Combining Change-in-Ratio, Index-Removal, and Removal Models for Estimating Population Size

Chiu-Lan Chen,[1] Kenneth H. Pollock,[2,*] and John M. Hoenig[3]

[1]Biomathematics Program, Department of Statistics,
North Carolina State University, Raleigh, North Carolina 27695, U.S.A.
[2]Department of Statistics, North Carolina State University,
Raleigh, North Carolina 27695, U.S.A.
[3]Virginia Institute of Marine Science, P.O. Box 1346,
Gloucester Point, Virginia 23062, U.S.A.

SUMMARY

There are three methods that can be used to estimate population size when survey data are collected just before and just after two or more known harvests: change-in-ratio, index-removal, and catch-effort (removal) methods. In this paper, we introduce a methodology that combines all three methods. We begin by modeling the survey and removal processes as a Poisson point process and a linear death process, respectively, and then we combine the two processes. The complete-data likelihood can be factored into three parts: the general likelihood function of the index-removal method, the general likelihood function of the change-in-ratio method, and the general likelihood function of the catch-effort method. We compute the maximum likelihood estimates using the Powell search algorithm. Monte Carlo simulations are used to demonstrate that the estimates from combining change-in-ratio, index-removal, and catch-effort methods are more precise than the estimates based on combining any two of them or only using a single method. An example based on snow crab data is presented to illustrate the methodology.

1. Introduction

Removal or harvest data constitute an important type of information about populations. Many techniques have been developed for the analysis of such data. In this paper, we focus on change-in-ratio, index-removal, and removal (catch-effort) methods for the estimation of initial population size. The population is assumed closed (i.e., constant) except for the removals.

The change-in-ratio (CIR) method was developed by Kelker (1940) in an intuitive manner to estimate the size of a harvested deer population. The population size was estimated from the knowledge of sex ratios before and after a differential kill by hunting. Chapman (1954, 1955) introduced the first stochastic model for a closed population with two subclasses and one removal. The method requires the assumptions that (1) the population is closed except for the removals, which are known exactly or can be estimated, and (2) all animals have the same probability of being captured in a survey. His model has been generalized to more than two subclasses (Otis, 1980; Udevitz and Pollock, 1991, 1995) and more than one removal (Chapman, 1955; Pollock et al., 1985; Udevitz and Pollock, 1991, 1995). Most of these models assume all animals have the same probability of being caught in each survey (Kelker, 1940; Chapman, 1955; Otis, 1980). Chapman and Murphy (1965) presented a two subclasses model that allows different sampling probabilities for each subclass but assumes the sampling probabilities remain constant over time. Pollock et al. (1985) provided a generalization in which the ratio of subclass sampling probabilities is constant over time. Udevitz and Pollock (1991) developed a general approach that incorporates all the previous CIR models.

---

* *Corresponding author's email address:* pollock@stat.ncsu.edu

Key words: Animal populations; Catch-effort method; Change-in-ratio method; Combining likelihoods; Index-removal method; Removal method; Sampling.

The index-removal (IR) method makes use of the decline in a measure of relative abundance due to a known removal. The relative abundance is measured in surveys before and after the removal. The estimator was first given by Petrides (1949). The method requires the assumptions that (1) the population is closed except for the removals, which are known exactly or can be estimated, and (2) all animals have the same probability of being captured in a survey and the capture probability is constant over time. Eberhardt (1982) showed that, if the survey sample counts are assumed to follow a Poisson distribution, then the IR estimator is also the maximum likelihood estimator. Routledge (1989) generalized the simple case (two surveys, one removal) to multisurveys with multiremovals.

If animals are sequentially removed from a closed population, we would expect to see a decline in the number caught per unit effort and the population size can be estimated from the observed decline. That is the basis from which removal or catch-effort methods were developed. This technique is now widely used in the study of fish populations. There are two main approaches for the estimation of population size from catch-effort data: least squares regression (Leslie and Davis, 1939; DeLury, 1947; Ricker, 1958) and maximum likelihood estimation (Seber, 1982; Gould, 1994). Bishir and Lancia (1996) state that weighted least squares regression estimators are just as precise as MLEs. The assumptions made for the catch-effort models are (1) the population is closed except for the removals, (2) removal is a Poisson process with regard to effort and the catchability coefficient is assumed constant over time, (3) all animals have the same probability of being removed in each removal, (4) the units of effort are independent, and (5) the removals from the population are known. When the effort expended in each removal is constant, the catch-effort method is often termed the removal method. However, this distinction is not universally accepted and we use the terms interchangeably. In the special cases of two and three removals, the MLE of initial population size has explicit form (Seber, 1982). For the general case, Moran (1951) and Zippin (1956, 1958) derived an iterative procedure for finding the maximum likelihood estimates of initial population size and capture probability.

Since the CIR, IR, and removal estimators all provide information on the population size, it is possible to design a study that combines these methods. Dawe, Hoenig, and Xu (1993) combined the CIR and IR estimates by computing a weighted mean. Routledge (1989) considered a series of indices of abundance alternating with a series of removals. He combined the removal and IR methods using a likelihood approach. However, no work has been done on combining CIR, IR, and removal methods. In this paper, we present a general approach to combining all three methods to get a better estimator of population size that utilizes all the data collected in a study. A more detailed version of the paper is found in Chen (1995).

2. Basic Structure of Data

The survey and removal processes are the two basic elements involved in the study. Here we consider a general case for a population that can be partitioned into r subclasses ($r > 1$) and s sampling periods ($s > 1$). A sampling period generally consists of a survey sample of the population followed by a removal. However, we show in Section 5 that some sampling periods can have a sequence of removals. The subclasses can be defined on factors such as sex, age, size, or marking of the individuals. Hunting regulations and hunter preferences often result in selective removals from wildlife populations, and fishing gear selectivity often causes a differential harvest of size classes for fish populations. Marking animals is another possible way of causing selective removals. In general, the harvest data are the typical data collected by the removal process and the method used for survey sampling is usually not related to the method used in the removal process. For example, a deer population might be surveyed by roadside counts, while the removal data are collected by regulated hunting. A fish population might be surveyed by research counts and the removal data collected by a commercial fishery.

Initially, f_1 units of sampling effort (e.g., net hours) are expended in the first survey, and a total of n_1 animals are counted with $x_{11}, x_{21}, \ldots, x_{r1}$ animals being seen in each subclass, respectively. The first survey is followed by a harvest with $R_{11}, R_{21}, \ldots, R_{r1}$ removals in each subclass, and a total f_1^* units of effort are expended in the first harvest. The survey–removal procedures are repeated until all s surveys and $s - 1$ removals are completed. The procedure can be illustrated in tabular form as

Survey 1	Removal 1	Survey 2	Removal 2	...	Removal $s-1$	Survey s
x_{11}	R_{11}	x_{12}	R_{12}		$R_{1,s-1}$	x_{1s}
x_{21}	R_{21}	x_{22}	R_{22}		$R_{2,s-1}$	x_{2s}
.	.	.	.		.	.
.	.	.	.		.	.
.	.	.	.		.	.
x_{r1}	R_{r1}	x_{r2}	R_{r2}		$R_{r,s-1}$	x_{rs}
n_1		n_2				n_s
f_1	f_1^*	f_2	f_2^*		f_{s-1}^*	f_s

3. General Sampling Model

We present the general probability models for survey and removal processes first, then introduce an idea to combine these two processes.

3.1 *General Probability Model for the Survey Process*

When sampling is with replacement and the sampling efforts are known, Udevitz and Pollock (1995) modeled the survey sampling process as a Poisson point process where each animal is counted at random with respect to increments of sampling effort and it is assumed that the encounter probabilities for each individual are independent. Under these assumptions, it can be shown that

$$(x_{ij} \mid f_j) \overset{\text{independent}}{\sim} \text{Poisson}(\gamma_{ij} f_j X_{ij}) \qquad i = 1, \ldots, r, \; j = 1, \ldots, s.$$

$X_{ij} = X_{i1} - \Sigma_{k=1}^{j-1} R_{ik}$, where the R_{ik} are known fixed quantities. X_{ij} is the number of type i animals in the population at time j. The γ_{ij} is the intensity function of the Poisson point process and it also can be interpreted as the probability of being captured by one unit of effort in the research survey. The intensity functions are assumed to be the same for all individuals of a given subclass during a given sampling period, but they may depend on subclass and sampling period. The likelihood is given by

$$P(x_{1j}, x_{2j}, \ldots x_{rj}, j = 1, \ldots, s \mid f_j, j = 1, \ldots, s)$$
$$= \prod_{j=1}^{s} \prod_{i=1}^{r} \Pr(x_{ij} \mid f_j)$$
$$= \prod_{j=1}^{s} \prod_{i=1}^{r} \frac{1}{x_{ij}!} (\gamma_{ij} f_j X_{ij})^{x_{ij}} e^{-(\gamma_{ij} f_j X_{ij})}. \tag{3.1}$$

Under the model, the initial subclass sizes ($X_{i1}, \; i = 1, 2, \ldots, r$) and intensity functions ($\gamma_{ij}, \; i = 1, \ldots, r, j = 1, \ldots, s$) are parameters of interest. There is a total of $rs + r$ unknown parameters, which exceeds the number of minimum sufficient statistics ($x_{ij}, \; i = 1, 2, \ldots, r, j = 1, 2, \ldots, s$). Thus, the parameters are not identifiable. In order to obtain identifiable parameters, we must impose at least r constraints on the γ_{ij}. Udevitz and Pollock (1991, 1995) imposed a constant probability ratio assumption to the $\gamma_{ij}, \; i = 1, \ldots, r, j = 1, \ldots, s$, that is, $\gamma_{ij} = \mu \lambda_i \beta_j$, where λ_i represents the ratio of the sampling probability for subclass i to that for subclass 1 at any given time; β_j represents the ratio of the sampling probability at time j to that at time 1 for any subclass (with $\lambda_1 = \beta_1 = 1$). Applying the constant probability ratio assumption to the likelihood (3.1), the likelihood becomes

$$\prod_{j=1}^{s} \prod_{i=1}^{r} \frac{1}{x_{ij}!} (\mu \lambda_i \beta_j f_j X_{ij})^{x_{ij}} e^{-(\mu \lambda_i \beta_j f_j X_{ij})}. \tag{3.2}$$

3.1.1 *General probability model for the index removal method.* The total number of encounters ($n_j = \Sigma_{i=1}^{r} x_{ij}$) is the sum of the independent Poisson random variables ($x_{ij}, i = 1, 2, \ldots, r$). This implies n_j is a Poisson random variable also. Thus,

$$(n_j \mid f_j) \overset{\text{independent}}{\sim} \text{Poisson}\left(\mu f_j \beta_j \sum_{i=1}^{r} \lambda_i X_{ij}\right) \qquad j = 1, \ldots, s,$$

where $X_{ij} = X_{i1} - \Sigma_{k=1}^{j-1} R_{ik}$. The likelihood for the total sample counts is

$$L_1 = \prod_{j=1}^{s} \left(\frac{1}{n_j!} \left(\mu\beta_j f_j \sum_{i=1}^{r} \lambda_i X_{ij} \right)^{n_j} e^{-(\mu\beta_j f_j \Sigma_{i=1}^{r} \lambda_i X_{ij})} \right). \tag{3.3}$$

If we assume all animals have the same probability of being caught by one unit of sampling effort regardless of class and survey (i.e., assume $\gamma_{ij} = \mu$ for $i = 1, 2, \ldots, r$, and $j = 1, 2, \ldots, s$), then L_1 ends up as the likelihood of the index removal method. If the capture probabilities for the subclasses are not all the same, the total population $(\Sigma_i X_{i1})$ is not estimable in L_1, but the problem of heterogeneity of capture probability can be solved by making separate estimates for each subclass provided that the sampling probabilities within each subclass are homogeneous.

3.1.2 *General probability model for the change-in-ratio method.* Udevitz and Pollock (1991) developed a general probability model that incorporates all of the previous CIR models for estimating the size of each subclass in a population. This generalization can be viewed as part of the likelihood of (3.2). The likelihood of (3.2) can be factored as

$$P(x_{1j}, x_{2j}, \ldots, x_{rj}, j = 1, \ldots, s \mid f_j, j = 1, \ldots, s)$$
$$= \Pr(x_{1j}, x_{2j}, \ldots, x_{r-1,j}, n_j, j = 1, \ldots, s \mid f_j, j = 1, \ldots, s)$$
$$= \prod_{j=1}^{s} \Pr(x_{1j}, \ldots, x_{r-1,j} \mid n_j, f_j)\Pr(n_j \mid f_j),$$

where

$$(x_{1j}, x_{2j}, \ldots, x_{r-1,j} \mid n_j, f_j) \overset{\text{independent}}{\sim} \text{multinomial} \left(n_j, \frac{\lambda_1 X_{1j}}{\sum_{i=1}^{r} \lambda_i X_{ij}}, \ldots, \frac{\lambda_{r-1} X_{r-1,j}}{\sum_{i=1}^{r} \lambda_i X_{ij}} \right)$$

$$(n_j \mid f_j) \overset{\text{independent}}{\sim} \text{Poisson} \left(\mu f_j \beta_j \sum_{i=1}^{r} \lambda_i X_{ij} \right), \qquad j = 1, \ldots, s.$$

If factored in this way, $\Pi_{j=1}^{s} \Pr(x_{1j}, \ldots, x_{r-1,j} \mid n_j, f_j)$ is the likelihood of Udevitz's general model, and $\Pi_{j=1}^{s} \Pr(n_j \mid f_j)$ is the likelihood of the IR method as shown in Section 3.1. Thus, the likelihood for the CIR method is

$$L_2 = \prod_{j=1}^{s} \left(\frac{n_j!}{x_{1j}! \cdots x_{rj}!} \prod_{i=1}^{r} \left(\frac{\lambda_i X_{ij}}{\sum_{i=1}^{r} \lambda_i X_{ij}} \right)^{x_{ij}} \right). \tag{3.4}$$

3.2 *General Probability Model for the Removal Process (and Method)*

We now propose a general probability model for the removal process. The model described in Section 3.1 assumes that the sampling is with replacement, and therefore it is reasonable to assume a Poisson process for the survey process. However, when animals are permanently removed from a finite population, the removal rates, which are related to the intensity functions of the Poisson process, are not a constant but related to the population size. If we assume the removal rates are linearly proportional to the population size, then the removal system can be modeled as a pure linear death process (Bhat, 1972). The assumptions of the linear death process are

(1) $\Pr(X_{ij}(w + \Delta w) = x - 1 \mid X_{ij}(w) = x) = x\eta_{ij}\Delta w + o(\Delta w)$
(2) $\Pr(X_{ij}(w + \Delta w) = x - z \mid X_{ij}(w) = x) = o(\Delta w), \ z > 1,$

where $X_{ij}(w)$ is the number of subclass i animals in the population at time j after expending w units of effort in the removal process. Note $X_{ij}(0) = X_{ij}$ and η_{ij} is the individual death rate, which is assumed to be the same for all individuals of a given subclass during a given removal period. In fisheries work, η_{ij} is referred to as the catchability coefficient. Under these assumptions, we have

$$(R_{ij} \mid f_j^*) \sim \text{binomial} \left(X_{ij}, 1 - e^{-\eta_{ij} f_j^*} \right).$$

If we impose the constant probability ratio assumption on the individual death rate as $\eta_{ij} = \nu\rho_i\tau_j$, then the full likelihood function for the removal process is

$$L_3 = \Pr(R_{1j}, \ldots, R_{rj}, j = 1, 2, \ldots, s - 1 \mid f_j^*, j = 1, 2, \ldots, s - 1)$$

$$= \prod_{j=1}^{s-1} \prod_{i=1}^{r} \binom{X_{ij}}{R_{ij}} \left(1 - e^{-\nu \rho_i \tau_j f_j^*}\right)^{R_{ij}} \left(e^{-\nu \rho_i \tau_j f_j^*}\right)^{X_{ij} - R_{ij}}.$$

In the case of $r = 1$, L_3 is the same likelihood function as the catch-effort method discussed by Seber (1982).

3.3 *Integration of the Removal and the Survey Sampling Processes*

The models described in Sections 3.1 and 3.2 treat removals as fixed quantities rather than random variables. If we model both sampling counts (x_{ij}) and removals (R_{ij}) as random variables, then the likelihood function of x_{ij} and R_{ij} can be factored into three parts, that is,

$$L = \Pr(x_{1j}, \ldots, x_{rj}, j = 1, \ldots, s; R_{1j}, \ldots, R_{rj}, j = 1, \ldots, s-1 \mid f_j, j = 1, \ldots, s;$$
$$f_j^*, j = 1, \ldots, s-1)$$
$$= \prod_{j=1}^{s} \Pr(n_j \mid R_{1j}, \ldots, R_{rj}; f_j, X_{i1}) \prod_{j=1}^{s} \Pr(x_{1j}, \ldots, x_{r-1j} \mid n_j, R_{ij}, \ldots, R_{rj}; X_{i1})$$
$$\times \prod_{j=1}^{s-1} \Pr(R_{1j}, \ldots, R_{rj} \mid f_j^*; X_{i1})$$
$$= L_1 \times L_2 \times L_3,$$

where L_1 is the general likelihood of the index-removal model, L_2 is the general change-in-ratio likelihood, and L_3 is the general likelihood function of catch-effort or removal models. This approach should lead to large gains in efficiency because of more complete use of the data collected.

3.4 *Numerical Methods*

We apply the Powell (1964) search algorithm, which is a method to locate the minimum of a function of several variables, to find the minimum of the negative of the log (likelihood function). In order to ensure that the global maximum is found, the use of several starting values is sometimes helpful. If the likelihood function is from the exponential family of distributions, the maximum likelihood problem can be reformulated as a weighted least squares problem as done by Udevitz and Pollock (1991). The MLEs can then be calculated using PROC NLIN (SAS, 1985). We have checked that the Powell method gives the same estimates as those calculated from PROC NLIN. Since the MLE is asymptotically efficient, the variance of the MLE can be approximated by inverting the information matrix (Casella and Berger, 1990).

4. Simulation Study

Monte Carlo simulation was used to investigate the performance of the index-removal, change-in-ratio, and removal methods and the gains of the combined methods. Here we considered two levels of population size in the simulation study. The first one is a case with small population size, and we assume the removal rate is proportional to the population size. The small population case may be more realistic for wildlife populations but not for fishery studies. Thus, we also considered a large population case. The simulation data were generated from the probability models described above under the following conditions.

Case 1: Small Population

(1) Population size $N_1 = 2000$, with $X_{11} = 1000$ and $X_{21} = 1000$ ($r = 2$).
(2) Number of survey samples $= 3$, number of removal samples $= 2$.
(3) $\lambda_1 = \lambda_2 = 1.0$; $\beta_1 = \beta_2 = \beta_3 = 1.0$.
(4) $\tau_1 = \tau_2 = 1.0$.
(5) $\mu = 0.001$, with effort chosen so that 0.1 of the population is expected to be seen in each survey.
(6) $\nu = 0.0001$, with effort chosen so that 0.1 or 0.3 of type-1 animals are removed in each removal ($\rho_1 = 1$).
(7) ρ_2 is chosen so that, in each removal, the proportion of type-2 animals removed is 0.2, 0.4, 0.6, 0.8.
(8) f and f^* were set equal to 1.

Case 2: Large Population

(1) Population size $N_1 = 6,700,000$, with $X_{11} = 3,500,000$ and $X_{21} = 3,200,000$.

(2) Number of survey samples $= 3$, number of removal samples $= 2$.
(3) $\lambda_1 = \lambda_2 = 1.0$; $\beta_1 = \beta_2 = \beta_3 = 1.0$.
(4) $\tau_1 = \tau_2 = 1.0$.
(5) $\mu = 0.00002$, with effort chosen so that 0.01 of the population is seen in each survey.
(6) $\nu = 0.0001$, with effort chosen so that 0.01 or 0.03 of type-1 animals are removed in each removal ($\rho_1 = 1$).
(7) ρ_2 is chosen so that the proportion of the type-2 animals removed is 0.02, 0.04, 0.06, 0.08.
(8) f and f^* were set equal to 1.

The simulation data were generated according to

$$(n_1) \sim \text{Poisson}(\mu(X_{11} + X_{21}))$$
$$(n_2) \sim \text{Poisson}(\mu(X_{11} + X_{21} - R_{11} - R_{21}))$$
$$(n_3) \sim \text{Poisson}(\mu(X_{11} + X_{21} - R_{11} - R_{21} - R_{12} - R_{22}))$$
$$(x_{11}) \sim \text{binomial}\left(n_1, \frac{X_{11}}{X_{11} + X_{21}}\right); \qquad x_{21} = n_1 - x_{11}$$
$$(x_{12}) \sim \text{binomial}\left(n_2, \frac{X_{11} - R_{11}}{X_{11} + X_{21} - R_{11} - R_{21}}\right); \qquad x_{22} = n_2 - x_{12}$$
$$(x_{13}) \sim \text{binomial}\left(n_3, \frac{X_{11} - R_{11} - R_{12}}{X_{11} + X_{21} - R_{11} - R_{21} - R_{12} - R_{22}}\right); \qquad x_{23} = n_3 - x_{13}$$
$$(R_{11}) \sim \text{binomial}(X_{11}, 1 - e^{-\nu})$$
$$(R_{21}) \sim \text{binomial}(X_{21}, 1 - e^{-\nu\rho_2})$$
$$(R_{12}) \sim \text{binomial}(X_{11} - R_{11}, 1 - e^{-\nu})$$
$$(R_{22}) \sim \text{binomial}(X_{21} - R_{22}, 1 - e^{-\nu\rho_2})$$

We model the sampling and removal processes using the stochastic models described in Section 3 and assume $\gamma_{ij} = \mu$ for $i = 1, 2$ and $j = 1, 2$; $\eta_{1j} = \nu$, and $\eta_{2j} = \nu\rho_2$ for $j = 1, 2, \ldots, s$. Then the identifiable parameters corresponding to each likelihood function are as given in Table 1. We set a lower bound for the estimate of the population size as the total removal $+ 1$ but set no upper bound in the Powell search algorithm. Plots of the likelihood functions indicate that the likelihood functions were unimodal in a broad region around the maximum. In the case where there is no strong selective removal, the Powell method may converge locally. In our estimation procedure for the simulations, we use the true parameter values as the initial values when searching for the MLEs. When analyzing actual data, it is advisable to try a number of initial values, including those obtained from analytical solutions for simple models such as the traditional CIR, IR, and removal estimators.

When 10% of type-1 animals and 20% of the type-2 animals are removed from a small population, the frequency distributions of $\hat{N}_1$ are skewed and have long right tails (see Chen, 1995). As the removal rate for type-2 animals increases, the overestimates tend to occur less often and the distributions are less skewed, except for the histograms of the removal estimates. For the large population case, the distribution patterns of $\hat{N}_1$ are similar to those found in the small population case. The standard errors as measured in the simulations are quite sensitive to extreme values and are therefore particularly unstable under the conditions where extreme values of the estimates are most likely. For example, we simulated 1000 data sets by Monte Carlo methods twice in the

Table 1
Identifiable parameters associated with the various likelihood
functions for the simulation study. For the index-removal
method, the population total only is estimated.

Likelihood			Parameters		
All	X_{11}	X_{21}	μ	ν	ρ_2
IR	$X_{11} + X_{21}$		μ		
CIR	X_{11}	X_{21}			
R	X_{11}	X_{21}		ν	ρ_2
IR/CIR	X_{11}	X_{21}	μ		
IR/R	X_{11}	X_{21}	μ	ν	ρ_2
CIR/R	X_{11}	X_{21}		ν	ρ_2

case where 10% of the type-1 animals and 20% of the type-2 animals are removed; the means and standard errors (in parentheses) of the MLE of the total population size for combining all three methods are 2152.72 (1053.13) and 2130 (750.24) for the two runs. Consequently, the mean and variance are unreliable indicators of the accuracy of the estimate. We chose the median and the proportions of the estimates within specified ranges of the true values to measure the location and the spread of the distribution.

The medians of the distributions based on the 1000 Monte Carlo simulations for the small and large population cases are given in Tables 2 and 3, respectively. In the simulation of the large population case, the medians are reasonably close to the true values. However, in the case of the small population, the medians were less than the true values when 10% of type-1 and 20% of type-2 animals are removed, except for the medians of the total population size of the removal estimates and the estimates of combining the IR and the removal methods, but they approached the true values as the removals increased. The percentage of estimates within a given distance of the true value are given in Tables 4 and 5. In both cases, the estimates from combining all three methods are uniformly more precise than the others. Comparing the fourth, fifth, and sixth columns to the seventh, eighth, and ninth columns of Tables 4 and 5, we also see the gains from combining any two of the CIR, IR, and removal methods compared to the use of a single method.

Theoretically, we expect that increasing the sampling proportion in the survey process improves the precision of the estimators, except for the removal estimate (which does not use the survey data). Therefore, we fixed the sampling proportion but selected a range of removal rates to investigate its influence on the properties of the estimator. First, we compared the estimates of the total population size for CIR, IR, and removal methods (Tables 4 and 5). Comparing the fourth and sixth columns, we found that both IR and removal estimates increase their efficiencies as the proportions of removals increase, and the IR method performs better than the removal method all

Table 2

Comparison of medians of the estimates obtained under index-removal, change-in-ratio, removal, and combined methods for the case of a small population ($N_1 = 2000$, $X_{11} = 1000$, $X_{21} = 1000$)

R_{1j}/X_{1j}	R_{2j}/X_{2j}	All	IR	CIR	R	IR/CIR	IR/R	CIR/R
Median of $\hat{N}_1$								
0.1	0.2	1958	1963	1956	2061	1995	2011	1915
	0.4	1982	1990	1965	1827	1979	1977	1976
	0.6	1998	2009	2016	1844	2007	2003	1999
	0.8	1997	2012	2001	1906	2005	2002	1990
0.3	0.4	1987	2003	2016	1990	2002	1993	1972
	0.6	1993	1999	1994	1981	1994	1992	1992
	0.8	1996	1990	2016	1990	2003	1990	2001
Median of $\hat{X}_{11}$								
0.1	0.2	977	—	995	774	995	910	946
	0.4	992	—	977	780	992	976	979
	0.6	998	—	1011	839	1006	1005	998
	0.8	997	—	999	908	1006	1003	990
0.3	0.4	989	—	1025	987	1002	995	986
	0.6	994	—	1002	977	995	986	994
	0.8	994	—	1016	995	1001	993	999
Median of $\hat{X}_{21}$								
0.1	0.2	978	—	971	964	995	974	963
	0.4	991	—	980	991	990	991	988
	0.6	998	—	1001	997	1002	997	999
	0.8	1001	—	1000	1000	999	1000	1000
0.3	0.4	995	—	994	994	1001	995	993
	0.6	999	—	999	1000	1000	1001	999
	0.8	1001	—	999	1001	999	1001	1000

Table 3
*Comparison of medians ($\times 10^6$) of the estimates obtained under index-removal,
change-in-ratio, removal, and combined methods for the case of a large
population ($N_1 = 6.7 \times 10^6$, $X_{11} = 3.5 \times 10^6$, $X_{21} = 3.2 \times 10^6$)*

R_{1j}/X_{1j}	R_{2j}/X_{2j}	All	IR	CIR	R	IR/CIR	IR/R	CIR/R
				Median of $\hat{N}_1$				
0.01	0.02	6.697	6.685	6.710	7.133	6.697	6.793	6.701
	0.04	6.700	6.681	6.724	6.694	6.697	6.652	6.687
	0.06	6.703	6.711	6.681	6.685	6.696	6.687	6.710
	0.08	6.696	6.700	6.685	6.706	6.704	6.711	6.687
0.03	0.04	6.704	6.704	6.692	6.741	6.701	6.720	6.707
	0.06	6.603	6.674	6.710	6.697	6.679	6.675	6.696
	0.08	6.701	6.706	6.712	6.705	6.703	6.718	6.708
				Median of $\hat{X}_{11}$				
0.01	0.02	3.500	—	3.510	3.648	3.499	3.550	3.501
	0.04	3.501	—	3.518	3.493	3.496	3.467	3.494
	0.06	3.500	—	3.490	3.488	3.497	3.498	3.504
	0.08	3.498	—	3.490	3.502	3.503	3.504	3.492
0.03	0.04	3.503	—	3.492	3.503	3.502	3.483	3.505
	0.06	3.496	—	3.510	3.474	3.488	3.474	3.499
	0.08	3.502	—	3.505	3.501	3.501	3.501	3.504
				Median of $\hat{X}_{21}$				
0.01	0.02	3.196	—	3.200	3.200	3.196	3.199	3.199
	0.04	3.198	—	3.200	3.197	3.197	3.188	3.194
	0.06	3.199	—	3.200	3.204	3.200	3.205	3.203
	0.08	3.196	—	3.191	3.196	3.199	3.194	3.195
0.03	0.04	3.200	—	3.200	3.214	3.201	3.195	3.204
	0.06	3.196	—	3.200	3.200	3.189	3.197	3.199
	0.08	3.204	—	3.200	3.207	3.201	3.206	3.205

the time. This is because it is not the rate of removal but rather the change in population composition (due to a differential harvest) that determines the precision. This result is consistent with the work of Paulik and Robson (1969). The pattern of performance of the CIR, IR, and removal estimators for the small population case is also revealed in the large population case (Table 5). The only case where the CIR method performs better than the IR method in the estimation of total population size is when 10% of type-1 and 80% of type-2 animals are removed from the small population. After examining the simulation results, we conclude that, if there is a strong selective removal resulting in a big change in subclass proportions before and after the removals, the CIR method often performs better than the IR and removal methods. Otherwise, we can get the most precise estimator of the total population size from the IR method.

5. Extension and Example

Sometimes it is impractical to obtain a midseason survey sample, perhaps due to a substantial increase in costs, but it is possible to collect survey samples before and after the fishery and to monitor a series of removals between the surveys. The following example, using a study of snow crabs in St. Mary's Bay, Newfoundland, Canada (Chen et al., 1998), shows how the combined method can be generalized in this way. Research sampling surveys were conducted before and after the snow crab fishery. Fifty-five fleets of traps were set at randomly selected locations in the first and the second surveys. At each location, two large-meshed and two small-meshed traps were set for 24 hours during each sampling period. Only male snow crabs with carapace width greater than or equal to 95 mm are harvested in Atlantic Canada. Fishermen were taught to return crabs with carapace less than 95 mm to the water as quickly as possible to protect the prerecruits. The snow crabs were classified into two size classes, with type 1 being males with 78 mm $\leq$ carapace width < 95 mm (sublegal size) and type 2 being males $\geq$95 mm (legal size). The sampling effort was

Table 4

Comparison of distribution of $\hat{N}_1$ estimated from index-removal, change-in-ratio, removal, and combined methods for the case of a small population

R_{1j}/X_{1j}	R_{2j}/X_{2j}	All	IR	CIR	R	IR/CIR	IR/R	CIR/R
0.1	0.2	18[a]	13	5	7	15	14	11
		46[b]	34	14	18	37	36	28
		80[c]	64	29	35	67	68	57
	0.4	39	25	18	7	15	25	32
		82	56	42	18	37	59	71
		97	88	77	35	67	89	95
	0.6	61	33	34	6	47	33	54
		97	71	72	16	88	72	94
		100	95	93	41	99	95	100
	0.8	76	43	50	15	64	52	66
		99	81	90	30	96	87	98
		100	98	100	55	100	98	100
0.3	0.4	59	44	10	38	44	56	41
		96	84	22	80	85	95	87
		100	98	42	94	99	100	99
	0.6	73	48	23	43	53	64	63
		99	93	61	84	95	95	97
		100	100	87	96	100	99	100
	0.8	90	68	48	55	74	81	76
		99	97	86	90	98	99	99
		100	100	98	97	100	100	100

[a] Percentage of estimates within 10% of true value.
[b] Percentage of estimates within 25% of true value.
[c] Percentage of estimates within 50% of true value.

measured as the number of locations selected in the survey. Thus, we define $f_1 = f_2 = 55$. Here, we only consider data for small-mesh traps. During the preseason sampling period, 7943 crabs were caught, of which 5519 were sublegal size and 2424 were legal-size crabs. In the postseason survey sample, 4846 crabs were classified as sublegal size and 1819 crabs were legal size. In 1992, the fishery in St. Mary's Bay occurred from September 1 to September 10, and it was determined that 58,425 sublegal crabs and 578,425 legal-size crabs were removed from the population. These data can be used to compute CIR, IR, and combined estimates. We also illustrate how the combined method can be implemented in a case of two surveys with two removals between the surveys. The removals for the first half and the second half of the season were not recorded. Purely for purposes of illustration of our general method, we divided the removal period into two periods with equal fishing efforts and made a guess at the fraction of the total removal occurring in each period. The data including the hypothetical removals can be summarized in the following:

	Fishery		
	t_1		t_2
$x_{11} = 5519$	$R_{11} = 29{,}982$	$R_{12} = 28{,}443$	$x_{12} = 4846$
$x_{21} = 2424$	$R_{21} = 305{,}330$	$R_{22} = 273{,}095$	$x_{22} = 1819$
$n_1 = 7943$			$n_2 = 6665$
$f_1 = 55$	$f_1^* = 1$	$f_2^* = 1$	$f_2 = 55$

Dawe et al. (1993) noted that the catchability of sublegal crabs is probably less than that of legal-size crabs. If we assume the ratio of the sampling probability for legal-size crabs to that of

Table 5
*Comparison of distribution of $\hat{N}_1$ estimated from index-removal, change-in-ratio,
removal, and combined methods for the case of a large population*

R_{1j}/X_{1j}	R_{2j}/X_{2j}	All	IR	CIR	R	IR/CIR	IR/R	CIR/R
0.01	0.02	27[a]	21	10	14	22	25	19
		62[b]	48	19	28	53	54	42
		91[c]	82	39	61	83	83	74
	0.04	55	34	22	17	42	37	45
		94	74	52	39	82	77	85
		100	97	83	69	98	97	99
	0.06	77	47	38	21	54	47	69
		100	88	77	40	94	89	99
		100	99	98	71	100	99	100
	0.08	94	58	50	25	70	60	90
		100	96	89	38	99	96	100
		100	100	100	66	100	100	100
0.03	0.04	64	48	9	44	49	62	47
		97	89	16	85	90	96	87
		100	99	37	98	99	100	99
	0.06	82	59	24	48	61	72	71
		100	100	55	90	98	99	99
		100	100	85	100	100	100	100
	0.08	94	72	38	49	78	80	90
		100	99	78	90	99	100	100
		100	100	98	99	100	100	100

[a] Percentage of estimates within 10% of true value.
[b] Percentage of estimates within 25% of true value.
[c] Percentage of estimates within 50% of true value.

sublegal crabs (λ_2) is constant over time, then λ_2 is part of the unknown parameters that need to be estimated. The minimum sufficient statistics and the unknown parameters associated with each likelihood function are given in Table 6. The likelihood functions of the CIR and IR methods (CIR[1] and IR[1] of Table 6) have the problem of lack of identifiability because the number of unknown parameters is greater than the number of sufficient statistics. If we assume $\lambda_2 = 1$ in L_2 (i.e., CIR[2] in Table 6), then the MLE under this likelihood function is the traditional CIR estimator, which gives a biased estimate for total population size, if the assumption of equal catchability fails (i.e.,

Table 6
*Minimum sufficient statistics and unknown parameters corresponding to each
likelihood function for the snow crab study. See text for meaning of notation.*

Likelihood	Minimum sufficient statistics	Parameters
All	$x_{11}, x_{12}, n_1, n_2, R_{11}, R_{21}, R_{12}, R_{22}$	$X_{11}, X_{21}, \mu, \lambda_2, \nu, \rho_2$
IR[1]	n_1, n_2	$X_{11} + X_{21}, \mu, \lambda_2$
CIR[1]	x_{11}, x_{12}	$X_{11}, X_{21}, \lambda_2$
R	$R_{11}, R_{21}, R_{12}, R_{22}$	$X_{11}, X_{21}, \nu, \rho_2$
IR[1]/CIR[1]	x_{11}, x_{12}, n_1, n_2	$X_{11}, X_{21}, \mu, \lambda_2$
IR[1]/R	$n_1, n_2, R_{11}, R_{21}, R_{12}, R_{22}$	$X_{11}, X_{21}, \mu, \lambda_2, \nu, \rho_2$
CIR[1]/R	$x_{11}, x_{12}, R_{11}, R_{21}, R_{12}, R_{22}$	$X_{11}, X_{21}, \lambda_2, \nu, \rho_2$
CIR[2]	x_{11}, x_{12}	$X_{11}, X_{21}, \lambda_2 = 1$
IR(L)	x_{21}, x_{22}	X_{21}, μ, λ_2
IR(S)	x_{11}, x_{12}	X_{11}, μ

Table 7
*Parameter estimates ($\times 10^6$) and their standard errors ($\times 10^6$) obtained under
CIR, index-removal, removal, and combined methods for the snow crab study*

	Total	Sublegal	Legal
All	3.406 (0.0717)	0.535 (0.0403)	2.871 (0.0592)
IR(S, L)	—	0.479 (0.0679)	2.318 (0.2160)
CIR^2	12.540 (2.8268)	8.716 (2.0102)	3.828 (0.8185)
R	3.475 (0.1035)	0.584 (0.0836)	2.892 (0.0609)
IR^1/CIR^1	2.797 (0.2263)	0.479 (0.0674)	2.318 (0.2160)
IR^1/R	3.475 (0.1020)	0.583 (0.0822)	2.892 (0.0608)
CIR^1/R	3.504 (0.0749)	0.620 (0.0432)	2.884 (0.0599)

$\lambda_2 \neq 1$). However, there is an exception; when removals are made from only one class, the estimate of population for that class is unbiased (Paulik and Robson, 1969). Thus, the MLE under CIR^2 should give a nearly unbiased estimate of the legal-size crab population because 91% of the removals are legal-size crabs. For the index-removal method, the problem of heterogeneity can be avoided by making separate estimates for each size class. We calculated separate index-removal estimates for the legal-size and sublegal-size populations. The estimates of initial population of legal- and sublegal-size snow crabs obtained under the traditional CIR estimator (CIR^2 in Table 6), index-removal estimators for legal- and sublegal-size crabs, removal, and combined methods are given in Table 7. The estimates of the population of legal-size crabs range from 2,318,000 to 3,828,000. The estimates of the population of sublegal-size crabs range from 479,000 to 620,000, except for the estimate using the CIR method alone. We believe that the traditional CIR estimate for the sublegal size crab population is biased, which accounts for the discrepancy. Comparing the standard errors of the estimates, we see the gains in efficiency obtained by combining all three methods. This example illustrates the flexibility of the general probability model presented in this paper.

6. Discussion

The CIR method is sometimes referred to as the survey–removal method. Thus, the survey and removal processes are two basic elements involved in the CIR framework. The CIR technique uses part of the observed data by modeling the survey process but treats the removals and total counts in the survey samples as fixed quantities (or, in other words by conditioning on them). It does not fully utilize all the data available in the CIR study, which suggests the idea of developing the combined method presented here. After modeling the survey and removal processes, we deduced that the catch-effort and removal models in the literature (Seber, 1982) are special cases of the removal model presented in this paper, and the IR method is a special case of the model developed in Section 3.1. The general probability model is extremely flexible. It can be applied to any number of subclasses, surveys, and removal samples. It also allows heterogeneity of capture probability among subclasses or over time by imposing the constant probability ratio assumption. We showed in an example that the general probability model can be easily modified to fit a situation with two surveys with a series of removals between the surveys.

If the proportions removed are small, then unreasonably large estimates are more likely to occur, which causes a skewed distribution with a long right tail for the removal and index-removal methods. The long right tail also occurs for the CIR method when the changes in subclass proportions are small. From the simulation results, the combined method will compensate for the deficiencies of each individual method. Thus, we suggest a full usage of the observed data to get a more accurate and precise estimator, provided, of course, that the assumptions can be met.

Catch-effort and removal methods are now widely used in the study of fish and small-mammal populations. Both these methods assume the catchability coefficient is constant over time and the same for all animals, but the assumption may easily be violated in practice. The most common source of the violation is due to heterogeneous capture probabilities of animals. For example, the catchability may vary with size. This is usually overcome by stratifying the sample according to size class and estimating each class separately, but this will reduce the precision of the catch-effort estimator. We suggest a survey sample be obtained before each removal or conducting preseason and postseason surveys and then applying the combined method to get a better overall estimator.

RÉSUMÉ

Trois méthodes peuvent être utilisées pour estimer l'effectif d'une population lorsque les données d'enquête sont collectées juste avant ou après deux ou plus prélèvements d'individus en nombres connus: la méthode des changements de proportions, la méthode des changements d'indices d'abondance par prélèvement, et la méthode de relation prélèvement-effort. Dans cet article nous introduisons une méthodologie qui combine les trois méthodes. Nous modélisons tout d'abord le processus d'enquête comme un processus ponctuel de Poisson et celui de prélèvement comme un processus de mort linéaire, et nous combinons ensuite les deux processus. La vraisemblance pour des données complètes peut être décomposée en trois facteurs: la fonction de vraisemblance générale de la méthode des changements d'indices par prélèvement, celle de la méthode des changements de proportions, et celle de la méthode de relation prélèvement-effort. Nous calculons les estimations du maximum de vraisemblance à l'aide de l'algorithme de recherche de Powell. Des simulations de Monte-Carlo sont utilisées pour démontrer que les estimations obtenues en combinant les trois méthodes sont plus précises que celles obtenues en combinant deux quelconques des méthodes, ou en n'en utilisant qu'une seule. Un exemple basé sur des données concernant le crabe *Chionoecetes opilio* est présenté pour illustrer la méthodologie.

REFERENCES

Bhat, U. N. (1972). *Elements of Applied Stochastic Processes.* New York: John Wiley and Sons.
Bishir, J. and Lancia, R. A. (1996). On catch-effort methods of estimating animal abundance. *Biometrics* **52,** 1457–1466.
Casella, G. and Berger, R. L. (1990). *Statistical Inference.* California: Wadsworth.
Chapman, D. G. (1954). The estimation of biological populations. *Annals of Mathematical Statistics* **25,** 1–15.
Chapman, D. G. (1955). Population estimation based on change of composition caused by a selective removal. *Biometrika* **42,** 279–290.
Chapman, D. G. and Murphy, G. I. (1965). Estimates of mortality and population from survey–removal records. *Biometrics* **21,** 921–935.
Chen, C. L. (1995). New developments in change-in-ratio estimation of population size. Ph.D. dissertation, North Carolina State University, Raleigh.
Chen, C. L., Hoenig, J. M., Dawe, E. G., Brownie, C., and Pollock, K. H. (1998). New developments in change-in-ratio and index-removal methods, with application to snow crab (*Chionoecetes opilio*). *Canadian Special Publication of Fisheries and Aquatic Sciences,* in press.
Dawe, E. G., Hoenig, J. M., and Xu, X. (1993). Change-in-ratio and index-removal methods for population assessment and their application to snow crab (*Chionoecetes opilio*). *Canadian Journal of Fisheries Aquatic Sciences* **50,** 1467–1476.
DeLury, D. B. (1947). On the estimation of biological populations. *Biometrics* **3,** 145–167.
Eberhardt, L. L. (1982). Calibrating an index using removal data. *Journal of Wildlife Management* **16,** 734–740.
Gould, W. R. (1994). New developments in catch-effort estimation of important demographic parameters. Ph.D. dissertation, North Carolina State University, Raleigh.
Kelker, G. H. (1940). Estimating deer population by differential hunting loss in the sexes. *Proceedings of the Utah Academy of Science, Arts and Letters* **17,** 65–69.
Leslie, P. H. and Davis, D. H. S. (1939). An attempt to determine the absolute number of rats on a given area. *Journal of Animal Ecology* **8,** 94–113.
Moran, P. A. P. (1951). A mathematical theory of animal trapping. *Biometrika* **38,** 307–333.
Otis, D. L. (1980). An extension of the change-in-ratio method. *Biometrics* **36,** 141–147.
Paulik, G. J. and Robson, D. S. (1969). Statistical calculations for change-in-ratio estimators of population parameters. *Journal of Wildlife Management* **33,** 1–27.
Petrides, G. A. (1949). View points on the analysis of open season sex and age ratios. *Transactions of the North American Wildlife Conference* **14,** 391–410.
Pollock, K. H., Lancia, R. A., Conner, M. C., and Wood, B. L. (1985). A new change-in-ratio procedure robust to unequal catchability of types of animal. *Biometrics* **41,** 653–662.
Powell, M. J. D. (1964). An efficient method for finding the minimum of a function of several variables without calculating derivatives. *Computer Journal* **7,** 155–162.

Ricker, W. E. (1958). Handbook of computations for biological statistics of fish populations. *Bulletin of the Fisheries Research Board of Canada* **119,** 1–300.

Routledge, R. (1989). The removal method for estimating natural populations: Incorporating auxiliary information. *Biometrics* **45,** 111–121.

SAS. (1985). *SAS User's Guide: Statistics,* Version 5. Cary, North Carolina: SAS Institute.

Seber, G. A. F. (1982). *The Estimation of Animal Abundance and Related Parameters.* New York: Oxford University Press.

Udevitz, M. S. and Pollock, K. H. (1991). Change-in-ratio estimates for populations with more than two subclasses. *Biometrics* **47,** 1531–1546.

Udevitz, M. S. and Pollock, K. H. (1995). Using effort information with change-in-ratio data for population estimation. *Biometrics* **51,** 471–481.

Zippin, C. (1956). An evaluation of the removal method of estimating animal populations. *Biometrics* **12,** 163–169.

Zippin, C. (1958). The removal method of population estimation. *Journal of Wildlife Management* **22,** 82–90.