L. Daniel Maxim
Everest Consulting Associates
Princeton Jct., NJ 08550
Harrison D. Weed
Leigh Harrington
Mathtech, Inc.
Arlington, VA 22209
Mary Kennedy
Consultant
Great Falls, VA 22066

Intensity Versus Extent of Coverage

Statistical models for the determination of an optimal balance between detection and sampling errors in aerial surveys are developed.

Abstract: *Statistical models for determination of an optimal balance between detection and sampling errors in aerial surveys are developed. Detection errors can often be reduced by the acquisition and exploitation of multiple images of the same scene. Sampling errors can be reduced by imaging a larger number of quadrats in the population. A fixed imagery budget implies a tradeoff between these errors.*

Models are presented here for various assumptions as to the independence (or lack of same) of detections in successive images and are illustrated with several numerical examples. Tables are presented to facilitate implementation.

Problem Definition

Considered here is the problem of estimating the total number of objects of interest in a geographical area subdivided into N quadrats or cells, each of a size covered by one image of a remote sensing system. If the objects can be detected with certainty and a total of N images can be acquired and exploited, a complete census is taken and (aside from clerical errors) the estimated total number of objects is correct. If the available image budget, B, is less than N, a statistical sample must be chosen and some *sampling error* will be attached to the estimated total number of objects. If, as well, inspection of imagery is not without error, a detection probability needs to be estimated and the observed objects in each cell need to be 'scaled up' by this detection probability. Even if all N cells were to be examined, the total number of objects would still be an estimate, reflecting some *detection error*.

For a fixed technology, the detection error in any cell can often be reduced (i.e., the detection probability increased) by an examination of multiple images of that cell or, in other words, by increasing the frequency of coverage or depth of exploitation. Consideration of detection error alone, therefore, would lead to a policy of acquiring multiple images of a given cell. Consideration of sampling error alone would lead to a policy of maximizing the number of cells examined or maximizing the extent of coverage. But, for a fixed imagery budget, a choice that increases the number of cells examined must also decrease the number of replications per cell. The problem, then, is to select the combination of the number of cells to be examined and the number of looks per cell that results in the smallest variance of the estimated total number of objects for a fixed imagery budget.

Solutions to this problem are presented assuming that detections are either independent or alternatively exhibit one of two forms of dependence. Convenient tables and charts are presented to determine the optimal number of looks per cell and characterize the sensitivity of this optimal solution.

Photogrammetric Engineering and Remote Sensing,
Vol. 47, No. 6, June 1981, pp. 789-797.

Preliminaries: Development of the Model

A geographical area consists of N cells or quadrats. Objects of interest are distributed among these cells. The numbers of objects in each cell are identically distributed with mean μ and variance σ^2 (or, if this assumption is untenable, the population can be subdivided into strata that have this property). Assume first that the probability of detection of any object in a single image of any cell is p, the single look detection probability, and further that detections are independent. (Later in the discussion this assumption will be relaxed.) If X_i is the number of objects observed in cell i, $i = 1 \ldots n$, then $y_i = X_i/p$ is an estimate of the true but unknown number of objects in the cell (see Johnson and Kotz (1969) for a description of the properties of this estimate). If n cells chosen at random are examined, then the average number of objects per cell, $\bar{y}$, and the estimated total number of objects, T, are given by

$$\bar{y} = \sum \frac{y_i}{n} = \sum \frac{X_i}{np}, \tag{1}$$

and,

$$T = N\bar{y} = \frac{N}{np} \sum X_i. \tag{2}$$

Now if l images of each of the n sample cells are examined, a total of ln images are required. The probability that an object is detected in at least one of l independent looks is

$$P = 1 - (1 - p)^l, \tag{3}$$

(To see this, note that the probability of a failure to detect on a single look is $(1 - p)$. If looks are independent, the probability that an object is not detected in any of l looks is $(1 - p)^l$ and, therefore, the probability of at least one detection in l looks is one minus this quantity, as shown in Equation 3.) Combining Equations 2 and 3, the estimate of the total number of objects, T, is given by

$$T = \frac{N}{n(1 - (1 - p)^l)} \sum X_i. \tag{4}$$

The variance of this estimate, σ_T^2, given a budget of $B = ln$ images, is shown in Appendix A to be given by

$$\sigma_T^2 = \frac{N^2}{B}\left[\frac{\mu l(1 - p)^l}{1 - (1 - p)^l} + l\sigma^2\right]. \tag{5}$$

Note that the boundary conditions on the parameters in Equation 5 satisfy known results. For example, when p is unity, the optimal choice for l is clearly one. Then $B = n$ and Equation 5 reduces to $\sigma_T^2 = N^2\sigma^2/n$, a well known formula in sampling theory in the absence of the finite population correction factor (e.g., see Cochran, 1963). When σ^2 is zero, all cells have the same number of objects, so there is nothing to be gained by increasing the extent of coverage and, therefore, l should be set equal to B, a consequence that follows from analysis of the properties of Equation 5. As p approaches zero, the variance of the estimated total becomes increasingly large, an intuitive result reflecting the increasing uncertainty of the estimator when detection probabilities are low.

Optimizing the Number of Looks

Inspection of the terms in brackets in Equation 5 reveals that the first term is a decreasing function of l while the second term is linearly increasing in l. Under appropriate circumstances, therefore, multiple looks will be optimal. These circumstances are readily determined. Note that the optimal number of looks will be unity whenever σ_T^2 based on two looks exceeds σ_T^2 for only one look per cell. This follows from the fact that the function can be shown to be convex. In terms of Equation 5 one look will be preferable to two looks whenever

$$\frac{2\mu(1 - p)^2}{1 - (1 - p)^2} + 2\sigma^2 \geq \frac{\mu(1 - p)}{1 - (1 - p)} + \sigma^2, \tag{6}$$

or, upon rearrangement, whenever

$$\frac{\sigma^2}{\mu} \geq \frac{p(1 - p)}{1 - (1 - p)^2}, \tag{7}$$

a critical value dependent upon p. Table 1 shows the critical value for σ^2/μ as a function of the single look detection probability p. Equation 7 agrees with intuition for it indicates that when σ^2/μ is large, i.e., sampling errors predominate, it is optimal to maximize cell coverage. However, when detection errors predominate, it is optimal to control these at the expense of sampling error. This point will be explored in greater detail later in the paper.

It is easy to show, using L'Hospital's rule, that the limit of Equation 7 as p approaches zero is 1/2. Thus, even if the single look detection probability is vanishingly small, multiple looks at quadrats can never be optimal if the distribution of the number of objects per quadrat has a ratio of σ^2/μ greater than 1/2. Such a specification excludes many discrete distributions. The Poisson distribution, for example, has $\sigma^2/\mu = 1.0$. As well, the discrete uniform distribution, i.e., $P(x) = 1/(r + 1)$, $x = 0, 1, \ldots, r$ has $\sigma^2/\mu \geq 0.5$ for all values of r. Similar statements can be made for the Negative Binomial, Logarithmic series, and Neyman Type A distributions. For multiple looks to be optimal, the distribution of the number of objects per quadrat must be more 'bunched' than any of these. Examples of discrete distributions for which $\sigma^2/\mu \leq 1/2$ are the Bernoulli (for values of $q \leq 1/2$), the binomial (for $q \leq 1/2$), and the geometric (for $p \geq 0.5$). Note: for a comprehensive list of discrete distributions and their properties, see Johnson and Kotz (1969).

TABLE 1. CIRCUMSTANCES WHERE $l^* > 1$

If p is	then multiple looks are profitable if σ^2/μ is less than
0.1	0.4737
0.2	0.4444
0.3	0.4118
0.4	0.3750
0.5	0.3333
0.6	0.2857
0.7	0.2308
0.8	0.1667
0.9	0.0909
1.0	0

For distributions where the ratio σ^2/μ is less than the critical value shown in Equation 7, Equation 5 will have an interior minimum. Disregarding the integer nature of l and treating Equation 5 as a continuous function, the optimal value of l, l^*, can be determined by setting the derivative of Equation 5 with respect to l equal to zero and solving. These steps produce the transcendental equation,

$$\frac{\sigma^2}{\mu} = \frac{-(1-p)^{l^*}(l^* \log_e(1-p) + 1 - (1-p)^{l^*})}{(1-(1-p)^{l^*})^2}, \quad (8)$$

which, though it does not have an analytical solution, clearly indicates that the optimal value of l is a function only of p and the ratio σ^2/μ. Equation 8 can be solved numerically using highly efficient root finding methods (e.g., Bolzano's method (Wilde, 1964)) to determine l^*. Equivalently, the original objective function (Equation 5) can be optimized directly for integer values of l by univariate search methods.

THE TRADEOFF ILLUSTRATED: A NUMERICAL EXAMPLE

Suppose the population, N, consists of 1600 quadrats; the budget, B, is 600 images; the single look detection probability is 0.5; σ is estimated to be 0.1; and the mean number of objects per quadrat, μ, is unity. Figure 1 shows how σ_T varies with l. For this case, the optimal number of looks, l^*, is nine and the corresponding value of σ_T, σ_T^*, is

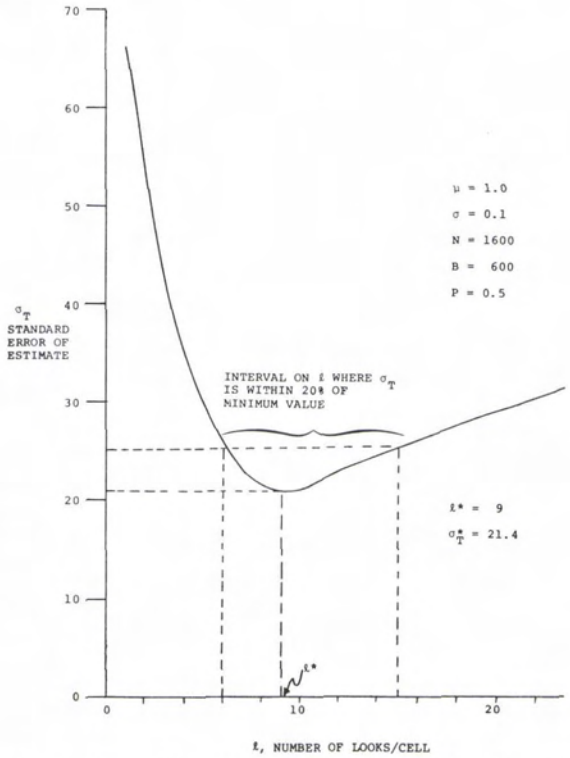

FIG. 1. Standard error of estimate, σ_T, as a function of the number of looks, l.

21.4. Though the sensitivity analysis shown in Table 2 suggests that the optimal choice of l is sensitive to the inputs σ, μ, and p, inspection of Figure 1 indicates that the penalty in σ_T of a suboptimal choice for l is modest. In this example, values of l varying by nearly a factor of three, i.e., from 6 to about 16, produce a σ_T within about 20 percent of the minimum value. Note also from Figure 1 that in this instance the function is not symmetrical. Overestimates of l^* have less effect on σ_T than underestimates of l^*, a point to be considered when inputs σ, μ, and p are estimates of perhaps considerable uncertainty.

Figure 2 shows how the optimal number of looks depends upon p for this example. The 'stairstep' appearance reflects the fact that l^* must be an integer.

TABLE 2. SENSITIVITY ANALYSIS ABOUT BASE CASE

PARAMETER	BASE CASE VALUE	RANGE OF VALUES OVER WHICH SOLUTION OPTIMAL		OPTIMAL VALUE OF l AT LEFT ENDPOINT	OPTIMAL VALUE OF l AT RIGHT ENDPOINT
μ	1	0.7	1.25	8	10
σ	0.1	0.085	0.12	10	8
p	0.5	0.483	0.525	10	8

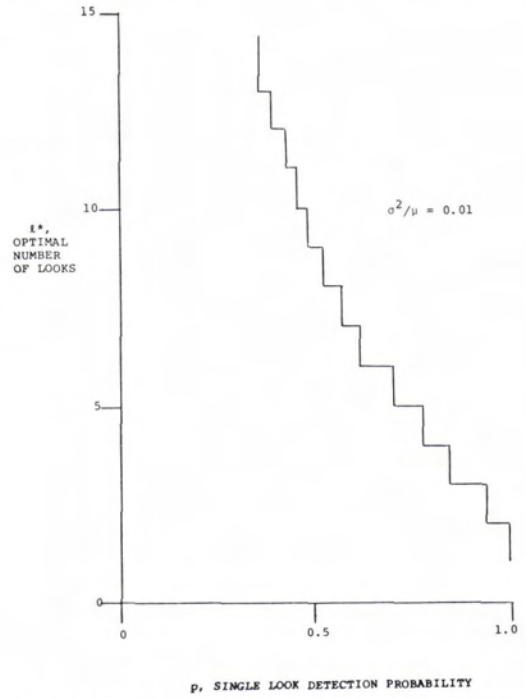

FIG. 2. The effect of detection probability on the optimal number of looks.

SAMPLE TABLES OF OPTIMAL SOLUTIONS

As was shown by Equation 8, l^* depends only upon p and the ratio of σ^2/μ. Thus, it is possible to create a simple and useful table to determine l^*. These results are displayed in Table 3. The entries show the optimal value of l, determined by numerical search techniques, corresponding to various values of p and σ^2/μ. Shown also in Table 3 are the endpoints of the range on l over which σ_T is within 20 percent of the minimum value (l, of course, must be greater than or equal to one). The entries under $p = 0.5$ and $\sigma^2/\mu = 0.01$ correspond to the numerical example just presented.

As can be seen from Table 3, *ceteris paribus*, the optimal number of looks

- decreases as the single look detection probability approaches unity and
- decreases as the 'noise to signal' ratio (σ^2/μ, of the distribution of number of objects) increases, or, in other words, as the sampling error increases.

Additionally, at least for small values of σ^2/μ, the number of looks varies with the single look detection probability in such a manner that the overall detection probability is roughly constant. Thus, for example, a total of 28 looks given a single look detection probability of 0.2 (optimal if σ^2/μ is 0.01) yields a detection probability of 0.998, almost exactly equivalent to 9 looks given a detection probability of 0.5. This rule breaks down as σ^2/μ approaches 0.5, and the optimal number of looks is equal to one regardless of detection probability.

Some further remarks are in order vis-a-vis the sensitivity of the solution, measured here as the width of the interval on l over which σ_T is within 20 percent of the optimal value:

- the function becomes highly sensitive when the detection error is small (i.e. p approaches unity) and/or the sampling error is small;
- overall, the detection error exerts more leverage on the sensitivity of the solution than does the sampling error; for example, l^* is 7 when σ^2/μ is 0.01 and p is 0.6 or when σ^2/μ is 0.25 and p is 0.2; however, the latter solution is much less sensitive to the choice on l; and, finally,
- as noted earlier for the specific numerical example, the '20 percent interval' is in general not symmetric about l^*; overestimates of l^* are less penalizing.

EXTENSIONS TO THE MODEL: DEPENDENT DETECTIONS

An important assumption of the foregoing model is that of the statistical independence of detections on successive looks at a quadrat. This assumption is reflected, for example, in Equation 3 specifying how the probability of detection increases with the number of looks. It is also reflected at several places in the derivation of Equation 5.

Experience and intuition suggest, however, that the assumption of independence is, at best, only approximately correct. The reason that an object is not detected in one look, for example, may relate to phenomena that confer independence (e.g., weather, interpreter oversight, or terrain 'shadow' given a "random" access path, etc.). Alternatively, other reasons for failure to detect may have been operative, such as

- if each cell is imaged with the same geometry on each access, then any fixed objects missed because of terrain masking will continue to escape detection (in this case, the detection probability of those objects will not depend upon the number of looks); or
- if the objects of interest are crops and the first looks failed to provide detection and identification because the crop had not yet reached a growth stage where detection was possible, then subsequent looks at a later date may have increasing probabilities of detection (alternatively, if the crop has been harvested, then detection probabilities may decrease, as reported, for example, by Heller and Johnson (1979) in noting the decrease in IR reflectance of wheat fields after harvest); or
- if the object to be detected might have some reason to be hidden—as, for example, an illicit whiskey still in Virginia (Franklin, 1980), a surface mining violation in eastern Kentucky (Maxim and Cullen, 1977), or the appearance of a new air-defense missile—then a reconnaissance attempt may stimulate (further) attempts at camouflage

TABLE 3. OPTIMAL VALUES FOR l

Ratio of Noise to Signal σ^2/μ	Single Look Detection Probability, p									
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
	67	32	20	14	10	8	6	4	3	1
0.005	48 112	23 53	14 33	10 23	8 17	6 13	5 9	4 7	3 5	1 1
	60	28	18	12	9	7	5	4	3	1
0.01	40 101	19 48	12 30	9 2	6 16	5 11	4 8	3 6	2 4	1 1
	49	23	14	10	7	6	4	3	2	1
0.025	29 89	14 40	9 25	6 17	5 13	4 10	3 7	2 5	2 4	2 1
	40	19	12	8	6	5	4	3	2	1
0.05	21 24	10 31	6 21	5 15	4 11	3 8	2 6	2 4	1 3	1 1
	30	14	9	6	5	4	3	2	1	1
0.10	11 61	6 28	4 18	3 12	2 9	2 7	1 5	1 4	1 3	1 1
	15	7	5	3	2	2	1	1	1	1
0.25	1 41	1 19	1 12	1 8	1 6	1 4	1 3	1 2	1 1	1 1
	1	1	1	1	1	1	1	1	1	1
0.50	1 22	1 10	1 6	1 4	1 3	1 2	1 2	1 1	1 1	1 1

(in this situation the detection probability on subsequent looks might be expected to be lower).

The above examples not only serve to indicate how dependence can arise but also point to a fundamental difficulty in any "generic" view of dependence, its lack of uniqueness. The situation is analogous to a challenge to the assumption of linearity of a mathematical function. The phrase "non-linear" admits a large class of functions. 'Lack of independence' admits the same large class of possible detection functions. The approach taken here is to consider some specific models for dependent looks. These models are 'reasonable' but are by no means exhaustive of the logical possibilities, given a failure of the independence assumption.

An obvious challenge point in a critical view of the independence assumption is the detection probability function specified in Equation 3. Logic suggests that attention be limited to the class of monotonic non-decreasing functions as, in the absence of the possibility of false positives, increased looks cannot *lower* the cumulative detection probability. This limiting condition is reached when the detection probability remains constant as the number of looks is increased. For this case the optimal number of looks must equal unity, because any increase in l must increase the sampling error while leaving the detection error unchanged, a counterproductive change.

An alternative detection probability model follows from the assumption that, if an object is not detected in the first image of a cell (an event with probability p), the probability of detection on each subsequent image is αp, where by hypothesis, $0 < \alpha < 1.0$. The detection function corresponding to Equation 3 is, under this assumption,

$$P = 1 - (1 - p)(1 - \alpha p)^{l-1}, \qquad (9)$$

and results in an obvious modification to Equation 5. Figure 3 shows the same numerical example as illustrated in Figure 1 for various values of α. The case $\alpha = 1$ corresponds exactly to the independent case. Where $\alpha < 1$, the standard error is identical to the base case for l equal to one, but for large l becomes approximately equivalent to the independent case if p is reduced by the factor α. Thus, for example, when $\alpha = 0.6$, the variance function approaches the independent case with $p = 0.6(0.5)$ or 0.3. Ultimately, as α continues to decrease, however, a point will be reached where the interior minimum of the variance function exceeds the variance corresponding to the case $l = 1.0$, and one look becomes optimal. For the numerical example shown in Figure 3, the critical value of α is approximately 0.129. That is, for values of α greater than 0.129, the optimal number of looks *increases* as α decreases (indeed, l^* is roughly equal to that corresponding to the solutions in Table 3 for $pl = \alpha p$). For values of α less than the critical value,

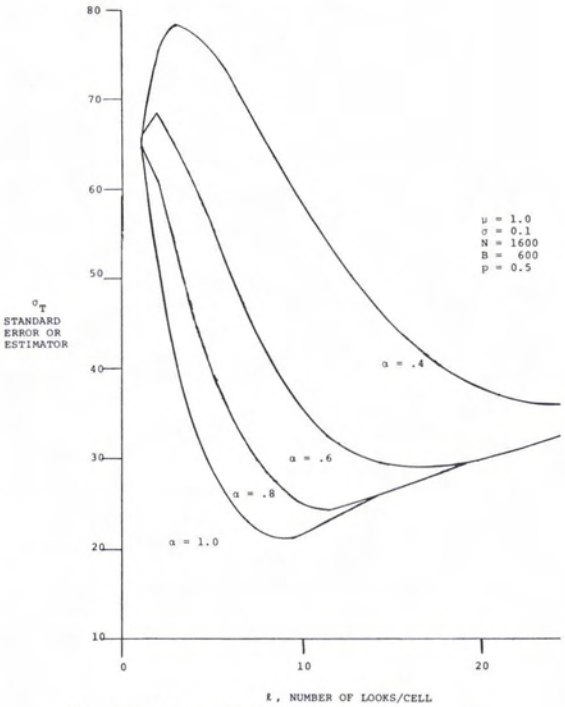

FIG. 3. Dependent detections, one view.

0.129 in this example, the optimal number of looks is unity. Table 4 shows results for this example. The critical value of α is a function of the ratio, σ^2/μ, and the single look detection probability, p. To illustrate, if p is increased in this example from 0.5 to 0.8, the computed critical value of α increases to approximately 0.25.

Optimization of the variance function for this model is most conveniently done by a numerical search of the objective function. In contrast to the case where detections were independent, this function is *not unimodal* as can be seen in Figure 3, however, so that efficient search techniques re-

TABLE 4. OPTIMAL VALUES FOR l, DEPENDENT DETECTIONS
$\mu = 1.0$, $\sigma = 0.1$, $N = 1600$, $B = 600$

VALUE OF α	l^*	$\sigma_T{}^*$
1.0	9	21.4
0.9	10	22.9
0.8	12	24.6
0.7	14	26.6
0.6	16	29.1
0.5	20	32.2
0.4	25	36.3
0.3	35	42.3
0.2	53	52.3
0.129	82	65.5
0.128	1	65.7
<0.128	1	65.7

quiring this property (e.g., Fibonacci methods (Wilde, 1964)) cannot be used.

Yet another dependent detection model may be appropriate in some circumstances. Suppose, for example, that multiple looks are always acquired with the same access geometry. Objects that are concealed from this look angle (e.g., masked by terrain shadow, located along tree lines, etc.) will not be detected regardless of number of looks. Even if the access geometry were to change, certain objects (e.g., wildlife in forests) might not be detectable. Such examples suggest an alternative model of the form:

- a certain fraction, θ, of the objects is "detectable" (i.e., not permanently obscured);
- detectable objects satisfy the assumptions of the independent detection model; but
- a fraction, $1 - \theta$, of the objects cannot be detected at all.

Mathematically, the probability that an object, selected at random, is detected in l looks is given by the equation,

$$P = \theta(1 - (1 - p)^l), \tag{10}$$

which should be compared with Equations 3 and 9. As before, this detection model can be inserted into the formula for the variance of the estimate of the total number of objects in N quadrats. Figure 4 shows the original example displayed in Figure 1 for the 'theta detection model' for various values of θ. As can be seen,

- the model is quite sensitive to the value of θ (if only 10 percent of the objects are hidden (i.e., $\theta = 0.9$), the best attainable value of σ_T increases nearly threefold from 21.4 to 57.7), and
- unlike the effect of α in the earlier model, the effect of decreasing θ (increasing the fraction of undetectable objects) is to decrease l^* (in this instance over 55 percent from 9 looks (the base case) to 4 looks), and, finally,
- the sensitivity of σ_T to a suboptimal choice for l increases as the fraction of undetectable objects increases.

These two models of dependent detections illustrate the difficulty of making categorical statements about 'lack of independence' other than it acts to increase the uncertainty of estimates or (if not properly accounted for) introduce bias in the estimates. In the first instance (the 'alpha' model) the effect was to increase l^*, whereas the presence of hidden objects acts to decrease l^*. A hybrid model would have α, θ pairs that leave l^* unchanged, though of course increasing the error of the estimate.

Summary and Conclusions

In remote sensing applications where only a sample of quadrats can be examined and where detection probabilities are less than unity, a tradeoff between detection and sampling errors

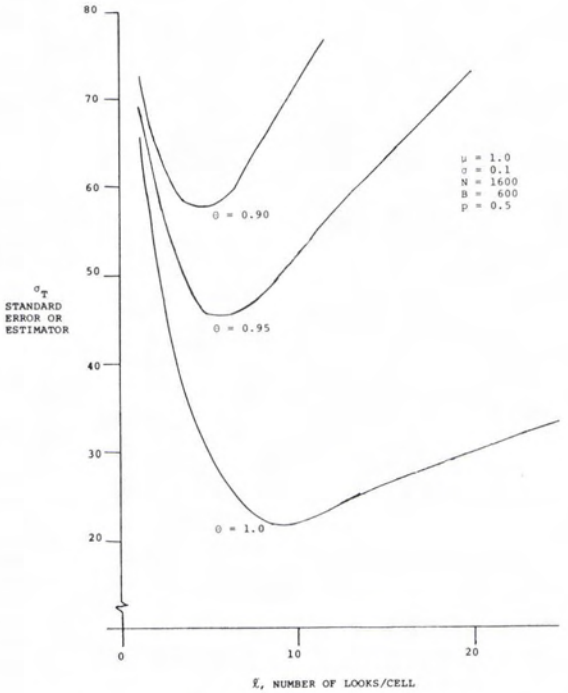

FIG. 4. Dependent detections, another view.

results. The model(s) presented here enable an optimal balance to be struck between these errors by selecting the frequency and extent of coverage to minimize the variance of the desired estimate.

If detection can be assumed to be independent, then the optimal number of looks is inversely proportional to both the detection probability and the noise/signal ratio (N/SR) of the distribution of objects per quadrat. There exists a critical noise/signal ratio which is a function of the detection probability such that, if the actual noise/signal ratio exceeds this threshold, multiple looks per cell cannot be optimal. For distributions with N/SR less than the critical value, tables are presented which enable determination of the optimal number of looks.

Two models that include dependent looks are developed and optimal solutions to the multiple look problem are presented. In one model, in which a failure to detect on the first look is associated with lower subsequent single look probabilities, the optimal number of looks increases over the case of the independent model. In another model, in which a certain fraction of objects are 'non-detectable,' the optimal number of looks decreases.

Aside from the specific *constructs* presented here, the general *concept* of trade-off between detection and sampling errors is central to many remote sensing applications and merits further inquiry. The choice of platform and resolution ver-

sus coverage issues are identical in principle and can be explored by related models. Green *et al.* (1977) for example, have measured the effect of imagery scale and film type on detection probabilities. Their results indicate generally that detection probabilities decrease as photographic scale decreases for scales smaller than 1:1000 (a consequence of increased image distortion). Yet, for a fixed imagery budget, the smaller the scale the smaller the sampling error: exactly the tradeoff discussed here (Appendix B shows how this is done). When updated to account for the effects of inflation, the cost model and data presented by Ulliman (1975) can be imbedded in the model framework so that a minimum cost survey of specified accuracy can be designed.

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: Derivation of the Variance of the Estimator

The estimated total number of objects is given by Equation 4 as

$$T = \frac{N}{n(1 - (1 - p)^l)} \sum X_i. \tag{A-1}$$

This appendix derives the variance of the estimator, σ_T^2, given in Equation 5 under the assumptions of constant detection probability, p, and independent detections. The desired result is

$$\sigma_T^2 = \frac{N^2}{B}\left[\frac{\mu l(1 - p)^l}{1 - (1 - p)^l} + l\sigma^2\right]. \tag{A-2}$$

Let Z_i be the true number of objects in cell i. The Z_i are assumed to be independent realizations of a random variable, say Z, with mean μ and variance σ^2. That is, $E(Z) = \mu$ and $V(Z) = \sigma^2$. For a given cell and fixed Z_i, the observed number of objects in cell i, X_i, is a binomial random variable, where Z_i is the number of trials and $P = 1 - (1 - p)^l$ is the probability of success on any trial.

Thus, using conditional probability notation,

$$E(X_i \mid Z_i) = Z_i P, \quad V(X_i \mid Z_i) = Z_i P(1 - P). \tag{A-3}$$

That is, assuming Z_i as given, the mean and variance of X_i are $Z_i P$ and $Z_i P(1 - P)$, respectively. The unconditional variance of X_i is obtained from the following relationship:

$$V(X_i) = E_z\left[V(X \mid Z)\right] + V_z\left[E(X \mid Z)\right], \tag{A-4}$$

where E_z and V_z denote expectation and variance with respect to the probability distribution of Z. Using this relationship,

$$V(X_i) = E_z\left[Z_i P(1 - P)\right] + V_z\left[Z_i P\right], \tag{A-5}$$

$$= \mu P(1 - P) + \sigma^2 P^2. \tag{A-6}$$

Returning now to the estimator, T, with $P = 1 - (1 - p)^l$, the variance of T is

$$\sigma_T^2 = \frac{N^2}{n^2 P^2} V(\sum X_i) \tag{A-7}$$

$$= \frac{N^2}{n^2 P^2} \sum V(X_i). \tag{A-8}$$

The last equation uses the fact that the X_i are assumed to be independent. Now, applying the previously derived result,

$$\sigma_T^2 = \frac{N^2}{n^2 P^2}\, n\left[\mu P(1 - P) + \sigma^2 P^2\right]. \tag{A-9}$$

Replacing P by $1 - (1 - p)^l$ and $n = B/l$,

$$\sigma_T^2 = \frac{N^2}{B}\left[\frac{\mu l(1 - p)^l}{1 - (1 - p)^l} + l\sigma^2\right]. \tag{A-10}$$

This is the desired result given in Equation A-2.

Appendix B: The Choice of Optimum Scale

This appendix modifies the formulae presented in the main paper to encompass the detection versus scale tradeoff. The logical starting point is Equation A-9, which specifies how the variance of the estimate of the total number of objects is related to the number of areas sampled, n, detection probability, P, etc. Denote by the subscript s the scale of the photo; e.g., P_s is the detection probability associated with images of scale, s. Let R_s be the relative number of quadrats that are contained in an image of scale, s. To compute R_s requires specification of a reference scale as a *numeraire*. For example, assume this reference is a scale of 1:1000, i.e., when s is 1:1000, $R_s = 1.0$. To continue, when s = 1:2000, $R_s = 4$, when s = 1:3000, $R_s = 9$, etc. Given these definitions, it follows from Equation A-9 that (to within a normalizing constant, ψ)

$$\sigma_T^2 = \frac{\psi}{R_s}\left[\frac{\mu(1 - P_s)}{P_s} + \sigma^2\right]. \tag{B-1}$$

Provided empirical detection probability versus scale curves can be developed, the above equation can be used to select the scale so as to minimize the variance of the estimator. As a concrete illustration, consider a portion of the data from Green *et al.* (1977) on detection of Imported Fire Ant (IFA) mounds summarized and reproduced in Table B1 following.

In August 1972, Green reported the average number of IFA mounds/acre as 79.7. No figure was given for the variance, σ^2, about this mean. Equation B-1 can be used to evaluate σ_T^2 at each photographic scale. The three choices are summarized below:

TABLE B1. DETECTION PROBABILITY VS. PHOTOGRAPHIC SCALE: EMPIRICAL DATA (B&W IR FILM AUGUST 1972)

PHOTOGRAPHIC SCALE	R_s	DETECTION PROBABILITY P_s
1:1000	1	0.324
1:2000	4	0.183
1:3000	9	0.099
1:4000	16	—

SOURCE: Green *et al.* (1977).

$$\text{at } s = 1{:}1000 \quad \left[79.7(0.676)/0.324 + \sigma^2\right]$$
$$\text{at } s = 1{:}2000 \quad 1/4\left[79.7(0.817)/0.183 + \sigma^2\right]$$
$$\text{at } s = 1{:}3000 \quad 1/9\left[79.7(0.901)/0.099 + \sigma^2\right].$$

In this instance, it is unnecessary to know σ^2 as it is easily shown that, for any $\sigma^2 > 0$, the scale 1:3000 produces the smallest value of $\sigma_{\hat{T}}^2$. This is an interesting result, since it implies a selection of scale where detection probabilities are relatively small, lending further insight to the recommendations of Green *et al.* (1977), who state:

> "... Photographic missions should be flown at the highest altitude *that still provides effective detection of ant mounds.* Increases in altitude cause a decrease in photographic scale but increase the area included in each photographic frame. Hence, the number of photographs and associated cost required for area coverage should be reduced by selecting a film type that would provide *accurate mound detection information* at the *highest altitudes possible to record accurately the presence of mounds.*" (Emphasis added)

In this case, accurate detection, measured in terms of the variance of the estimated total number of mounds does not necessarily correspond to the altitude where detection probabilities are the highest. Rather, it corresponds to the altitude where the variance of the estimate is smallest.

REFERENCES

Cochran, W. G., 1963. *Sampling Techniques* (second edition) New York: John Wiley & Sons.

Franklin, B. A., 1980. Revenuers (and their TNT) visit Corn Liquor Country, *The New Times,* Tuesday, January 22, 1980.

Green, L. R., 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. 1051-1057.

Heller, R. C., and K. A. Johnson, 1979. Estimating Irrigated Land Acreage from Landsat Imagery, *Photogrammetric Engineering and Remote Sensing,* Vol. 45, No. 10, pp. 1379-1386.

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

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, pp. 1009-1025.

Ulliman, J. J., 1975. Cost of Aerial Photography, *Photogrammetric Engineering and Remote Sensing,* Vol. 41, No. 4, pp. 491-497.

Wilde, Douglass J., 1964. *Optimum Seeking Methods,* Englewood Cliffs, N.J.: Prentice Hall.

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