

Optimal allocation of sample sizes between regular banding and radio-tagging for estimating annual survival and emigration rates

MARLINA D. NASUTION¹, CAVELL BROWNIE² & KENNETH H. POLLOCK², ¹Family Health International, USA and ²Department of Statistics, North Carolina State University, USA

ABSTRACT *Many authors have shown that a combined analysis of data from two or more types of recapture survey brings advantages, such as the ability to provide more information about parameters of interest. For example, a combined analysis of annual resighting and monthly radio-telemetry data allows separate estimates of true survival and emigration rates, whereas only apparent survival can be estimated from the resighting data alone. For studies involving more than one type of survey, biologists should consider how to allocate the total budget to the surveys related to the different types of marks so that they will gain optimal information from the surveys. For example, since radio tags and subsequent monitoring are very costly, while leg bands are cheap, the biologists should try to balance costs with information obtained in deciding how many animals should receive radios. Given a total budget and specific costs, it is possible to determine the allocation of sample sizes to different types of marks in order to minimize the variance of parameters of interest, such as annual survival and emigration rates. In this paper, we propose a cost function for a study where all birds receive leg bands and a subset receives radio tags and all new releases occur at the start of the study. Using this cost function, we obtain the allocation of sample sizes to the two survey types that minimizes the standard error of survival rate estimates or, alternatively, the standard error of emigration rates. Given the proposed costs, we show that for high resighting probability, e.g. 0.6, tagging roughly 10-40% of birds with radios will give survival estimates with standard errors within the minimum range. Lower resighting rates will require a higher percentage of radioed birds. In addition, the proposed costs require tagging the maximum possible percentage of radioed birds to minimize the standard error of emigration estimates.*

Correspondence: M. D. Nasution, Family Health International, PO Box 13950, Research Triangle Park, NC 27709, USA. E-mail: MNasution@fhi.org

1 Introduction

To obtain more information in studying a particular animal population, biologists conduct multiple surveys using more than one marking technique. They obtain several data sets corresponding to the different types of marks because of the use of multiple marks. Rather than analysing each data set separately, a comprehensive analysis should be carried out on the combined data sets in order to obtain as much information as possible about the biology and sampling processes. Constructing a joint likelihood for the combined data sets may permit the estimation of a greater number of parameters and can increase the precision of estimates compared with those from the separate analyses (Burnham, 1993; Lebreton *et al.*, 1995; Tsai, 1996; Barker, 1997; Catchpole *et al.*, 1998; Nasution *et al.*, 2001).

In studies where dispersal or movement between sites are of interest, the use of more than one type of mark could have important advantages. Given resighting data from all sites, movement and survival can be separately estimated only if it is assumed that all movement occurs at the beginning of the survival period (Brownie *et al.*, 1993). If in addition to the resighting data, telemetry data are obtained at all sites at more frequent intervals, a combined analysis of both data types should permit separate estimation of survival and movement under less restrictive assumptions.

Developing the appropriate joint analysis for studies with more than one type of mark is therefore an important area of research (Schwarz & Seber, 1999). A related problem is to determine the most cost effective allocation of resources to the two (or more) types of recapture survey. For example, if the study will involve both regular banding and radio tagging surveys followed by a combined analysis of the resulting data, it is necessary to consider allocation of funds to cover both resighting and relocation efforts, as well as the processes of capturing and marking new animals using a leg band and/or a radio tag. Since the costs of marking and recapture, as well as the recapture rates, are not the same for different types of marks, the allocation of resources between radio tags and leg bands should be considered carefully.

To optimize the allocation of the different mark types, a clear objective is required (Barker *et al.*, 1993). When the main reason for tagging animals is to obtain information about survival and movement parameters, we could use, for example, the precision of survival and/or movement estimates to develop optimality criteria. When our main interest is estimation of survival rates, assuming a fixed total cost, the allocation of animals will be optimal if the variance of survival estimates corresponding to this allocation is minimum. Alternatively, if emigration is the parameter of greatest interest, assuming a fixed total cost, the allocation of animals will be optimal if the variance of emigration estimates corresponding to this allocation is minimum.

In this paper, we investigate the optimal allocation of leg bands and radio tags in a study that will include both resighting and radio-telemetry surveys, where the availability of telemetry data enables us to consider either the precision of true survival estimates or the precision of emigration estimates as the optimality criterion. For simplicity, we consider a study in a single location, so that emigration means dispersal or movement away from the study site. Since the costs associated with radio tags and monitoring using radios are high compared to the cost of leg bands and subsequent resighting surveys, we assume that the cheaper leg bands are placed on all marked birds while radios are placed on relatively few birds. We

assume annual resighting surveys and consider different sampling frequencies for the telemetry surveys, as well as different costs, in studying optimal allocation.

2 Cost function

To construct a cost function for a study involving both resighting and telemetry surveys, there are numerous practical issues relating to the timing and location of the surveys that must be considered. To be specific, we base the likelihood and cost function on the sampling procedures described by Bennetts *et al.* (1999) and Nasution *et al.* (2001) for a study on a Snail Kite population in the Florida Everglades. The following simplifying assumptions are made.

- Telemetry and resighting surveys are conducted independently in the same area at the start of each year. Also, additional telemetry surveys may be conducted at regular intervals within each year.
- All newly marked birds are released at the beginning of the study.
- All birds are marked as adults.

Note that, although we use a specific likelihood, the general principles of constructing a cost function can be applied broadly.

2.1 Cost function for the regular banding survey

In a regular banding study, the activities related to the cost are:

- capture of unmarked birds at each sampling time. The corresponding costs are a function of the worker-hours needed per capture as well as equipment costs (e.g. nets).
- banding newly captured birds. Costs are related to worker-hours needed to apply leg bands and the costs of the bands.
- resighting the birds at each sampling time. Costs are related to equipment needed for the survey and to the time spent resighting birds.

For this situation, we propose the following cost function:

$$C^s = c_1^{0s} N_s + c_1^{1s} N_s + \sum_{i=2}^{K_s+1} c_i^{2s} \tag{1}$$

where:

- K_s = the total number of annual resighting occasions,
- N_s = the number of newly banded birds released at time 1,
- c_1^{0s} = cost of capturing a bird at sampling time 1,
- c_1^{1s} = banding cost per bird at sampling time 1,
- c_i^{2s} = cost of the resighting survey (including equipment and vehicle costs) at sampling time i , $i = 2, \dots, K_s + 1$

It is certainly possible to consider increasing the effort of resighting by modifying the last part of equation (1) as a function of N_s ; however, it is beyond the scope of this paper. Note that, under this function, the capture and banding costs depend on the number of birds that are banded, but the resighting cost is assumed to be based on a predetermined limit of the effort to be expended and is therefore independent of the number of marked birds.

2.2 Cost function for the radio-tagging survey

The activities related to the cost of a radio-tagging study are:

- capture of unmarked birds at each sampling time. The corresponding costs are a function of the worker-hours needed per capture as well as equipment costs (e.g. nets).
- tagging newly captured birds. Costs are related to worker-hours and to the cost of the radio tags and other equipment. Radios are assumed to last for all years of the radio tagging survey.
- relocation of both live and dead birds at each sampling time. Costs are related to the time spent tracking birds with radio equipment, including worker-hours and vehicle operating costs.

Under these activities, the cost function proposed is:

$$C^r = (C_{1,1}^{0r} + c_{1,1}^{1r})N_r + \sum_i \sum_j (c_{i,j}^{2r} + c_{i,j}^{3r}(E(m_{i,j}^r + d_{i,j}^r))) \tag{2}$$

where:

- $\sum_i \sum_j$ = sum over relocating occasions at year i and sampling time j within year i such that $j = 2, \dots, k$ if $i = 1$ or $j = 1, \dots, k$ if $i = 2, \dots, K_r$ or $j = 1$ if $i = K_r + 1$,
- K_r = the total number of years in which marked birds are tracked using radio telemetry,
- k = the number of relocating efforts per year,
- N_r = the number of new releases at sampling time 1,
- $E(m_{i,j}^r)$ = the expected number of radioed birds relocated alive at the j th sample in year i ,
- $E(d_{i,j}^r)$ = the expected number of radioed birds relocated dead at the j th sample in year i ,
- $c_{1,1}^{0r}$ = cost of capturing a bird at sampling time 1,
- $c_{1,1}^{1r}$ = cost per bird for radio tags at sampling time 1,
- $c_{i,j}^{2r}$ = fixed equipment costs per telemetry survey at the j th sample in year i ,
- $c_{i,j}^{3r}$ = cost per bird of relocating the bird at the j th sample in year i .

Again, capture and tagging costs are assumed to depend on the number tagged (N_r) and, in contrast to the resighting survey, relocation costs also depend on the number of tagged birds because an effort is made to locate the signal of each individual radio tag.

If each bird receives only one type of mark (i.e. either a leg band or a radio tag), the cost function in the combined surveys is the sum of equations (1) and (2). The more common procedure is to place the cheaper leg bands on all birds and, in addition, to place radio tags on a subset of these. In this case, the total number of marked birds is N_s and the capture costs do not depend on the type of tag received (i.e. $c_{1,1}^{0r} = c_1^{0s} = c_1^0$). The cost function for the combined surveys is therefore:

$$C = c_1^0 N_s + c_1^{1s} N_s + c_{1,1}^{1r} N_r + \sum_i c_i^{2s} + \sum_i \sum_j (c_{i,j}^{2r} + c_{i,j}^{3r}(E(m_{i,j}^r + d_{i,j}^r))) \tag{3}$$

Using this combined cost function, given a fixed total cost C , we will be able to write N_s as a function of N_r or vice versa.

3 Optimal allocation for estimation of true annual survival and emigration

3.1 Likelihood function

Given resighting data from a banding study only, survival rate estimators are actually estimators of ‘apparent survival’, where apparent survival represents being alive and present in the study area. An important advantage of telemetry data is that (under certain assumptions about the emigration process) true survival (ϕ) and fidelity (F) can be estimated separately. To consider optimal allocation of resources to resighting and telemetry surveys, we focus on separately estimating true survival and emigration rates ($1 - F$), assuming it is of interest to distinguish between these processes. Our objective is, therefore, to find the optimal allocation of resources by minimizing the standard error of the estimator of true annual survival in the combined surveys of resighting and radio tagging given a specific cost. Alternatively, the same approach can be used with minimizing the standard error of the estimator of emigration.

We consider studies involving resighting and radio tagging surveys where all birds are leg-banded and some of the birds are radio-tagged. For all cases, we consider a single release of marked birds followed by two yearly resighting surveys ($K_r = 2$). With respect to the telemetry surveys, we consider three situations each with marked birds followed for two years ($K_r = 2$) and with the number of surveys per year as follows:

- Case 1:** yearly radio-telemetry survey ($k = 1$).
- Case 2:** six-monthly radio-telemetry survey ($k = 2$).
- Case 3:** three-monthly radio-telemetry survey ($k = 4$).

In theory, we can allow more frequent telemetry surveys, but computations are tedious for $k \geq 4$. However, certain consistent findings should carry over to situations with more frequent telemetry surveys.

To develop a likelihood for the combined data, we also assume that the probability of radio failure is 0, that dead birds are relocated at the end of the sampling interval and that resighting and telemetry efforts do not interfere with each other. In addition, we assume that emigration is permanent and occurs at the start of each sampling interval.

Modifying the likelihood in Nasution *et al.* (2001) to account for emigration and true survival, the joint likelihood for data from the combined surveys is:

$$L = L_1 \times L_2 \times L_3 \tag{4}$$

where L_1 and L_3 pertain to the birds that receive both radio tags and bands (radioed birds) and L_2 pertains to the birds that receive leg bands only (non-radioed birds). Here,

$$L_1 = \prod_{i=1}^{K_r} \prod_{j=1}^k \left\{ \frac{R_{i,j}^r j!}{\prod_{i',j'} [m_{i,j,i',j'}^r d_{i,j,i',j'}!]} \left[\prod_{i',j'} \pi_{i,j,i',j'}^{m_{i,j,i',j'}^r} \pi_{d_{i,j,i',j'}} \right] \chi_{i,j}^{r(R_{i,j}^r - r_{i,j}^r)} \right\} \tag{5}$$

with:

$R_{i,j}^r$ = the number of radioed birds released after the j th sample in year i including newly and previously marked birds ($i = 1, \dots, K_1, j = 1, \dots, k$),

- $r_{i,j}^r$ = the total number of radioed birds later relocated by radio, or by resighting, of the $R_{i,j}^r$ birds released after the j th sample in year i ($i = 1, \dots, K_r, j = 1, \dots, k$). The birds may be relocated alive or dead,
- $m_{i,j,i',j'}^r$ = the number of radioed birds released, or relocated by radio in the j th sample in year i and next relocated alive in the j th sample in year i' ,
- $d_{i,j,i',j'}$ = the number of radioed birds released or relocated by radio in the j th sample in year i and next relocated dead in the j th sample in year i' ,
- $\pi_{li,j,i',j'}$ = the probability that a radioed bird last relocated at year i , time j is next relocated alive at year i' , time j' ; $i' = i, \dots, K_r + 1$, and $j' = j + 1, \dots, k$ if $i' = i$ or $j' = 1, \dots, k$ if $i' > i$ or $j' = 1$ if $i' = K_r + 1$,
- $\pi_{di,j,i',j'}$ = the probability that a radioed bird last seen at year i , time j is next found dead at year i' , time j' ; $i' = i, \dots, K_r$, and $j' = j + 1, \dots, k$ if $i' = i$ or $j' = 1, \dots, k$ if $i' > i$ or $j' = 1$ if $i' = K_r + 1$,

and

- $\chi_{i,j}^r$ = the probability a radioed bird released at sampling time j is never found again (i.e. never relocated again by radio and never resighted) ($i = 1, \dots, K_r, j = 1, \dots, k$),

$$L_2 = \prod_{i=1}^{K_s} \left\{ \frac{R_i^{s*}!}{\prod_{j=i+1}^{K_s+1} m_{i,j}^{s*}!} \left[\prod_{j=i+1}^{K_s+1} \pi_{si,j}^{m_{i,j}^{s*}} \right] \chi_i^{s(R_i^{s*} - r_i^{s*})} \right\} \tag{6}$$

with:

- R_i^{s*} = the number of non-radioed birds released or resighted at year i ($i = 1, \dots, K_s$),
- r_i^{s*} = the total number of non-radioed birds resighted after year i of the R_i^{s*} birds released at i ($i = 1, \dots, K_s$),
- $m_{i,j}^{s*}$ = the number of non-radioed birds released or resighted at year i that are next resighted at year j ($i = 1, \dots, K_s, j = i + 1, \dots, K_s + 1$),
- $\pi_{si,j}$ = the probability any banded bird alive at year i is next resighted at year j , $i = 1, \dots, K_s, j = i + 1, \dots, K_s + 1$,
- χ_i^s = the probability a non-radioed bird released or resighted at year i is never resighted again ($i = 1, \dots, K_2$),

and

$$L_3 = \prod_{i=2}^{\min(K_r, K_s) + 1} \frac{m_{i,i,1}^r!}{m_{i,i,i,1}^{rs}! (m_{i,i,i,1}^r - m_{i,i,i,1}^{rs})!} \pi_{rsi}^{m_{i,i,i,1}^{rs}} (1 - \pi_{rsi})^{(m_{i,i,i,1}^r - m_{i,i,i,1}^{rs})} \tag{7}$$

where:

- $m_{i,i,i,1}^r$ = the total number of radioed birds relocated by radio at the start of year i ,
- $m_{i,i,i,1}^{rs}$ = the total number of the $m_{i,i,i,1}^r$ radioed birds relocated by radio at the start of year i that are also resighted at the start of year i ($i = 2, \dots, K_r$),
- π_{rsi} = the probability a bird alive at year i is resighted at year i given it is also relocated at i .

Assuming that survival is constant within years and that emigration occurs with the same probability following each telemetry survey, then the cell probabilities $\pi_{li,j,i',j'}$, $\pi_{di,j,i',j'}$, $\pi_{si,j}$ and π_{rsi} are functions of the following parameters.

TABLE 1. Selected cell probabilities for the product-multinomial likelihood with $k = 4$

Parameter	When		Cell probability
	last seen	next seen	
$\pi_{n,1,1,2}$	1,1	1,2	$\phi_1^{1/4} F_1^{1/4} p_{r1,2}$
$\pi_{n,1,1,3}$	1,1	1,3	$\phi_1^{1/4} F_1^{1/4} q_{r1,2} \phi_1^{1/4} F_1^{1/4} p_{r1,3}$
$\pi_{n,1,1,4}$	1,1	1,4	$\phi_1^{1/4} F_1^{1/4} q_{r1,2} \phi_1^{1/4} F_1^{1/4} q_{r1,3} \phi_1^{1/4} F_1^{1/4} p_{r1,4}$
$\pi_{n,1,1,2,1}$	1,1	2,1	$\phi_1^{1/4} F_1^{1/4} q_{r1,2} \phi_1^{1/4} F_1^{1/4} q_{r1,3} \phi_1^{1/4} F_1^{1/4} q_{r1,4} \phi_1^{1/4} F_1^{1/4} p_{r2,1}$
\vdots			
$\pi_{d1,1,1,2}$	1,1	1,2	$(1 - \phi_1^{1/4}) F_1^{1/4} p_{r1,2}$
$\pi_{d1,1,1,3}$	1,1	1,3	$\phi_1^{1/4} F_1^{1/4} q_{r1,2} (1 - \phi_1^{1/4}) F_1^{1/4} p_{r1,3}$
$\pi_{d1,1,1,4}$	1,1	1,4	$\phi_1^{1/4} F_1^{1/4} q_{r1,2} \phi_1^{1/4} F_1^{1/4} q_{r1,3} (1 - \phi_1^{1/4}) F_1^{1/4} p_{r1,4}$
$\pi_{d1,1,1,2,1}$	1,1	2,1	$\phi_1^{1/4} F_1^{1/4} q_{r1,2} \phi_1^{1/4} F_1^{1/4} q_{r1,3} \phi_1^{1/4} F_1^{1/4} q_{r1,4} (1 - \phi_1^{1/4}) F_1^{1/4} p_{r2,1}$
\vdots			
$\pi_{r1,2}$	1	2	$\phi_1 F_1 p_{r2}$
$\pi_{r1,3}$	1	3	$\phi_1 F_1 q_{r2} \phi_2 F_2 p_{r3}$
$\pi_{r2,3}$	2	3	$\phi_2 F_2 p_{r3}$
π_{rs2}	2	2	p_{r2}
π_{rs3}	3	3	p_{r3}

- ϕ_i = the probability that a bird survives from year i to year $i + 1$,
- $1 - F_i$ = the probability that a bird emigrates out of the study area between year i and year $i + 1$,
- $p_{ri,j}$ = the probability a radioed bird present in the study area at sampling time j in year i is relocated by radio at sampling time j ,
- p_{si} = the probability that a bird alive and in the study area at year i is resighted at year i .

Table 1 displays selected cell probabilities $\pi_{li,j,i',j'}$, $\pi_{di,j,i',j'}$, $\pi_{si,j}$ and π_{rsi} for case 3 in which there are $k = 4$ telemetry surveys per year.

To simplify computations, we further assume that survival rates, emigration rates, resighting rates and relocation rates are constant across years. As a result, the cell probabilities in Table 1 will have $\phi_i = \phi$, $F_i = F$, $p_{ri,j} = p_r$ and $p_{si} = p_s$, and the joint likelihood in equation (4) will be a function of $\theta = (\phi, F, p_r, p_s)^t$. As described in Nasution *et al.* (2001), maximum likelihood estimates (MLEs) of θ must be obtained using a numerical optimization procedure such as that in program SURVIV (White, 1992).

3.2 Determining optimal sample sizes in the combined survey

Our objective is to minimize $V(\hat{\phi})$ where $\hat{\phi}$ is the MLE of ϕ obtained from the likelihood in equation (4) as described above. Alternatively, our objective could be to minimize $V(1 - \hat{F}) = V(\hat{F})$ obtained from the same likelihood. It is not possible to obtain an explicit expression for $V(\hat{\phi})$ or for $V(\hat{F})$, and so it is not possible to obtain an explicit solution for the sample sizes N_r and N_s that minimize $V(\hat{\phi})$ or $V(\hat{F})$ for a fixed cost. Instead, we use a computational approach that involves (i) specifying values for all costs and parameters in equation (3), (ii) determining a range of sample sizes (N_r, N_s) that satisfy the cost equation, and (iii) calculating $V(\hat{\phi})$ for the series of (N_r, N_s) pairs in order to find the values that minimize $V(\hat{\phi})$. If optimizing with respect to \hat{F} , then $V(\hat{F})$ would be computed in (iii).

TABLE 2. Numbers of radio-tagged birds and banded birds, N_r and N_s , that satisfy the cost equation for cases 1, 2 and 3*

Case								
1			2			3		
N_r	N_s	N_r/N_s	N_r	N_s	N_r/N_s	N_r	N_s	N_r/N_s
20	1488	0.013	20	1386	0.014	20	1158	0.017
40	1359	0.029	40	1245	0.032	40	993	0.040
60	1230	0.049	60	1103	0.054	60	828	0.072
80	1100	0.073	80	961	0.083	80	662	0.121
100	971	0.103	100	820	0.122	100	497	0.201
120	842	0.142	120	678	0.177	120	332	0.361
140	713	0.196	140	537	0.261	140	167	0.835
160	583	0.274	160	395	0.404			
180	454	0.396	180	254	0.708			
200	325	0.615						

*See Section 3.2 for specific costs

(i) Based roughly on the study in Bennetts *et al.* (1999), we selected the following values of costs: $C = 100\,000$, $c_1^0 = 40$, $c_{1,1}^{1r} = 250$, $c_1^{1s} = 4.5$, $c_i^{2s} = 12000$, $c_{i,j}^{2r} = 2000$ and $c_{i,j}^{3r} = 30$, $\forall i, j$. It is also necessary to specify values for the parameters ϕ , F and p_r as these are needed to obtain $E(m_{i,j}^r + d_{i,j}^r)$ in the cost equation. The values used in each of the cases 1, 2 and 3 were $\phi = 0.49$, $F = 0.95$ and $p_r = 0.9$. The resighting probability p_s does not appear in the cost equation, but is required to compute $V(\hat{\phi})$ in (iii) below. We used two values of p_s , specifically $p_s = 0.2$ and $p_s = 0.6$, with each of cases 1, 2 and 3.

(ii) We then use equation (3) to obtain an expression for N_s as a function of N_r given by

$$N_s = \frac{(C - \sum_i c_i^{2s} - c_{1,1}^{1r} N_r - \sum_i \sum_j c_{i,j}^{2r} - \sum_i \sum_j c_i^{3r} (E(m_{i,j}^r + d_{i,j}^r)))}{(c_1^{1s} + c_1^0)} \tag{8}$$

Based on practical considerations, we consider values of N_r ranging from 20 to 200 and, using equation (8) with the costs and parameters specified in (i), we solve for the corresponding values of N_s . Resulting values of (N_r, N_s) for cases 1, 2 and 3 are given in Table 2, together with the ratio N_r/N_s , which represents the fraction of banded birds that also receive a radio.

(iii) Under certain regularity conditions, the MLE $\hat{\theta}$ will asymptotically be $N(\theta, \mathbf{V}(\hat{\theta}))$ with $\mathbf{V}(\hat{\theta}) = \mathbf{I}(\theta)^{-1}$ (Azzalini, 1996) and the asymptotic variance of $\hat{\phi}$ is the leading diagonal element of $\mathbf{V}(\hat{\theta})$. The asymptotic variance of (\hat{F}) is the second diagonal element of $\mathbf{V}(\hat{\theta})$. Catchpole & Morgan (1997) give the general expression for the expected information matrix $\mathbf{I}(\theta)$ for a multinomial distribution, which can be generalized for a product multinomial distribution, i.e.

$$\mathbf{I}(\theta) = \mathbf{D} \mathbf{\Pi}^{-1} \mathbf{D}' \tag{9}$$

with $\mathbf{\Pi} = \text{diag}(\boldsymbol{\mu})$ and

$$\mathbf{D} = \begin{pmatrix} \frac{\partial \mu_j}{\partial \theta_i} \end{pmatrix}_{i=1, \dots, 4, j=1, \dots, n_{cell}}$$

Here, μ is a vector consisting of the expected values of all possible outcomes in the product-multinomial distribution and $n_{cell} = 3K_r k + 2(K_s + \min(K_r, K_s)) + \sum_{i=1}^{K_2} (K_2 - i) + 2\sum_{j=1}^{K_r k} (K_r k - j)$.

For example, for case 1, $n_{cell} = 17$,

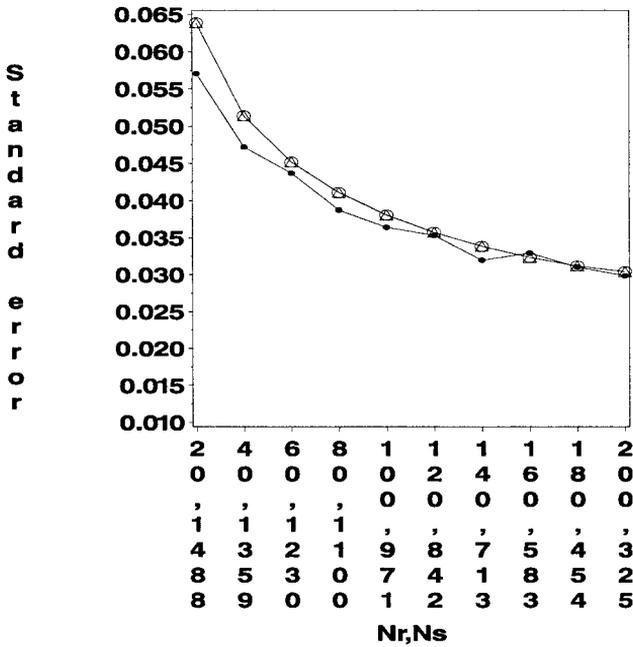
$$\begin{aligned} \mu^r = & (R_{1,1}^r \pi_{1,1,2,1}, R_{1,1}^r \pi_{1,1,3,1}, R_{1,1}^r \pi_{d1,1,2,1}, R_{1,1}^r \pi_{d1,1,3,1}, \\ & R_{1,1}^r (1 - \pi_{1,1,2,1} - \pi_{1,1,3,1} - \pi_{d1,1,2,1} - \pi_{d1,1,3,1}), \\ & R_{2,1}^r \pi_{1,2,1,3,1}, R_{2,1}^r \pi_{d2,1,3,1}, R_{2,1}^r (1 - \pi_{1,2,1,3,1} - \pi_{d2,1,3,1}), \\ & m_{r,\dots,2,1}^r p_{s2}, m_{r,\dots,2,1}^r (1 - p_{s2}), m_{r,\dots,3,1}^r p_{s3}, m_{r,\dots,3,1}^r (1 - p_{s3}), \\ & R_1^s \pi_{s1,2}, R_1^s \pi_{s1,3}, R_1^s (1 - \pi_{s1,2} - \pi_{s1,3}), R_2^s \pi_{s2,3}, R_2^s (1 - \pi_{s2,3}) \end{aligned}$$

To obtain $V(\hat{\phi})$ or $V(\hat{F})$ for case 1, elements of μ and \mathbf{D} must be calculated and $\mathbf{I}(\theta)^{-1}$ obtained using, for example, the IMSL (1987) FORTRAN subroutine LINRG to invert the 4×4 matrix $\mathbf{I}(\theta)$.

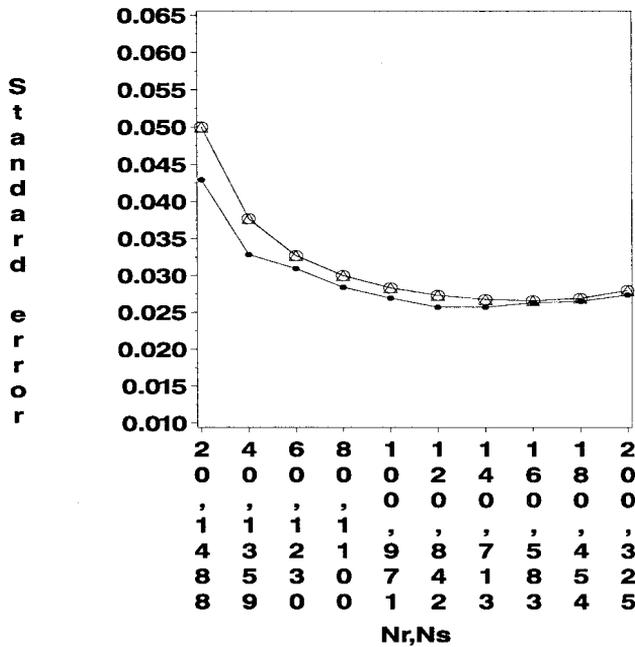
For cases 2 and 3, the number of possible outcomes is 33 and 89, respectively, and obtaining $\mathbf{I}(\theta)$ using equation (9) is tedious. An alternative approach for computing $V(\hat{\phi})$ or $V(\hat{F})$ is to calculate the expected frequencies corresponding to the elements of μ and to feed these as ‘data’ to program SURVIV. These expected counts are not integers and SURVIV must be modified to accept non-integer data. The numerical algorithm in SURVIV is a quasi-Newton iterative method (Rustagi, 1994) that produces MLEs and an estimate of $\mathbf{V}(\hat{\theta})$. When applied to the expected counts, SURVIV produces the same values for $V(\hat{\phi})$ and $V(\hat{F})$ as derived using equation (9).

Standard errors, $SE(\hat{\phi}_i)$, computed directly from equation (9) and by using SURVIV on the expected counts, are graphed against (N_r, N_s) in Fig. 1 for case 1. From this figure, we can see that using either $\mathbf{I}(\theta)^{-1}$ from equation (9) or the quasi-Newton iterative solution from SURVIV produces the same curve. When the resighting rate is high, i.e. $p_s = 0.6$, the standard error curve is almost flat for a wide range of (N_r, N_s) values. As long as the N_r values are not too low, the standard error of survival estimates will be close to the minimum. In practice, when the resighting rate is high enough, the optimal strategy will be to put radios on roughly 10% to 40% of the total marked birds. Although figures for cases 2 and 3 are not shown, the curves under these cases give similar patterns. When the resighting rate is low ($p_s = 0.2$), the standard errors of survival estimators tend to decrease as N_r increases. A steep decrease occurs at the lowest values of N_r (i.e. when N_r/N_s is very low). Figure 1 shows results for N_r values between 20 and 200; however, the minimum standard error is reached at $N_r = 211$, $N_s = 254$. In practice, when the resighting rate is low, the optimal strategy will be to put radios on a higher percentage of birds, say 40% to 80% of the total marked birds.

If estimating emigration is the objective, the above procedure would focus on $V(1 - \hat{F}) = V(\hat{F})$. Thus, Fig. 2 displays the standard error of \hat{F} as a function of (N_r, N_s) . Regardless of the values of p_s , we see that $SE(\hat{F})$ decreases as N_r increases. Further attempts to compute $SE(\hat{F})$ for higher N_r values continue to give lower $SE(\hat{F})$. Although the figures are not shown, the pattern remains the same under cases 2 and 3. Additional computations were carried out with the same functions and parameter values but with $F = 0.5$ instead of $F = 0.95$ (results are not shown). Again, the optimal strategy for estimating F in practice is to put radios on the maximum possible number of birds.

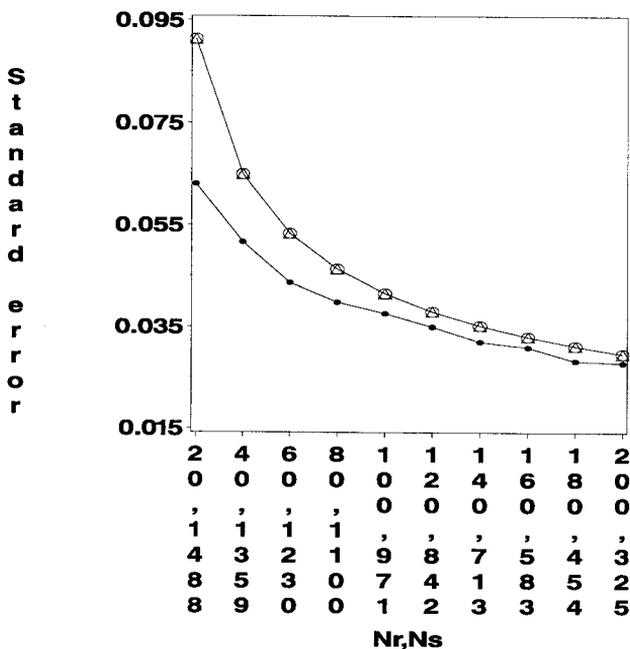


(a)

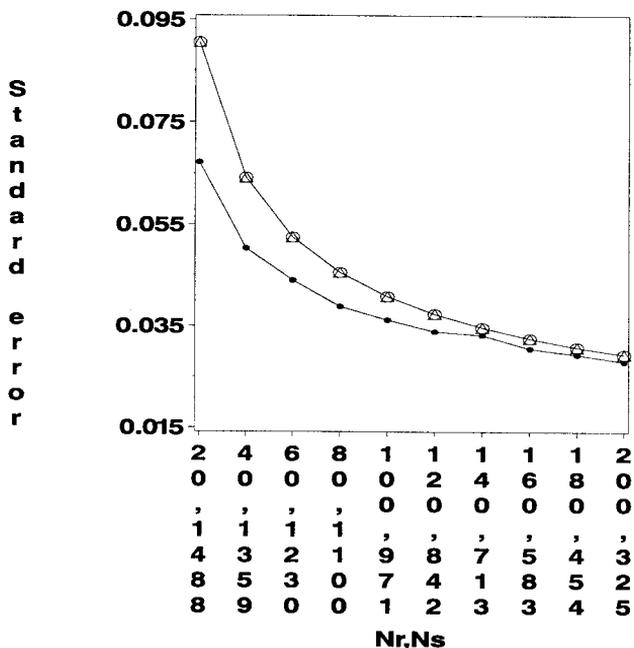


(b)

FIG. 1. Survival standard error curves under case 1. (a) Case 1, $\phi = 0.49$, $F = 0.95$, $p_r = 0.9$, $p_s = 0.2$; (b) case 1, $\phi = 0.49$, $F = 0.95$, $p_r = 0.9$, $p_s = 0.6$ large sample approx.: circle (SURVIV), triangle ($\mathcal{G}(\theta)^{-1}$); simulation: dot ($1000 \times$) $C = 100\,000$, $c_{1,1}^{0r} = 40$, $c_{1,1}^{1r} = 250$, $c^{2r} = 2000$, $c^{3r} = 30$, $c_1^{0s} = 40$, $c_1^{1s} = 4.5$ and $c^{2s} = 12\,000$.



(a)



(b)

FIG. 2. Emigration standard error curves under case 1. (a) Case 1, $\phi = 0.49$, $F = 0.95$, $p_r = 0.9$, $p_s = 0.2$; (b) case 1, $\phi = 0.49$, $F = 0.95$, $p_r = 0.9$, $p_s = 0.6$ large sample approx.: circle (SURVIV), triangle ($\mathcal{F}(\theta)^{-1}$); simulation: dot ($1000 \times$) $C = 100\,000$, $c_{1,1}^{0r} = 40$, $c_{1,1}^{1r} = 250$, $c^{2r} = 2000$, $c^{3r} = 30$, $c_1^{0s} = 40$, $c_1^{1s} = 4.5$ and $c^{2s} = 12\,000$.

TABLE 3. The large sample standard error of \hat{p}_r *

N_r	Case 1	Case 2	Case 3
	$p_r = 0.9$	$p_r = 0.9$	$p_r = 0.9$
20	0.102	0.056	0.036
40	0.072	0.040	0.025
60	0.059	0.032	0.021
80	0.051	0.028	0.018
100	0.046	0.025	0.016
120	0.042	0.023	0.015
140	0.039	0.021	0.014
160	0.036	0.020	
180	0.034	0.019	
200	0.033		

*when $p_s = 0.6$; $se(\hat{p}_s)$ when $p_s = 0.2$ is similar.

TABLE 4. The large sample standard error of \hat{p}_s

N_r	Case 1		Case 2		Case 3	
	$p_s = 0.2$	$p_s = 0.6$	$p_s = 0.2$	$p_s = 0.6$	$p_s = 0.2$	$p_s = 0.6$
20	0.029	0.033	0.030	0.034	0.032	0.037
40	0.027	0.033	0.027	0.034	0.029	0.037
60	0.026	0.033	0.026	0.035	0.028	0.038
80	0.025	0.034	0.025	0.035	0.028	0.040
100	0.024	0.034	0.025	0.036	0.028	0.041
120	0.024	0.034	0.025	0.037	0.031	0.044
140	0.024	0.035	0.026	0.038	0.040	0.050
160	0.025	0.036	0.028	0.039		
180	0.026	0.037	0.032	0.043		
200	0.029	0.039				

The standard errors of \hat{p}_r and \hat{p}_s are presented in Tables 3 and 4. Standard errors for \hat{p}_r decrease when N_r increases, with the sharpest decrease occurring when N_r is relatively small. In contrast, the standard errors of \hat{p}_s seem to be relatively constant although there is a slight tendency to be higher for larger N_r , smaller N_s .

4 Simulation study to assess variance approximations

Up to now we have discussed optimization based on using a large sample approximation to compute $V(\hat{\phi})$ or $V(\hat{F})$. However, we do not know whether these approximations are accurate for the sample sizes (N_r, N_s) studied in Section 3. Thus, it is important to validate the results derived from the large sample approximation and the only way to do this is by using Monte Carlo simulation (Fishman, 1996). The likelihood, parameter values and values of (N_r, N_s) used in the simulation are the same as for cases 1, 2 and 3 in Section 3. Note that there are two sets of parameter values corresponding to p_s considered for each case. The outline for the simulation is as follows.

- (1) Generate a data set from the product multinomial distribution presented in equations (4) to (7) for a fixed pair of (N_r, N_s) from Table 2 and a fixed set of parameter values.

- (2) Calculate the estimates and standard errors using program SURVIV.
- (3) Repeat Steps 1 and 2 for $n = 1000$ replications.
- (4) Repeat Steps 1, 2 and 3 for the other possible values of (N_r, N_s) and p_s .

The product-multinomial data are obtained by applying the RNMTN routine in the FORTRAN-IMSL library. The RNMTN routine generates pseudo-random numbers from a K -variate multinomial distribution with parameters N and $\Pi = (\pi_1, \dots, \pi_K)$ where K and N are positive. Each cell frequency x_i in the multinomial is successively generated from a *Binomial* (M_i, p_i) distribution where

$$M_1 = N, p_1 = \pi_1, M_i = N - \sum_{j=1}^{i-1} x_j \text{ and } p_i = \frac{\pi_i}{1 - \sum_{j=1}^{i-1} \pi_j} \text{ for } i = 2, \dots, K - 1.$$

The variance of the MLE $\hat{\phi}$ is calculated from the simulation results for a given set of (N_r, N_s) and p_s using the unbiased variance estimator $\hat{V}(\hat{\phi}) = (1/(n - 1)) \sum (\hat{\phi}_i - \hat{\phi})^2$. Note that here the $\hat{\phi}_i$ are the survival estimates obtained from the $n = 1000$ simulation data sets.

The Monte Carlo estimates of $SE(\hat{\phi})$, obtained as $\sqrt{\hat{V}(\hat{\phi})}$, are presented in Fig. 1 for case 1. From this figure, we can see that the simulation results generally agree well with the large sample standard errors. However, when the number of radioed birds is very low, the large sample approximation overestimates the small sample standard error. Although the figures are not shown for cases 2 and 3, the large sample approximation improves when the number of telemetry sampling times per year increases for small values of N_r .

Similarly for the Monte Carlo estimates of $SE(\hat{F})$ obtained as $\sqrt{\hat{V}(\hat{F})}$, presented in Fig. 2 for case 1. We can see that the large sample approximation overestimates the small sample standard error unless N_r is high. For small values of N_r , the large sample approximation improves when the number of telemetry sampling times per year increases (figures are not shown). The large sample approximation is less accurate for $V(\hat{F})$ than for $V(\hat{\phi})$, at least in part, because the value of F is close to 1.

5 Discussion

To investigate the optimal allocation of sample sizes between banding and radio-tags, we have considered a simple cost function and simple study designs with a single release of newly marked birds, only two years of follow up and at most four telemetry surveys per year. Even in the simplest case (case 1) it was not possible to obtain an analytical solution with explicit formulas for the optimal (N_r, N_s) . Results in Fig. 1 are therefore specific to the parameter values and costs assumed. Certain qualitative trends may hold more generally. For example, we showed that if the resighting rate is high, e.g. 0.6, tagging roughly 10-40% of the banded birds with radios gave survival estimators with standard errors close to the minimum. Given similar costs to those used, lower resighting rates, e.g. 0.2, will require a higher percentage of radioed birds (roughly between 40% and 80% of the total marked birds) to achieve the standard errors close to the minimum. Under the proposed costs, the maximum possible percentage of radioed birds is needed to minimize the standard error of emigration estimates regardless of the resighting rates.

In this paper, we used the precision of annual survival estimates assuming constant survival rates across time as the optimality criterion. However, there are other possible optimality criteria that may be considered. For example, since radio

telemetry data make it possible to estimate within-year variation in survival, the power to detect seasonal effects on survival could be a useful optimality criterion.

The cost function considered in this paper is limited to the situation where all the birds are released at the start of the year. In 1989, Pollock introduced the staggered entry design under which radioed birds enter the study at several times during the year. It would be useful to build a cost function that allows newly marked animals to enter the study following each telemetry survey and determine the optimal allocation of sample sizes for this situation.

There are many factors to be considered in constructing a cost function for the combined surveys. Accurate information about costs in both surveys is important. Modifications that produce more flexible and more realistic cost functions should be considered. For example, instead of considering the costs needed in each survey separately, a cost function that incorporates the ratio of costs for the two surveys could be more useful to biologists. It may also be possible to build in a relationship between costs and resighting and relocation rates.

The scope of this research is limited to the allocation of sample sizes when resighting and radio telemetry surveys are carried out on a single population. It would be useful to study optimal allocation of resources in more complex situations, for example, where resighting and telemetry data are obtained from several sites in order to estimate true survival and movement between sites. Also of interest is the situation where resighting and telemetry data are obtained from different age classes of birds, so that the optimization is related to the precision of age-dependent survival or fidelity estimates. Optimal allocation can also be considered for combinations of other data types, such as the combined analysis of resighting and recovery data considered by Burnham (1993).

Acknowledgements

The authors thank Jim Hines for providing us with the version of SURVIV that accepts non-integer cell frequencies. Thanks also go to Rob Bennetts who provided us with information regarding costs.

REFERENCES

- AZZALINI, A. (1996) *Statistical Inference based on the Likelihood* (New York, Chapman & Hall).
- BARKER, R. J. (1997) Joint modelling of live-recapture, tag-resight, and tag-recovery data, *Biometrics*, 53, pp. 666-677.
- BARKER, R. J., SAUER, J. R. & LINK, W. A. (1993) Optimal allocation of point-count sampling effort, *The Auk*, 110, pp. 752-758.
- BENNETTS, R. E., DREITZ, V. J., KITCHENS, W. M., HINES, J. E. & NICHOLS, J. D. (1999) Annual survival of Snail Kites in Florida: comparisons between radio-telemetry and capture-resighting data, *The Auk*, 116, pp. 435-447.
- BROWNIE, C., HINES, J. E., NICHOLS, J. D., POLLOCK, K. H. & HESTBECK, J. B. (1993) Capture-recapture studies for multiple strata including non-Markovian transitions, *Biometrics*, 49, pp. 1173-1187.
- BURNHAM, K. P. (1993) A theory for combined analysis of ring-recovery and recapture data. In: J. D. LEBRETON & P. NORTH (Eds) *Marked Individuals in the Study of Bird Population*, pp. 199-213 (Basel, Birkhauser Verlag).
- CATCHPOLE, E. A., FREEMAN, S. N., MORGAN, B. J. T. & HARRIS, M. P. (1998) Integrated recover/recapture data analysis, *Biometrics*, 54, pp. 33-46.
- CATCHPOLE, E. A. & MORGAN, B. J. T. (1997) Model selection in ring-recovery models using score tests, *Biometrics*, 52, pp. 664-672.
- FISHMAN, G. S. (1996) *Monte Carlo: Concepts, Algorithms and Applications* (New York, Springer).

- IMSL, INC (1987) *User's Manual STAT/Library FORTRAN Subroutines for Statistical Analysis* (Texas, IMSL).
- LEBRETON, J. D., MORGAN, B. J. T., PRADEL, R. & FREEMAN, S. N. (1995) A simultaneous analysis of dead-recovery and live-recapture data, *Biometrics*, 51, pp. 1418-1428.
- NASUTION, M. D., BROWNIE, C., POLLOCK, K. H. & BENNETTS, R. E. (2001) Joint survival analysis of resighting and radio-telemetry capture-recapture data for wild animals, *Journal of Agricultural, Biological and Environmental Statistics*, to appear.
- POLLOCK, K. H. (1989) Survival analysis in telemetry studies: the staggered entry design, *Journal of Wildlife Management*, 53, pp. 7-15.
- RUSTAGI, J. S. (1994) *Optimization Techniques in Statistics* (Boston, Academic Press).
- SCHWARZ, C. J. & SEBER, G. A. F. (1999) Estimating animal abundance: review III, *Statistical Science*, 14, pp. 427-456.
- TSAI, K. (1996) Survival analysis for telemetry data in wildlife studies, PhD Thesis, North Carolina State University, Raleigh.
- WHITE, G. C. (1992) *PC SURVIV User's Manual Version 1.4* (Colorado State University, Department of Fisheries and Wildlife Biology).