An Empirical Hierarchical Bayes Approach to Closed Model Capture-Recapture Abundance Estimation

Tomoharu Eguchi
Department of Mathematical Sciences
Montana State University

21 October 2003

A writing project submitted in partial fulfillment
of the requirements for the degree

Master of Sciences in Statistics

APPROVAL

of a writing project submitted by

Tomoharu Eguchi

This writing project has been read by the writing project director and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the Statistics faculty.

Oct 24, 2003

Date

Robert J Boik

Robert J Boik

WRITING PROJECT DIRECTOR

AN EMPIRICAL HIERARCHICAL BAYES APPROACH TO CLOSED MODEL CAPTURE-RECAPTURE ABUNDANCE ESTIMATION

ABSTRACT

A hierarchical Bayesian procedure was developed to make inferences on abundance of a closed population from the robust-design capture-mark-recapture data. I made the assumption that capture probabilities were random samples from a common hyper distribution. Data for one primary period were used for making the inference on abundance of the population whereas the data for other primary periods were used for making the inference on the hyperparameters. I used beta-binomial and beta-multinomial models for the likelihood function. Simulation analyses were used to test the performance of the proposed method and to consider the effects of temporal sampling design. Maximum likelihood estimates for the same simulated data sets were obtained using software MARK for comparison. A frequent short-term sampling provided more precise posterior distribution of abundance than infrequent long-term sampling. The performance of the proposed method was comparable to the maximum likelihood method. When sample sizes were small, however, the proposed method performed better than the maximum likelihood method.

Introduction

Estimating the abundance of a population is important in understanding biological processes and successful management of the population. Numerous methods have been developed for estimating the abundance of a population (Seber 1982, Schwarz and Seber 1999, Williams et al. 2001). Depending on the sampling process and underlying model, these methods can be grouped into one of three broad categories: counts, distance-based, and capture-mark-recapture (Williams et al. 2001). In this chapter, I develop a Bayesian method for estimating the abundance of a closed population using capture-mark-recapture (CMR) data. Discussion on other methods can be found elsewhere (e.g., Buckland et al. 1993, Williams et al. 2001).

In CMR methods, animals are caught in a series of samples. Animals in the first sample are marked individually and released into the population. In subsequent samples, new animals are marked individually and released, whereas recaptured animals are recorded and released. Consequently, capture histories of individual animals are available at the completion of a study. Marks on animals may be either artificial or natural. When feasible, natural marks are preferable to artificial marks and tags because the marking procedure does not affect the behavior of animals, the mortality due to direct handling of animals is absent, and loss of marks is usually less of an issue.

Existing methods for estimating the abundance of a population via CMR experiments require large numbers of recaptures to obtain precise results. In the user manual for their population analysis software, Arnason et al. (1998) stressed the importance of a carefully designed study and the use of prior knowledge of the population:

> *"Failure to use every scrap of prior information about the population of interest, or to carry out a proper sampling program that fully exploits this information, can have disastrous consequences. More often than not, a poor experiment, even if it has deployed prodigious amounts of equipment and man-hours, can lead to data that are difficult or impossible to analyse, and estimates that are vague, misleading, or even absurdly biased (Arnason et al. 1998; Chapter 6.2)."*

In many statistical models for capture-mark-recapture analyses, abundance is not incorporated in these models explicitly. The abundance during a period, however, is estimated by using the following relationship:

$$E[N_t] = \frac{n_t}{\hat{p}_t}, \tag{1}$$

where n_t is the number of animals caught during the t^{th} primary period and $\hat{p}_t$ is the estimated capture probability for the t^{th} primary period from a model. In these models, other population parameters, such as survival, immigration, emigration, and birth rates, also are incorporated. Using the maximum likelihood approach, estimation procedures involve maximizing a multidimensional surface with multiple parameters, which may not exhibit prominent peaks. Consequently, capture probabilities sometimes are estimated imprecisely, resulting in imprecise estimates of abundance.

In this chapter, a statistical method is proposed for making an inference on the abundance of a closed population using data from a CMR study with the robust design (Pollock 1982). In this method, the entire experiment is divided into a series of short time periods (primary periods). The population is assumed closed to additions and deletions within a primary period. Within each primary period, individuals are caught during multiple secondary occasions (Figure 1).

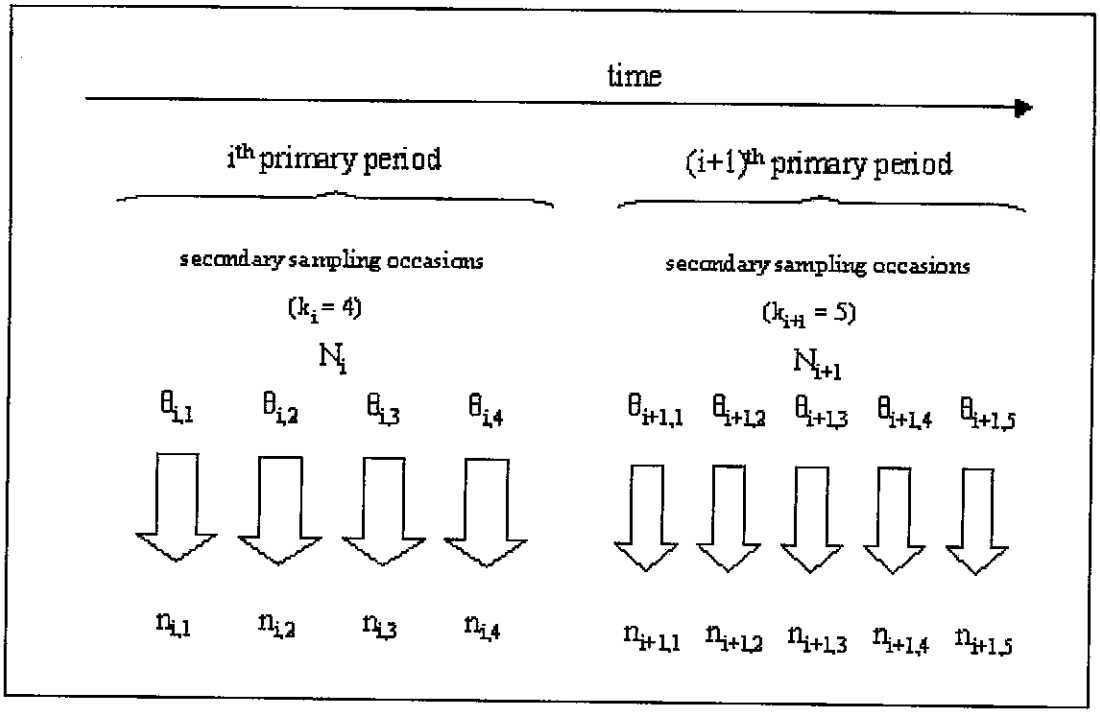

Figure 1. A schematic diagram of Pollock's robust design. Within each primary period, multiple samples are collected during secondary occasions with occasion-specific capture probability ($\theta_{i,j}$)

In many ecological studies and management situations, the inference on abundance for one primary period is desired. In the proposed method, a data set is divided into two groups: one primary period for which the abundance is inferred and all other primary periods from which supporting information on the capture probabilities is obtained. The assumption under this approach is that capture probabilities may vary between secondary occasions, but they are similar for the duration of the study. Consequently, the use of data from all other primary periods provides information on capture probabilities for the primary period of interest. Capture probabilities at a sampling location are considered random samples from a beta distribution, defined by two hyperparameters. The joint distribution of these hyperparameters is found by using all primary periods except for the primary period of interest. The inference on abundance

for the primary period is made by using the distribution of the hyperparameters, observed data for the primary period, and a probabilistic model. I consider beta-binomial and multinomial models. Although methods for estimating mortality, immigration and emigration rates, and movement rates have been proposed for analyzing data from the robust design (Kendall et al. 1995, Schwarz and Stobo 1997, Lindberg et al. 2001, Kendall and Bjorkland 2001), I only consider the inference for the abundance.

I use the hierarchical approach for the following two reasons: (1) capture probabilities are affected by factors independent of target animals, e.g., weather, personnel, and geographic features, and (2) capture histories and the number of individuals caught during a sampling occasion are affected by capture probabilities and the total number of available individuals. By using the hierarchical approach, I can indirectly incorporate the factors independent of target animals into the analysis. Inference for the abundance is made by using the relationship among the numbers of individuals caught, capture probabilities, and the total number of individuals available for capture.

In this study, I first derive the model and computational procedure. To test the performance of the proposed method, I use simulated data. I also compare the results to maximum likelihood estimates from freely available software (MARK; White and Burnham 1999).

Methods

The proposed inference procedure contains two steps. In the first step, I use the capture-recapture data for all but one primary period to make inference on the hyperparameters of capture probabilities. The posterior distribution of these

hyperparameters is used to determine whether or not the data provide enough information about capture probabilities. If the posterior distribution is informative, i.e., if it is not diffuse over the parameter space, the posterior distribution is used in the second step, in which the posterior distribution of abundance in the primary period of interest is computed. If the posterior distribution from the first step is diffuse, a uniform distribution on the hyperparameters of capture probabilities is used in the second step, because no useful information on capture probabilities can be obtained from other primary periods. For making the inference on the abundance, I consider two likelihood functions, beta-binomial and multinomial. The former is appropriate when few animals are recaptured, whereas the latter is applicable when a large number of animals are recaptured. Performances of these two models are compared.

The posterior distribution of abundance is computed for a primary period, whereas the data for other primary periods are used for computing the posterior distribution of hyperparameters for capture probabilities. Although the primary period of interest, i.e., the primary period for which the posterior distribution of abundance is computed, can be any primary period, often the abundance estimate is required for the most recent primary period. Consequently, in the following analyses I make inferences on the abundance for the last primary period, whereas the data for all other primary periods are used for the inference on the hyperparameters of capture probabilities.

Assumptions

I made the following assumptions for building capture-recapture models for making the inference for the abundance of a population.

(1) The population is assumed closed to immigration, emigration, births, and deaths during each primary period, which is a short time period, e.g., 15 days, 30 days, etc. Consequently, the population size (N_t) remains constant during the t^{th} primary period ($t = 1, 2, \ldots, T$). Individuals in the population are caught and recaptured on secondary occasions ($j = 1,\ldots,k_t$) within a primary period (Figure 1). Captured individuals are marked, and recaptured individuals are identified and released into the population. No animals are killed in the process. All previously marked individuals are identified without errors.

(2) All individuals in the population have equal capture probability ($\theta_{t,j}$) during the j^{th} sampling occasion of the t^{th} primary period regardless of their capture history.

(3) Captures are mutually independent events. Let $c_{t,j,l}$ = event that an individual l is caught during the j^{th} secondary occasion of the t^{th} primary period. I assume that $\{c_{t,j,l}\}$ are mutually independent for all t, j, and l.

(4) At least two sampling occasions exist for each primary period ($\min(k_t) > 1$).

(5) Capture probabilities ($\theta_{t,j}$) are exchangeable among all sampling occasions for the entire experiment (Figure 2). In other words, the joint density of capture probabilities is not dependent of the order of captures.

(6) Capture probabilities during the j^{th} secondary occasion of the t^{th} primary period ($\theta_{t,j}$) are mutually independent for all t and j.

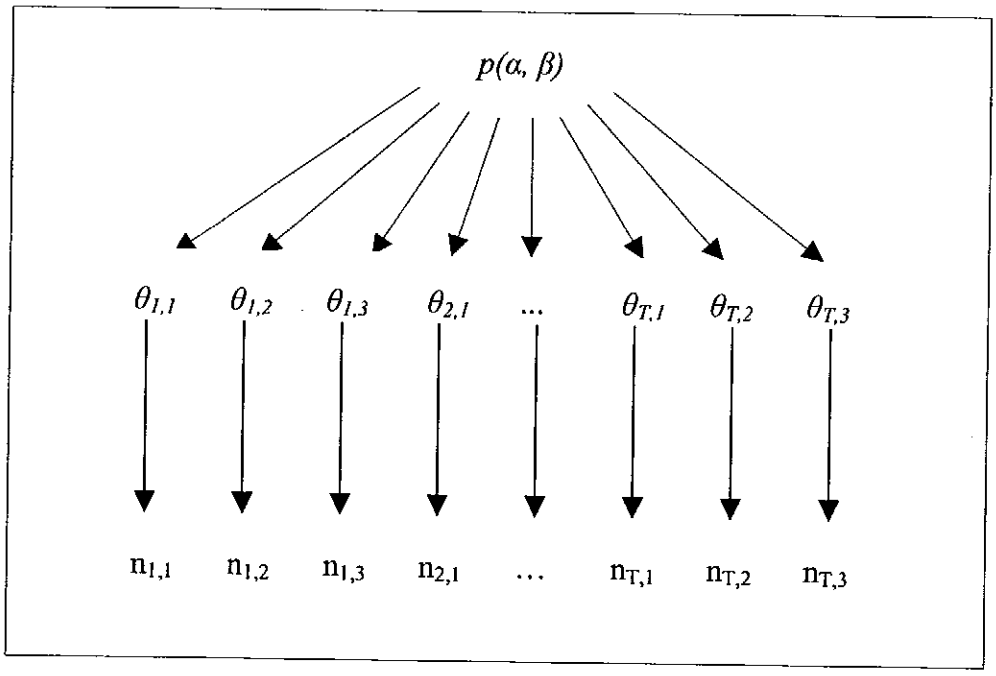

Figure 2. The structure of the hierarchical model for the capture-recapture model.

Models

Notation

N_t = population size during the t^{th} primary period ($t = 1,\ldots,T$),

$\theta_{t,j}$ = capture probability during the j^{th} secondary occasion of the t^{th} primary period ($j = 1, \ldots,k_t$, $t = 1,2,\ldots,T$),

$U_{t,j}$ = set of all individuals in the t^{th} primary period ($t = 1,2,\ldots,T$) that are caught for the first time on the j^{th} secondary occasion ($j = 1, \ldots,k_t$),

$u_{t,j}$ = the number of individuals in $U_{t,j}$,

$n_{\omega}^{(t)}$ = the number of individuals with a particular capture history ω during the t^{th} primary period, where ω is a non-empty subset of 0's and 1's of length k_t that indicates recapture histories for the t^{th} primary period. $\omega \in \Omega^{(k_t)}$,

where $\Omega^{(k_t)}$ is a non-empty set of 0's and 1's that indicates all possible recapture histories of length k_t.

$n_{t,j}$ = the total number of individuals caught during the j^{th} secondary occasion of the t^{th} primary period.

For all equations I use a bold face font for vectors and regular font for scalars.

Likelihood function for capture probabilities

Recapture processes can be modeled by describing recaptures as a function of capture probabilities conditional on the first capture. First, I present an example with three secondary occasions ($k = 3$). Because I assume that capture probabilities for secondary occasions are independent of each other (assumption 7) and abundances are independent among primary periods (assumption 1), I describe a model for a primary period and omit the subscript for primary periods. The general likelihood function for $k \geq 2$ is presented in the subsequent section.

With three secondary occasions, there are six possible recapture histories: $\Omega^{(3)} = \{111, 110, 101, 100, 011, 010\}$, where $\{001, 000\}$ are omitted because individuals that belong to these capture histories have not been recaptured. The individuals that belong to the first four capture histories in $\Omega^{(3)}$, i.e., $\{111, 110, 101, 100\}$, are elements of U_1, whereas those that belong to the last two, i.e., $\{011, 010\}$, are elements of U_2. The likelihood function for the numbers of individuals that belong to the first four capture histories ($\mathbf{n}_{1..} = [n_{111}, n_{110}, n_{101}, n_{100}]$), conditional on the individuals caught for the first time in the first occasion (u_1) and recapture probabilities for occasions 2 and 3 (θ_2 and θ_3), is:

$$l\left(\mathbf{n}_{1..} \mid u_1, \theta_2, \theta_3\right) = \frac{u_1!}{n_{111}!n_{110}!n_{101}!n_{100}!} \times \left(\theta_2\theta_3\right)^{n_{111}} \left[\theta_2\left(1-\theta_3\right)\right]^{n_{110}} \left[\left(1-\theta_2\right)\theta_3\right]^{n_{101}} \left[\left(1-\theta_2\right)\left(1-\theta_3\right)\right]^{n_{100}}. \tag{2}$$

The likelihood function for the numbers of individuals that belong to the last two capture histories ($\mathbf{n}_{01.} = \left[n_{011}, n_{010}\right]$), conditional on the individuals caught for the first time in the second occasion (u_2) and the recapture probability for the third occasion (θ_3), is:

$$l\left(\mathbf{n}_{01.} \mid u_2, \theta_3\right) = \frac{u_2!}{n_{011}!n_{010}!}\left(\theta_3\right)^{n_{011}}\left(1-\theta_3\right)^{n_{010}}. \tag{3}$$

Because captures are independent events (assumption 4), the likelihood function for these three secondary occasions is a product of the two likelihood functions:

$$\begin{aligned} l_1\left(\mathbf{n}_{\omega\in\Omega^{(3)}} \mid u_1, u_2, \theta_2, \theta_3\right) &= l\left(\mathbf{n}_{1..} \mid u_1, \theta_2, \theta_3\right) \times l\left(\mathbf{n}_{01.} \mid u_2, \theta_3\right) \\ &= \frac{u_1!u_2!}{\left(\prod_{\omega\in\Omega^{(3)}} n_\omega!\right)} \times \\ &\quad \theta_2^{n_{111}+n_{110}}\left(1-\theta_2\right)^{n_{101}+n_{100}} \theta_3^{n_{111}+n_{101}+n_{011}}\left(1-\theta_3\right)^{n_{110}+n_{100}+n_{010}}. \end{aligned} \tag{4}$$

Equation (4) is proportional to a product of two binomial probability mass functions and can be rewritten as:

$$\begin{aligned} l_1\left(\mathbf{n}_{\omega\in\Omega^{(3)}} \mid u_1, u_2, \theta_2, \theta_3\right) &= \frac{u_1!u_2!}{\left(\prod_{\omega\in\Omega^{(3)}} n_\omega!\right)} \theta_2^{n_{11.}}\left(1-\theta_2\right)^{u_1-n_{11.}} \theta_3^{n_{..1}}\left(1-\theta_3\right)^{u_1+u_2-n_{..1}} \\ &= \frac{\prod_{j=1}^{2} u_j!}{\left(\prod_{\omega\in\Omega^{(3)}} n_\omega!\right)} \prod_{i=2}^{3} \theta_i^{n_{\omega..j}}\left(1-\theta_i\right)^{\sum_{j=1}^{i} u_j - n_{\omega..j}}, \end{aligned} \tag{5}$$

where $\omega_{j,i}$ represents the capture history with ones in digits j and i ($j < i$) and $\omega_{j,i} \in \Omega^{(k)}$. For example, $\omega_{1,2}$ represents {111 and 110} whereas $\omega_{.,3}$ represents {111, 101, 011}. $n_{\omega_{j,i}}$ represents the observed number of individuals that belong to the capture history $\omega_{j,i}$.

The general likelihood for $k \geq 2$ can be constructed in a similar fashion. Recapture data are summarized into a vector of recapture histories ($\mathbf{n}_\omega$) and a vector of u_j's ($\mathbf{u}$). The likelihood function for k secondary occasions is a product of (k-1) multinomial probability mass functions with capture probabilities $\boldsymbol{\theta}$, the numbers of individuals sighted for the first time ($\mathbf{u}$), and the number of individuals that belong to capture histories $\omega_{j,i}$, similar to equation (5):

$$l_1\left(\mathbf{n}_{\omega\in\Omega^{(k)}} \mid \mathbf{u},\boldsymbol{\theta}\right) = \prod_{j=1}^{k-1} Mult\left(\mathbf{n}_{\omega_j\in\Omega^{(k)}} \mid \mathbf{u},\boldsymbol{\theta}\right), \tag{6}$$

where n_{ω_j} = the number of individuals that belong to $\omega_{.,j}$. This likelihood, however, is proportional to a product of binomial probability mass functions as is shown in the example with $k = 3$:

$$l_1\left(\mathbf{n}_{\omega\in\Omega^{(k)}} \mid \mathbf{u},\boldsymbol{\theta}\right) = \frac{\prod_{i=1}^{k-1} u_i!}{\prod_{\omega\in\Omega^{(k)}} n_\omega!} \prod_{i=2}^{k} \theta_i^{n_{\omega_{.,i}}} \left(1-\theta_i\right)^{\sum_{j=1}^{i-1} u_j - n_{\omega_{.,i}}}. \tag{7}$$

With an assumption that all primary periods are independent of each other, the likelihood function for the entire experiment with T primary periods is a product of l_1 (7) for all primary periods:

$$l\left(\left\{\mathbf{n}_{\omega\in\Omega^{(k_t)}}\right\}_{t=1}^{T} \mid \left\{\mathbf{u}^{(t)}\right\}_{t=1}^{T}, \left\{\boldsymbol{\theta}^{(t)}\right\}_{t=1}^{T}\right) = \prod_{t=1}^{T}\left[\frac{\prod_{i=1}^{k_t-1} u_i^{(t)}!}{\prod_{\omega\in\Omega^{(k_t)}} n_\omega^{(t)}!} \prod_{i=2}^{k_t} \theta_{t,i}^{n_{\omega_{.,i}}^{(t)}} \left(1-\theta_{t,i}\right)^{\sum_{j=1}^{i-1} u_j^{(t)} - n_{\omega_{.,i}}^{(t)}}\right], \tag{8}$$

where superscript and subscript (t) index primary periods. To increase the legibility of equations, I omit braces and Ω's in the subsequent equations. It is assumed that all ω's for each primary occasion are appropriate subsets of the corresponding $\Omega^{(k_t)}$'s and vectors contain all primary periods and secondary occasions.

I construct a hierarchical model for the capture probabilities. In this approach, I create a population of capture probabilities from which each capture probability is drawn randomly (Figure 2). Capture probabilities are independently and identically distributed (i.i.d.) according to a distribution. For computational convenience, I use a beta distribution:

$$\theta_{t,i} \overset{i.i.d}{\sim} Beta(\alpha, \beta), \tag{9}$$

where the subscript t indexes primary periods and the subscript i indexes secondary occasions. The two parameters for the beta distribution (α and β) are hyperparameters of the model that describe the distribution of $\theta_{t,i}$.

The joint posterior distribution of parameters and hyperparameters conditional on the observed data, $p(\boldsymbol{\theta}, \alpha, \beta \mid \mathbf{n}_\omega, \mathbf{u})$, is:

$$\begin{aligned} p(\boldsymbol{\theta}, \alpha, \beta \mid \mathbf{n}_\omega, \mathbf{u}) &= \frac{p(\boldsymbol{\theta}, \alpha, \beta, \mathbf{n}_\omega \mid \mathbf{u})}{p(\mathbf{n}_\omega \mid \mathbf{u})} \\ &= \frac{p(\alpha, \beta)\, p(\boldsymbol{\theta} \mid \alpha, \beta)\, p(\mathbf{n}_\omega \mid \mathbf{u}, \boldsymbol{\theta})}{p(\mathbf{n}_\omega \mid \mathbf{u})}, \end{aligned} \tag{10}$$

where $p(\alpha, \beta)$ is the unconditional joint probability density function of (α, β), $p(\boldsymbol{\theta} \mid \alpha, \beta)$ is the probability density function of $\boldsymbol{\theta}$ conditional on α and β, which is a product of independent beta distributions:

$$p(\boldsymbol{\theta} \mid \alpha, \beta) = \prod_{t=1}^{T} \prod_{i=2}^{k_t} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta_{t,i}^{\alpha-1} (1-\theta_{t,i})^{\beta-1}, \tag{11}$$

$p(\mathbf{n}_\omega \mid \mathbf{u}, \boldsymbol{\theta})$ is the probability mass function of $\mathbf{n}_\omega$ conditional on $\mathbf{u}$ and $\boldsymbol{\theta}$ (8), and

$$p(\mathbf{n}_\omega \mid \mathbf{u}) = \int \ldots \int p(\boldsymbol{\theta}, \alpha, \beta, \mathbf{n}_\omega \mid \mathbf{u}) d\boldsymbol{\theta} d\alpha d\beta . \tag{12}$$

Using the product of beta distributions (11) and the likelihood function (8), the joint conditional posterior distribution of $(\boldsymbol{\theta}, \alpha, \beta)$ is:

$$p(\boldsymbol{\theta}, \alpha, \beta \mid \mathbf{n}_\omega, \mathbf{u}) = \frac{p(\alpha,\beta)\prod_{t=1}^{T}\left[\prod_{i=2}^{k_t}\left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta_{t,i}^{\alpha-1}\left(1-\theta_{t,i}\right)^{\beta-1}\right]\left[\frac{\prod_{l=1}^{k_t} u_l^{(t)}!}{\prod_{\omega^{(t)}} n_{\omega^{(t)}}^{(t)}!}\prod_{i=2}^{k_t}\left[\theta_{t,i}^{n_{\omega_{\cdot,i}}^{(t)}}\left(1-\theta_{t,i}\right)^{\sum_{j=1}^{i-1}u_j^{(t)}-n_{\omega_{\cdot,i}}^{(t)}}\right]\right]\right]}{p(\mathbf{n}_\omega \mid \mathbf{u})} . \tag{13}$$

Rearranging θ's:

$$p(\boldsymbol{\theta}, \alpha, \beta \mid \mathbf{n}_\omega, \mathbf{u}) = \frac{p(\alpha,\beta)\prod_{t=1}^{T}\left[\left[\frac{\prod_{l=1}^{k_t} u_l^{(t)}!}{\prod_{\omega^{(t)}} n_{\omega^{(t)}}^{(t)}!}\right]\prod_{i=2}^{k_t}\left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta_{t,i}^{\alpha+n_{\omega_{\cdot,i}}^{(t)}-1}\left(1-\theta_{t,i}\right)^{\beta+\sum_{j=1}^{i-1}u_j^{(t)}-n_{\omega_{\cdot,i}}^{(t)}-1}\right]\right]}{p(\mathbf{n}_\omega \mid \mathbf{u})} . \tag{14}$$

Equation (14) can be simplified by multiplying the kernel of the numerator by

$$\frac{\Gamma\left(\alpha+\beta+\sum_{j=1}^{i-1}u_j^{(t)}\right)}{\Gamma\left(\alpha+n_{\omega_{\cdot,i}}^{(t)}\right)\Gamma\left(\beta+\sum_{j=1}^{i-1}u_j^{(t)}-n_{\omega_{\cdot,i}}^{(t)}\right)}\frac{\Gamma\left(\alpha+n_{\omega_{\cdot,i}}^{(t)}\right)\Gamma\left(\beta+\sum_{j=1}^{i-1}u_j^{(t)}-n_{\omega_{\cdot,i}}^{(t)}\right)}{\Gamma\left(\alpha+\beta+\sum_{j=1}^{i-1}u_j^{(t)}\right)} = 1 . \tag{15}$$

The numerator becomes:

$$p(\boldsymbol{\theta},\alpha,\beta,\mathbf{n}_{\omega}\,|\,\mathbf{u}) = p(\alpha,\beta)\times$$
$$\prod_{t=1}^{T}\left[\frac{\prod_{l=1}^{k_t} u_l^{(t)}!}{\prod_{\omega} n_{\omega}^{(t)}!}\prod_{i=2}^{k_t}\left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma\left(\alpha+n_{\omega,i}^{(t)}\right)\Gamma\left(\beta+\sum_{j=1}^{i-1}u_j^{(t)}-n_{\omega,i}^{(t)}\right)}{\Gamma\left(\alpha+\beta+\sum_{j=1}^{i-1}u_j^{(t)}\right)}\right]\right]\times$$
$$\prod_{t=1}^{T}\prod_{i=2}^{k_t}\left[Beta(\alpha+n_{\omega,i}^{(t)},\beta+\sum_{j=1}^{i-1}u_j^{(t)}-n_{\omega,i}^{(t)})\right]. \tag{16}$$

Consequently, the joint posterior distribution of $(\boldsymbol{\theta},\alpha,\beta)$ conditional on $\mathbf{n}_{\omega}$ and $\mathbf{u}$ is

$$p(\boldsymbol{\theta},\alpha,\beta\,|\,\mathbf{n}_{\omega},\mathbf{u}) = cp(\alpha,\beta)\times$$
$$\prod_{t=1}^{T}\left[\prod_{i=2}^{k_t}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma\left(\alpha+n_{\omega,i}^{(t)}\right)\Gamma\left(\beta+\sum_{j=1}^{i-1}u_j^{(t)}-n_{\omega,i}^{(t)}\right)}{\Gamma\left(\alpha+\beta+\sum_{j=1}^{i-1}u_j^{(t)}\right)}Beta(\alpha+n_{\omega,i}^{(t)},\beta+\sum_{j=1}^{i-1}u_j^{(t)}-n_{\omega,i}^{(t)})\right], \tag{17}$$

where

$$\frac{1}{c} = \int_0^{\infty}\int_0^{\infty} p(\alpha,\beta)\prod_{t=1}^{T}\prod_{i=2}^{k_t}\left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma\left(\alpha+n_{\omega,i}^{(t)}\right)\Gamma\left(\beta+\sum_{j=1}^{i-1}u_j^{(t)}-n_{\omega,i}^{(t)}\right)}{\Gamma\left(\alpha+\beta+\sum_{j=1}^{i-1}u_j^{(t)}\right)}\right]d\alpha d\beta. \tag{18}$$

To compute the joint posterior distribution of (α,β), θ's are integrated over the entire support:

$$p(\alpha,\beta\,|\,\mathbf{n}_{\omega},\mathbf{u}) = \int_0^1\int_0^1\cdots\int_0^1 p(\boldsymbol{\theta},\alpha,\beta\,|\,\mathbf{n}_{\omega},\mathbf{u})\,d\theta_{1,1}d\theta_{1,2}\cdots d\theta_{T,k_t} \tag{19}$$

Because θ's in equation (17) appear only in the beta distributions and all θ's are assumed independent, the product of the integrals is 1:

$$p(\alpha,\beta\,|\,\mathbf{n}_\omega,\mathbf{u}) = cp(\alpha,\beta)\prod_{t=1}^{T}\left[\prod_{i=2}^{k_t}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma\left(\alpha+n_{\omega,i}^{(t)}\right)\Gamma\left(\beta+\sum_{j=1}^{i-1}u_j^{(t)}-n_{\omega,i}^{(t)}\right)}{\Gamma\left(\alpha+\beta+\sum_{j=1}^{i-1}u_j^{(t)}\right)}\right]. \quad (20)$$

To compute the conditional probability distributions of parameters and hyperparameters, however, the prior distribution on hyperparameters (i.e., $p(\alpha, \beta)$) needs to be defined. One could construct a probability distribution for the hyperparameters from available information, outside the dataset ($\mathbf{n}_\omega$ and $\mathbf{u}$). In this report, however, I assume that there is no such available information on the capture probabilities. A diffuse non-informative prior distribution, therefore, is defined for the hyperparameters. I follow the method of Gelman et al. (1995).

The hyperparameters (α, β) are re-parameterized in terms of $\text{logit}\left(\alpha/(\alpha+\beta)\right) = \ln\left(\alpha/\beta\right)$, which is the logit of the mean, and $\ln(\alpha+\beta)$, which is the natural logarithm of the 'sample size' in the beta distribution for θ (Gelman et al. 1995; p. 130). I use a bivariate uniform prior distribution on $\left(\alpha/(\alpha+\beta), 1/\sqrt{\alpha+\beta}\right)$, which yields a proper posterior distribution (Gelman et al. 1995; p. 131). By multiplying by the appropriate Jacobian (Appendix A), this hyperprior distribution on the original scale is:

$$p(\alpha,\beta) \propto (\alpha+\beta)^{-\frac{5}{2}}, \quad (21)$$

and on the natural logarithm scale (Appendix B):

$$p\left(\ln\left(\frac{\alpha}{\beta}\right), \ln(\alpha+\beta)\right) \propto \alpha\beta(\alpha+\beta)^{-\frac{5}{2}}, \quad (22)$$

or

$$p(A,B) \propto \frac{e^{A-\frac{B}{2}}}{\left(1+e^{A}\right)^{2}}, \quad (23)$$

where $A = \ln(\alpha/\beta)$ and $B = \ln(\alpha+\beta)$.

The joint posterior distribution of $(A, B) = \left(\ln(\alpha/\beta), \ln(\alpha+\beta)\right)$ is obtained numerically via the Metropolis algorithm (Metropolis et al. 1953, Hastings 1970). The Metropolis algorithm is a Markov chain simulation, which creates samples from a specified target distribution (Gelman et al. 1995). For each data set, I use a minimum of four independent chains with 15,000 steps within each chain. The starting point for each chain is a random point in the range $-3 < A < -1$ and $1 < B < 4$. The jumping distribution for the first chain is Unif(-0.1, 0.1). To eliminate the initial sojourn steps, I discard the first 7,500 steps from each chain. The ratio between the two parameters of the jumping distribution for the next sequence is proportional to the covariance structure of the remaining 7,500 samples for A and B.

The scale reduction factor ($\sqrt{\hat{R}}$) of Gelman et al. (1995) is used to determine the convergence of a posterior distribution from a total of 30,000 pairs of A and B (7,500 times four chains). The scale reduction factor is the square root of the ratio between the marginal posterior variance and within-sequence variance:

$$\sqrt{\hat{R}} = \sqrt{\frac{\widehat{\text{var}}^{+}(\psi \mid y)}{W}} \tag{24}$$

where

$$\widehat{\text{var}}^{+}(\psi \mid y) = \frac{n-1}{n}W + \frac{1}{n}B, \tag{25}$$

$$W = \frac{1}{J}\sum_{j=1}^{J}\left[\frac{1}{n-1}\sum_{i=1}^{n}\left(\psi_{i,j} - \bar{\psi}_{.,j}\right)^2\right], \tag{26}$$

$$B = \frac{n}{J-1}\sum_{j=1}^{J}\left(\bar{\psi}_{.,j} - \bar{\psi}_{..}\right)^2, \tag{27}$$

$$\bar{\psi}_{.,j} = \frac{1}{n}\sum_{i=1}^{n}\psi_{i,j}\text{, and} \tag{28}$$

$$\bar{\psi}_{.,.} = \frac{1}{J}\sum_{j=1}^{J}\bar{\psi}_{.,j}\,. \tag{29}$$

The numerator in the square root of (24) is the estimated marginal posterior variance, which overestimates the marginal posterior variance under the assumption that the starting distribution is appropriately overdispersed but unbiased as n $\rightarrow \infty$. For any finite n, however, the within-sequence variance W should be an underestimate of $\text{var}(\psi \mid y)$ because the individual sequences have not had time to range over all of the target distribution. Consequently, $\sqrt{\hat{R}}$ declines to 1 as n $\rightarrow \infty$ (Gelman et al. 1995; p. 331).

If the maximum scale reduction factor is less than or equal to 1.2, the posterior samples are retained (Gelman et al. 1995). If the convergence statistic is greater than 1.2, another sequence of four chains is computed. A maximum of four independent sequences (i.e., 16 independent chains) are computed. To avoid the serial correlation among posterior samples, a total of 7,500 random samples with replacement are drawn from these retained posterior samples and used in the inference for the abundance. These 7,500 posterior samples are transformed back into the original scale (α, β). If the posterior distribution does not converge after four independent sequences, the uniform prior distribution is used for the hyperdistribution, i.e., (α, β) = (1,1).

To obtain the joint probability distribution of hyperparameters for the capture probabilities during the t^{th} primary period, all data except for the t^{th} primary period are used. The prior distribution of hyperparameters for the t^{th} primary period is denoted by $p(\alpha^{(-t)}, \beta^{(-t)})$, emphasizing the deletion of data for the t^{th} primary period. These

hyperparameters, then, are used in the following analysis for modeling the abundance for the t^{th} primary period.

To make an inference on the abundance, the relationship among the abundance within the t^{th} primary period, capture probabilities, and the number of animals caught during each secondary occasion needs to be modeled. I consider two models: beta-multinomial and beta-binomial. Although both models are based on the same set of assumptions, the beta-binomial model does not use the information from recaptures. The beta-multinomial model does use the information from the number of individuals that are recaptured.

Beta-multinomial model

The observed numbers of individuals that belong to unique capture histories during the t^{th} primary period ($\mathbf{n}^{(t)}_{\omega\in\Omega}$) are modeled with a multinomial likelihood function: $\mathbf{n}^{(t)}_{\omega\in\Omega} \sim mult\left(N^{(t)}, \boldsymbol{\theta}^{(t)}\right)$, where $N^{(t)}$ is the total number of catchable individuals in the population during the t^{th} primary period and $\boldsymbol{\theta}^{(t)}$ is a vector of capture probabilities with length k_t, where k_t is the number of secondary occasions within the t^{th} primary period. The full likelihood function is:

$$p\left(\mathbf{n}^{(t)}_{\omega\in\Omega} \mid N^{(t)}, \boldsymbol{\theta}^{(t)}\right) = \frac{N^{(t)}!}{\prod_{\omega\in\Omega} n^{(t)}_{\omega}!} \prod_{\omega\in\Omega} \theta_{\omega}^{n_{\omega}} \tag{30}$$

where θ_{ω} is the appropriate product of capture probabilities for the capture history ω. For example, for $k_t = 3$, $\Omega = \{111, 110, 101, 100, 011, 010, 001, 000\}$ and $n_{000} = N - (n_{111}+n_{110}+n_{101}+n_{100}+n_{011}+n_{010}+n_{001}) = N - n^{(t)}$, indicating the total number of individuals caught during the t^{th} primary period. The likelihood function for $k_t = 3$ is:

$$
\begin{aligned}
p\left(\mathbf{n}_{\omega\in\Omega}^{(t)} \mid N^{(t)}, \boldsymbol{\theta}^{(t)}\right) = & \frac{N^{(t)}!}{n_{111}!n_{110}!n_{101}!n_{100}!n_{011}!n_{010}!n_{001}!(N^{(t)}-n^{(t)})!} \times \\
& \left(\theta_1\theta_2\theta_3\right)^{n_{111}} \left(\theta_1\theta_2\left(1-\theta_3\right)\right)^{n_{110}} \left(\theta_1\left(1-\theta_2\right)\theta_3\right)^{n_{101}} \times \\
& \left(\theta_1\left(1-\theta_2\right)\left(1-\theta_3\right)\right)^{n_{100}} \left(\left(1-\theta_1\right)\theta_2\theta_3\right)^{n_{011}} \times \\
& \left(\left(1-\theta_1\right)\theta_2\left(1-\theta_3\right)\right)^{n_{010}} \left(\left(1-\theta_1\right)\left(1-\theta_2\right)\theta_3\right)^{n_{001}} \times \\
& \left(\left(1-\theta_1\right)\left(1-\theta_2\right)\left(1-\theta_3\right)\right)^{N^{(t)}-n^{(t)}}
\end{aligned} \tag{31}
$$

Rearranging θ's results in the following form:

$$
\begin{aligned}
p\left(\mathbf{n}_{\omega\in\Omega}^{(t)} \mid N^{(t)}, \boldsymbol{\theta}^{(t)}\right) = & \frac{N^{(t)}!}{n_{111}!n_{110}!n_{101}!n_{100}!n_{011}!n_{010}!n_{001}!(N^{(t)}-n^{(t)})!} \times \\
& \left(\theta_1\right)^{n_{(1)}} \left(\theta_2\right)^{n_{(2)}} \left(\theta_3\right)^{n_{(3)}} \times \\
& \left(1-\theta_1\right)^{N^{(t)}-n_{(1)}} \times \left(1-\theta_2\right)^{N^{(t)}-n_{(2)}} \left(1-\theta_3\right)^{N^{(t)}-n_{(3)}}
\end{aligned} \tag{32}
$$

where $n_{(j)}$ is the sum of the number of individuals that belong to capture histories with 1 in the j^{th} digit. For example, $n_{(2)} = n_{111}+n_{110}+n_{011}+n_{010}$. In general, the likelihood function for a primary period with k secondary occasions can be expressed as

$$
p\left(\mathbf{n}_{\omega\in\Omega}^{(t)} \mid N^{(t)}, \underline{\theta}^{(t)}\right) = c\frac{N^{(t)}!}{(N^{(t)}-n^{(t)})!}\prod_{j=1}^{k}\left(\theta_j\right)^{n_{(j)}}\left(1-\theta_j\right)^{N^{(t)}-n_{(j)}}, \tag{33}
$$

where $c = 1/n_{111}!n_{110}!n_{101}!n_{100}!n_{011}!n_{010}!n_{001}!$, which does not depend on $N^{(t)}$ and $\boldsymbol{\theta}^{(t)}$.

Under the assumption of independent capture probabilities among capture occasions, the joint distribution of capture probabilities during the t^{th} primary period is the product of beta distributions:

$$
p\left(\boldsymbol{\theta}^{(t)} \mid \boldsymbol{\alpha}^{(t)}, \boldsymbol{\beta}^{(t)}\right) = \prod_{j=1}^{k_t}\frac{\Gamma\left(\alpha_j^{(t)}+\beta_j^{(t)}\right)}{\Gamma\left(\alpha_j^{(t)}\right)\Gamma\left(\beta_j^{(t)}\right)}\left(\theta_j^{(t)}\right)^{\alpha_j^{(t)}-1}\left(1-\theta_j^{(t)}\right)^{\beta_j^{(t)}-1} \tag{34}
$$

The joint posterior probability distribution of the abundance, capture probabilities, and the hyperparameters conditional on the recapture data during the t^{th} primary period is proportional to the following:

$$p\left(N^{(t)},\boldsymbol{\theta}^{(t)},\boldsymbol{\alpha}^{(t)},\boldsymbol{\beta}^{(t)}\mid\mathbf{n}_{\omega}^{(t)}\right)\propto p\left(N^{(t)}\right)p\left(\boldsymbol{\alpha}^{(t)},\boldsymbol{\beta}^{(t)}\right)p\left(\boldsymbol{\theta}^{(t)}\mid\boldsymbol{\alpha}^{(t)},\boldsymbol{\beta}^{(t)}\right)p\left(\mathbf{n}_{\omega}^{(t)}\mid N^{(t)},\boldsymbol{\theta}^{(t)}\right) \tag{35}$$

Assuming that there is no prior information on the abundance, I consider the uniform distribution for the prior distribution of the abundance. If there is information on abundance, however, it should be used to construct a probability distribution for the abundance. The joint prior distribution of the hyperparameters is computed numerically during the first stage, where recapture data for all primary periods except for the t^{th} primary period are used for computing the joint distribution, which I denote $p_{\alpha^{(-t)},\beta^{(-t)}}\left(\alpha,\beta\mid\mathbf{n}_{\omega}^{(-t)},\mathbf{u}^{(-t)}\right)$. I use this joint distribution for the hyperparameters for all capture probabilities within the primary period:

$$p\left(\alpha_j,\beta_j\right)=p_{\alpha^{(-t)},\beta^{(-t)}}\left(\alpha_j,\beta_j\mid\mathbf{n}_{\omega}^{(-t)},\mathbf{u}^{(-t)}\right) \tag{36}$$

for all $j = 1, \ldots, k_t$. The joint prior distribution for the hyperparameters of capture probabilities is the product of (36):

$$\begin{aligned}p\left(\boldsymbol{\alpha}^{(t)},\boldsymbol{\beta}^{(t)}\right)&=p\left(\alpha_1^{(t)},\beta_1^{(t)}\right)p\left(\alpha_2^{(t)},\beta_2^{(t)}\right)\cdots p\left(\alpha_{k_t}^{(t)},\beta_{k_t}^{(t)}\right)\\&=\prod_{j=1}^{k_t}p_{\alpha^{(-t)},\beta^{(-t)}}\left(\alpha_j^{(t)},\beta_j^{(t)}\mid\mathbf{n}_{\omega}^{(-t)},\mathbf{u}^{(-t)}\right)\end{aligned} \tag{37}$$

Consequently, using the equation (34):

$$p\left(\boldsymbol{\alpha}^{(t)},\boldsymbol{\beta}^{(t)}\right)p\left(\boldsymbol{\theta}\mid\boldsymbol{\alpha}^{(t)},\boldsymbol{\beta}^{(t)}\right)=\prod_{j=1}^{k_t}p_{\alpha^{(-t)},\beta^{(-t)}}\left(\alpha_j^{(t)},\beta_j^{(t)}\mid\mathbf{n}_{\omega}^{(-t)},\mathbf{u}^{(-t)}\right)\prod_{j=1}^{k_t}p\left(\theta_j^{(t)}\mid\alpha_j^{(t)},\beta_j^{(t)}\right) \tag{38}$$

Substituting(33), (34), and (38) into (35):

$$p\left(N^{(t)},\boldsymbol{\theta}^{(t)},\boldsymbol{\alpha}^{(t)},\boldsymbol{\beta}^{(t)}\mid\mathbf{n}_{\omega}^{(t)}\right)\propto p\left(N^{(t)}\right)\prod_{j=1}^{k_t}p_{\alpha^{(-t)},\beta^{(-t)}}\left(\alpha_j^{(t)},\beta_j^{(t)}\mid\mathbf{n}_{\omega}^{(-t)},\mathbf{u}^{(-t)}\right)\times$$
$$\left[\prod_{j=1}^{k_t}\frac{\Gamma\left(\alpha_j^{(t)}+\beta_j^{(t)}\right)}{\Gamma\left(\alpha_j^{(t)}\right)\Gamma\left(\beta_j^{(t)}\right)}\left(\theta_j^{(t)}\right)^{\alpha_j^{(t)}-1}\left(1-\theta_j^{(t)}\right)^{\beta_j^{(t)}-1}\right]\times$$
$$\left[\frac{c\Gamma\left(N^{(t)}+1\right)}{\Gamma\left(N^{(t)}-n^{(t)}+1\right)}\prod_{j=1}^{k_t}\left(\theta_j^{(t)}\right)^{n_{(j)}^{(t)}}\left(1-\theta_j^{(t)}\right)^{N^{(t)}-n_{(j)}^{(t)}}\right] \tag{39}$$

Rearranging terms results in the following:

$$p\left(N^{(t)},\boldsymbol{\theta}^{(t)},\boldsymbol{\alpha}^{(t)},\boldsymbol{\beta}^{(t)}\mid\mathbf{n}_{\omega}^{(t)}\right)\propto\frac{cp\left(N^{(t)}\right)\Gamma\left(N^{(t)}+1\right)}{\Gamma\left(N^{(t)}-n^{(t)}+1\right)}\times\prod_{j=1}^{k_t}p_{\alpha^{(-t)},\beta^{(-t)}}\left(\alpha_j^{(t)},\beta_j^{(t)}\mid\mathbf{n}_{\omega}^{(-t)},\mathbf{u}^{(-t)}\right)$$
$$\prod_{j=1}^{k_t}\frac{\Gamma\left(\alpha_j^{(t)}+\beta_j^{(t)}\right)}{\Gamma\left(\alpha_j^{(t)}\right)\Gamma\left(\beta_j^{(t)}\right)}\left(\theta_j^{(t)}\right)^{n_{(j)}^{(t)}+\alpha_j^{(t)}-1}\left(1-\theta_j^{(t)}\right)^{N^{(t)}-n_{(j)}^{(t)}+\beta_j^{(t)}-1} \tag{40}$$

The posterior distribution of the abundance ($N^{(t)}$) is computed by integrating over $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\theta}$. Using the following relationship,

$$\int_0^1\theta_j^{n_j+\alpha_j^{(t)}+1}\left(1-\theta_j\right)^{N-n_j+\beta_j^{(t)}+1}d\theta_j=\frac{\Gamma\left(n_j+\alpha_j^{(t)}\right)\Gamma\left(N^{(t)}-n_j+\beta_j^{(t)}\right)}{\Gamma\left(N^{(t)}+\alpha_j^{(t)}+\beta_j^{(t)}\right)}, \tag{41}$$

capture probabilities ($\boldsymbol{\theta}$) are integrated out from (40). To improve the legibility of the equation, I omit the superscripts on $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$:

$$p\left(N^{(t)},\boldsymbol{\alpha},\boldsymbol{\beta}\mid\mathbf{n}_{\omega}\right)\propto\frac{cp\left(N^{(t)}\right)\Gamma\left(N^{(t)}+1\right)}{\Gamma\left(N^{(t)}-n^{(t)}+1\right)}\prod_{j=1}^{k_t}p_{\alpha^{(-t)},\beta^{(-t)}}\left(\alpha_j^{(t)},\beta_j^{(t)}\mid\mathbf{n}_{\omega}^{(-t)},\mathbf{u}^{(-t)}\right)\times$$
$$\prod_{j=1}^{k_t}\frac{\Gamma\left(n_j+\alpha_j\right)\Gamma\left(N^{(t)}-n_j+\beta_j\right)\Gamma\left(\alpha_j+\beta_j\right)}{\Gamma\left(N^{(t)}+\alpha_j+\beta_j\right)\Gamma\left(\alpha_j\right)\Gamma\left(\beta_j\right)} \tag{42}$$

Integrating out $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ results in the posterior distribution of the abundance:

$$p\left(N^{(t)} \mid \mathbf{n}_{\omega}\right) \propto \int_0^{\infty} \cdots \int_0^{\infty} p\left(N^{(t)}, \alpha_j, \beta_j \mid \mathbf{n}_{\omega}\right) d\alpha_1 \cdots d\alpha_{k_t} d\beta_1 \cdots d\beta_{k_t} \tag{43}$$

Beta-binomial model

The number of captured animals during the j^{th} secondary occasion is modeled with an independent beta-binomial likelihood function: $n_j^{(t)} \overset{i.i.d}{\sim} Beta-Bin(N^{(t)}, \alpha_j^{(t)}, \beta_j^{(t)})$ for $j = 1,2,\ldots, k_t$. In other words, the number of animals caught during the j^{th} occasion is a binomial random variable with parameters $N^{(t)}$ and $\theta_j^{(t)}$; $n_j^{(t)} \sim Bin\left(N^{(t)}, \theta_j^{(t)}\right)$, whereas the capture probability during the j^{th} occasion is a beta random variable with parameters $\alpha_j^{(t)}$ and $\beta_j^{(t)}$; $\theta_j^{(t)} \sim Beta\left(\alpha_j^{(t)}, \beta_j^{(t)}\right)$. The combination of these distributions result in the beta-binomial distribution; $n_j^{(t)} \sim Beta-Bin(N^{(t)}, \alpha_j^{(t)}, \beta_j^{(t)})$.

The likelihood function for the t^{th} primary period with k_t secondary occasions, therefore, is a product of k_t independent beta-binomial probability mass functions:

$$p\left(\mathbf{n}_t \mid N^{(t)}, \boldsymbol{\alpha}^{(t)}, \boldsymbol{\beta}^{(t)}\right) = \prod_{j=1}^{k_t} \frac{\Gamma\left(N^{(t)}+1\right)\Gamma\left(\alpha_j^{(t)}+n_j^{(t)}\right)\Gamma\left(N^{(t)}+\beta_j^{(t)}-n_j^{(t)}\right)\Gamma\left(\alpha_j^{(t)}+\beta_j^{(t)}\right)}{\Gamma\left(n_j^{(t)}+1\right)\Gamma\left(N_j^{(t)}-n_j^{(t)}+1\right)\Gamma\left(\alpha_j^{(t)}+\beta_j^{(t)}+N^{(t)}\right)\Gamma\left(\alpha_j^{(t)}\right)\Gamma\left(\beta_j^{(t)}\right)}, \tag{44}$$

where $\mathbf{n}^{(t)} = \left(n_1^{(t)}, \ldots, n_{k_t}^{(t)}\right)$, $\boldsymbol{\alpha}^{(t)} = \left(\alpha_1^{(t)}, \ldots, \alpha_{k_t}^{(t)}\right)$, and $\boldsymbol{\beta}^{(t)} = \left(\beta_1^{(t)}, \ldots, \beta_{k_t}^{(t)}\right)$. The inference on the abundance during the t^{th} primary period, $N^{(t)}$, is made by using the beta-binomial model (44), prior distributions for $\left(\boldsymbol{\alpha}^{(t)}, \boldsymbol{\beta}^{(t)}\right)$ and $N^{(t)}$, and the observed data $\left(\mathbf{n}^{(t)} = \left(n_1^{(t)}, \ldots, n_{k_t}^{(t)}\right)\right)$ via a Bayesian procedure.

The joint posterior probability distribution of $N^{(t)}$, $\boldsymbol{\alpha}^{(t)}$ and $\boldsymbol{\beta}^{(t)}$ can be expressed by:

$$p\left(N^{(t)}, \boldsymbol{\alpha}^{(t)}, \boldsymbol{\beta}^{(t)} \mid \mathbf{n}^{(t)}\right) \propto p\left(N^{(t)}\right) \prod_{j=1}^{k_t} p\left(\alpha_j^{(t)}, \beta_j^{(t)}\right) p\left(n_j^{(t)} \mid N^{(t)}, \alpha_j^{(t)}, \beta_j^{(t)}\right). \tag{45}$$

The posterior distribution of the abundance is computed by integrating out α's and β's over the entire parameter space:

$$p\left(N^{(t)} \mid \mathbf{n}^{(t)}\right) \propto \int_0^\infty \cdots \int_0^\infty p\left(N^{(t)}, \alpha_j^{(t)}, \beta_j^{(t)} \mid \mathbf{n}^{(t)}\right) d\alpha_1^{(t)} \cdots d\alpha_{k_t}^{(t)} d\beta_1^{(t)} \cdots d\beta_{k_t}^{(t)} \qquad (46)$$

Computation of the posterior distribution

To compute the posterior marginal distribution of the abundance from either model ((43) or (46)), I use a uniform prior distribution for $N^{(t)}$ under the assumption that no information is available, i.e., $p(N^{(t)}) \sim UNIF(N_{\min}^{(t)}, N_{\max}^{(t)})$, where $N_{\min}^{(t)}$ and $N_{\max}^{(t)}$ are positive integers, where $N_{\min}^{(t)} < N_{\max}^{(t)}$, and denote possible minimum and maximum population sizes. For the prior distribution of the hyperparameters, the posterior distribution from the previous analysis, i.e., $p\left(\alpha^{(-t)}, \beta^{(-t)}\right)$, is used. The posterior probability distribution of $N^{(t)}$ conditional on data can be computed by numerically integrating the hyperparameters $\left(\boldsymbol{\alpha}^{(t)}, \boldsymbol{\beta}^{(t)}\right)$ from the joint posterior distribution.

Because no algebraic solution is available for (43) or (46), I compute the posterior distribution directly using the following steps:

(1) Compute $n_j^{(t)}$ for j = 1,2,..., k_t, i.e., the number of individuals caught per secondary occasion within the t^{th} primary period.

(2) For the primary period t, find the joint prior distribution for $\left(\alpha_j^{(t)}, \beta_j^{(t)}\right)$ by using equation (20) from the previous section, i.e., $p\left(\alpha^{(-t)}, \beta^{(-t)} \mid \mathbf{n}_\omega, \mathbf{u}\right)$.

(3) For each possible $N^{(t)}$ between $N_{\min}^{(t)}$ and $N_{\max}^{(t)}$:

   a. For each pair of (α, β) from the posterior distribution $p\left(\alpha^{(-t)}, \beta^{(-t)} \mid \mathbf{n}_\omega, \mathbf{u}\right)$ compute the natural logarithm of unnormalized posterior density by using the natural logarithm of the likelihood function (44) for each secondary occasion and sum over all secondary occasions (j = 1,2,...,k_t).

b. Transform the unnormalized log-posterior to anti-log scale.

c. Numerically integrate over all (α, β) pairs for computing $cp(N_t \mid \mathbf{n}_t)$, where c is a constant.

d. Repeat the previous steps (a) through (c) for all possible $N^{(i)}$ values.

(4) Normalize the posterior by setting the total probability $p(N^{(i)} \mid \mathbf{n}^{(i)})$ under the curve to be one.

If information on abundance is available and a probability distribution can be assigned to the abundance, it can be used in step (3). I use a common prior distribution for capture probabilities for all secondary occasions within the primary period of interest because there is no information, outside the data for this period, to distinguish them.

To summarize the inference on the abundance for the t^{th} primary period, the mode, median, and mean of the posterior distribution are reported along with the quantiles of interest. To conduct the analyses, computer programs were written in Matlab® (MathWorks).

Simulations

To evaluate the validity of the proposed method and effects of changes in the temporal sampling design and the underlying parameters on posterior distributions, I conducted a series of Monte Carlo simulations. A total of seven combinations of parameters were tested (Table 1). The true hyperparameters that created simulated data are denoted (α_0, β_0), whereas the hyperparameters in the models are denoted with either no subscripts or subscripts other than 0.

Table 1. Combinations of temporal sampling design and true hyperparameters for Monte Carlo simulations. T = the number of primary periods, k = the number of secondary occasions within each primary period, (α_0, β_0) = set of the true hyperparameters for capture probabilities.

Case	T	k	(α_0, β_0)
A	3	4	(20, 2)
B	3	4	(2, 20)
C	3	4	(3, 3)
D	6	4	(2, 20)
E	12	4	(2, 20)
F	3	6	(2, 20)
G	3	8	(2, 20)

Simulation case A was used for validating the inference procedures. With consistently high capture probabilities, posterior distributions should center on the true abundance and their widths should be narrow.

To determine the effectiveness of using the two-step approach, I analyzed datasets for simulation cases A, B, and C with the two-step approach and when the hyperparameters in the analyses were set at constants ($(\alpha, \beta) = (1, 1)$). The latter approach is equal to assigning a non-informative prior distribution on capture probabilities. If the two-step approach is not useful, posterior distributions from both approaches should be the same.

To determine the effects of true capture probabilities on posterior distributions of abundance, datasets for simulation cases A, B, and C were analyzed using the two models with the two-step approach. Sampling distributions of true capture probabilities in these simulation cases corresponded to consistently high (case A), consistently low (case B), and variable (case C) capture probabilities. Corresponding beta distributions are shown in Figure 3. The numbers of primary periods and secondary occasions were held constant for these simulation cases.

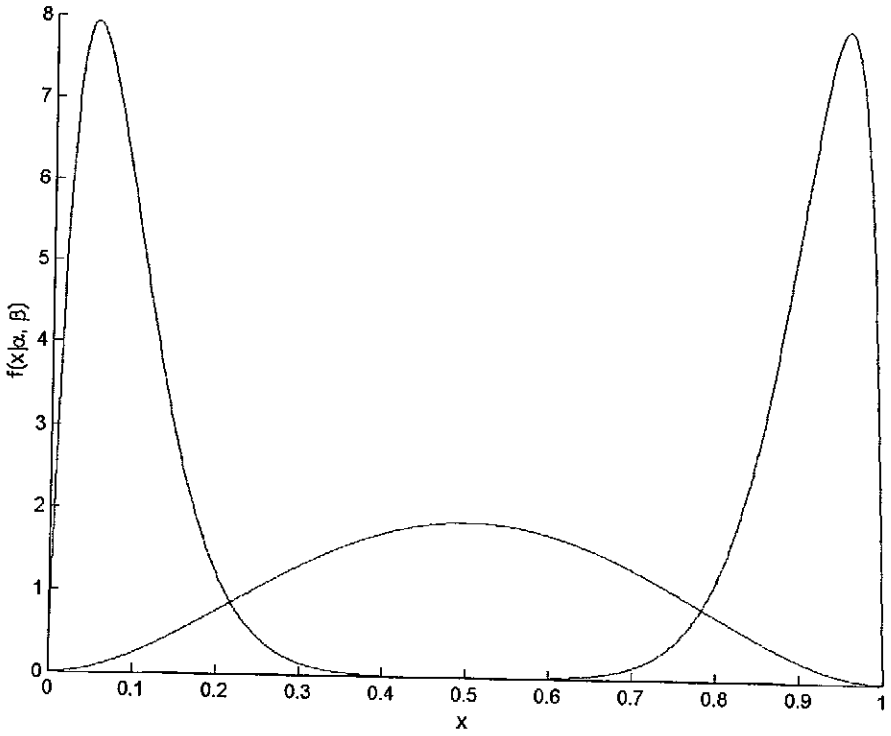

Figure 3. Probability density functions for three beta distributions; (mode from left to right) Beta(2, 20), Beta(3, 3), and Beta(20, 2).

To determine effects of temporal sampling designs on posterior distributions of abundance, the numbers of primary periods and secondary occasions were changed systematically, while keeping the true hyperparameters constant. Effects of the number of primary periods on posterior distributions were determined by comparing posterior distributions from three simulation cases. For these simulations, the number of primary periods was increased from three (case B) to six (case D) and twelve (case E), while the number of secondary sampling occasions per primary period and the true hyperparameters were held constant (Table 1).

Effects of the number of secondary sampling occasions on posterior distributions were determined by comparing posterior distributions from three simulation cases. For

these simulations, the number of secondary occasions for each primary period was increased from four (case B) to six (case F) and eight (case G), while the number of primary periods and the true hyperparameters were held constant (Table 1).

To create a simulated dataset, the number of individuals in a population for a primary period was drawn from a uniform distribution with the minimum 50 and maximum 500; $N^{(i)} \sim$UNIF(50, 500). The capture probability for a sampling occasion was drawn from a beta distribution with the appropriate hyperparameters for the simulation case ($\theta_j^{(i)} \sim Beta(\alpha_0, \beta_0)$; Table 1). For each sampling occasion, each individual in the population was caught with the true capture probability ($\theta_j^{(i)}$) for the sampling occasion (i.e., a capture of an individual during a sampling occasion is a Bernoulli event with probability $\theta_j^{(i)}$). This process resulted in sequences of zeros and ones for all individuals in the population. These sequences, then, were summarized into the sufficient statistics for the models; the number of individuals caught for each sampling occasion, the number of recaptured individuals for each sampling occasion, and the number of new individuals captured for each sampling occasion. For each simulation case, 500 independent datasets were simulated.

These simulated datasets were analyzed using the proposed method. I used UNIF(10, 1500) for the prior distribution of the abundance. By defining the prior distribution, an implicit assumption of the maximum possible population size was made for each analysis. The implication of the violation of the assumption is discussed in the discussion section.

Posterior distributions were summarized in the following statistics: widths of 95% posterior intervals, the proportion of 95% posterior intervals that included the true abundance, point estimators (mode $\hat{N}_{\text{mode}}$, mean $\hat{N}_{\text{mean}}$, and median $\hat{N}_{\text{med}}$ of the posterior distribution), differences between point estimates and true abundances $\left(\Delta = \hat{N} - N\right)$, and

the ratio between the width of the 95% posterior intervals (97.5 percentile – 2.5 percentile) and the mode. The ratio is analogous to the coefficient of variation, in which variability is expressed in the same scale regardless of the size of a point estimate. I hereafter call this ratio the relative precision.

These summary statistics provided several criteria for finding a better model (beta-binomial vs. beta-multinomial) and approach (i.e., two-step vs. fixed hyperparameters). A model or approach is considered better than the other when widths of posterior intervals are narrow, point estimates are close to the true abundances, the relative precision is consistently small and unaffected by the true abundance, and the posterior intervals include the true abundance with claimed probability.

To quantify the variability of point estimators, I report the slope, intercept, and coefficient of determination (R^2) of a simple linear regression between the true abundance and a point estimator. If the point estimator provides accurate estimates for a case of simulated datasets, the regression line should have a slope of one, an intercept of zero, and a coefficient of determination (R^2) = 1.0. The slope and intercept of a regression line indicate the discrepancy between point estimates and the true abundance. An imprecise point estimator would result in a smaller value of the coefficient of determination.

To compare the performance of the proposed method to the maximum likelihood, I analyzed the same simulated datasets using software MARK (White and Burnham 1999). In MARK, 'Closed Captures' (CC) and 'Huggins Closed Captures' (HCC) models were used for the analysis. I used these models to compute the abundance estimate for the last (t^{th}) primary period of each dataset. The difference between these models is whether the abundance (N) is treated as an explicit parameter (CC) or as a derived parameter (HCC). Consequently, the number of parameters was greater for the

CC model than for the HCC model. For example, for a period with four capture occasions, the CC model contains five parameters (four capture/recapture probabilities and abundance), whereas the HCC model contains four parameters (four capture/recapture probabilities), where I set capture probabilities and recapture probabilities equal. Outputs of MARK were summarized in the point estimates (MLE) of the abundance and the 95% confidence intervals. To compute a comparable statistic to the relative precision defined previously, I computed the ratio between the width of 95% confidence interval and maximum likelihood estimate for each dataset. It is called the relative precision of MLE in this chapter.

Results

Effectiveness of the two-step approach

To determine the effectiveness of the two-step approach, three simulation sets (A, B, and C) were analyzed using two approaches: the two-step and when hyperparameters were held constant ($\alpha = 1, \beta = 1$). For this analysis, I only consider the beta-multinomial model because the beta-binomial likelihood function is independent of data when $\alpha = 1$ and $\beta = 1$. For simulation case A, point estimates from the two approaches indicated no apparent differences (Figure 4).

When the hyperparameters were fixed, point estimators did not overestimate the true abundance ($-2 \leq \Delta_{mode} \leq 0$, $-2 \leq \Delta_{median} \leq 0$, $-2 \leq \Delta_{mean} \leq 0$). For all estimators, point estimates were less than the true abundance for 36% (180/500) of the simulations. For the remainder, point estimates equaled the true abundance. Widths of 95% posterior intervals were one or two. Only 65% (325/500) of the posterior intervals included true

abundance. For all estimators, the estimated intercepts of the regression lines were 0.48, slopes were 1.00, and coefficients of determinations were 1.0 (Table 2).

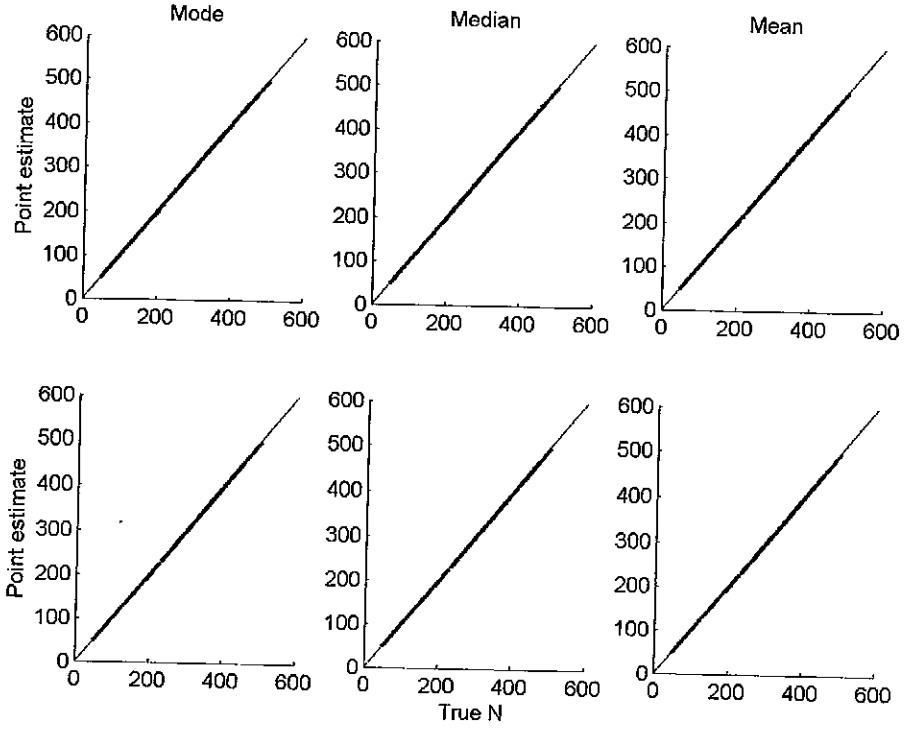

Figure 4. Relationships between point estimates from the beta-multinomial model and the true abundance for simulation case A. The upper three figures are results for the constant hyperparameters (1, 1), whereas for the lower three figures are for the two-step approach.

For the two-step approach, point estimators did not overestimate the true abundance ($-3 \leq \Delta_{mode} \leq 0$, $-3 \leq \Delta_{median} \leq 0$, $-3 \leq \Delta_{mean} \leq 0$). Point estimates were less than the true abundance for approximately 50% (253/500) of the simulations. For the remainder, point estimates equaled the true abundance. Widths of 95% posterior intervals ranged from one to three. Only 54% (270/500) of the posterior intervals included true abundance. For all estimators, the estimated intercept of the regression lines were 0.92, slopes were 0.99, and coefficients of determinations were 1.0 (Table 2).

Table 2. Estimated parameters for the linear regression analysis between the true abundance and point estimates for simulation case A for the beta-multinomial model.

Approach	Point estimates	Intercept	Slope	r^2
Fixed parameters	Mode	0.48	1.00	1.000
	Median	0.48	1.00	1.000
	Mean	0.48	1.00	1.000
Two-step	Mode	0.92	0.99	1.000
	Median	0.92	0.99	1.000
	Mean	0.92	0.99	1.000

Relative precisions were similar for the two approaches. The narrow posterior intervals caused the relative precision to be an inverse function of the true abundance ($1/N$, $2/N$, or $3/N$; Figure 5). When the hyperparameters were fixed, the relative precision ranged from 0.002 to 0.03 with the median of 0.004, whereas for the two-step approach, it ranged from 0.002 to 0.024 with a median of 0.004 (Figure 5).

For simulation case B, point estimators were less accurate than for simulation case A (Figure 6). When the hyperparameters were fixed, point estimators were variable around the true abundance ($-280 \leq \Delta_{mode} \leq 258$, $-253 \leq \Delta_{median} \leq 387$, $-235 \leq \Delta_{mean} \leq 414$). The majority of modes (453/500), medians (380/500), and means (342/500) were less than the true abundances. Consequently, the intercepts of least squares regression lines were negative (Table 3). Coefficient of determinations were approximately 0.80.

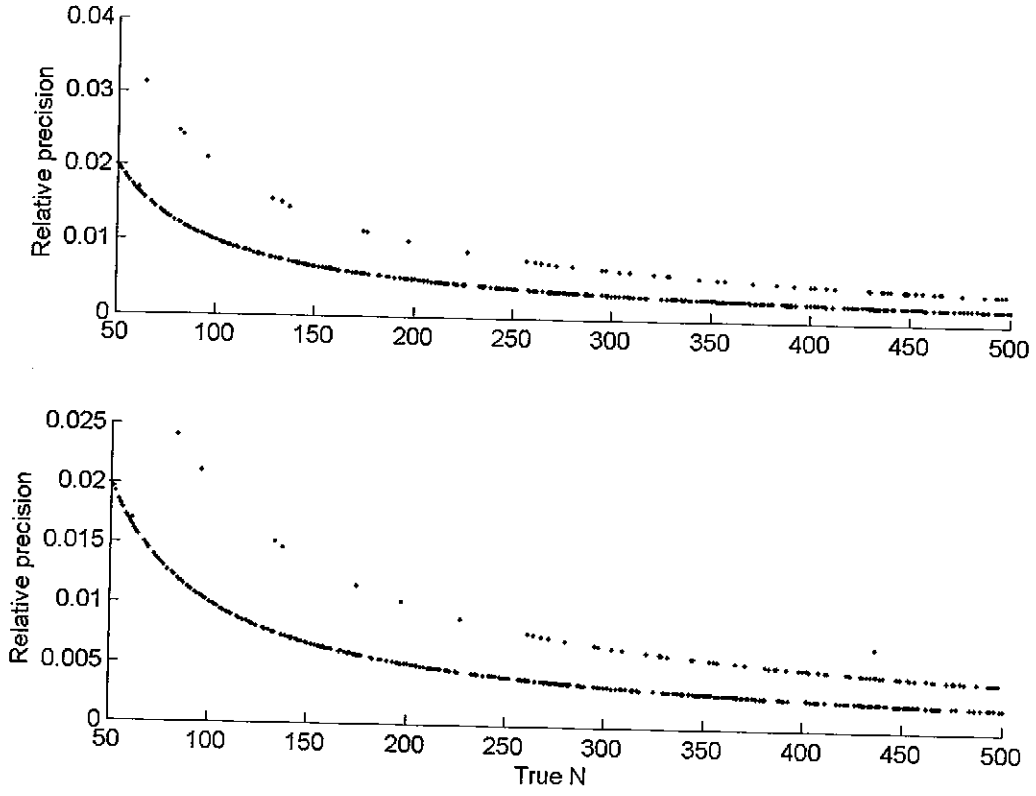

Figure 5. The relationships between the true abundance and relative precision for simulation case A when the beta-multinomial model was used. The upper figure is the result for the constant hyperparameters (1, 1), whereas for the lower figure is for the two-step approach.

When the same datasets were analyzed using the two-step approach, the variability in point estimators was similar to when the hyperparameters were fixed ($-264 \leq \Delta_{mode} \leq 346$, $-236 \leq \Delta_{median} \leq 479$, $-211 \leq \Delta_{mean} \leq 500$; Figure 6). Approximately 62% (310/500) of modes, 48% (244/500) of medians, and 40% (202/500) of means were less than the true abundances. The intercept of the regression line for the mode was negative, whereas they were positive for median and mean. Coefficient of determinations ranged between 0.70 and 0.78 (Table 3).

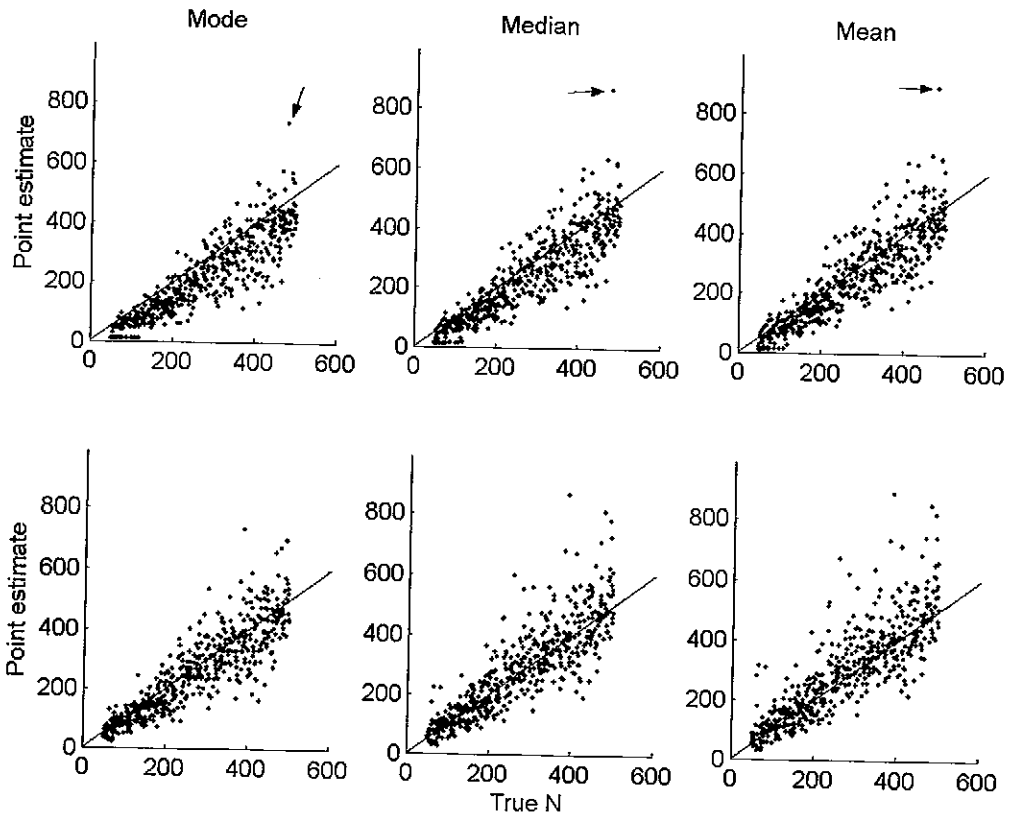

Figure 6. Relationships between point estimates and true abundance for simulation case B. The multinomial model was used for computing the posterior distributions. The top three panels are results for the fixed hyperparameters, whereas the bottom three panels are for the two-step approach. Solid lines indicate 45-degree lines. An arrow indicates one extreme value.

Table 3. Estimated parameters for the linear regression analysis between the true abundance and the point estimates for simulation case B when the beta-multinomial model was used.

Approach	Point estimates	Intercept	Slope	r^2
Fixed parameters	Mode	-35.1	0.93	0.803
	Median	-28.42	0.98	0.791
	Mean	-22.42	1.01	0.779
Two-step	Mode	–9.03	0.98	0.781
	Median	4.36	1.02	0.741
	Mean	17.05	1.03	0.707

Posterior intervals for simulation case B were wider than for simulation case A. When the hyperparameters were fixed, widths of 95% posterior intervals ranged from 11 to 1045, with the median of 215. True abundances were included in approximately 88% (438/500) of the posterior intervals. For the two-step approach, they ranged from 37 to 1208, with a median of 248, and approximately 94% (472/500) of the posterior intervals included the true abundances.

The relative precision for simulation case B was uncorrelated with the true abundance (Figure 7). When the hyperparameters were fixed, the relative precision ranged from 0.29 to 23.7, with the median of 1.2. Several extreme values were found for small population sizes. The relative precisions for the two-step approach ranged from 0.30 to 14.44 with the median of 1.1.

For simulation case B, the additional information from the past periods increased the proportion of posterior intervals that included the true abundance and the range of relative precision. Widths of posterior distributions, however, were comparable for both approaches.

For simulation case C, both approaches performed approximately the same (Figure 8). For the two-step approach, however, hyperparameters did not converge for 14 datasets. When the hyperparameters were fixed, the variability of point estimators was small ($-33 \leq \Delta_{mode} \leq 18$, $-31 \leq \Delta_{median} \leq 19$, $-30 \leq \Delta_{mean} \leq 20$; Figure 8). The two-step approach did not improve the variability of point estimators ($-35 \leq \Delta_{mode} \leq 17$, $-33 \leq \Delta_{median} \leq 18$, $-32 \leq \Delta_{mean} \leq 18$). Estimated regression parameters were comparable between the two approaches, where intercepts were approximately -1, slopes were 1, and coefficient of determinations were approximately 1.0 (Table 3).

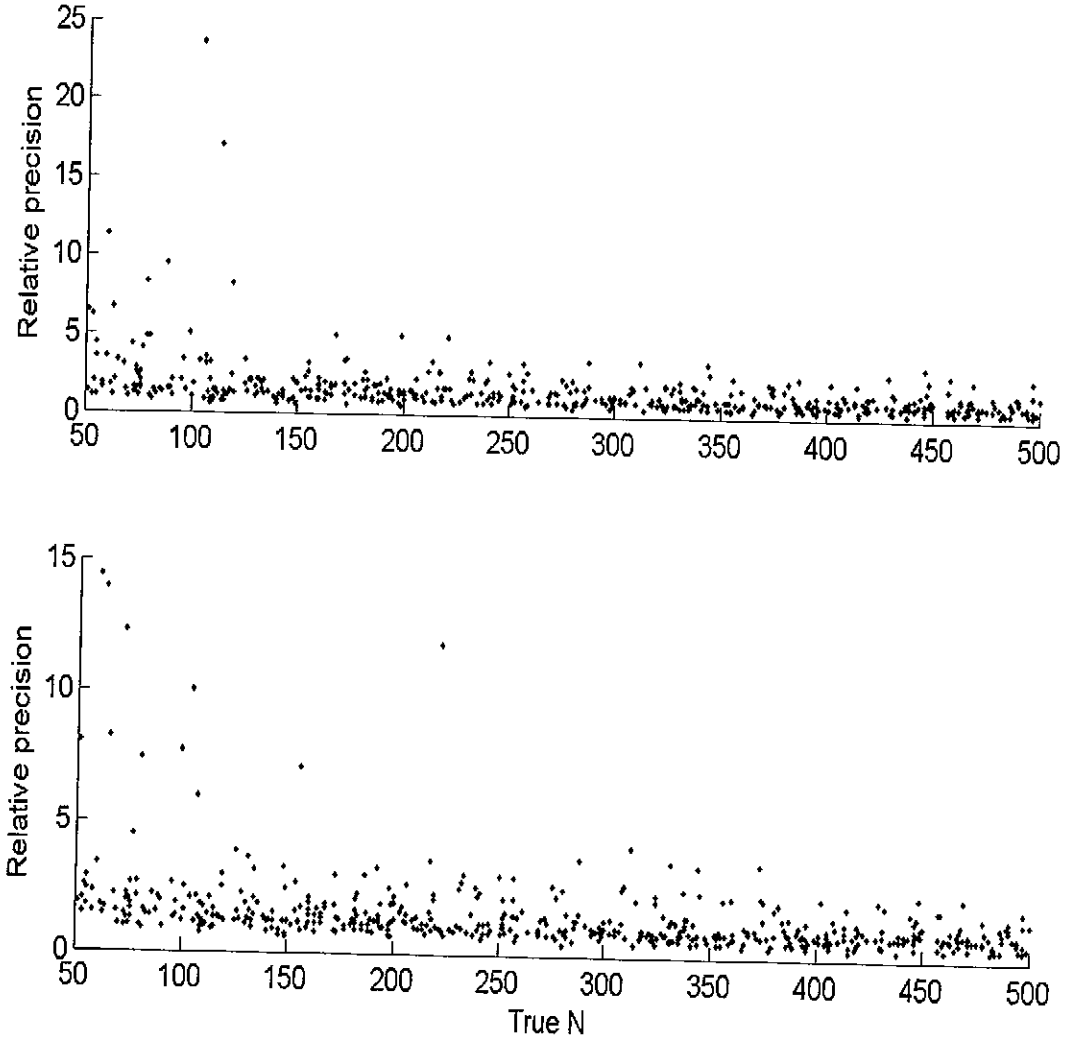

Figure 7. The relationships between the true abundance and relative precision for simulation case B, when the beta-multinomial model was used. The upper figure is the result for the constant hyperparameters (1, 1), whereas for the lower figure is for the two-step approach.

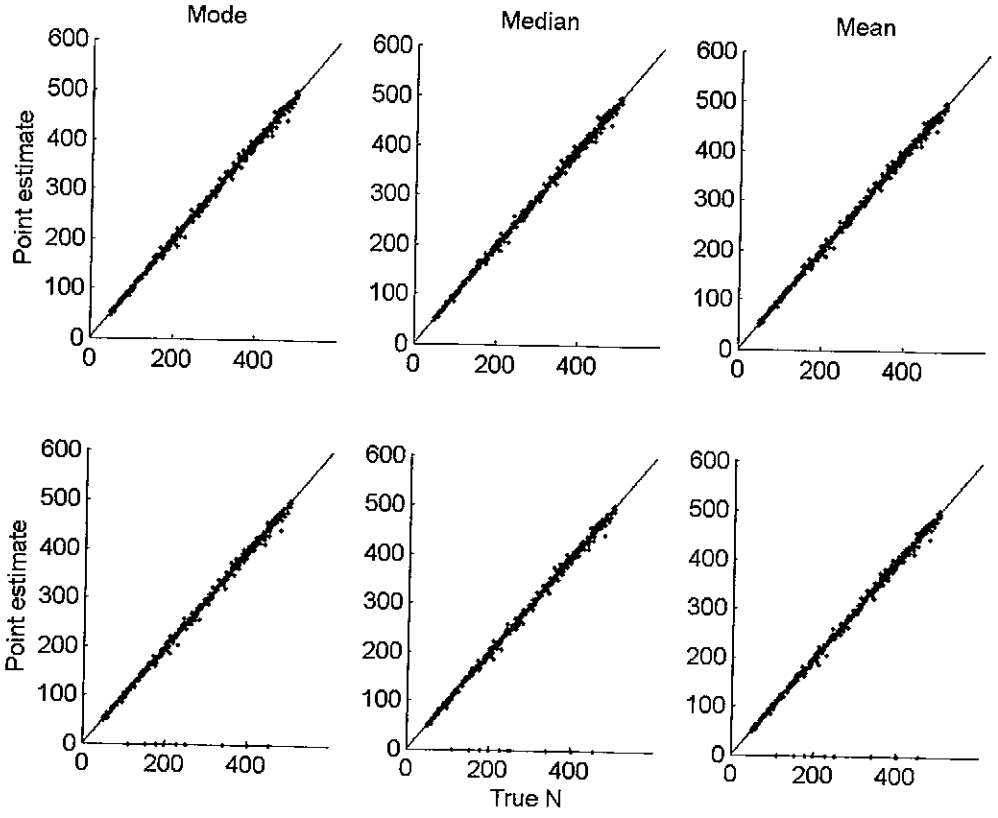

Figure 8. Relationships between point estimates and true abundance for simulation case C. The multinomial model was used for computing the posterior distributions. The top three panels are results for the fixed hyperparameters, whereas the bottom three panels are for the two-step approach.

Posterior intervals for simulation case C for both approaches were comparable. When the hyperparameters were fixed, widths of 95% posterior intervals ranged from 2 to 85, with the median of 16. Approximately 97% (483/500) of the posterior intervals included true abundances. For the two-step approach, they ranged from 3 to 76, with the median of 16. Approximately 96% (467/486) of the posterior intervals included true abundances.

Table 4. Estimated parameters for the linear regression analysis between the true abundance and the point estimates for simulation case C, using the beta-multinomial model. For intercept, slope, and r^2, values in parentheses are for the analysis when the hyperparameters were fixed.

Model	Point estimates	Intercept	Slope	r^2
Beta-multinomial	Mode	(-1.46) -1.69	(1.00) 1.00	(0.998) 0.998
	Median	(-0.63) -1.10	(1.00) 1.00	(0.998) 0.998
	Mean	(-0.22) -0.65	(1.00) 1.00	(0.998) 0.998

The relative precisions for both approaches also were comparable (Figure 9). When the hyperparameters were fixed, relative precisions ranged from 0.008 to 0.389 with a median of 0.064. For the two-step approach, they ranged from 0.01 to 0.353 with the median of 0.068.

For simulation case C, results were similar when numerical convergence was reached in the two-step approach. When the convergence was not reached, the fixed-parameter approach performed better than the two-step approach.

The proposed two-step approach provided additional information for making the inference on abundance (Table 5). For simulation case B, approximately 95% posterior distributions included the true abundance for the two-step approach, whereas 88% included the true abundance for the fixed-parameter approach. For simulation cases A and C, no apparent differences were found between the two approaches. When the hyperparameters did not converge in simulation case C, however, the fixed-parameter approach performed better than the two-step approach.

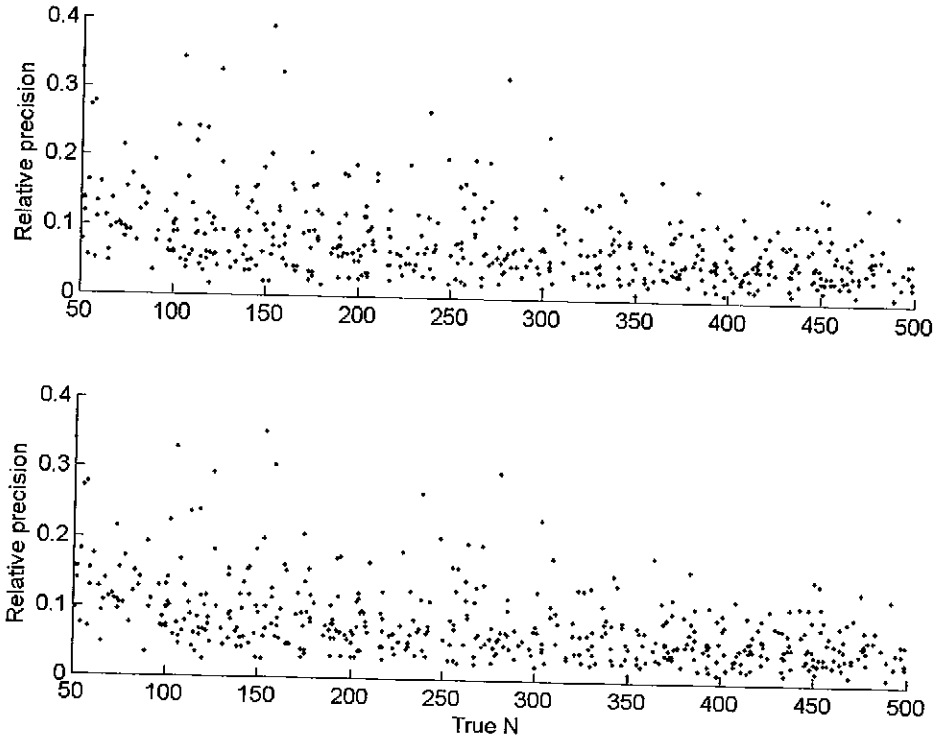

Figure 9. The relationships between the true abundance and relative precision for simulation case C when the beta-multinomial model was used. The upper figure is the result for the constant hyperparameters (1, 1), whereas for the lower figure is for the two-step approach.

Table 5. A summary of the comparison between the two-step and fixed-hyperparameter approaches. Values are minimum and maximum, whereas medians are in parentheses.

Simulation case	Fixed-hyperparameters	Two-step
A		
Widths of 95% PI	1 - 2	1 – 3
% true N captured	65%	54%
Relative precision	0.002 – 0.03 (0.004)	0.002 – 0.024 (0.004)
B		
Widths of 95% PI	11 – 1045 (215)	37 – 1208 (248)
% true N captured	88%	94%
Relative precision	0.29 – 23.7 (1.2)	0.30 – 14.44 (1.10)
C		
Widths of 95% PI	2 – 85 (16)	3 – 76 (16)
% true N captured	97%	96%
Relative precision	0.008 – 0.389 (0.064)	0.01 – 0.353 (0.068)

Effects of capture probabilities on the posterior distribution of abundance

To assess the effects of true capture probabilities on posterior distributions of abundance, results from the two-step approach were compared among simulation cases A, B, and C. Posterior distributions of abundance from the proposed method also were compared to results from MARK.

With consistently high capture probabilities (case A), estimators from both models were precise (Figure 10). For the beta-binomial model, the variability of point estimators was small ($-50 \leq \Delta_{mode} \leq 33$, $-40 \leq \Delta_{median} \leq 35$, $-37 \leq \Delta_{mean} \leq 38$). Approximately 20% (102/500) of modes, 46% (228/500) of medians, and 59% (294/500) of means were greater than the true abundance.

Slopes of the regression lines between point estimators and true abundance were approximately one and intercepts were near zero (Table 6). Widths of posterior intervals ranged from 4 to 130, with a median of 37. Approximately 98% (491/500) of posterior intervals included the true abundance.

Table 6. Estimated parameters for the linear regression analysis between the true abundance and the point estimates for simulation case A, using the two-step approach.

Model	Point estimates	Intercept	Slope	r^2
Beta-binomial	Mode	0.22	0.98	0.993
	Median	0.42	1.00	0.995
	Mean	0.75	1.00	0.995
Multinomial	Mode	0.92	0.99	0.999
	Median	0.92	0.99	0.999
	Mean	0.92	0.99	0.999
Closed capture	MLE	-0.53	1.00	1.000
Huggins closed	MLE	-0.53	1.00	1.000

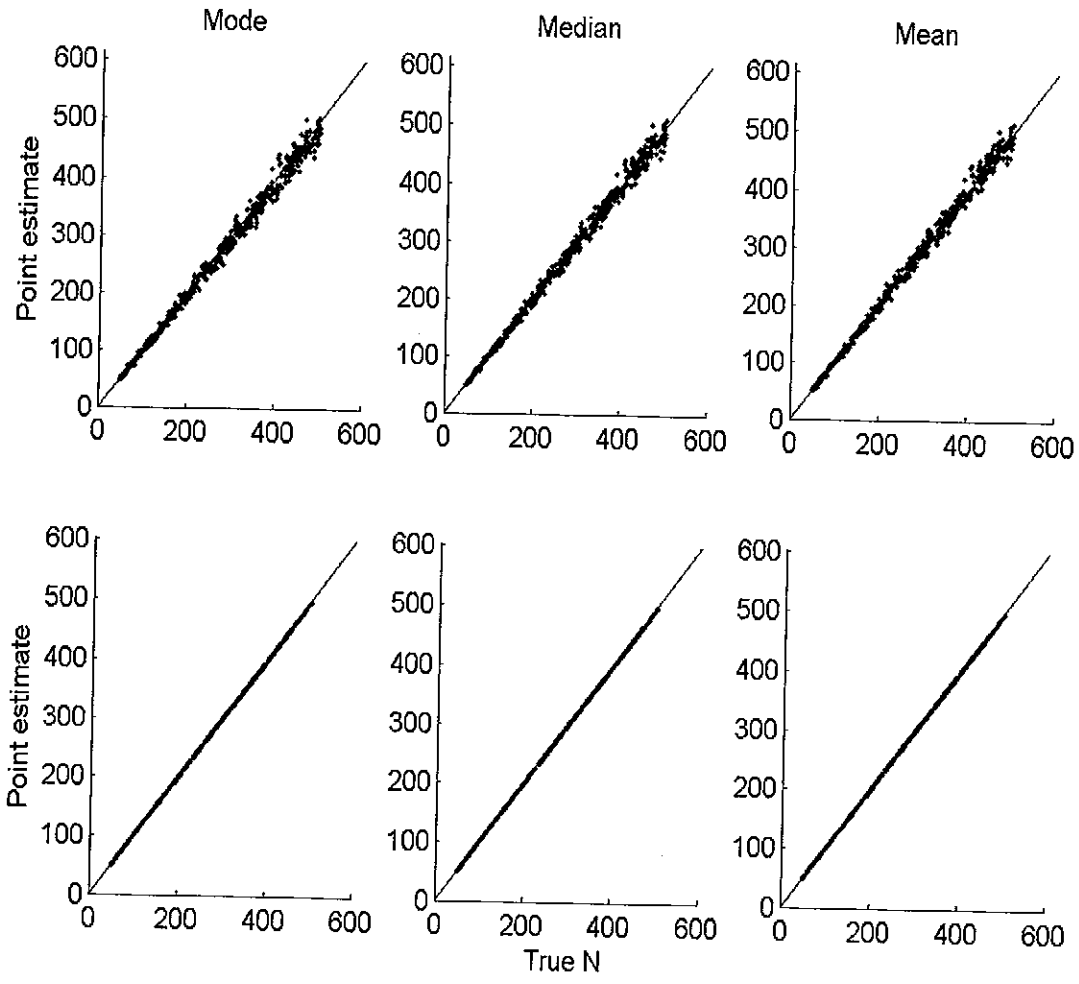

Figure 10. Relationships between three point estimates and the true population size for simulation case A. Solid lines indicate the 45-degree lines. The upper three panels are for the beta-binomial model whereas the lower three panels are for the multinomial model.

The relative precision was uncorrelated with the true abundance and ranged from 0.06 to 0.3 with the median of 0.14 (Figure 11). Results for the beta-multinomial model were discussed in the previous section.

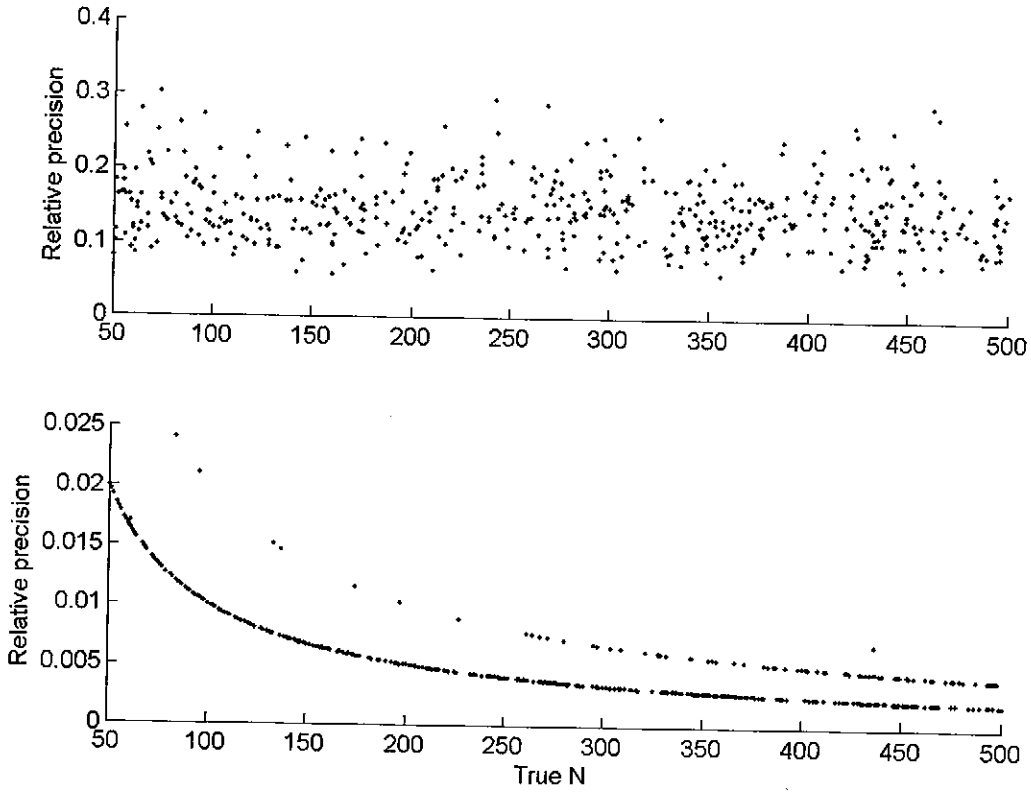

Figure 11. The relationship between the relative precision and the true population size for simulation case A. The upper panel is for the beta-binomial model whereas the lower panel is for the multinomial model.

When datasets for simulation case A were analyzed using MARK, MLE's underestimated the true abundance. For both models, all estimates were less than true abundances (Figure 12). Consequently, the intercepts of regression lines were negative (Table 6). For the closed capture model, widths of 95% confidence intervals ranged from 0 to 0.004, with the median of 0.00. None of the confidence intervals, however, included true abundances. Widths of confidence intervals were positively correlated with the true abundance. The relative precision of MLE was 0.00 for all results because of the extremely narrow confidence intervals. For the Huggins closed model, widths of 95%

confidence intervals ranged from 0 to 3.1, with a median of 0.002 (Figure 12). Only 12% (60/500) of confidence intervals included true abundances. The relative precision of MLE ranged from 0.00 to 0.027 with the median of 0.002.

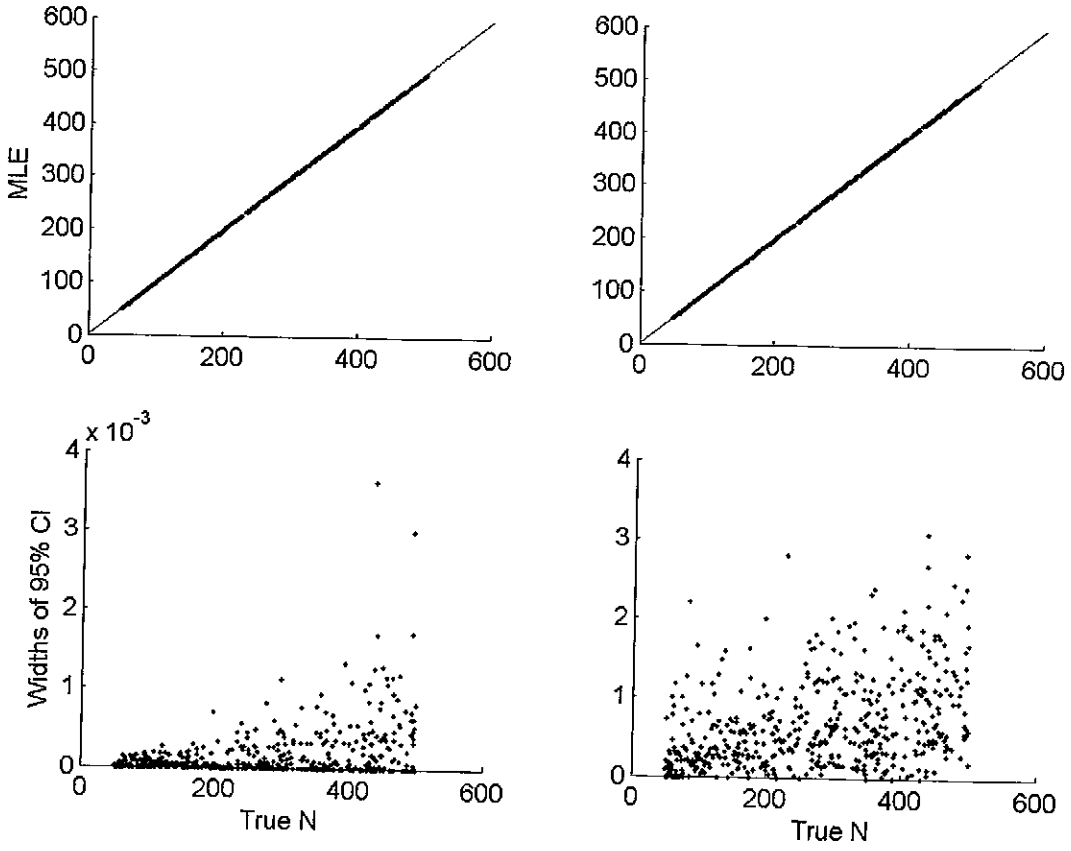

Figure 12. Relationships between the true abundance and maximum likelihood estimates (upper figures) and widths of 95% confidence intervals (lower figures) from MARK for simulation case A. The solid line indicates the 45-degree line. Left panels are for the closed capture model, whereas right panels are for the Huggins closed model.

When datasets for simulation case B were analyzed using the proposed method, point estimators were less precise than for simulation case A (Figure 13). For the beta-binomial model, point estimators were imprecise and inaccurate ($-422 \leq \Delta_{mode} \leq 1029$, -

$315 \leq \Delta_{median} \leq 751$, $-209 \leq \Delta_{mean} \leq 707$). Approximately 28% (138/500) of modes, 60% (299/500) of medians, and 74% (370/500) of means were greater than the true abundances. Consequently, the estimated intercepts of regression lines between true abundance and medians and means were positive.

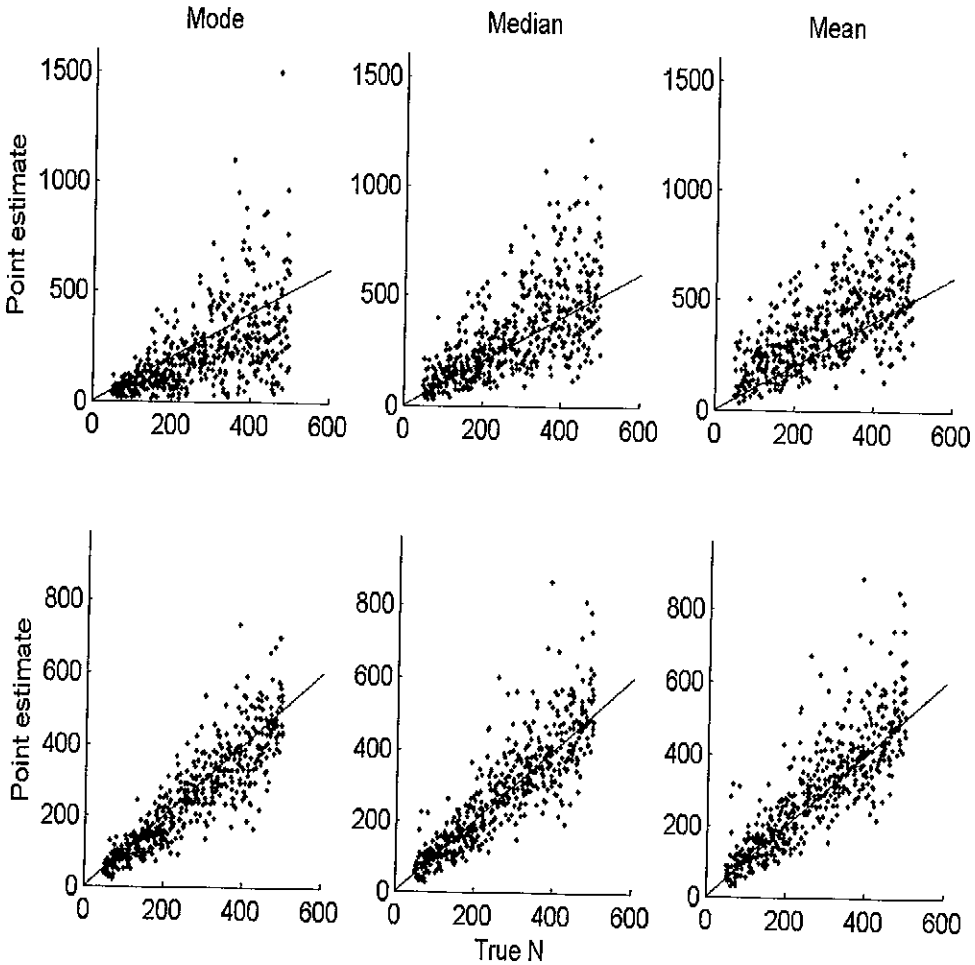

Figure 13. Relationships between three point estimates and the true population size for simulation case B. Solid lines indicate the 45-degree lines. Upper three panels represent the results for the beta-binomial model, whereas lower three panels represent the results for the multinomial model

Table 7. Estimated parameters for the linear regression analysis between the true abundance and the point estimates for simulation case B.

Model	Point estimates	Intercept	Slope	r^2
Beta-binomial	Mode	3.12	0.82	0.335
	Median	26.27	1.07	0.454
	Mean	90.73	1.06	0.457
Multinomial	Mode	-9.03	0.98	0.781
	Median	4.36	1.02	0.741
	Mean	17.05	1.03	0.707
Closed Captures	MLE	-14.63	1.17	0.404
Huggins Closed	MLE	3.37	1.10	0.518

The estimated slopes of the regression lines were approximately 1. Coefficient of determinations were approximately 0.4 (Table 7). Widths of the posterior intervals ranged from 54 to 1374, with the median of 1010. Approximately 99% (493/500) of the posterior intervals included the true abundance. The relative precision ranged from 0.6 to 90.1 with a median of 4.3 (Figure 14). The combination of low capture probabilities and few capture occasions resulted in the small number of recaptures within each primary period. Consequently, posterior distributions were imprecise. Results for the beta-multinomial model were discussed in the previous section.

When datasets for simulation case B were analyzed using MARK, unreasonable estimates were obtained for 11 datasets in the closed capture model and 12 datasets in the Huggins closed capture model. Maximum likelihood estimates for these datasets were greater than 10,000. These estimates, therefore, were deleted from the following analyses. For both models, maximum likelihood estimates were similar to those from the two-step multinomial model (Figure 15).

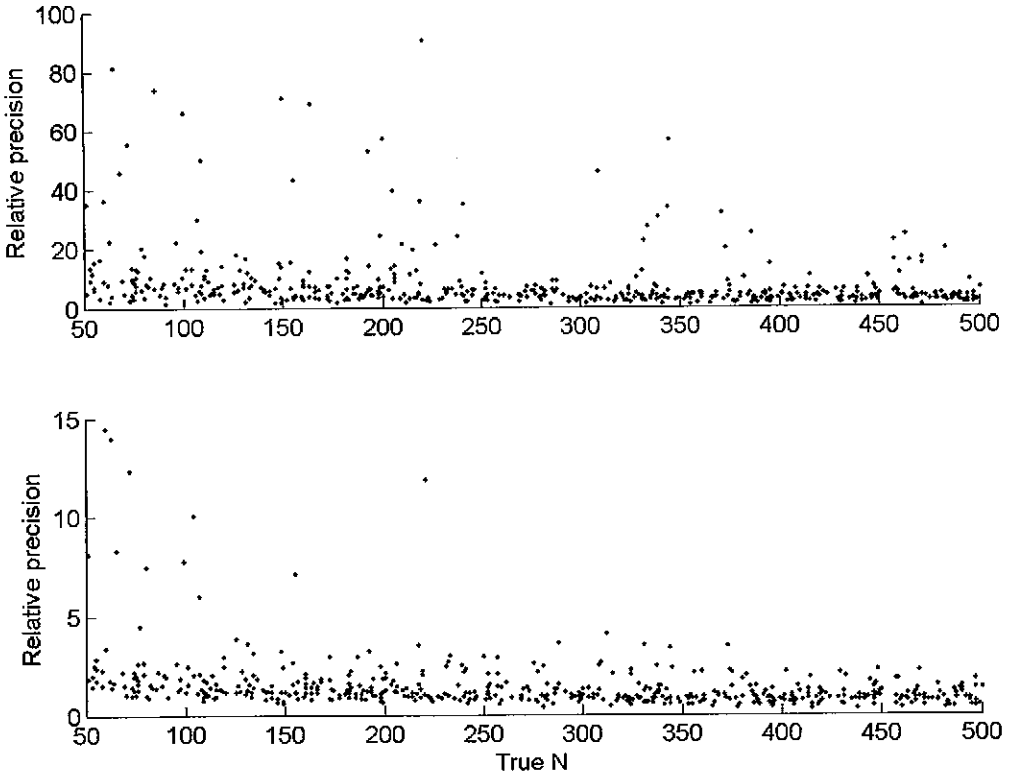

Figure 14. The relationship between relative precision and true population size for simulation case B. Upper panel is for the beta-binomial model, whereas the lower panel is for the multinomial model.

For the closed capture model, large estimates were found in several datasets ($-191.1 < \Delta_{mle} < 2831.1$; Figure 15). Approximately 48% (235/489) of estimates were greater than true abundances. The estimated intercept of the regression line was –14.6, the slope was approximately 1, and the coefficient of determination was approximately 0.4 (Table 7). Widths of 95% confidence intervals ranged from approximately 41 to 16146, with a median of 290 (Figure 15). True abundances were included in approximately 94% (463/489) of the confidence intervals. The relative precision of MLE ranged from 0.29 to 4.88 with the median of 1.16.

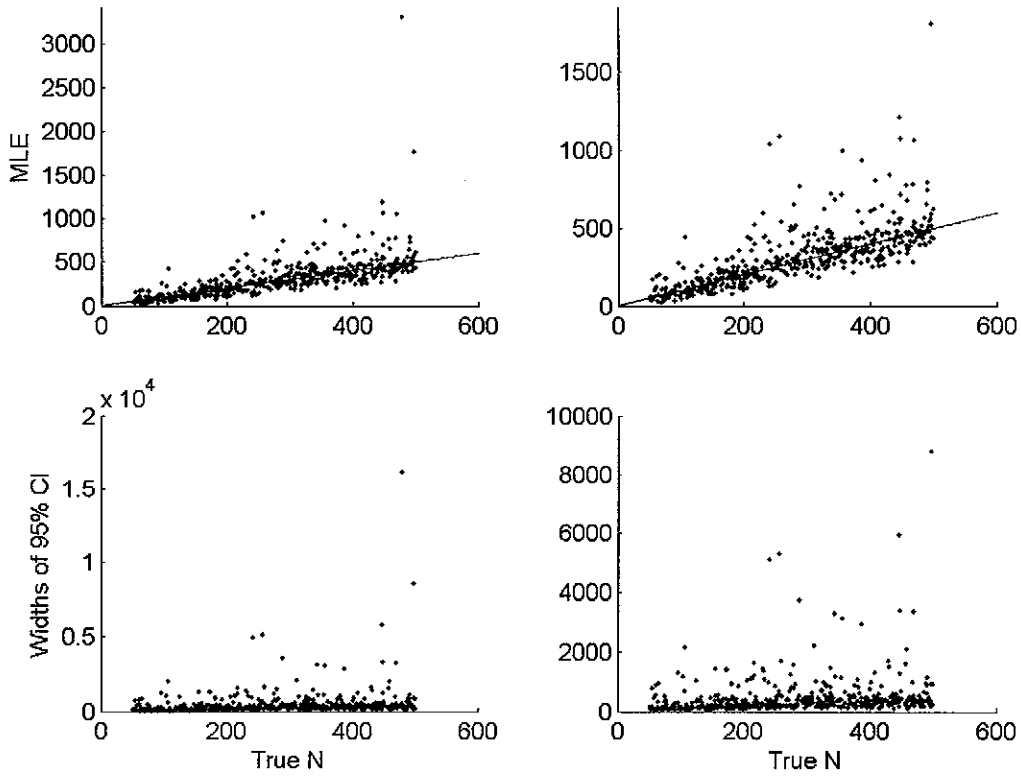

Figure 15. Relationships between true abundance and maximum likelihood estimates and widths of 95% confidence intervals from MARK for simulation case B. The solid line indicates the 45-degree line. Left panels are for closed capture model, whereas right panels are for Huggins closed capture model.

For the Huggins closed model, large estimates were found in several datasets ($-186.4 < \Delta_{mle} < 1315.3$; Figure 15). Approximately 53% (258/488) of MLE's were greater than the true abundance. The estimated intercept of the regression line was 3.4, the slope was 1.1, and the coefficient of determination was approximately 0.52 (Table 7). Widths of 95% confidence intervals ranged from approximately 43 to 8781, with a median of 296 (Figure 15). True abundances were included in approximately 95% (465/488) of

confidence intervals. The relative precision of the MLE's ranged from 0.29 to 4.85 with a median of 1.16.

For simulation case C, posterior distributions of the hyperparameters did not converge in 14 datasets. For these 14 datasets, no posterior distributions of abundances were computed. For the beta-binomial model, the point estimators were variable ($-452 \leq \Delta_{mode} \leq 254$, $-197 \leq \Delta_{median} \leq 323$, $-462 \leq \Delta_{mean} \leq 371$; Figure 16). The variability was positively correlated with the true abundance (Figure 16). Approximately 32% (156/486) of modes, 57% (278/486) of medians, and 71% (345/486) of means were greater than the true abundances.

Although the differences between the point estimates and true abundances were large, the slopes of regression lines were approximately 1.0 (Table 8). The coefficient of determinations ranged between 0.65 and 0.78 (Table 8). Widths of 95% posterior intervals ranged from 38 to 1066, with a median of 339. Approximately 97% (470/486) of the posterior intervals included the true abundance. Relative precision for the beta-binomial model was uncorrelated with the true abundance and ranged from 0.46 to 6.22 with the median of 1.42 (Figure 17). Results for the beta-multinomial model were discussed in the previous section.

Table 8. Estimated parameters for the linear regression analysis between the true abundance and the point estimates for simulation case C, using the hierarchical approach.

Model	Point estimates	Intercept	Slope	r^2
Beta-binomial	Mode	-4.36	0.91	0.654
	Median	-2.42	1.07	0.779
	Mean	0.68	1.12	0.693
Multinomial	Mode	2.75	0.98	0.964
	Median	3.36	0.98	0.964
	Mean	3.81	0.98	0.964
Closed captures	MLE	-1.33	1.00	0.998
Huggins closed	MLE	-0.62	1.00	0.998

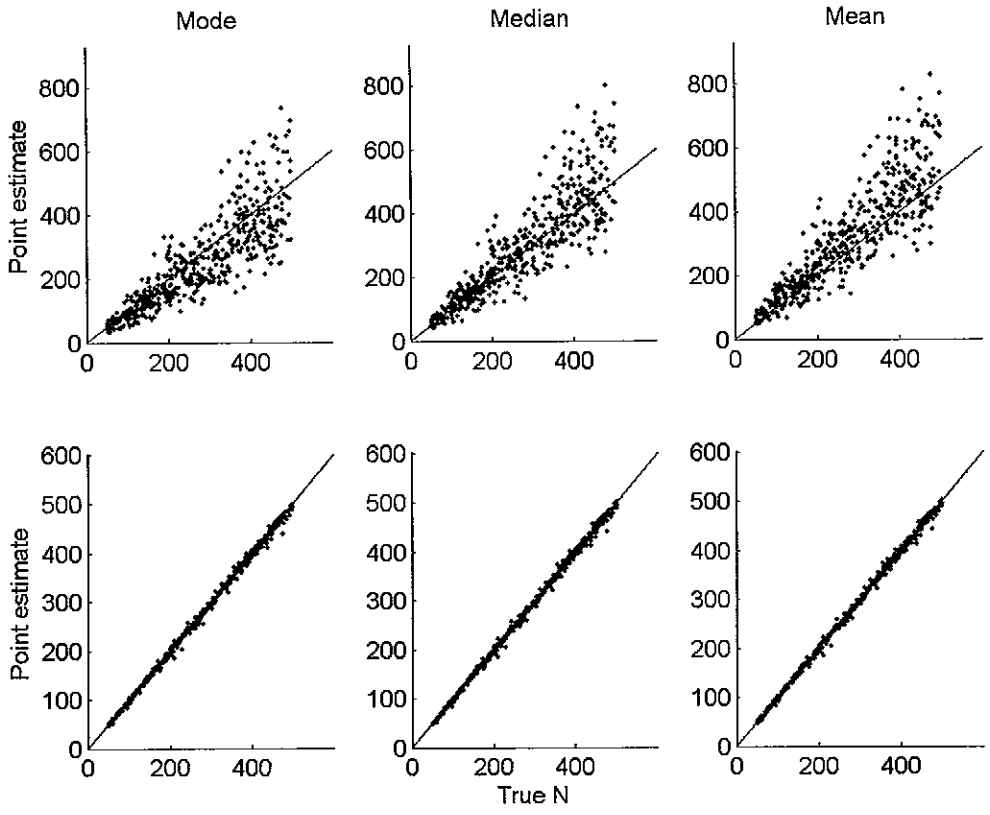

Figure 16. Relationships between true population size and three point estimates for simulation case C. The upper three panels are for the beta-binomial model, whereas the lower three panels are for the multinomial model. The solid lines indicate the 45-degree lines.

The numerical convergence of hyperparameters was critical for obtaining precise estimates for simulation case C. Because of the assumed broad distribution of the true hyperdistribution of capture probabilities, combined with infrequent sampling, not enough information was available for computing joint posterior distributions for the hyperparameters in some data sets. The diffuse prior distribution of the hyperparameters dominated the posterior distribution of the hyperparameters.

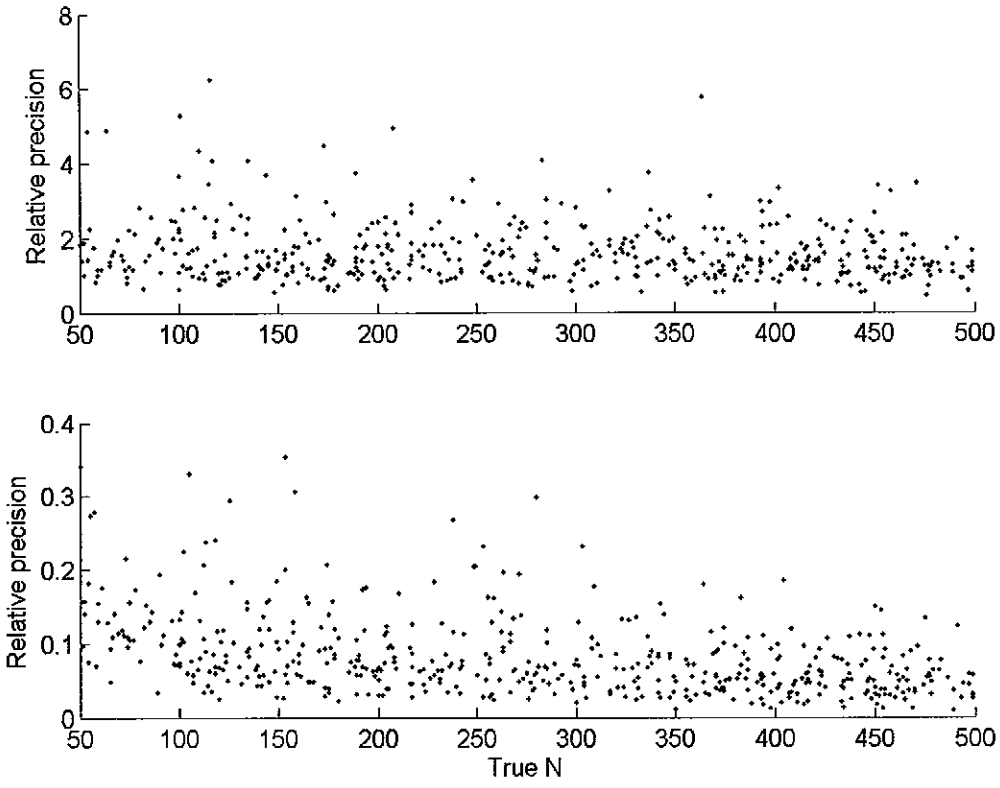

Figure 17. Relative precision as a function of true population size for simulation case C. Upper panel is for the beta-binomial model, whereas the lower panel is for the multinomial model.

When the same datasets for simulation case C were analyzed using MARK, maximum likelihood estimators were accurate for both models (Figure 18). For the closed capture model, variability of the estimator was similar to that of the multinomial model ($-31.2 \leq \Delta_{mle} \leq 18.68$). Approximately 33% (165/500) of estimates were greater than true population sizes. The estimated intercept of the regression was –1.0, the slope was one, and the coefficient of determination was greater than 0.99 (Table 8). Widths of 95% confidence ranged from 0.0 to 86.7, with a median of 15.5, and 476 (95%) included

the true abundance. The relative precision of MLE ranged from 0.00 to 0.39 with the median of 0.07.

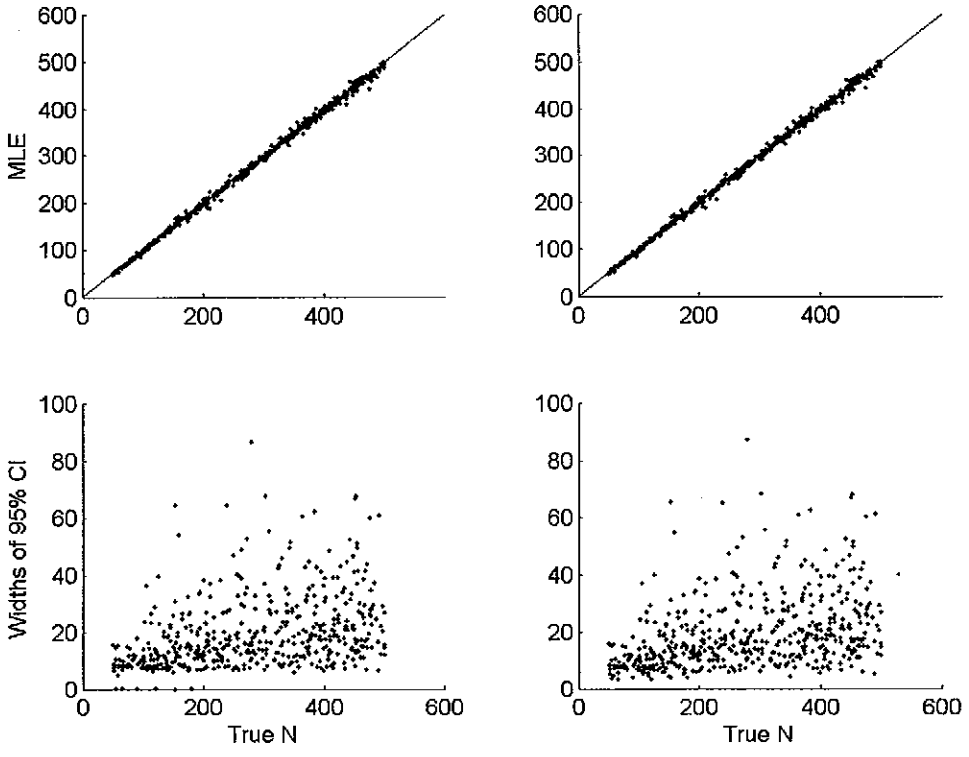

Figure 18. Relationships between true abundance and MLE and the widths of 95% confidence intervals for simulation set C. Left panels are for the closed capture model, whereas the right panels are for the Huggins closed model. The solid line indicates the 45-degree line.

For the Huggins closed model, the variability in the estimator was similar to that of the closed model and the multinomial model ($-30.2 \leq \Delta_{mle} \leq 19.6$). Approximately 39% (193/500) of estimates were greater than true abundances. The estimated intercept was –0.62, the slope was one, and the coefficient of determination was greater than 0.99 (Table 8). Widths of 95% confidence intervals ranged from 3.3 to 87.4 with the median of 15.7. Approximately 96% (478/500) of confidence intervals included the true

abundance. The relative precision of MLE ranged from 0.01 to 0.39 with a median of 0.07.

Results for the three simulation cases (A, B, and C) indicated that greater capture probabilities resulted in more precise estimates of abundance. The comparison among these simulation cases indicated that the average capture probability affected the precision of posterior distributions more than the variance of capture probabilities for both models. When capture probabilities were variable, however, the joint posterior distributions of hyperparameters did not converge in some datasets.

In all simulation cases, results from the proposed method were comparable to those from MARK, except for simulation case A (Table 9). For simulation case A, posterior intervals of the Bayes hierarchical beta-multinomial model and confidence intervals of the two maximum likelihood methods were narrow (0 to 3). Consequently, true abundances were included only in a fraction of the confidence intervals (12% of HCC and 0 for CC) and the half (54%) of the posterior intervals from the Bayes hierarchical beta-multinomial model. For the Bayes hierarchical beta-binomial model, however, posterior intervals were wider (4 – 130) and true abundances were included in 98% of the posterior intervals. For simulation case B, the hierarchical beta-multinomial model resulted in the narrowest intervals around point estimates and the smallest relative precision. For simulation case C, widths of posterior and confidence intervals and the proportion of true abundances that were included in the intervals were similar for the maximum likelihood methods and the Bayes hierarchical beta-multinomial model. The Bayes approach, however, provided the most precise estimates (Table 9).

Table 9. A summary of the comparison among the hierarchical beta-binomial, the hierarchical beta-multinomial, closed captures, and Huggins closed captures models. Minimum, maximum, and median (in parentheses) are reported.

Simulation case		Beta-Binomial	Beta-Multinomial	Closed capture	Huggins closed capture
A					
	Widths	4 – 130 (37)	1 – 3 (1)	0.0 – 0.004 (0.0)	0.0 – 3.1 (0.51)
	% true N	98%	54%	0%	12%
	Precision	0.06 – 0.3 (0.14)	0.002 – 0.024 (0.004)	0.00 – 0.00 (0.00)	0.00 – 0.27 (0.002)
B					
	Widths	54 – 1,374 (1,010)	37 – 1,208 (248)	41 – 16,146 (290)	43 – 8,781 (296)
	% true N	99%	94%	94%	95%
	Precision	0.6 - 90.1 (4.3)	0.30 – 14.44 (1.10)	0.29 – 4.88 (1.16)	0.29 – 4.85 (1.16)
C					
	Widths	38 – 1,066 (339)	3 – 76 (16)	0.0 – 86.7 (15.5)	3.3 – 87.4 (15.7)
	% true N	97%	96%	95%	96%
	Precision	0.46 – 6.22 (1.42)	0.01 – 0.353 (0.068)	0.29 – 4.88 (1.16)	0.29 – 4.85 (1.16)

Effects of the number of primary periods on the posterior distribution of abundance

To determine the effects of the number of primary periods on the precision of posterior distributions for abundance, posterior distributions for three simulation cases were compared. The number of primary periods was increased from three (case B), to six (case D) and twelve (case E), while keeping the number of secondary occasions and the true hyperparameters constant (Table 1). No analyses were conducted using MARK in this analysis because the number of primary periods did not affect the estimates for the

last primary period when the closed capture or the Huggins closed capture model was used.

For simulation case D, point estimators were similar to those for simulation case B (Figure 13 vs. Figure 19). For the beta-binomial model, the variability of point estimators was large ($-349 \leq \Delta_{mode} \leq 673$, $-327 \leq \Delta_{median} \leq 648$, $-298 \leq \Delta_{mean} \leq 632$) and increased with the true abundance (Figure 19). The variability, however, was less than that for simulation case B (Figure 13). Approximately 35% (348/500) of modes, 60% (301/500) of medians, and 74% (372/500) of means were greater than the true population sizes.

Estimated intercepts of the regression lines were positive, the slopes were approximately one, and coefficient of determinations were approximately 0.5 (Table 10). Widths of the 95% posterior intervals ranged from 43 to 1309, with a median of 836. True abundance was included in 96% (480/500) of posterior intervals. The relative precision ranged from 0.84 to 48.6 (median = 3.1; Figure 20).

Table 10. Estimated parameters for the linear regression analysis between the true abundance and the point estimates for simulation case D, using the hierarchical approach.

Model	Point estimates	Intercept	Slope	r^2
Beta-binomial	Mode	6.57	0.91	0.486
	Median	26.57	1.08	0.508
	Mean	65.49	1.12	0.510
Multinomial	Mode	-5.89	0.95	0.784
	Median	4.40	0.99	0.773
	Mean	12.77	1.00	0.758

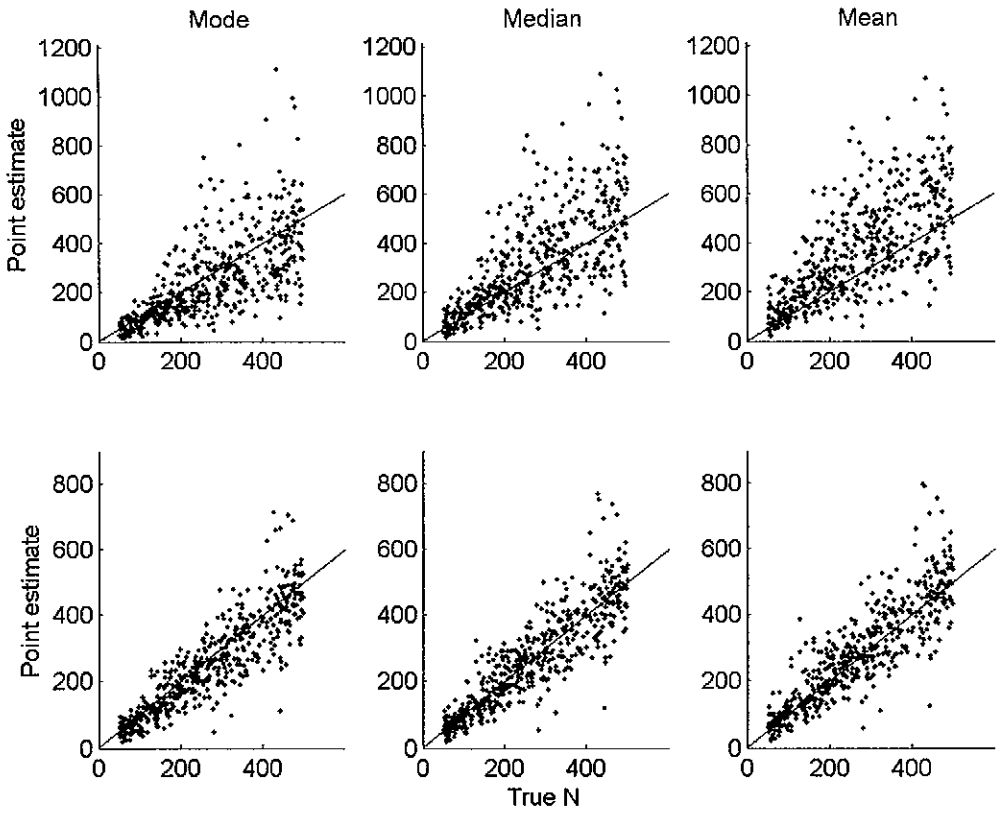

Figure 19. Relationships between true abundances and point estimates for simulation case D. The upper three panels are for the beta-binomial model whereas the lower three panels are for the multinomial model.

For the multinomial model, the variability in the point estimators was less than that for the beta-binomial model ($-331 \leq \Delta_{mode} \leq 286$, $-322 \leq \Delta_{median} \leq 342$, $-316 \leq \Delta_{mean} \leq 371$; Figure 19). The variability was less than that of simulation case B (Figure 13). Approximately 35% (173/500) of modes, 48% (240/500) of medians, and 55% (277/500) of means were greater than the true population sizes.

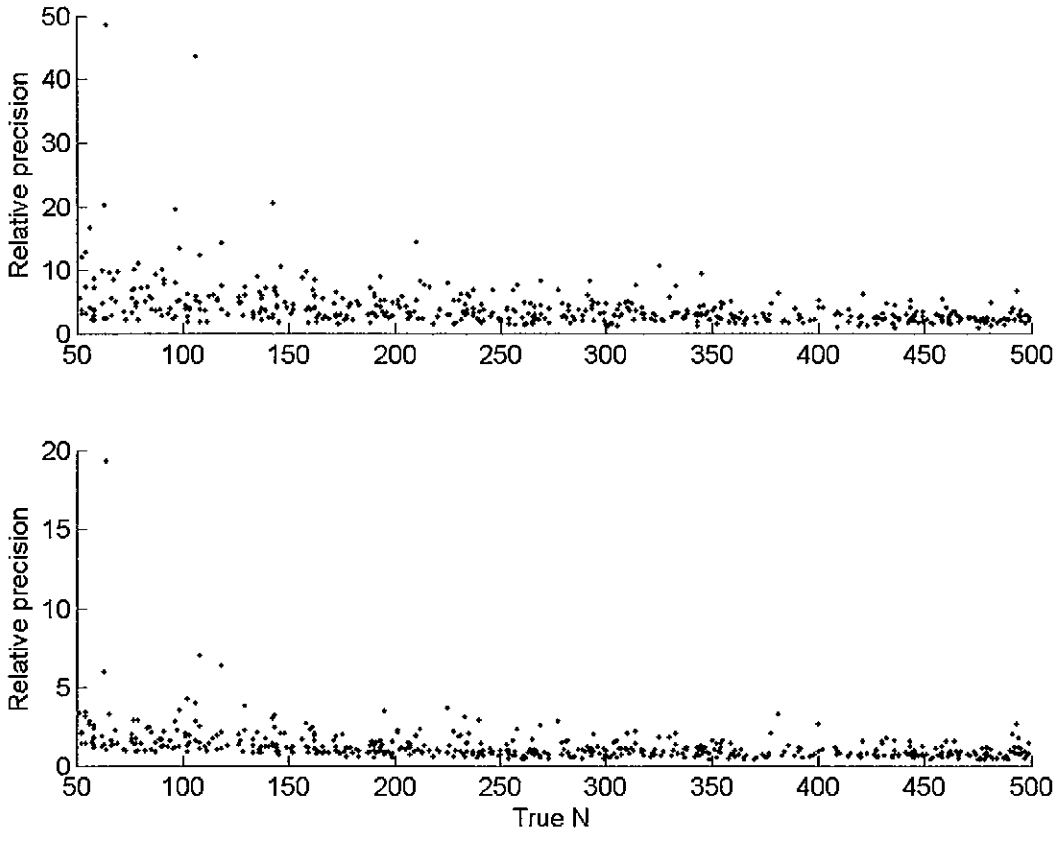

Figure 20. The relationship between relative precision and true abundance for simulation case D. The upper panel is for the beta-binomial model whereas the lower panel is for the multinomial model.

The estimated intercepts of the regression analyses were positive for medians and means, whereas it was negative for the mode. The slopes of regression lines were approximately one and the coefficients of determinations were between 0.76 and 0.78 (Table 10). Widths of 95% posterior intervals ranged from 32 to 959, with a median of 223. Approximately 92% (460/500) of the posterior intervals included the true abundance. The relative precision ranged from 0.39 to 19.3 (median = 1.0; Figure 20).

As the number of primary periods was increased to 12 (simulation case E), the variability in point estimators decreased from simulation cases B and D. For the beta-binomial model, the variability in point estimators increased with the true abundance ($-275 \leq \Delta_{mode} \leq 502$, $-238 \leq \Delta_{median} \leq 564$, $-233 \leq \Delta_{mean} \leq 573$; Figure 21). The variability, however, was less than for simulation case D (Figure 19). Approximately 42% (212/500) of modes, 67% (334/500) of medians, and 78% (390/500) of means were greater than the true population sizes.

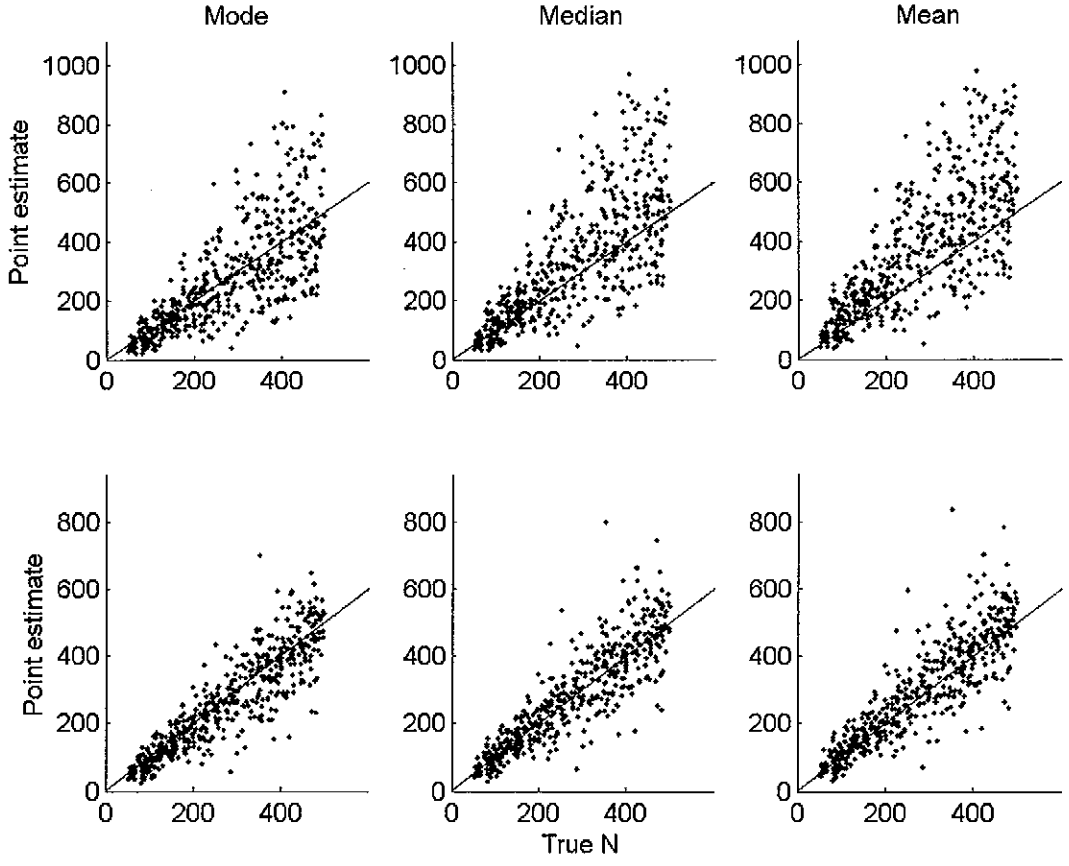

Figure 21. Relationships between the true population size and three point estimates for simulation case E. The upper three panels are for the beta-binomial model, whereas the lower three panels are for the multinomial model.

Estimated intercepts of the regression analyses were positive for medians and means, whereas it was negative for the mode. Estimated slopes of the regression lines were approximately one and coefficients of determinations were approximately 0.57 (Table 10). Widths of 95% posterior intervals ranged from 83 to 1236, with a median of 671. Approximately 95% (477/500) of the posterior intervals included the true abundance. The relative precision ranged from 1.05 to 27.9 (median = 2.5; Figure 22). The outlier in the relative precision (27.9) was the result of recapturing only one individual in the primary period when the true population was small (N = 65). The outlier was caused by the wide posterior interval for the dataset, when the mode was 27. For the remainder of datasets, the relative precision was less than 10.

For the multinomial model, the variability in point estimators for simulation case E was less than that for case D ($-260 \leq \Delta_{mode} \leq 347$, $-242 \leq \Delta_{median} \leq 447$, $-234 \leq \Delta_{mean} \leq 484$; Figure 21). Approximately 41% (205/500) of modes, 54% (272/500) of medians, and 59% (297/500) of means were greater than the true population size. Estimated intercepts of the regression analyses were positive for medians and means, whereas it was negative for the mode.

Table 11. Estimated parameters for the linear regression analysis between the true abundance and the point estimates for simulation case E, using the hierarchical approach.

Model	Point estimates	Intercept	Slope	r^2
Beta-binomial	Mode	-1.76	1.01	0.566
	Median	12.57	1.16	0.569
	Mean	36.51	1.22	0.576
Multinomial	Mode	-7.47	0.99	0.795
	Median	2.37	1.02	0.782
	Mean	9.32	1.03	0.771

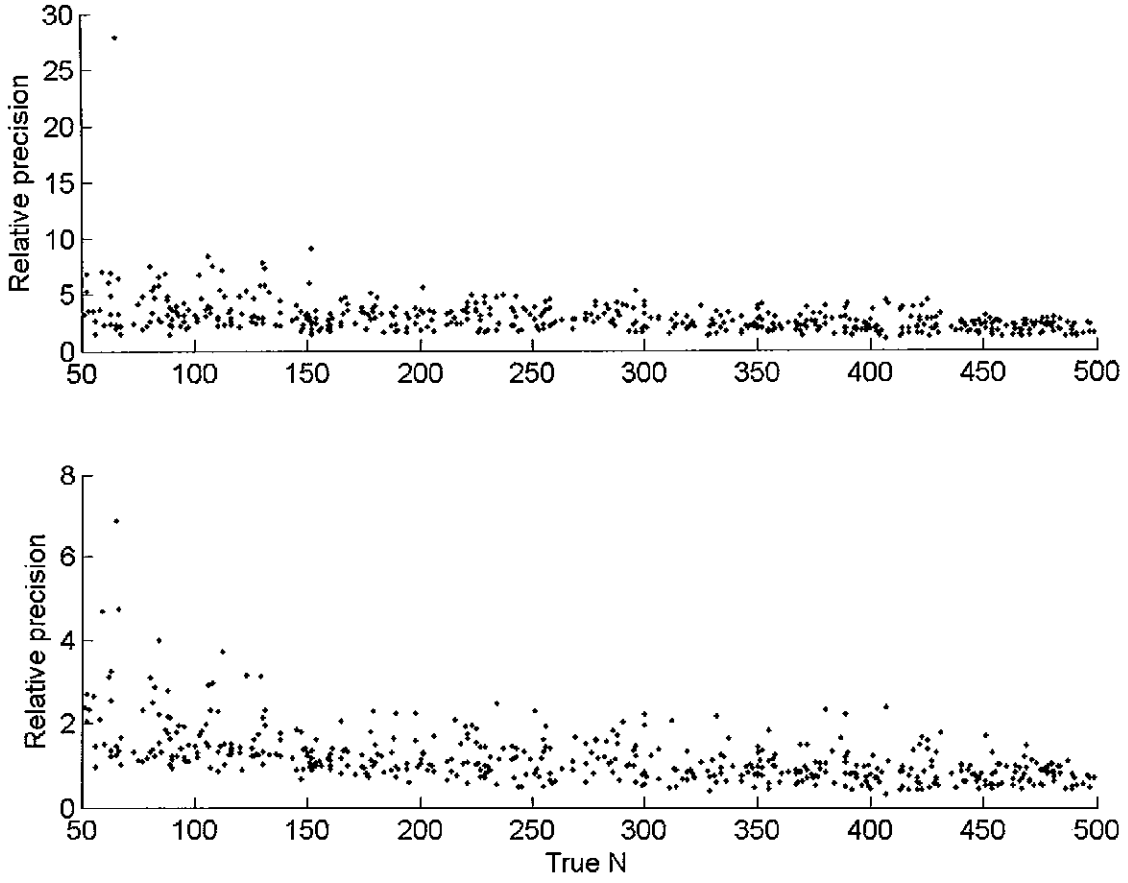

Figure 22. Relationships between the relative precision and the true abundance for simulation case E. The upper figure is for the beta-binomial model, whereas the lower figure is for the multinomial model.

Estimated slopes of the regression lines were approximately one and the coefficients of determinations were approximately 0.8 (Table 11). Widths of 95% posterior intervals ranged from 55 to 1035, with a median of 233. Approximately 93% (466/500) of the posterior intervals included the true abundance. The relative precision ranged from 0.33 to 6.88 (median = 1.01). Although the relative precision improved as the number of primary periods increased from three (Figure 14) to six and twelve (Figure

20 and Figure 22), no prominent improvement in the precision of posterior distributions of abundance was found when the number of primary periods increased.

Table 12. A summary of the comparison among the hierarchical beta-binomial, the hierarchical beta-multinomial, closed captures, and Huggins closed captures models. Values are minimum and maximum, whereas medians are in parentheses.

Simulation case		Beta-Binomial	Beta-Multinomial
B			
	Widths	54 – 1374 (1010)	37 – 1208 (248)
	% true N	99%	94%
	Precision	0.6 - 90.1 (4.3)	0.30 – 14.44 (1.10)
D			
	Widths	43 – 1,309 (836)	32 – 959 (223)
	% true N	96%	92%
	Precision	0.84 – 48.6 (3.1)	0.39 – 19.3 (1.0)
E			
	Widths	83 – 1,236 (671)	55 – 1,035 (233)
	% true N	95%	93%
	Precision	1305 – 27.9 (2.5)	0.33 – 6.88 (1.01)

Effects of the number of secondary occasions on the posterior distribution of abundance

To determine the effects of the number of secondary occasions on the precision of posterior distributions of abundance, posterior distributions from three simulation cases were compared. In these cases, the number of secondary occasions was increased from four (case B) to six (case F) and eight (case G), while keeping the number of primary

periods and the true hyperparameters constant (Table 1). These datasets also were analyzed using MARK.

For simulation case F, results from the beta-binomial model indicated the variability in the point estimators was large and increased with the true abundance ($-306 \leq \Delta_{mode} \leq 787$, $-291 \leq \Delta_{median} \leq 778$, $-281 \leq \Delta_{mean} \leq 766$; Figure 23). The variability, however, was less than for simulation case B (Figure 13). Approximately 35% (174/500) of modes, 56% (280/500) of medians, 70% (351/500) of means were greater than the true abundance.

Table 13. Estimated parameters for the linear regression analysis between the true abundance and the point estimates for simulation case F, using the hierarchical approach.

Model	Point estimates	Intercept	Slope	r^2
Beta-binomial	Mode	-3.17	0.96	0.542
	Median	6.71	1.11	0.558
	Mean	27.29	1.17	0.567
Multinomial	Mode	-1.32	0.97	0.904
	Median	5.65	0.98	0.900
	Mean	10.02	0.98	0.895
Closed capture	MLE	9.37	0.98	0.865
Huggins closed	MLE	12.53	0.98	0.861

The estimated intercepts of the regression lines were positive for medians and means, whereas it was negative for the mode (Table 13). Estimated slopes of the regression lines were approximately one and the coefficients of determinations were approximately 0.55. Widths of 95% posterior intervals ranged from 58 to 1272, with the median of 628. Approximately 97% (486/500) of the posterior intervals included the true abundance. The relative precision ranged from 0.76 to 17.16 with a median of 2.57.

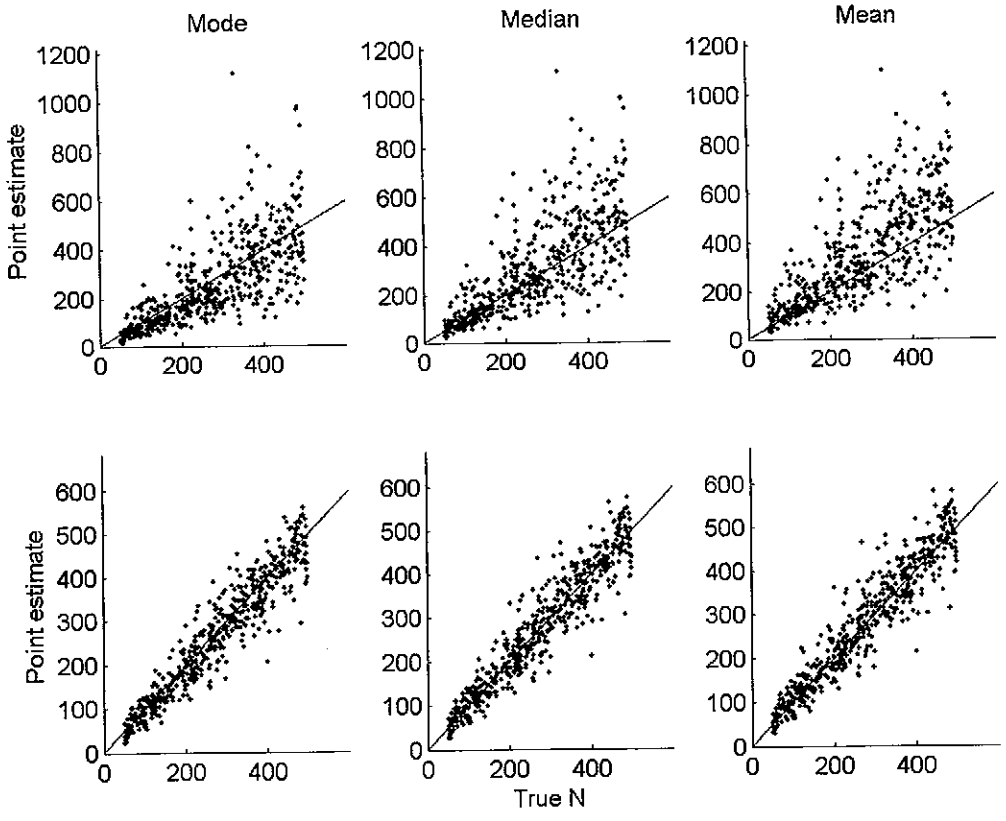

Figure 23. Relationships between the true abundance and three point estimates in simulation case F. The solid lines indicate 45-degree lines. Upper three panels were for the beta-binomial model, whereas the lower three panels were for the multinomial model.

For the beta-multinomial model, the variability in the point estimators was less than for the beta-binomial model ($-195 \leq \Delta_{mode} \leq 127$, $-189 \leq \Delta_{median} \leq 168$, $-186 \leq \Delta_{mean} \leq 196$; Figure 23). The variability, also, was less than for simulation case B (Figure 13). Approximately 41% (204/500) of modes, 50% (250/500) of medians, and 55% (276/500) of means were greater than true abundances.

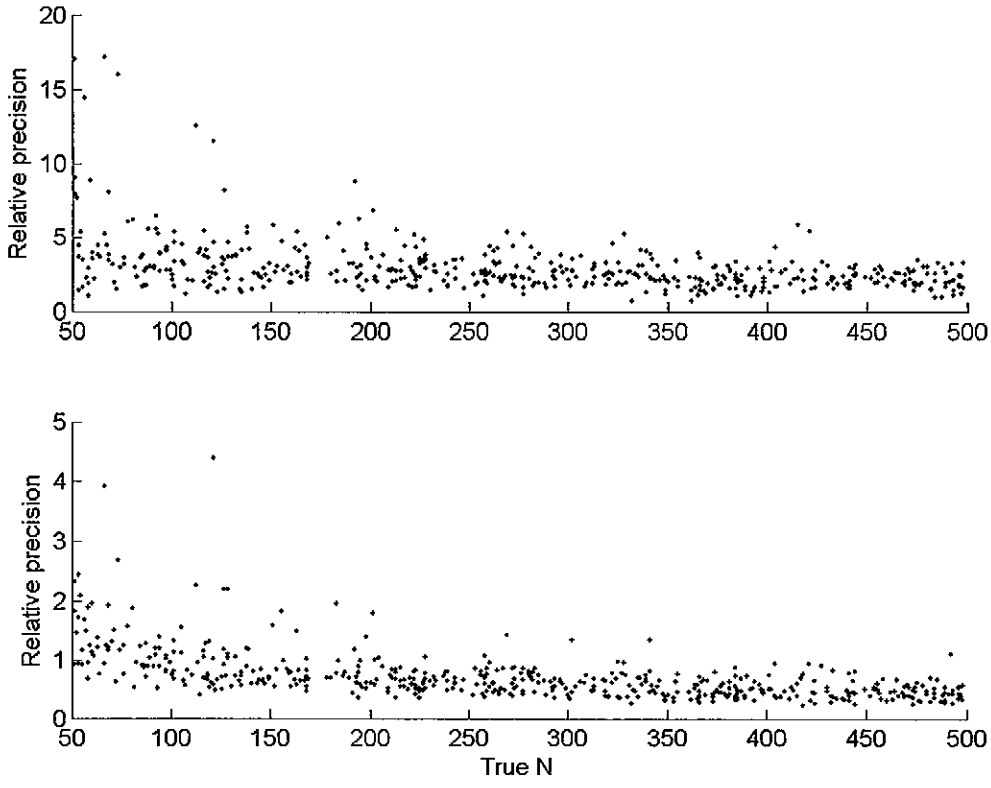

Figure 24. The relationship between the relative precision and the true abundance for simulation case F. The upper figure is for the beta-binomial model, whereas the lower figure is for the multinomial model.

The estimated intercepts of the regression lines were positive for the median and mean, whereas it was negative for the mode (Table 13). Estimated slopes of the regression lines were approximately one and the coefficients of determinations were 0.9 (Table 13). Widths of 95% posterior intervals ranged from 35 to 571, with a median of 146. Approximately 94% (470/500) of the posterior intervals included the true abundance. The relative precision ranged from 0.23 to 4.39 with the median of 0.62. For both models, large values of the relative precision were found for small population sizes (Figure 24).

When the same datasets for simulation case F were analyzed using MARK, the variability of maximum likelihood estimators were small for both models (Figure 25). Results from one dataset were discarded because these estimates (MLE's > 10,000) were unreasonable. For the closed capture model, the variability in the estimator was less than for simulation case B ($-135.0 < \Delta_{mle} < 311.2$; Figure 25). Approximately 52% (258/499) of estimates were greater than the true population sizes.

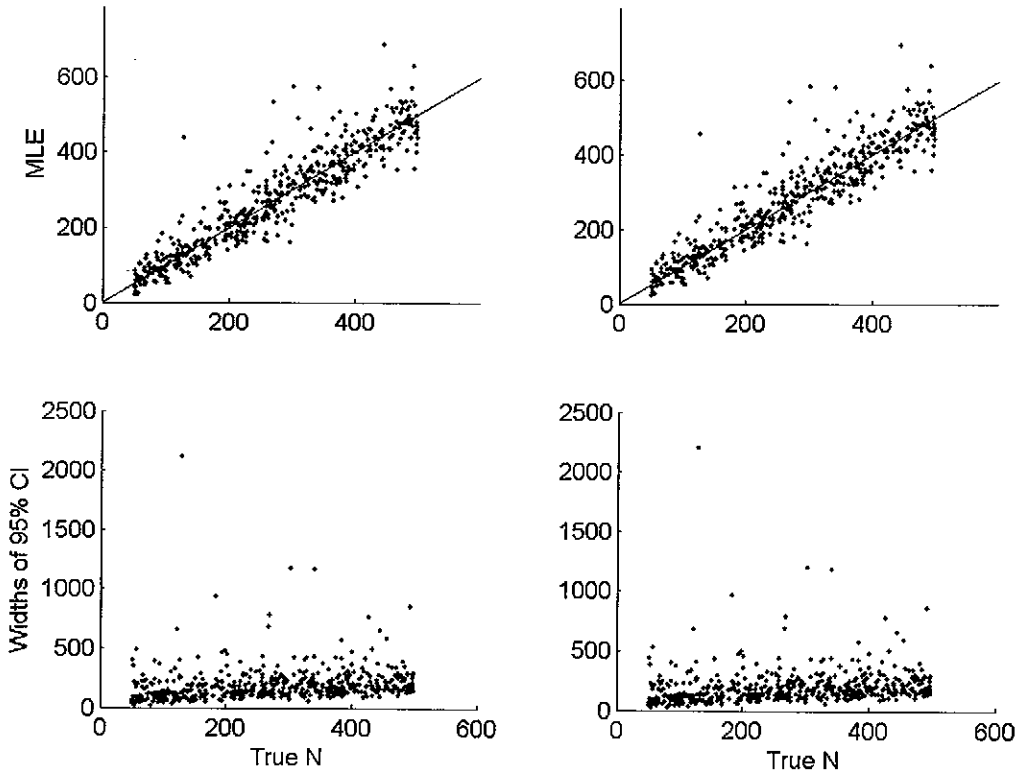

Figure 25. Relationships between true abundance and MLE and the widths of 95% confidence intervals for simulation set F. Left panels are for the closed capture model, whereas right panels are for the Huggins closed model. Solid lines indicate 45-degree lines.

The estimated intercept of the regression line was approximately 9, the slope was approximately one, and the coefficient of determination was 0.87 (Table 12). Widths of 95% confidence intervals ranged from 23.9 to 2118.4, with a median of 162.4. Approximately 93% (466/499) of confidence intervals included the true abundance. The relative precision of MLE ranged from 0.22 to 4.83 with the median of 0.65.

For the Huggins closed model, results were almost identical to those from the closed capture model. The variability of the estimator was less than for simulation case B ($-132.8 < \Delta_{mle} < 328.5$; Figure 25). Approximately 55% (275/499) of estimates were greater than true abundances.

The estimated intercept of the regression line was 13, the slope was approximately one, and the coefficient of determination was 0.86 (Table 12). Widths of 95% confidence intervals ranged from 24.5 to 2202.4, with a median of 164.5. Approximately 93% (466/499) confidence intervals included the true abundance. The relative precision of MLE ranged from 0.22 to 4.83 with a median of 0.65. When the number of secondary occasions within each primary period was increased to 8 (simulation case G), the variability of point estimators for the beta-multinomial model decreased. For the beta-binomial model, however, the variability of the estimators was less than for simulation case F ($-273 \leq \Delta_{mode} \leq 526$, $-243 \leq \Delta_{median} \leq 591$, $-220 \leq \Delta_{mean} \leq 598$; Figure 26). Approximately 42% (209/500) of modes, 57% (285/500) of medians, and 67% (335/500) of means were greater than the true abundance. The estimated intercept of the regression lines were positive, the slopes were approximately one, and the coefficients of determinations were approximately 0.6 (Table 14). Widths of 95% posterior intervals ranged from 32 to 1152, with a median of 448. The true abundances were included in 475 (95%) posterior intervals. The relative precision ranged from 0.84 to 7.20 (median = 1.9).