purposes.

In this paper we propose a method based upon capture-recapture statistics to estimate detection probabilities. The method requires only an independent examination of aerial imagery by two (or more) photo-interpreters. With data from three or more observers, checks upon the assumptions of the method are possible as is the development of more sophisticated and exact models of the detection process.

Background: Detection Probabilities

The detection probability (as used in this paper, more properly the detection and identification probability), p_d, is defined as the probability that an object which is successfully imaged (e.g., not obscured by cloud cover or masked by terrain shadow in the case of photographic systems) is detected and correctly identified. Numerically, it is equal to the fraction of objects in a sampling frame or quadrat that are correctly interpreted. Logic and experience suggest that detection probabilities are a function of the type and scale of imagery (platform, sensor, and film), season, terrain, size, shape, and location of the object (e.g., near tree lines versus exposed), sun angle, observation geometry, photo-interpreter experience, extent to which reliable signatures exist, etc. (Green *et al.*, 1977; Egbert and Ulaby, 1972; Heath, 1973; McDonnell and Lewis, 1978; Williman and French, 1977). In principle, this functional dependence can be determined given sufficient ground truth data. In practice, this functional dependence can often be ignored or subsumed in an overall detection probability. To simplify the exposition, it is assumed that some average detection probability can be used to characterize the objects in sample quadrats—what is termed the *homogeneous detection* assumption. This assumption will be relaxed in a later section.

As stated earlier, given adequate ground truth data, the estimation of p_d is straightforward. It is simply the fraction of objects known to exist that are correctly detected and identified. If N is the number of objects in a sample quadrat and if x objects are detected and identified, the estimate of p_d, $\hat{p}_d$, is given by (circumflexes are used to denote estimates)

$$\hat{p}_d = x/N. \quad (1)$$

Thus, for example, if 50 objects out of 100 were detected, the estimate of p_d would be 0.5. This estimate has certain desirable statistical properties: it is a minimum variance unbiased estimator and is asymptotically normally distributed. The standard error of this estimate is well known (Johnson and Kotz, 1969, p. 51) and is given by

$$\sigma_{\hat{p}_d} = \left[\frac{p_d(1 - p_d)}{N}\right]^{1/2} \quad (2)$$

Though there may be circumstances where p_d is of intrinsic interest, in the context of the problem described, the real utility of knowledge of p_d is that it enables unbiased estimates to be made of the true but unknown number of objects in a sample quadrat, i.e., "reversing" the logic of Equation 1 to compute an estimate $\hat{N}$ of N assuming that p_d is known.

Specifically, it can be shown that, if p_d is known, and if detections are independent and satisfy some other assumptions of a more technical nature (Johnson and Kotz, 1969, p. 56), the best estimator for N, $\hat{N}$, given that x objects are observed in the cell is, simply,

$$\hat{N} = x/p_d, \quad (3)$$

and the standard error of this estimator, $\sigma_{\hat{N}}$, is

$$\sigma_{\hat{N}} = \left[\frac{N(1 - p_d)}{p_d}\right]^{1/2}, \quad (4)$$

which can be calculated by substituting $\hat{N}$ as an estimate of N. Thus, for example, if x = 50 objects were detected in a cell for which an appropriate value of p_d was 0.5, the estimate of the true number of objects (from Equation 3) is 100, with standard error (from Equation 4) of 10, so that an approximate 95 percent confidence interval on N is from 80 to 120.

But what if, as is the case for this problem, there are no ground truth data from which to estimate p_d? Surely the estimated number of objects in a cell should not be set equal to the observed number, x, since this amounts to assuming a value of unity for p_d. Yet, some estimate of p_d must be made, either from facts or by fiat. One set of approaches for avoiding fiat estimates is presented here. It is based upon a method for estimation of animal populations from capture-recapture data. The basic approach is next described.

Capture-Recapture Statistics

Capture-recapture methods were developed originally and are now widely employed for estimation of mobile animal populations. The basic idea is simple and is perhaps most easily explained by example (Feller, 1957). Suppose that 1000 fish caught in a lake are tagged and released. After some time a new catch of 800 is made, and it is found that 80 of these are tagged. What conclusions can be reached regarding the size of the fish population? It is assumed that the time interval between catches is sufficiently large so that the catches are independent and that the fish population is constant (i.e., no births, deaths, or migration). A simple answer is to note from the second catch that approximately 10 percent of the fish are tagged (i.e., 80/800). But from the first catch it is known that 1,000 were tagged, so if 1,000 is approximately 10 percent of the population, the population size is approximately 10,000.

More generally, if N_1 fish are tagged from the first catch and if N_{12} tagged fish are found in the second catch of size N_2, the estimate of the population size, $\hat{N}$, is given by

$$\hat{N} = N_1 N_2 / N_{12}. \quad (5)$$

This intuitive result can be shown to be correct by rigorous methods (McDonnell and Lewis, 1978) provided the assumption of independence is justified (Bailey, 1951; Chapman, 1951; Wittes *et al.*, 0000; Fienberg, 1972) for some extensions). Note that it is not required that N_1 and N_2 be the same. The fraction of fish caught in each of the trials (or probability of catch in each trial) can be estimated as shown below and illustrated with the above numerical example:

$$\hat{p}_1 = N_1/\hat{N} = 0.1 \quad (6)$$
$$\hat{p}_2 = N_2/\hat{N} = 0.08.$$

The Analogy

The above situation is analogous to the problem of aerial detection. Suppose that one photo interpreter (PI) discovers (catches) N_1 objects (fish) in a given quadrat (lake). These objects are marked on an overlay (tagged). Suppose a second PI were to examine (fish in) the same quadrat and discover (catch) N_2 objects (fish), without, however, knowing those objects identified by the first PI. On comparison, it is found that N_{12} objects were discovered by both PIs. What, then, is an estimate of the total number of objects in the cell? The analogy between fish and objects is clear. Provided that the detections can be regarded as statistically independent or approximately so, Equations 5 and 6 can be used to estimate the total number of objects in the quadrat and derivatively the detection probability for each PI. Additionally, it must be assumed that there are no false-positives in the detection process (i.e., detecting an object that is not there). If the cells were believed to be representative of the area at large and the detection probability of each observer were to remain constant (i.e., no further learning takes place), then the detection probabilities so determined could be applied to future cells to be examined by each PI. In a nutshell, this is the central idea of the paper. Subsequent sections will illustrate the application of this approach and provide some extensions suggested by the data.

An Illustration

Table 1 shows illustrative imagery readout from three observers. Note from the entries in this table that PI #1 found 60 objects (N_1 = 60), PI #2 found 89 objects (N_2 = 89), etc. Similarly, PIs #1 and #2 found 48 objects in common (N_{12} = 48), etc. Finally, 47 detections (N_{123} = 47) were common to all PIs.

Preliminary Data Analysis

This section contains a preliminary analysis of the data in Table 1 in terms of the capture-recapture concept. Using data from observers 1 and 2, an estimate of the number of objects in the quadrat (from Equation 5) is

$$\hat{N} = N_1 N_2 / N_{12} = (60)(89)/48 = 111,$$

while a similar estimate derived from N_1 and N_3 is 126 and from N_2 and N_3 is 120. It is possible to derive a statistical estimate of N based upon data from all three PIs, i.e., it is a quadratic in terms of N_1, N_2, N_3, N_{12}, N_{13}, N_{23}, and N_{123} (Johnson and Kotz, 1977)

$$\begin{aligned} &N^2(N_{12} + N_{13} + N_{23} - N_{123}) \\ &- N(N_1N_2 + N_1N_3 + N_2N_3) \\ &+ N_1N_2N_3 = 0, \end{aligned} \quad (7)$$

which can be solved as $\hat{N}$ for N. Some properties of this estimate are developed in Appendix A. For the data in Table 1, solution of Equation 7 yields the estimate $\hat{N}$ = 126. Given $\hat{N}$ = 126, the estimated detection probabilities for each of the observers and the average detection probability are

$$\begin{aligned} &\hat{p}_1 = N_1/\hat{N} = 60/126 = 0.48 \quad \text{and the average} \\ &\hat{p}_2 = N_2/\hat{N} = 89/126 = 0.71 \quad \hat{p}, \bar{p} = 0.69. \\ &\hat{p}_3 = N_3/\hat{N} = 109/126 = 0.87 \end{aligned}$$

Challenges to the Simple Model

The fundamental assumption in the use of the capture-recapture model for estimation of the

Table 1. Data from an Imagery Experiment

		THAT WERE ALSO DETECTED BY OBSERVER		
		1	2	3
NUMBER OF	1	60		
OBJECTS SEEN	2	48	89	
BY OBSERVER	3	52	81	109
		NUMBER SEEN BY ALL OBSERVERS WAS 47		

number of objects and detection probabilities is that detections are independent. Put another way, it assumes that the probability that a given object is detected by one observer is not a function of whether or not it was detected by another observer. While it is possible to argue against this assumption of independence on *a priori* grounds, it is ultimately an empirical question, one partially answerable from the data. To test the hypothesis of independence in the absence of ground truth requires data from three (or more) observers as provided in Table 1.

If p_i is used to denote the detection probability of the i^{th} observer, then under the assumption of independence, the expected number of objects common to all observers should be given by (this is the simple multiplication rule for independent events)

$$\hat{N}_{123} = Np_1p_2p_3. \tag{8}$$

Using the data from Table 1, for example, and the computed estimates of detection probabilities, $\hat{p}_i$, this estimate is

$$\hat{N}_{123} = (126)(0.48)(0.71)(0.87) \cong 37.$$

The observed value (see Table 1) for N_{123} is 47, some 27 percent higher.

The fact that the observed number of detections common to all observers, N_{123}, is greater than that computed under the independence assumption suggests that there is at least a modest degree of *dependence* and specifically that the conditional probability of detection by one observer given a detection by another is *greater* than the overall probability of detection.

It can be shown that, for the two PI case (see Appendix B), the consequences of assuming independence when such positive dependence is present are that

- estimates of the number of objects will be biased downward, that is, the estimate will *understate* the actual number of objects; and
- detection probabilities will be *overstated*.

The three PI case is more complicated but generally similar. As noted earlier, the data from Table 1 suggest some modest dependence, so estimates based upon this assumption will likely have the above faults. Given this finding, one choice is simply to disregard the dependence and accept a somewhat biased estimate. Another choice is to develop models which explicitly incorporate some form of dependence and use these to produce estimates. One such approach is outlined and explored below.

An Alternative Detection Model

Earlier it was assumed that all objects could be characterized by a single (average) detection probability. Such an assumption materially simplified the analysis, but may be unwarranted in practice. In view of the data, a logical next step is to assume that there are two (or more) types of objects,

- objects that are virtually certain to be detected, termed s-objects or "sures" in what follows, and
- objects whose detection is "probabilistic," termed p-objects in what follows.

Now s-objects could arise for a number of reasons, e.g.,

- these objects were used for development of keys and were recognized by the observer, or
- these objects were of a size or so located as to facilitate detection.

For convenience, it is assumed that the detection probability is close to unity for s-objects for each observer, i.e., $p_1 = p_2 = p_3 \cong 1.0$. How would such a phenomenon affect the analysis? First, all other things being equal, it would act to increase (i.e., overstate) *apparent* detection probabilities and, hence, the expected number of objects detected. Second, depending upon the extent of "sures," the relative number of objects discovered by various sets of observers would be altered. In turn, this would affect the outcome of various tests of independence because one component of the detection process would be dependent, the "sures." That is, given that an object were discovered by one PI, the probability that another observer would likewise detect it is higher than the unconditional likelihood of detection by the second observer. This statement can be proven in rigorous terms using Bayes' Theorem. The outline of the argument is as follows: The effective detection probability of any object by a PI is a function of the detection probability, p_i, for p-objects and the fraction of objects for which detection is certain. Now, suppose that an object were detected by one PI. This could be because the object was detected "probabilistically" or because it was an s-object and detection was certain. It is the case, however, that the fraction of s-objects among those detected by any observer is greater than the proportion of s-objects in the population. To see that this is true, suppose that there were 100 objects in a quadrat of which 40 were s-objects. Suppose also that the probability of detection of the p-objects were 0.5. Now all s-objects would be detected, but only 50 percent of the 60 p-objects or 30 p-objects would be in the sample. The proportion of s-objects *among those detected* would be 40 out of 70 or 57 percent rather than the 40 percent in the population. Thus, if an object is detected by one PI, it increases the odds that it is an s-object. Because this is so, it increases the likelihood that it will be detected by another PI above what it would have been in the absence of this knowledge, i.e., $p\,(2|1) > p_2$.

TABLE 2. DEBITING OUT SURES: AN EXAMPLE ASSUMING THERE ARE 35 SURES IN THE SAMPLE

THIS → BECOMES → THIS

		Observer 1	Observer 2	Observer 3				Observer 1	Observer 2	Observer 3	
Observer	1	60			All	Observer	1	25			All
	2	48	89		$\boxed{47}$		2	13	54		$\boxed{12}$
	3	52	81	109			3	17	46	74	

So much for the concept. How can this concept be reduced to practice and how well does this model fit the data? These questions are next explored.

A Computational Approach and Example

Deriving estimates and testing hypotheses for the case of *heterogeneous detection* is basically a straightforward if tedious extension of the previous method. It is perhaps best illustrated with a numerical example. To maintain the continuity of presentation, the same data from Table 1 will be analyzed. Note first that, since the number of objects observed by any combination of PIs can be written as the sum of two numbers—those objects certain to be detected, S, and those detections which are probabilistic, denoted by U—the sures must be subtracted out of the raw data matrix to derive estimates of the number of p-objects in the sample. These data can be analyzed as before to estimate the total number of p-objects in the quadrat. Then the "sures" can be added to estimate the total of all objects in the cell.

Let S be the number of s-objects in the quadrat. How S is estimated will be made clear later in the discussion, but assume that it is known. The estimated number of p-objects in any category, denoted by U_1, U_2, U_{12}, etc., is given by subtracting S from each of the appropriate object counts. To be concrete, consider the data contained in Table 1. If S were known to be 35, then the number of detections that would have resulted without the presence of sures would be as shown in Table 2 following. The data in Table 2 (right) can be analyzed as before, resulting in an estimate of the total number of p-objects $\hat{U}$ as 96. Detection probabilities for p-objects are

$$\hat{p}_1 = U_1/\hat{U} = 25/96 = 0.26 \quad \text{and the average}$$
$$\hat{p}_2 = U_2/\hat{U} = 54/96 = 0.56 \quad \hat{p},\ \bar{p} = 0.53.$$
$$\hat{p}_3 = U_3/\hat{U} = 74/96 = 0.77$$

and are, of course, smaller than those estimated earlier. What then of estimates of other quantities? How well do the data fit the model? Table 3 summarizes the necessary computations. The estimated number of objects identified by all three observers, for example, is the sum of the estimated number of p-objects, given by $\hat{U}\hat{p}_1\hat{p}_2\hat{p}_3 = 11$, plus the number of s-objects (35, by assumption) for a total of 47, substantially closer to the true value than that obtained assuming independence. Similar remarks can be made for the other estimated quantities, $\hat{N}_{12}$, $\hat{N}_{13}$, and $\hat{N}_{23}$. As can be seen, the choice of $S = 35$ produces a close match between observed and expected counts.

TABLE 3. A COMPARISON BETWEEN OBSERVED AND ESTIMATED VALUES, ASSUMING S = 35

QUANTITY ESTIMATED		FORMULAE FOR ESTIMATION	ESTIMATED VALUE (ROUNDED TO NEAREST INTEGER)	OBSERVED VALUE
Total Objects in Quadrat	$\hat{N}$	$\hat{U} + S$	131	N/A
Objects Detected By Each Observer	N_1	$U_1 + S$	Correct by	60
	N_2	$U_2 + S$	definition	89
	N_3	$U_3 + S$	of U_i	109
Objects Detected by Various Pairs of Observers	$\hat{N}_{12}$	$\hat{U}\hat{p}_1\hat{p}_2 + S$	49	48
	$\hat{N}_{13}$	$\hat{U}\hat{p}_1\hat{p}_3 + S$	54	52
	$\hat{N}_{23}$	$\hat{U}\hat{p}_2\hat{p}_3 + S$	77	81
Objects Detected by All Three Observers	$\hat{N}_{123}$	$\hat{U}\hat{p}_1\hat{p}_2\hat{p}_3 + S$	46	47
		Residual Sum of Squares	27.53	

A measure of the agreement between the actual counts and those "postdicted" by the heterogeneous detection model is the sum of squared deviations or residual sum of squares (RSS). It is obtained by adding the square of the differences between the observed and estimated counts. For $S = 35$, the computed value of RSS is 27.53.

The computed value of RSS is a function of the assumed number of s-objects in the sample. Figure 1 shows how the computed RSS varies with the assumed value of S over the range from zero, the smallest possible value, to N_{123}, the largest possible value. Note that, by assumption, s-objects are certain to be detected by all observers and, thus, cannot exceed the number N_{123} common to all PIs. The estimate $\hat{S}$ of S is that value which minimizes the RSS. In this example, $\hat{S} = 35$ as can be seen by inspection of Figure 1. Alternate definitions of goodness of fit (e.g., a chi-squared criterion) have also been explored. Suffice it to say that the estimates are generally insensitive to the criterion function.

Adjustment Factors with Sures

The heterogeneous detection model has a somewhat different "scale-up" rule than for the simple model shown in Equation 3. Its development is sketched below.

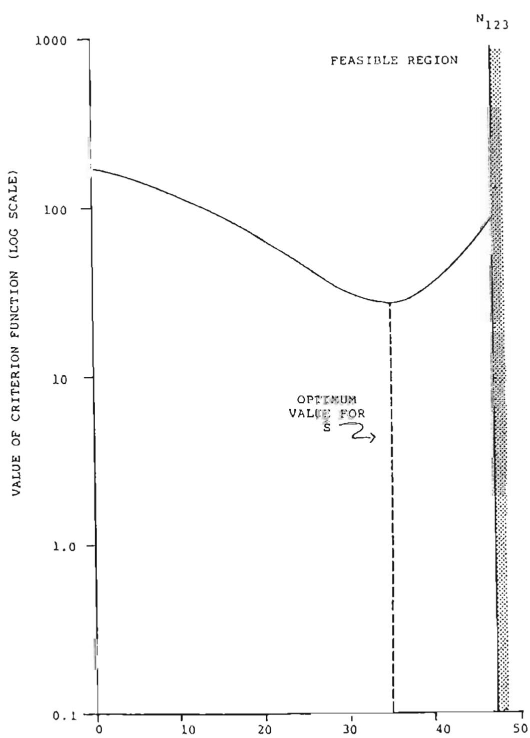

Fig. 1. Results of numerical search to estimate S, using an RSS criterion function.

Let x be the observed number of objects in a quadrat. Then $x - S$ is the number in the sample that are p-objects and, if p_d is the assumed detection probability for p-objects, $(x - S)/p_d$ equals the estimated number, U, of p-objects in the cell. The estimated total number of objects in the cell, $\hat{N}_t$, is the sum of p-objects and s-objects, or

$$\hat{N}_t = U + S.$$

But, by assumption, the number of s-objects is proportional to the total number of objects, i.e.,

$$S = \theta N_t,$$

so that neglecting the distinction between N_t and $\hat{N}_t$ and combining the above leads to

$$\hat{N}_t = \frac{x - \theta \hat{N}_t}{p_d + \theta \hat{N}_t},$$

which, after rearrangement, becomes

$$\hat{N}_t = \frac{x}{p_d[1 + \theta(1/p_d - 1)]} \text{ or } \hat{N}_t = \frac{x}{\theta + p_d(1 - \theta)}. \quad (9)$$

Equation 9 parallels the simple scale-up formula presented earlier in Equation 3. This result, in addition to being useful in its own right, also helps to explain the sensitivity analysis results as shown in Figure 2. These curves show how the estimates $\hat{U}$, $\hat{N}$, and $\bar{p}_d$ vary with the assumed value of S. While it is true that $\bar{p}_d$ is relatively sensitive to S, so too is $\theta = S/N_t$ in a compensating manner. Thus, the estimate of $\hat{N}$ of the total number of objects in a quadrat is not strongly dependent upon the assumed value of S (a pleasing result). Other sensitivity analyses are detailed in the Appendix.

Further Extensions

A different method of analyzing capture-recapture data, with dependent detection probability, is possible using what are termed log-linear models. Such models are discussed very thoroughly in Fienberg (1972), and are extremely general. However, it is sometimes difficult to interpret these models in a meaningful way. Moreover, our experience suggests that models of the form presented here provide both an adequate and a parsimonious representation of detection logic.

A Summing Up

It is probably still too early to assess the overall utility of the methodology outlined in this paper. Our experience suggests it to be a fruitful approach for the application circumstances envisioned. In the absence of ground truth it offers a logical, if less than perfect, basis for extrapolation and brings to mind the proverb, "In the land of the blind, the one-eyed man is king." Even in cases where ground truth can be obtained, it offers a useful first-order method to develop approximate detection probabilities.

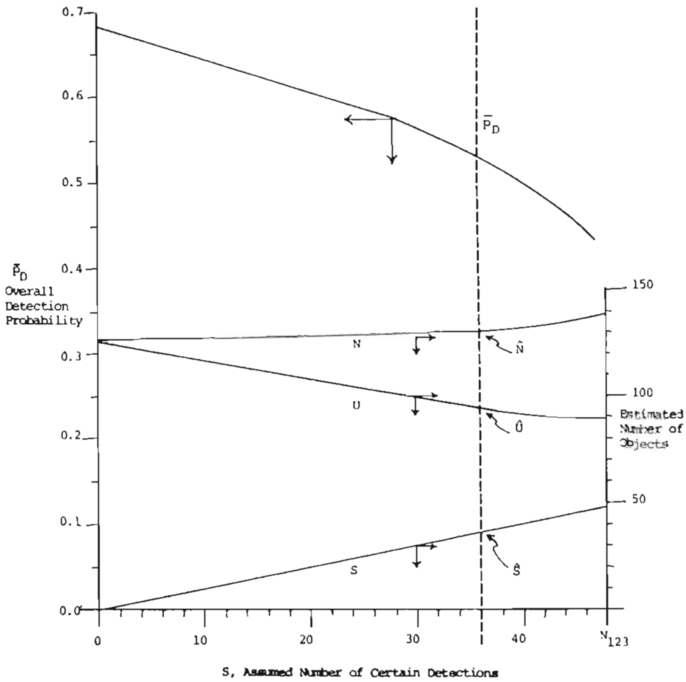

FIG. 2. Sensitivity of various estimates to the assumed number of certain detections.

ACKNOWLEDGMENTS

The authors wish to thank the referees and the editors for their useful comments and suggestions on an earlier draft of this paper.

APPENDIX A

The solution to the second degree equation, Equation 7 in the main text, is given by the well known quadratic formula,

$$\hat{N} = [-b + \sqrt{b^2 - 4ac}]/2a \qquad \text{(A-1)}$$

where

$$a = N_{12} + N_{13} + N_{23} - N_{123}$$
$$b = -(N_1N_2 + N_1N_3 + N_2N_3), \text{ and}$$
$$c = N_1N_2N_3.$$

In the above formula, the root corresponding to $[-b - \sqrt{b^2 - 4ac}]/2a$ has been discarded.

LEVEL CURVES

Some geometric insights into the nature of solutions to Equation A-1 can be obtained by examination of level curves for $\hat{N}$ in the coefficient space. It can be seen from the quadratic equation itself (e.g., holding $\hat{N}$ fixed at a particular value ψ) that level curves are in fact planes in a b c space which obey the relation,

$$\psi^2 a + \psi b + c = 0. \qquad \text{(A-2)}$$

Figure A1 shows an isometric rendering of three iso-$\hat{N}$ planes corresponding to values for $\hat{N}$ of 127, 200, and 300. The geometry of the situation is actually somewhat more complex than is depicted on Figure A1, however, because all combinations of values of a, b, and c are not logically possible when dealing with capture-recapture data.

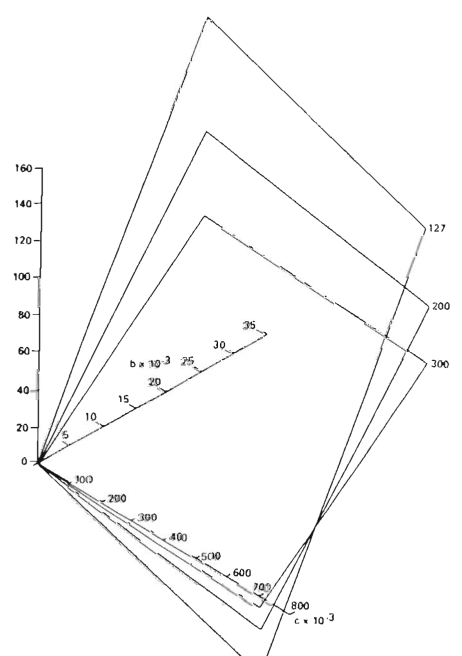

FIG. A1. Level curves for N in coefficient space

ENTER CONSTRAINTS

To illustrate, consider first the relationship between the coefficients b and c. Given actual data, of course, the relationship is fixed and known. But *a priori* it can be shown to lie within certain bounds. For a fixed value of c, the largest feasible value for b can be obtained by solving the optimization problem:

$$\begin{aligned}&\text{Max } (b) \\ &\text{where} \\ &b = -(N_1N_2 + N_1N_3 + N_2N_3) \\ &\text{Subject to } N_1N_2N_3 = c.\end{aligned} \tag{A-3}$$

By forming the Lagrangian and taking partial derivatives, it can be shown that the maximizing solution to Equation A-3 occurs when $N_1 = N_2 = N_3 = c^{1/3}$, and hence from Equation A-3, $b \leq -3c^{2/3}$. Minimizing Equation A-3 leads to an unbounded solution where one of the N_i goes to zero, assuming that the product $N_1N_2N_3$ is a constant. In this case the N_i values must be non-negative integers and, if the product is non-zero, each must be at least unity. With this constraint a minimizing solution is to set N_1 and N_2 to one and N_3 to c, whence the smallest value for b is $-(2c + 1)$. It follows then that

$$-(2c + 1) \leq b \leq -3c^{2/3}. \tag{A-4}$$

For the example given in the main text, $a = 134$, $b = -21{,}581$, and $c = 582{,}060$. Note that inequality Equation A-4 is satisfied, i.e., $-1{,}164{,}181 \leq -21{,}581 \leq -20{,}913$.

Similar bounds or constraints can be derived between/among the other combinations of coefficients. Table A1 shows the resulting optimization problems and derived bounds of the relationships between a and b, b and c, and a and c. Shown also in Table A1 is a numerical illustration of each of these bounds taken from the same example.

Figure A2 attempts to capture the geometry of the level curves with the constraints derived in Table A2 superimposed. In this exhibit the iso-$\hat{N}$ plane corresponding to $\hat{N} = 127$ (the numerical example) is shown. For clarity other iso-$\hat{N}$ planes are omitted. The numerical data fall close to the b vs c constraint and well away from others. All lines are shown in bold, i.e., hidden lines are not dotted.

Yet other limits upon the choice of values for a, b, and c and the various N values arise if the solution for $\hat{N}$ is to be well behaved. For example, in order that the solution to Equation A-1 has real roots, the condition $b^2 - 4ac \geq 0$ must obtain. The value of a must be greater than zero if $\hat{N}$ is to be non-infinite, etc. Table A2 provides a convenient summary of applicable constraints upon the various data entries and derived coefficients.

SENSITIVITY ANALYSIS

It is of interest to note the sensitivity of the computed estimate, $\hat{N}$, to the values of the individual data elements or assumed quantities (in the case of the heterogeneous detection model). For the homogeneous detection model, the partial derivatives of $\hat{N}$ with respect to each of the data ele-

TABLE A1. OPTIMIZATION PROBLEMS USED TO DEVELOP CONSTRAINTS AMONG THE PARAMETERS

CONSTRAINT	MATHEMATICAL STATEMENT	MINIMIZING SOLUTION	MAXIMIZING SOLUTION
I. a vs b	Min or Max $[N_{12} + N_{13} + N_{23} - N_{123}]$ Subject to: $N_1N_2 + N_1N_3 + N_2N_3 = -b$ $N_{12} \leq \text{Min } [N_1,N_2]$ $N_{13} \leq \text{Min } [N_2,N_3]$ $N_{23} \leq \text{Min } [N_2,N_3]$ $N_{123} \leq \text{Min } [N_{12},N_{13},N_{23}]$	$a = 0$ (0)	$a = (-3b)^{1/2}$ (254.45)
II. b vs c	Min or Max $[N_1N_2 + N_1N_3 + N_2N_3]$ Subject to: $N_1N_2N_3 = c$	$b = -2c - 1$ (−1,164,121)	unbounded non-integer $b = -3c^{2/3}$ integer (−20,913)
III. a vs c	Min or Max $[N_{12} + N_{13} + N_{23} - N_{123}]$ Subject to: $N_1N_2N_3 = c$ $N_{12} \leq \text{Min } [N_1,N_2]$ $N_{13} \leq \text{Min } [N_1,N_3]$ $N_{23} \leq \text{Min } [N_2,N_3]$ $N_{123} \leq \text{Min } [N_{12},N_{13},N_{23}]$	$a = 0$ (0)	$a = 3(c)^{1/3}$ (250.48)

NOTE: VALUES IN PARENTHESES CORRESPOND TO BOUNDS COMPUTED FROM TEXT EXAMPLE
$a = 134$, $b = -21{,}581$ and $c = 582{,}060$

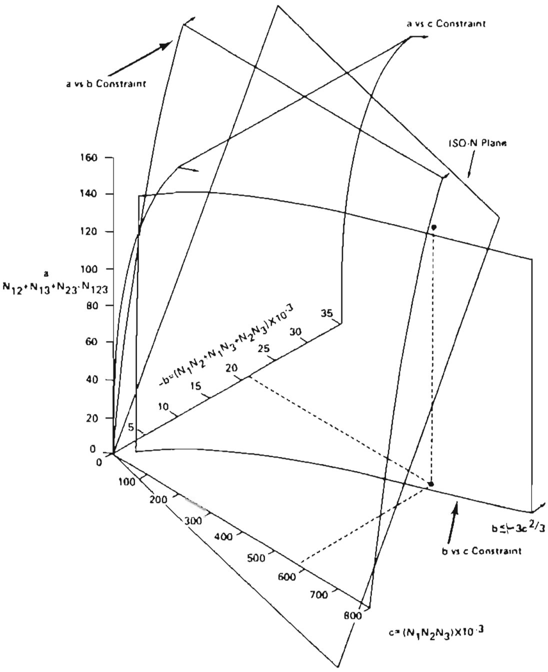

FIG. A2. Applicable constraints upon the coefficients superimposed

ments evaluated in the feasible region described in Table A2 indicate that (other inputs fixed)

- $\hat{N}$ is an increasing function of N_i,
- $\hat{N}$ is a decreasing function of N_{ij}, and
- $\hat{N}$ is an increasing function of N_{ijk},

all intuitively plausible results. These results are borne out by a numerical example of sensitivity analysis shown in Table A3. The base case for comparison is the example given in the main body of the text. Table A3 shows the effect of a 10 percent one-at-a-time change to each of the input quantities in terms of the change to $\hat{N}$.

For the case of the heterogeneous detection model, the situation is somewhat more complex. A change to any of the data inputs will not only alter the estimate of $\hat{N}$, but also the estimate of S. The same sensitivity analyses described above for the homogeneous detection model have been conducted assuming that sures are relevant. Each of the inputs behaves as before with respect to the estimate $\hat{N}$, but the pattern of change for $\hat{S}$ is less clear cut. The magnitudes of percentage changes to $\hat{N}$ for a fixed change in the inputs are also somewhat higher (approximately 30 percent in this case) than for the case of homogeneous detection.

APPENDIX B

To show that a positive dependence between two PIs will cause a downward bias in the estimate(s),

$$\hat{N} = \frac{N_1 N_2}{N_{12}},$$

note that such a dependence implies that if one PI detects the object then the second PI will have a greater probability of detecting the object too. Intuitively, then, positive dependence will increase

TABLE A2. SUMMARY OF AVAILABLE CONDITIONS FOR WELL BEHAVED ESTIMATES (Homogeneous or Heterogeneous Detection)

CONSTRAINT	SOURCE
$0 \le N_{12} \le \text{Min}\,[N_1, N_2]$ $0 \le N_{13} \le \text{Min}\,[N_1, N_3]$ $0 \le N_{23} \le \text{Min}\,[N_2, N_3]$ $0 \le N_{123} \le \text{Min}\,[N_{12}, N_{13}, N_{23}]$	Logical constraints From Physical Problem
$N_{12} + N_{13} + N_{23} - N_{123} > 0$	Non-infinite estimate, $\hat{N}$
$N_i \ge 1 \quad \forall\, i$	Detection probabilities Not all zero and $\hat{N}$ determinate
$(N_{12} + N_{13} + N_{23} - N_{123}) \le \dfrac{(N_1N_2 + N_1N_3 + N_2N_3)^2}{4N_1N_2N_3}$	Estimating equation produces non-imaginary estimate
$3(N_1N_2N_3)^{2/3} \le (N_1N_2 + N_1N_3 + N_2N_3) \le [2(N_1N_2N_3) + 1]$ $0 \le (N_{12} + N_{13} + N_{23} - N_{123}) \le 3(N_1N_2N_3)^{1/3}$ $0 \le (N_{12} + N_{13} + N_{23} - N_{123}) \le (3(N_1N_2 + N_1N_3 + N_2N_3))^{1/2}$	Derived constraints See Table A1
$0 \le S \le N_{123}$	Sures cannot exceed those in common

TABLE A3. ILLUSTRATIVE SENSITIVITY ANALYSIS

Base Case Data: Shown Below

Base Case Results: Homogeneous Detection $\hat{N} = 127$

Heterogeneous Detection $\hat{S} = 35$, $\hat{U} = 96$, $\hat{N} = 131$

A 10% INCREASE IN THIS FACTOR	WOULD RAISE IT		AND ALTER FINAL ESTIMATES AS FOLLOWS: Homogeneous Detection		Heterogeneous Detection*			
	From	To	$\hat{N}$	% Change	$\hat{S}$	$\hat{U}$	$\hat{N}$	% Change in $\hat{N}$
N_1	60	66	134	5.5	32	108	140	6.9
N_2	89	98	138	8.7	37	109	146	11.5
N_3	109	120	139	9.5	38	108	146	11.5
N_{12}	48	53	121	−4.7	35	88	123	−6.1
N_{13}	52	57	121	−4.7	30	92	122	−6.9
N_{23}	81	89	117	−7.9	21	96	117	−10.7
N_{123}	47	48*	128		37	97	134	

* 10% increase would violate constraint $N_{123} \leq N_{12}$ so N_{123} set equal to constraint.

N_{12}, the number seen by both. Thus, $\hat{N}$ will be smaller, on the average, than when detections are independent. More formally, conditioned on given sample sizes N_1 and N_2, and assuming $N_{12} \neq 0$,

$$E(\hat{N}) = N_1 N_2 E\left(\frac{1}{N_{12}}\right).$$

It can be shown that $E(1/N_{12})$ is a decreasing function of p_{12}, where p_{12} is the probability that both PI #1 and PI #2 detect the object. Thus $E(\hat{N}|p_{12} > p_1p_2) < E(\hat{N}|p_{12} = p_1p_2)$, the latter being the case where both PIs detect the object assuming independence. Since $\hat{p}_i = N_i/\hat{N}$, the above result also implies that the estimated detection probability for each PI will be overstated, on average, whenever there is a positive dependence.

REFERENCES

Bailey, N. T. J., 1951. On Estimating the Size of Mobile Populations from Recapture Data, *Biometrika*, Vol. 38, p. 293-306.

Benson, A. S., W. C. Draeger, and L. R. Pettinger, 1971. Ground Data Collection and Use, *Photogrammetric Engineering and Remote Sensing*, Vol. 37, No. 11, pp. 1159-1166.

Bishop, Y. M. M., S. E. Fienberg, and P. W. Holland, 1975. *Discrete Multivariate Analysis*, MIT Press, Cambridge, Mass.

Chapman, D. G., 1951. *Some Properties of the Hypergeometric Distribution with Applications to Zoological Sample Censuses*, University of California Publications in *Statistics*, Vol. 1, pp. 131-160.

Egbert, D. D., and F. T. Ulaby, 1972. Effect of Angles on Reflectivity, *Photogrammetric Engineering and Remote Sensing*, Vol. 38, No. 6, pp. 556-564.

Feller, W., 1957. *An Introduction to Probability Theory and Its Application*, Volume 1, Second Ed., John Wiley and Sons, New York.

Fienberg, S. E., 1972. The Multiple Recapture Census for Closed Populations and Incomplete 2k Contingency Tables, *Biometrika*, Vol. 59, No. 3, pp. 591 et seq.

Green, L. K., J. K. Olson, W. G. Hart, and M. R. Davis, 1977. Aerial Photographic Detection of Imported Fire Ant Mounds, *Photogrammetric Engineering and Remote Sensing*, Vol. 43, No. 8, pp. 1051-1058.

Heath, G. R., 1973. Hot Spot Determination, *Photogrammetric Engineering and Remote Sensing*, Vol. 39, No. 11, pp. 1205-1214.

Heller, R. C., 1968. Previsual Detection of Ponderosa Pine Trees Dying from Bark Beetle Attack, *Proceedings of the 5th Symposium on Remote Sensing of the Environment*, Ann Arbor, University of Michigan, pp. 287-434.

Huddleston, H. F., and E. H. Roberts, 1968. Use of Remote Sensing for Livestock Investories, *Proceedings of the 5th Symposium on Remote Sensing of the Environment*, Ann Arbor, University of Michigan, pp. 307-323.

Johnson, N. J., and S. Kotz, 1969. *Discrete Distributions*, Houghton Mifflin Co., Boston.

———, 1977. *Urn Models and Their Applications*, John Wiley & Sons, N.Y.

Maxim, L. D., and D. E. Cullen, 1977. A Cost Model for Remote Inspection of Ground Sites, *Photogrammetric Engineering and Remote Sensing*, Vol. 43, No. 8, 1977, pp. 1007-1025.

McDonnell, M. J., and H. J. Lewis, 1978. Ship Detection from Landsat Imagery, *Photogrammetric Engineering and Remote Sensing*, Vol. 44, No. 3, pp. 297-302.

Williman, J. J., and D. W. French, 1977. Detection of Oak Wilt with Color in Aerial Photography, *Photogrammetric Engineering and Remote Sensing*, Vol. 43, No. 10, pp. 1267-1272.

Wittes, J. T., *et al.*, 19??. Capture-Recapture Methods for Assessing the Completeness of Case Ascertainment When Using Multiple Information Sources, *J. Chron. Dis.*, Vol. 27, pp. 25-36.

(Received 10 May 1980; revised and accepted 24 November 1980)