WILEY

Partial stratification in two-sample capture–recapture experiments

W. A. Lasantha Premarathna[1] (ORCID) | Carl James Schwarz[1] | Thomas S. Jones[2]

[1]Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, British Columbia, Canada

[2]Division of Fish and Wildlife, Minnesota Department of Natural Resources, Aitkin, MN, U.S.A.

Correspondence
W. A. Lasantha Premarathna, Department of Statistics and Actuarial Science, Simon Fraser University, Room SC K10545, 8888 University Drive, Burnaby, British Columbia, V5A 1S6, Canada.
Email: wpremara@sfu.ca

Funding information
NSERC; Minnesota Department of Natural Resources, Great Lakes Indian Fish and Wildlife Commission; Sport Fish Restoration Act

Capture heterogeneity is known to cause bias in estimates of abundance in capture–recapture experiments. This heterogeneity is often related to observable fixed characteristics of the animals such as sex. If this information can be observed for each handled animal at both sample occasions, then it is straightforward to stratify (e.g., by sex) and obtain stratum-specific estimates. However, in many fishery experiments, it is difficult to sex all captured fish because morphological differences are slight or because of logistic constraints. In these cases, a subsample of the captured fish at each sample occasion is selected, and additional and often more costly measurements are made, such as sex determination through sacrificing the fish. Our data now consist of two types of marked animals: animals whose value of the stratification variable is unknown, and subsamples at each occasion where the value of the stratification variables are determined. In this paper, we develop and apply new methods for these types of experiments. Furthermore, given the relative costs of sampling for a simple capture and for processing the subsample, optimal allocation of effort for a given cost can be determined. We also develop methods to account for additional information (e.g., prior information about the sex ratio) and for supplemental continuous covariates such as length. These methods are applied to a problem of estimating the size of the walleye population in Mille Lacs Lake, MN.

KEYWORDS
abundance, Bayesian analysis, capture heterogeneity, capture–recapture, partial stratification, survey design and analysis

1 | INTRODUCTION

Capture–recapture is a method used to estimate the abundance of an animal population. Many methods have been used to estimate the parameters relating to both "closed" and "open" animal populations (Schwarz & Seber, 1999), and new methods are rapidly developing. The simplest among the capture–recapture methods is the Lincoln–Petersen method (Williams, Nichols, & Conroy, 2002). Under the assumption that the population is closed (the number of individuals is not changing through birth, death, immigration, or emigration), animals do not lose their marks; their marks are correctly recorded and animals act independently. The two-sample Lincoln–Petersen estimate (maximum likelihood) for abundance is $\hat{N} = n_1 n_2 / m$. Here, n_1 is the total number of animals captured, marked, and released in the first sample occasion, n_2 is the total number of animals captured in the second sample occasion, and m is the number of marked animals captured in the second sample occasion. Various sample protocols have been proposed to justify the Lincoln–Petersen

estimate and obtain the associated standard error. The simple two-sample capture–recapture models for estimating population abundance of a closed population also assume that all animals have the same capture probability at each sample occasion. However, this assumption is not satisfied in most applications, and heterogeneity in capture probability is expected.

The estimate of population abundance of a closed population can be biased (negatively or positively) when there is heterogeneity in capture probability. For example, consider a population of fish in a lake, and the population consists of two categories of fish: male and female. Capture probabilities for each category may vary between categories and also between the first and second sample occasions. If the sex was known at each sample occasion, the problem is straightforward. Simply stratify the population into male and female and use the Lincoln–Petersen estimate on each sex separately to calculate the estimates of abundance for each sex. Then, add these stratum-specific estimates to get the estimate of total population abundance. However, in some cases, each sampled fish cannot be sexed due to various reasons such as time constraints, total resources allocated for the study, or difficulty in sexing. In the situations where all the fish cannot be sexed, a partial stratification is done (a subsample of the captured fish at each sampling occasion is selected and sex determined).

We develop new methods for two-sample capture–recapture experiments using partial stratification and apply these methods to estimate the abundance of walleye in Mille Lacs Lake, MN, in 2013. In these types of experiments, there is a cost to capture a fish in each occasion and to categorize a fish from the subsample. Furthermore, given the relative costs of sampling for simple capture and for stratifying the subsample, the optimal allocation of effort for a given cost is determined. First, we develop a method using maximum likelihood to estimate the population abundance and to find the optimal allocation of effort for given a cost. Then, we develop a method using a Bayesian approach when prior information is available. Finally, we develop a method using individual continuous covariates. R programming language (R Development Core Team, 2008) was used for analysis of Mille Lacs Lake walleye data and for simulation studies discussed in this paper.

2 | MLE APPROACH TO PARTIAL STRATIFICATION IN TWO-SAMPLE CAPTURE–RECAPTURE EXPERIMENTS

2.1 | Notation

Let $t(1, 2)$ be the capture occasion. We consider that the population can be stratified into k categories. The following notation is used to represent an animal at each sample occasion:

- 0 - animal is not captured
- U - animal is captured but not stratified
- C - animal is captured and identified as category C, where $C = \{Category_1, Category_2 \ldots, Category_k\}$
  In practice, we use a unique letter to represent each category in the population. For example, if we stratify a population as male and female, then $C = \{M, F\}$.

All the animals in the population are represented by a capture history. For example, the capture history $U0$ represents an animal caught but not stratified in the first sample occasion and not captured in the second sample occasion. CC is the capture history that represents an animal captured and identified to category at both sample occasions. In general, the possible capture histories for partial stratification in a two-sample capture–recapture model are $U0, UU, 0U, C0, CC, 0C$, and 00. Note that the capture history 00 is unobservable. Because there are k number of capture histories related to each of the representations $C0, CC$, and $0C$, there are $3k + 4$ total capture histories.

Let $P_{U0}, P_{UU}, P_{0U}, P_{C0}, P_{CC}, P_{0C}$, and P_{00} denote the probabilities of the capture histories $U0, UU, 0U, C0, CC, 0C$, and 00, respectively.

Costs related to capture and stratification are denoted as follows:

- C_0 − total cost available to perform two-sample capture–recapture study
- c_1 − cost to capture an animal at the first sample occasion
- c_1^* − cost to stratify an animal at the first sample occasion
- c_2 − cost to capture an animal at the second sample occasion
- c_2^* − cost to stratify an animal at the second sample occasion
- c_f − fixed cost for the study regardless of the sample size

2.2 ⎜ Statistics

n_{U0} — number of animals with capture history $U0$
n_{UU} — number of animals with capture history UU
n_{0U} — number of animals with capture history $0U$
n_{C0} — number of animals with capture history $C0$
n_{CC} — number of animals with capture history CC
n_{0C} — number of animals with capture history $0C$
n — total number of animals captured in the study

$$n = n_{U0} + n_{UU} + n_{0U} + \sum_C n_{C0} + \sum_C n_{CC} + \sum_C n_{0C}$$

n_1 — total number of animals captured at the first sample occasion

$$n_1 = n_{U0} + n_{UU} + \sum_C n_{C0} + \sum_C n_{CC}$$

n_1^* — subsample size at the first sample occasion

$$n_1^* = \sum_C n_{C0} + \sum_C n_{CC}$$

n_2 — total number of animals captured at the second sample occasion

$$n_2 = n_{UU} + n_{0U} + \sum_C n_{CC} + \sum_C n_{0C}$$

n_2^* — subsample size at the second sample occasion

$$n_2^* = \sum_C n_{0C}$$

2.3 ⎜ Model parameters

P_{tC} — capture probability of animals belonging to category C at sample occasion t
λ_C — proportion of category C animals in the population; $\sum_C \lambda_C = 1$
θ_t — subsample proportion at sample occasion t
N — population abundance
N_C — population abundance of category C; $N_C = N \times \lambda_C$

As defined above, the total number of model parameters relating to p_{tC}, λ_C, θ_t, and N is $3k + 3$. However, there are only $3k + 2$ parameters to be estimated with the constraint $\sum_C \lambda_C = 1$.

2.4 ⎜ Model development

2.4.1 ⎜ Sampling protocol

Consider an animal population that can be divided into nonoverlapping categories where the stratification variable has been determined. At the first sample occasion, a random sample of size n_1 is captured. Then, a subsample of size n_1^* is selected from n_1, and the stratum is determined for all animals in the subsample. All captured animals are tagged, usually with a unique tag number. All captured animals are released to the population after marking. Again, some time later, another sample of animals of size n_2 is captured randomly from the population. The animals captured at the second sample occasion contain animals captured and marked at the first occasion (some of them might be stratified and some might not be stratified), as well as animals not captured at the first occasion. One of the requirements here is that some of the stratified subsamples at the first occasion are recaptured. Also, animals must not be sacrificed to determine stratification membership at the first sample occasion. Again, a subsample of size n_2^* is selected from the captured sample at the second sample occasion including only animals not marked at the first occasion. A pictorial view of the sampling protocol is given in Figure A.1 in the Supporting information, available on the journal's website.

2.4.2 ⎜ Model assumptions

In addition to the assumptions related to the standard capture–recapture experiments, some additional assumptions about the subsample are required, as follows:

- The population is **closed** (geographically and demographically). The number of individuals does not change during the study through birth or immigration and/or death or emigration.
- The population can be divided into nonoverlapping categories.
- Mark status is correctly identified at each sample occasion.
- Marks are not lost between sample occasions.
- Capture and marking do not affect subsequent catchability of an animal.
- The subsample at each occasion is a random sample of animals that are not marked.
- The category of each animal in the subsamples is successfully identified.
- Animal captures are independent.

2.4.3 | Probability statements of capture history

Probability expressions for capture histories including the history 00 can be explicitly given using the parameters λ_C, θ_t, and p_{tC} as follows:

$$
\begin{aligned}
P_{U0} &= \sum_C \lambda_C \, p_{1C} \, (1 - \theta_1) \, (1 - p_{2C}) \\
P_{UU} &= \sum_C \lambda_C \, p_{1C} \, (1 - \theta_1) \, p_{2C} \\
P_{0U} &= \sum_C \lambda_C \, (1 - p_{1C}) \, p_{2C} \, (1 - \theta_2) \\
P_{C0} &= \lambda_C \, p_{1C} \, \theta_1 (1 - p_{2C}) \\
P_{CC} &= \lambda_C \, p_{1C} \theta_1 \, p_{2C} \\
P_{0C} &= \lambda_C \, (1 - p_{1C}) \, p_{2C} \theta_2 \\
P_{00} &= \sum_C \lambda_C \, (1 - p_{1C}) \, (1 - p_{2C}),
\end{aligned}
$$

where $P_{U0} + P_{UU} + P_{0U} + \sum_C P_{C0} + \sum_C P_{CC} + \sum_C P_{0C} + P_{00} = 1$.

2.4.4 | Likelihood

Under the model assumptions, the number of animals related to each of the capture histories has a multinomial distribution with an unknown index, as follows:

$$
\begin{aligned}
L = \frac{N!}{n_{U0}! \quad n_{UU}! \quad n_{0U}! \quad \prod_C n_{C0}! \quad \prod_C n_{CC}! \quad \prod_C n_{0C}! \quad (N-n)!} \quad \times \\
(P_{U0})^{n_{U0}} \times (P_{UU})^{n_{UU}} \times (P_{0U})^{n_{0U}} \quad \times \\
\prod_C (P_{C0})^{n_{C0}} \times \prod_C (P_{CC})^{n_{CC}} \times \prod_C (P_{0C})^{n_{0C}} \times (P_{00})^{N-n}.
\end{aligned}
\tag{1}
$$

We use the logit-link functions to reparameterize the model parameters p_{tC}, λ_C, θ_t, and the log-link function for N.

2.4.5 | Model constraints

Constraints on the parameters in the likelihood defined in Equation 1 (e.g., equal capture probability across time or category) are implemented using design matrices. Offset values also allow additional constraints (e.g., fixed at certain values) to be placed on real parameter estimates in the model likelihood.

For example, the vector of parameters $\{p_{tC}\}$ can be constrained by using the design matrix $\mathbf{X}$ and offset vector as follows, where β is the corresponding vector of beta parameters:

$$
\text{logit}(p_{tC}) = \mathbf{X}\beta + \text{offset}.
$$

Consider a model with two categories (say M and F) in which all individuals can be successfully assigned to a category. Then, there are four capture probabilities relating to the two categories and two sample occasions. If there is a restriction

on parameters such that $p_{1M} = p_{1F}$ and $p_{2M} = p_{2F}$, then we use design matrices and offsets as follows:

$$\text{logit}\left(\begin{bmatrix} p_{1M} \\ p_{1F} \\ p_{2M} \\ p_{2F} \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

2.4.6 | Parameter estimation

Parameters are estimated numerically using maximum likelihood. With the use of the logit-link and log-link functions, the parameter estimates and variance–covariance matrix are obtained in logit and log scales. The delta method (Ver Hoef, 2012) is used to obtain the variance–covariance matrix on the back-transformed scale.

Closed-form solutions under the maximum likelihood method are not available, but moment estimates can be derived (Davidson & Solomon, 1974; see section B.3 in the Supporting information).

2.4.7 | Model specification, selection, and assessment

Model specification follows the following notation used for naming models in the computer program MARK (White & Burnham, 1999):

$c * t$ = parameter varies by category and time
t = parameter varies by time but not by category
c = parameter varies by category but not by time
$\cdot$ = parameter does not vary by time or by category
$c + t$ = additive model: variation to be parallel between categories across time.

For example, the model defined by $\{p(c * t)\theta(t)\lambda(c)\}$ is the model in which capture probabilities vary by category and time, subsample proportions vary by time, and category proportions vary by category. The model defined by $\{p(c * t)\theta(t)\lambda(0.4)\}$ is the same as the previous model, but the category proportions are fixed such that the category one proportion is 0.4 and the category two proportion is 0.6.

Large numbers of different models can be defined for a given data set using the appropriate design matrices and offset vectors. Model selection is based on Akaike information criterion (AICc; Burnham & Anderson, 2004).

There is no guarantee that the model that has the smallest AICc value actually fits the data well. Therefore, we need to check whether the selected model can describe the data adequately. Model fitness can be assessed using two methods. One way to assess the specified models for a data set is through residual plots using the standardized residual (Dupuis & Schwarz, 2007) for all observable capture histories as $(O_i - E_i)/\sigma_i$, where O_i and E_i are the observed and expected counts for the capture history i, and σ_i is its standard deviation where $\sigma_i \approx \sqrt{E_i}$. Model fitness can also be assessed using the parametric bootstrap goodness-of-fit test using the deviance statistics and the Tukey statistics (Brooks, Catchpole, & Morgan, 2000). Identification of violation of assumptions and a poor model fit are illustrated in the Supporting information, available on the journal's website through goodness-of-fit plots using simulated data.

2.4.8 | Planning experiments

In this partial stratified two-sample capture–recapture study, there is a cost to capturing an animal at each sample occasion, a cost to identify the category of the captured animal in the subsamples, and also a fixed cost regardless of the sample size. If there is a fixed amount of funds (C_0) to be used in the study, then the objective is to find the optimal number of animals to capture at both sample occasions and the optimal sizes of the subsamples to be categorized so that the variance of the estimated population abundance ($\text{Var}(\hat{N})$) is minimized.

The total cost (C) of the experiment can be considered as a linear function of sample sizes and is given by

$$C = c_f + n_1 c_1 + n_1^* c_1^* + n_2 c_2 + n_2^* c_2^* \leq C_0, \tag{2}$$

where

$$n_1^* \leq n_1 \quad and \quad n_2^* \leq n_2 - E\left(n_{UU} + \sum_C n_{CC}\right).$$

Numerical optimization methods are used to find the optimal allocation of n_1, n_2, n_1^*, and n_2^* with respect to the linear constraint defined in Equation 2 such that $\text{Var}(\hat{N})$ is minimized. The following packages were used for numerical optimization in R programing language (R Development Core Team, 2008). General-purpose optimization with L-BSGS-B

method (Byrd, Lu, Nocedal, & Zhu, 1995) was used for numerical optimization. Simulated annealing method with GenSA package (Xiang, Gubian, Suomela, & Hoeng, 2013) can also be used for optimization of the conditional likelihood. Rsolnp package (Ghalanos & Theussl, 2015) was used for general nonlinear optimization with constraints.

It is also important to assess the performance of the model that is going to be used in the experiment before data collection. Power, bias, and precision assessments are performed using the method based on expected values given by Devineau, Choquet, and Lebreton (2006).

Results of the power analysis for partial stratification in two-sample capture–recapture experiments with two categories in the population are given in the Supporting information.

3 | EXAMPLE: ANALYSIS OF MILLE LACS LAKE WALLEYE DATA (MLE APPROACH)

3.1 | Sampling protocol

The study took place on Mille Lacs Lake, MN, in 2013. Walleyes were captured on the spawning grounds. Almost all the fish were sexed in the first sample occasion. All the captured fish were tagged and released, and recapture occurred 3 to 6 weeks later using gill nets. At the time of recapture, all of the fish could not be sexed by external examination. From a sample of fish captured at the second sample occasion with no tags, a random sample was selected and sexed internally. Figure 1 shows the sampling protocol of the walleye data.

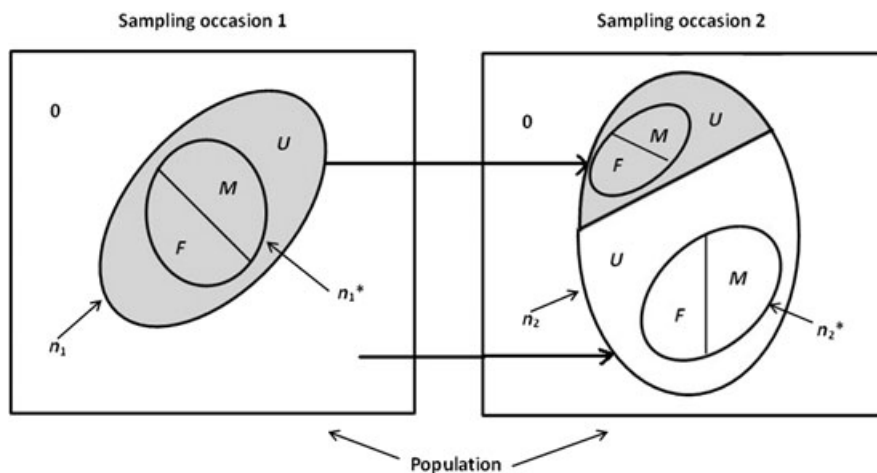Capture histories for walleye from Mille Lacs Lake are provided in Table 1.

FIGURE 1 Sampling protocol: walleye sampled in a two-sample capture–recapture experiment (Mille Lacs Lake, MN, in 2013). The notations U and 0 represent a fish in each sampling time as described in Section 2.1, and $C \equiv \{M, F\}$ where M represents a fish that is captured and identified as a male and F represents a fish that is captured and identified as a female. Shaded areas represent marked fish. n_1, n_1^*, n_2, and n_2^* are as in Section 2.2

TABLE 1 Capture histories for walleye sampled in a two-sample capture–recapture experiment (Mille Lacs Lake, MN, in 2013)

Capture History	Number of Fish
$U0$	42
UU	1
$M0$	5,071
MM	40
$F0$	1,555
FF	32
$0M$	41
$0F$	237
$0U$	3,058

TABLE 2 Model comparison for walleye example using MLE method

Model	np	$\hat{N}$ '000s	$s.e.(\hat{N})$ '000s	AICc	ΔAICc	AICc Weights	Deviance Statistics Observed	Deviance Statistics p-value	Tukey Statistics Observed	Tukey Statistics p-value
$\{p(c*t)\theta(t)\lambda(c)\}$	8	209.1	27.0	77.1	0.0	0.93	0.5	0.85	0.1	0.85
$\{p(c*t)\theta(t)\lambda(0.5)\}$	7	210.6	24.8	82.4	5.3	0.07	7.8	0.01	1.9	0.01
$\{p(c+t)\theta(t)\lambda(c)\}$	7	356.9	48.6	528.3	451.2	0.00	453.6	0.00	116.4	0.00
$\{p(t)\ \theta(t)\lambda(c)\}$	6	314.7	36.2	537.6	460.5	0.00	464.9	0.00	119.2	0.00
$\{p(c)\ \theta(t)\lambda(c)\}$	5	400.2	54.6	1,657.6	1,580.5	0.00	1,584.9	0.00	428.1	0.00
$\{p(.)\ \theta(t)\lambda(c)\}$	3	352.7	40.7	1,666.9	1,589.8	0.00	1,596.2	0.00	430.7	0.00

Note. np = number of parameters; AICc = Akaike information criterion.

3.2 | Best fitted model

Six different models were fitted as given in Table 2. According to the AICc criteria, the best model for the walleye data is the model $\{p(c*t)\theta(t)\lambda(c)\}$. This conclusion is confirmed by examining residual plots (Figure B.1 in the Supporting information), the deviance and Tukey statistics for observed data (Table 2), parametric bootstrap p-values, and the corresponding histograms for 1,000 parametric bootstrap samples (Figure B.2 in the Supporting information) for each of the top four models.

The estimates for all the parameters and their standard errors using the best MLE model $\{p(c*t)\theta(t)\lambda(c)\}$ for the walleye data are shown in Table 3. Because there is substantial capture heterogeneity between the two sexes, the Lincoln–Peterson method is biased and overestimates the population abundance (N) with a larger standard error compared with the estimate given by the model $\{p(c*t)\theta(t)\lambda(c)\}$. Usually, the effect of ignoring heterogeneity is to negatively bias estimates. However, we did not observe a negative bias in the Lincoln–Peterson estimate because the heterogeneity observed was mixed heterogeneity. Pure heterogeneity, which results in negative biases (Pollock, Nichols, Brownie, & Hines, 1990), occurs when individuals from one category of the study animal (e.g., males) are more likely to be observed in both the initial sampling occasion and subsequent occasions. Mixed heterogeneity occurs when one category is more likely to be sampled in the first occasion, but another category is more likely to be sampled in a subsequent occasion. In the case of Mille Lacs Lake walleyes, males were more likely to be observed at the first sample occasion and less likely at the second sample occasion. Conversely, females were less likely at the first sample occasion and more likely at the second sample occasion. Mixed heterogeneity sometimes results in positive biases in Lincoln–Peterson estimates.

TABLE 3 Estimates for the parameters under MLE method, Bayesian method, and individual covariate method using models $\{p(c*t)\theta(t)\lambda(c)\}$, $\{p(c*t)(0,0.3)\theta(t)(0.7,1)(0,0.3)\lambda(c)(20,40)\}$ and $\{p(\text{lenght}*\text{category}*\text{time}+\text{lenght}^2*\text{category}*\text{time}),\theta(t),\lambda(c)\}$, respectively

Parameter	MLE model Estimate	MLE model SE	Bayesian model Posterior Mean	Bayesian model SD	Covariate model Estimate	Covariate model SE
p_{1M}	0.076	0.015	0.082	0.013		
p_{1F}	0.011	0.002	0.013	0.002		
p_{2M}	0.008	0.001	0.008	0.001		
p_{2F}	0.020	0.004	0.023	0.003		
λ_M	0.323	0.060	0.336	0.042	0.293	0.068
λ_F	0.677	0.060	0.664	0.042	0.707	0.068
θ_1	0.994	0.001	0.993	0.001	0.993	0.001
θ_2	0.083	0.005	0.083	0.005	0.083	0.005
N	209,067	27,032	192,229	22,923	279,230	57,034
N_M	67,527	13,335	64,327	9,990	81,833	25,398
N_F	141,540	24,771	127,902	18,631	197,397	44,625
$\hat{N}_{LP}$	314,673	36,232				

Note. $\hat{N}_{LP}$ in the last row is the Lincoln–Peterson estimate for population abundance. The model $\{p(t)\theta(t)\lambda(c)\}$ is the model related to two-sample Lincoln–Peterson method.

Approximate closed-form solutions for the parameter estimates using the best model for the walleye data are given in the Supporting information, available on the journal's website.

3.3 │ Optimal allocation of sampling effort

We consider the problem of how to optimally allocate sampling effort at the two sample occasions. We need to find optimal values for n_1, n_1^*, n_2, and n_2^* such that the variance for the estimate of abundance is minimized (i.e., minimize $\text{Var}(\hat{N})$).

The following costs were considered for the analysis of optimal allocation: $c_f = 0$, $c_1 = 4$, $c_1^* = 0.4$, $c_2 = 6$, $c_2^* = 0.4$, and $C_0 = 90,000$. In this study, cost can be considered as the time in minutes. $C_0 = 90,000$ is the total number of minutes (i.e., 1,500 h) available for this study. Then, the total cost (C) of the experiment can be considered as a linear function of sample sizes and is given by Equation 2.

Optimal allocation is carried out before collecting the data with suitable guesstimates for the parameters using previous studies or the researcher's experience. For the Mille Lacs Lake walleye data, we found the estimates for the parameters using the MLE method for the best fitted model $\{p(c * t)\,\theta(t)\lambda(c)\}$ (Table 3). We considered these estimates as our guesstimates for the optimal allocation to see how much the precision of the estimate of population abundance can be improved compared with the current allocation. We used the guesstimates for the parameters as follows considering the model $\{p(c * t)\,\theta(t)\lambda(c)\}$:

$$N = 209,000, \qquad \lambda_M = 0.33, \qquad r_1 = 6.5, \qquad r_2 = 0.4,$$

where r_1 is the guesstimate of the ratio of p_{1M}/p_{1F}, and r_2 is the guesstimate of the ratio of p_{2M}/p_{2F}. It is difficult to give guesstimates for capture probabilities for each category at each sample occasion. However, in practice, we can give a ratio of the capture probabilities at each sample occasion. For example, if detection probability of males is half that of females at the first sample occasion, then the ratio r_1 is 0.5. The guesstimates of the ratios r_1 and r_2 given above are calculated using the estimates for capture probabilities given in Table 3 under the MLE model.

Optimal allocations of sample sizes and subsample sizes produced by numerical methods for the given costs are $n_1 = 8,929$, $n_1^* = 8,908$, $n_2 = 8,359$, and $n_2^* = 1,412$. At these optimal values, $\text{SE}(\hat{N})$ is 13,657. This standard error is 50% lower than the $\text{SE}(\hat{N})$ obtained using the MLE method (Table 3).

Different solutions are available for n_1, n_2, n_1^*, and n_2^* at optimal allocation. Conditional contour plots were used to see these different solutions. A conditional contour plot in this situation is a contour plot for standard error of $\hat{N}$ where two values of n_1, n_2, n_1^*, or n_2^* are fixed at the optimal values. Conditional contour plots for standard error of $\hat{N}$ when n_1^* and n_2^* are fixed at the optimal values and when n_1 and n_2 are fixed at the optimal values are given in Figures B.4 and B.5, respectively, in the Supporting information. These contour plots show that many solutions are possible for optimal allocation.

3.4 │ Precision of the estimates when additional information is available

Now, consider how well the precision of the estimate of population abundance can be improved in the presence of additional information. For example, we can compare the precision of the estimate under the best fitted model with the model where the sex ratio is known.

In the case of walleye in Mille Lacs Lake, let the first model be $\{p(c * t)\theta(t)\lambda(c)\}$ and the second model be $\{p(c*t)\theta(t)\lambda(\text{MLE})\}$. The first model is the best fit model to the walleye data, and the second model can be considered as the model where the sex ratio is known and fixed at the MLE values obtained from the first model. Therefore, λ_M is fixed at 0.323 (i.e., λ_F is 0.677) for the second model. The first model has eight parameters to be estimated, and the second model has only seven parameters to be estimated because male and female proportions are fixed. Estimates for the parameters produced by the model $\{p(c *t)\theta(t)\lambda(\text{MLE})\}$ are the same as the estimates given in the Table 3 under the model $\{p(c * t)\theta(t)\lambda(c)\}$. However, the standard errors for the population estimates produced by these two models are different. Standard errors for the estimates of the parameters N, N_M, and N_F are 27,032, 13,335, and 24,771 under the model $\{p(c * t)\theta(t)\lambda(c)\}$ and 26,103, 8,431, and 17,632 under the model $\{p(c*t)\theta(t)\lambda(\text{MLE})\}$. According to these standard errors, there is little improvement of the precision of the estimate of population abundance (N) when the second model is used. However, the precision of the estimates of population abundance for each category N_M and N_F is considerably better (around 35% improvement) when the second model is used compared with the best fitted model. Standard errors for all the estimates produced by these two models are given in Table B.2 in the Supporting information.

4 | BAYESIAN ANALYSIS

The previous analysis shows that prior knowledge about the parameters in the model may have a substantial impact on the precision of the estimate of the population abundance and especially on the precision of the estimates for each category. This motivates the use of a Bayesian analysis approach to develop models of partial stratification in two-sample capture–recapture experiments. In this section, we use Bayesian methods to develop such models and apply them to the walleye data from Mille Lacs Lake.

4.1 | Model development

4.1.1 | Prior selection

We consider sensible prior distributions for each of these parameters. Because the parameters p_{tC} and θ_t are probabilities between 0 and 1, beta prior distributions are suitable for them. Because we reparameterize these parameters using the logit-link function, we use normal distributions on the logit scale as prior distributions with suitable means and variances to closely represent these beta distributions. The Dirichlet prior distribution is used for the category proportions because all of the category proportions sum to 1. Because population abundance (N) is an unknown large value, we use a normal flat prior distribution (Gelman, Carlin, Stern, & Rubin, 2004) on the log scale, as follows:

$$\text{logit}(p_{tC}) \sim N\left(\mu_{tC}, \ \sigma_{tC}^2\right)$$

$$\text{logit}(\theta_t) \sim N\left(\mu_t, \ \sigma_t^2\right)$$

$$\lambda_C \sim \text{Dirichlet}(m, n), \text{where} \, m, n > 0$$

$$\log(N) \sim N\left(\mu_N, \sigma_N^2\right)$$

We use informative priors based on expert opinion or prior data. For example, let prior belief for capture probability for a particular category at a certain sample occasion be located in the interval (0,0.3) with an average around 0.15. A Beta(2, 10) would be a good prior distribution to represent this information. Because we model the capture probabilities on the logit scale, $N(-2, 1)$ would closely represent this prior information. However, when informative priors are not available, $U(0, 1)$ can be used. Because we use priors on the logit scale, $N(0, 1.78)$ distribution is considered as the prior distribution because it closely represents $U(0, 1)$ on the logit scale.

4.1.2 | Model specification, selection, and assessment

Models are defined using the notation described in Section 2.4.7 and also using prior information. For example, the model defined by $\{p(c * t)(0, 0.3)\theta(t)(0.7, 1)(0, 0.3)\lambda(c)(2, 4)\}$ is the model where the capture probabilities vary by category and time, subsample proportions vary by time, and category proportions vary by category. Prior information is given within the parenthesis for each parameter. $p(c * t)(0, 0.3)$ means prior belief for all the capture probabilities that are in the interval (0,0.3). In this case, $N(-2, 1)$ prior distributions are used on the logit scale. $\theta(t)(0.7, 1)(0, 0.3)$ means two different prior beliefs are used for subsample proportions at two sampling occasions. The subsample proportion at the first sample occasion is in the interval (0.7,1), and the subsample proportion at the second sample occasion is in the interval (0,0.3). $N(2, 1)$ and $N(-2, 1)$ prior distributions on the logit scale are suitable for the subsample proportions at these two occasions, respectively. $\lambda(c)(2, 4)$ means that the prior distribution of category proportions is Dirichlet(2, 4). In this case, we need to specify only one normal prior distribution on the logit scale because there are only two categories; defining one category automatically also defines the second. A $N(-0.7, 0.95)$ on the logit scale is a good prior distribution for the first category proportion. The model defined by $\{p(c * t)(0, 0.3)\theta(t)(0.7, 1)(0, 0.3)\lambda(0.5)\}$ is the same model described above except that the category proportions are fixed at 0.5.

Model selection is performed using the deviance information criterion (DIC) for finite sample sizes where the posterior distributions of the model have been obtained by Markov chain Monte Carlo (MCMC) simulations using the Metropolis–Hastings method. Two methods are considered for the DIC calculation, the first using pD (Spiegelhalter, Best, Carlin, & Van Der Linde, 2002) and the second using pv (Gelman et al., 2004). The model that has the smallest DIC value is considered the best fitted model for the data.

For a likelihood $p(y|\theta)$, we define the deviance as $D(\theta) = -2\log(p(y|\theta))$, where y are the data and θ are the unknown parameters. Posterior mean deviance is defined as $\bar{D}$, where $\bar{D} = E[D(\theta)]$. Then, we define pD and pv as follows:

$$pD = \bar{D} - D(\bar{\theta}) \qquad pv = \frac{1}{2}\text{Var}(D(\theta)),$$

where $\bar{\theta}$ is the expectation of θ. Then, the DIC values are

$$DIC = pD + \bar{D} \quad or \quad DIC = pv + \bar{D}.$$

Model assessment of fit is done using the Bayesian p-value (Brooks et al., 2000). Two discrepancy statistics, (a) deviance statistics and (b) Freeman–Tukey (FT) statistics, are used. Bayesian p-value close to 0.5 implies that the distribution of the discrepancy statistics for the observed and the simulated data is similar and that the given model can describe the data well.

4.2 | Example: Bayesian analysis of Mille Lacs Lake walleye data

For the Mille Lacs walleye data given in Table 1, a MCMC Metropolis–Hastings method was used with 3 chains, 60,000 iterations for burn-in samples, and 100,000 iterations for post-burn-in samples. The MCMC output was thinned by a factor of 50. This produced a sample of 6,000 (3 chains each with 2,000) to approximate the posterior distribution.

Table C.1 in the Supporting information shows the DIC values and related information for six different models denoted by $M1$ to $M6$ using two methods; one involving the value pD and the other involving the value pv as described in Section 4.1.2. Informative prior distributions for capture probabilities and subsample proportions were considered for the first three models $M1$ to $M3$. Under the pv method, the model with the lowest DIC value is $M1 \equiv \{p(c * t)(0, 0.3)\theta(t)(0.7, 1)(0, 0.3)\lambda(c)(20, 40)\}$. This model also has a DIC value very close to the lowest value under the pD method.

Figure C.1 in the Supporting information shows the Bayesian p-value goodness-of-fit plots King, Morgan, Gimenez, & Brooks, 2010 using two discrepancy statistics, deviance statistics and Freeman–Tukey statistics, for the best four models according to the DIC criteria. Bayesian p-value using both discrepancy statistics for the model $M1 \equiv \{p(c * t)(0, 0.3)\theta(t)(0.7, 1)(0, 0.3)\lambda(c)(20, 40)\}$, which was selected from the DIC criteria, was close to 0.5. Hence, the selected model fit the Mille Lacs Lake walleye data well.

Trace plots, autocorrelation plots, and potential scale reduction plots ($\hat{R}$ plots) of the beta parameters for the best fitted model show that the convergence of the posterior distributions is rapid and that dependence between iterations is small. Trace plots for beta parameters are given in Figure C.2 in the Supporting information.

The results obtained by the MCMC Metropolis–Hastings simulations for the best fitted model $\{p(c * t)(0, 0.3)\theta(t)(0.7, 1)(0, 0.3)\lambda(c)(20, 40)\}$ are given in Table C.2 in the Supporting information. Means and standard deviations of the posterior distributions produced from this best fitted Bayesian model are given in Table 3. The means of the posterior distributions for all the parameters were similar to the estimates of the MLE model $\{p(c * t)\theta(t)\lambda(c)\}$ given in Table 3 with improved standard deviation for the parameters. Variance estimates reveal that the Bayesian method gives better precision over the MLE method when estimating population abundance N and population category totals N_M and N_F. Bayesian analysis provides better models if we can provide informed prior distributions for the parameters. Consider the two models $M1$ and $M3$ given in Table C.1 in the Supporting information. The prior distributions are the same for these two models except for the category proportions. Model $M1$ has a Dirichlet$(20, 40)$ prior and the model $M3$ has a Dirichlet$(2, 4)$ prior. Model $M1$ fits better than model $M3$ because the Dirichlet$(20, 40)$ prior distribution has less variance than the Dirichlet$(2, 4)$.

Posterior distributions for the population abundance and category totals are given in Figure 2. Posterior distributions for the capture probabilities, category proportions, and the subsample proportions are given in Section C.5 in the Supporting information, available on the journal's website.

5 | SIMULATION STUDIES FOR MLE AND BAYESIAN METHODS

Bayesian methods produced estimates for population abundance and category totals with improved precision compared with the MLE method when walleye data were used. The sample size (n) of walleye data is about 10,000, which is a
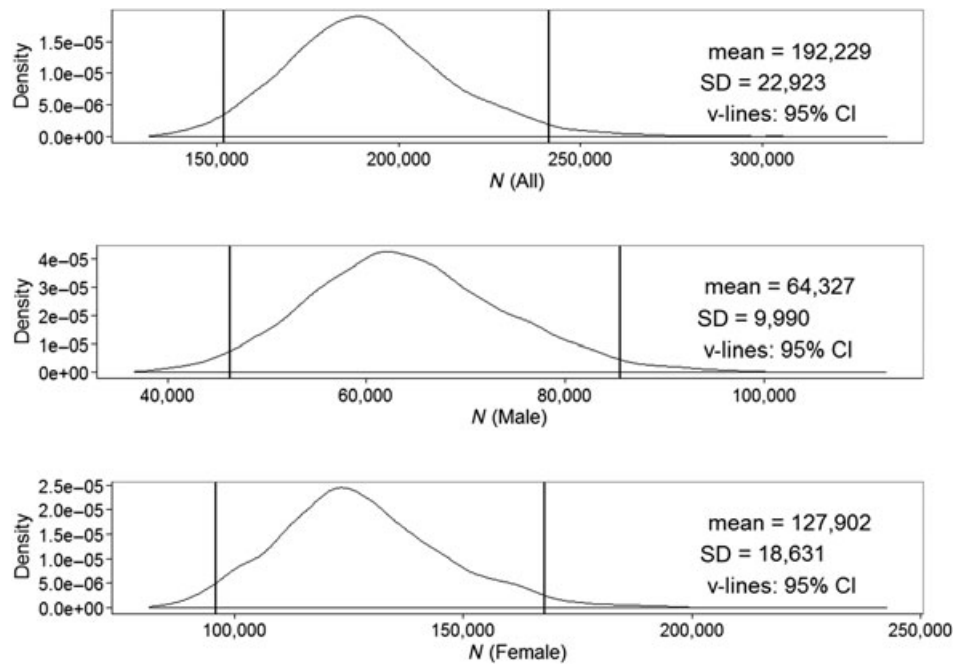
FIGURE 2 Posterior distributions for population abundance (N) and for each category total for males (N_M) and females (N_F) using a sample of 6,000 (3 chains each with 2,000) observations from the posterior distribution using the best model $\{p(c*t)(0, 0.3)\theta(t)(0.7, 1)(0, 0.3)\lambda(c)(20, 40)\}$. Vertical bars are the 95% credible interval

TABLE 4 Simulation study considering a large population and a small population where 1,000 samples were simulated from each population

| Parameter | Samples from a large population | | | | | Samples from a small population | | | | |
| | True value | MLE Method | | Bayesian Method | | True value | MLE Method | | Bayesian Method | |
		Mean	SE	Mean	SE		Mean	SE	Mean	SE
p_{1M}	0.07	0.069	0.010	0.075	0.009	0.15	0.151	0.029	0.153	0.017
p_{1F}	0.01	0.010	0.002	0.011	0.001	0.10	0.106	0.043	0.111	0.019
p_{2M}	0.01	0.010	0.001	0.010	0.001	0.25	0.253	0.045	0.253	0.024
p_{2F}	0.02	0.020	0.004	0.022	0.003	0.15	0.155	0.057	0.174	0.020
λ_M	0.35	0.352	0.057	0.359	0.028	0.6	0.588	0.111	0.615	0.025
λ_F	0.65	0.648	0.057	0.641	0.028	0.4	0.411	0.111	0.384	0.025
θ_1	0.9	0.900	0.004	0.900	0.004	0.9	0.898	0.026	0.899	0.021
θ_2	0.2	0.200	0.007	0.199	0.006	0.4	0.400	0.035	0.399	0.024
N	212,000	218,797	30,126	198,174	20,616	1,000	1,073	289	967	80
N_M	74,000	76,301	12,028	70,669	8,239	600	611	118	594	64
N_F	138,000	142,496	28,520	127,505	15,434	400	462	271	372	40

Note. The model $\{p(c*t)\theta(t)v(t)\lambda(c)\}$ was fitted to samples under the MLE method. The model $\{p(c*t)(0, 0.3)\theta(t)(0.7, 1)(0, 0.3)\lambda(c)(20, 40)\}$ was fitted to the samples from the large population, and the model $\{p(c*t)(0, 0.3)\theta(t)(0.7, 1)(0.25, 0.0.55)\lambda(c)(30, 20)\}$ was fitted to the samples from the small population under the Bayesian method. Under both the MLE and Bayesian methods, the SE matches the actual mean SD of estimates over the simulated samples

large sample. To investigate the effect of sample size, we conducted two simulation studies to test the performance of MLE and Bayesian model development.

We simulated 1,000 samples (total sample size was around 10,000) from a large population with a true population size of 212,000. We also simulated 1,000 samples (total sample size was around 300) from a small population with a true population size of 1,000. Results of simulation studies from both situations are given in Table 4. Estimates of the parameters are similar to the true parameter values under both MLE and Bayesian methods. However, the Bayesian method produced improved precision for population abundance and category totals for both large and small populations.

6 | ANALYSIS WITH INDIVIDUAL COVARIATES

In this section, we develop a method to estimate the abundance of a closed population with partial stratification in two-sample capture–recapture experiments using observable individual covariates. For example, the individual covariates for a fish population can be age, weight, length, etc. To develop the method with partial stratification in capture–recapture experiments, we used methods described by Huggins (1989, 1991) and Alho (1990) for estimating population abundance in capture–recapture experiments in a closed population with heterogeneous capture probabilities.

6.1 | Likelihood

Let the population consist of $i = 1, 2, \ldots, N$ individuals. Then, the likelihood is

$$L^* = \prod_{i=1}^{N} (P_{U0i})^{\Delta_{U0i}} \, (P_{UUi})^{\Delta_{UUi}} \, (P_{0Ui})^{\Delta_{0Ui}} \, \prod_C (P_{C0i})^{\Delta_{C0i}} \, \prod_C (P_{CCi})^{\Delta_{CCi}} \, \prod_C (P_{0Ci})^{\Delta_{0Ci}} \times (P_{00i})^{\left(1 - \sum_h \Delta_{hi}\right)}, \tag{3}$$

where

$$\Delta_{hi} = \begin{cases} 1 & \text{; if the } i\text{th individual captured has a capture history in the set } h, \text{ where} \\ & \qquad\qquad h = \{U0, UU, 0U, C0, CC, 0C\} \\ 0 & \text{; otherwise,} \end{cases}$$

$$P_{U0i} = \sum_C \lambda_C \, p_{1Ci} \, (1 - \theta_1) \, (1 - p_{2Ci})$$

$$P_{UUi} = \sum_C \lambda_C \, p_{1Ci} \, (1 - \theta_1) \, p_{2Ci}$$

$$P_{0Ui} = \sum_C \lambda_C \, (1 - p_{1Ci}) \, p_{2Ci} \, (1 - \theta_2)$$

$$P_{C0i} = \lambda_C \, p_{1Ci} \, \theta_1 \, (1 - p_{2Ci})$$

$$P_{CCi} = \lambda_C \, p_{1Ci} \, \theta_1 \, p_{2Ci}$$

$$P_{0Ci} = \lambda_C \, (1 - p_{1Ci}) \, p_{2Ci} \, \theta_C$$

$$P_{00i} = \sum_C \lambda_C \, (1 - p_{1Ci}) \, (1 - p_{2Ci})$$

and

$$P_{U0i} + P_{UUi} + P_{0Ui} + \sum_C P_{C0i} + \sum_C P_{CCi} + \sum_C P_{0Ci} + P_{00i} = 1.$$

Capture probabilities for each individual depend on the individual covariates, and those probabilities can be represented by p_{tCi} where $t = \{1, 2\}$. We consider a logistic model for the capture probabilities. For example, if the individual covariate is length, then the capture probabilities are given by

$$\text{logit}(p_{tCi}) = \log\left(\frac{p_{tCi}}{1 - p_{tCi}}\right) = \beta_0 + \beta_1 \, (\text{length}_i) + \beta_2 \, (\text{length}_i)^2 + \beta_3 \, (\text{category}) + \beta_4 \, (\text{time}).$$

In this capture probability formula, time takes values 0 and 1, where it represents the sample time 1 and 2, respectively. If there are two categories in the population, then the variable category takes values 0 and 1 representing each respective category. When there are more than two categories in the population, more indicator variables have to be defined accordingly.

We can write the likelihood given in Equation (3) by letting $i = 1, 2, \ldots, n$ as the captured individuals in the experiment and $i = n + 1, \ldots, N$ for animals not captured in the experiment, as follows:

$$L^* = \prod_{i=1}^{n} \left((P_{U0i})^{\Delta_{U0i}} \, (P_{UUi})^{\Delta_{UUi}} \, (P_{0Ui})^{\Delta_{0Ui}} \, \prod_C (P_{C0i})^{\Delta_{C0i}} \, \prod_C (P_{CCi})^{\Delta_{CCi}} \, \prod_C (P_{0Ci})^{\Delta_{0Ci}} \right) \times \left(\prod_{i=n+1}^{N} (P_{00i})^{\left(1 - \sum_h \Delta_{hi}\right)} \right). \tag{4}$$

Now, we condition on the captured individuals because we have covariate information only on the individuals captured at least once in the study. Therefore, $(1 - P_{00i})$ denotes the probability that the ith individual is captured at least once in the study for $i = 1, 2, \ldots, n$. Then, the conditional likelihood can be written as

$$L = \prod_{i=1}^{n} \left(\frac{(P_{U0i})^{\Delta_{U0i}} (P_{UUi})^{\Delta_{UUi}} (P_{0Ui})^{\Delta_{0Ui}} \prod_C (P_{C0i})^{\Delta_{C0i}} \prod_C (P_{CCi})^{\Delta_{CCi}} \prod_C (P_{0Ci})^{\Delta_{0Ci}}}{1 - P_{00i}} \right). \tag{5}$$

The conditional likelihood involves only the captured individuals, and the maximum conditional estimates of the parameters can be found.

The Horvitz–Thompson estimator (Huggins, 1991) gives an estimate for population abundance N, as follows:

$$\hat{N}(\beta) = \sum_{i=1}^{n} \frac{1}{\phi_i(\beta)} \qquad \text{with} \quad \phi_i(\beta) = 1 - P_{00i},$$

where ϕ_i is the probability for the ith individual captured at least once in the study for $i = 1, 2, \ldots, n$, and β is the parameter vector associated with the model.

As described by Huggins (1991), when β is estimated from the data by $\hat{\beta}$, the variance of $\hat{N}(\beta)$ is

$$\text{Var}(\hat{N}(\hat{\beta})) = s^2(\hat{\beta}) + \hat{\mathbf{D}}^T \hat{\mathbf{I}}^{-1} \hat{\mathbf{D}},$$

where

$$s^2(\hat{\beta}) = \sum_{i=1}^{n} \phi_i(\hat{\beta})^{-2} [1 - \phi_i(\hat{\beta})],$$

$\hat{\mathbf{I}}$ is a matrix of second derivatives (Hessian matrix) of the conditional log-likelihood evaluated at $\hat{\beta}$, and $\hat{\mathbf{D}}$ is a vector calculated as follows:

$$\left. \frac{dN(\beta)}{d\beta} \right|_{\hat{\beta}} = \sum_{i=1}^{n} \phi_i(\hat{\beta})^{-2} \left. \frac{d\phi_i(\beta)}{d\beta} \right|_{\hat{\beta}}.$$

6.2 | Model fitting and parameter estimation for Mille Lacs Walleye Data

Now, we consider the individual lengths as additional information available for all captured and recaptured walleyes in both sampling occasions for the Mille Lacs Lake walleye data given in Table 1. The distribution of length for males (M), females (F), and for the walleyes whose sex is not determined (U) in the captured samples in both sampling occasions is given in Figure D.1 in the Supporting information.

Seven different models were fitted to walleye data as shown in Table D.1 in the Supporting information. According to the AICc criteria, the best model for the walleye data is the model $\{p(\text{length} * \text{category} * \text{time} + \text{length}^2 * \text{category} * \text{time}), \theta(t), \lambda(c)\}$. In the best fitted model, $\theta(t)$ means subsample proportions vary by time, $\lambda(c)$ means category proportions vary by category, and $p(\text{length} * \text{category} * \text{time} + \text{length}^2 * \text{category} * \text{time})$ represents the capture formula. The capture probabilities for each individual i for each category C at time t can be found by this capture formula as follows:

$$\begin{aligned}
\text{logit}(p_{tCi}) &= \log \left(\frac{p_{tCi}}{1 - p_{tCi}} \right) \\
&= \beta_0 + \beta_1 (\text{length}_i) + \beta_2 (\text{category}) + \beta_3 (\text{time}) + \beta_4 (\text{length}_i)^2 + \\
&\quad \beta_5 (\text{length}_i)(\text{category}) + \beta_6 (\text{length}_i)(\text{time}) + \beta_7 (\text{category})(\text{time}) + \\
&\quad \beta_8 (\text{category})(\text{length}_i)^2 + \beta_9 (\text{time})(\text{length}_i)^2 + \\
&\quad \beta_{10} (\text{length}_i)(\text{category})(\text{time}) + \beta_{11} (\text{category})(\text{time})(\text{length}_i)^2.
\end{aligned}$$

Motivation to use this complicated capture formula with quadric length can be explained by considering the size selectivity of the gill nets used to capture walleyes (Figure D.1 in the Supporting information). Very large and very small walleyes are difficult to capture. The length frequency of captured walleyes displays a unimodal distribution suggestive of a quadratic relationship of catchability versus length.

Estimates of the parameters were obtained by maximizing the conditional log likelihood function. Table 3 gives the estimates for the parameters and their standard errors for subsample proportions, category proportions, and population abundance using the best fitted covariate model $\{p(\text{length} * \text{category} * \text{time} + \text{length}^2 * \text{category} * time), \theta(t), \lambda(c)\}$
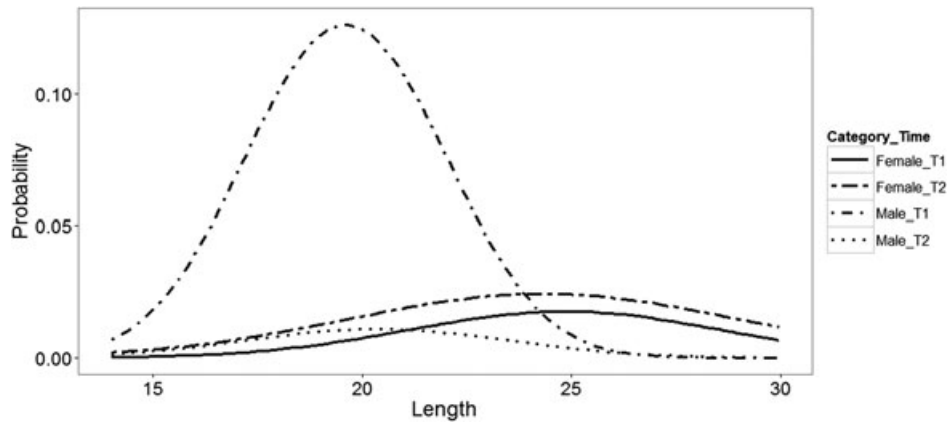
FIGURE 3 Estimated capture probabilities versus the individual covariate length (inches) of males and females at capture time 1 and 2 using the model $\{p(\text{length} * \text{category} * \text{time} + \text{length}^2 * \text{category} * \text{time}), \theta(t), \lambda(c)\}$

with the individual covariate length. Estimates and the standard errors for $\beta_0, \beta_1, \ldots, \beta_{11}$ are given in section D.3 in the Supporting information. The estimate of population abundance N is 279, 230 (SE = 57,034). Variance estimates of population abundance N and category totals N_M and N_F under the individual covariate model are higher compared with the methods discussed in Sections 3 and 4. Higher estimates are a result of residual individual heterogeneity having been modeled. Also, the individual covariate model is highly parameterized compared with the models used in the previous sections. Estimates of category proportions under the individual covariate model are $\lambda_M = 0.293$ (SE = 0.068) and $\lambda_F = 0.707$ (SE = 0.068). These estimates are similar to the estimates found without the individual covariate in Sections 3 and 4 using MLE and Bayesian methods, respectively.

Figure 3 shows that estimated capture probabilities for male and female walleyes against length at sample occasions 1 and 2. According to this figure, male walleyes have a much higher capture probability at sample occasion 1 than female walleyes. At the second occasion, females have a higher capture probability compared with males. Moreover, the catchability of males between 17 and 21 inches in length was very high at the first occasion. Virtually, all walleyes longer than 27 inches were female.

The estimated capture probabilities in Figure 3 can be well explained by using the summary distribution of length of walleyes in Figure D.1 in the Supporting information. Therefore, we can expect that the fitted model using the individual covariate length fits the walleye data well.

7 | DISCUSSION

Capture heterogeneity can cause bias in estimates of abundance in two-sample capture–recapture experiments for closed populations. In fishery experiments, this heterogeneity is often related to characteristics such as sex or size of the fish.

If there is heterogeneity in capture probabilities in each category, the simplest way to address the heterogeneity is to stratify it into different categories and use the stratum-specific estimates. Then, the sum of these stratum-specific estimates gives the estimate for the overall abundance. However, complete stratification is not always possible in experiments. In such cases, we can do a partial stratification where a subsample of the captured animals is selected at each sampling occasion and assigned to categories.

The optimal allocation procedures help us allocate effort for a given cost among the various costs of the study to obtain the estimate of the abundance with smallest variance. It also shows that many different solutions are possible at the optimal allocation.

The Bayesian method provides improved precision for the estimates of population abundance and category totals because it allows us to use expert prior knowledge on the parameters of the model. We observed that the Bayesian solutions are sensitive to the choice of prior distributions. It is important to use informative prior distributions, especially on category proportions (sex ratios) in order to fit a better model. Having good knowledge of the sex ratio is important to obtain estimates of the population abundance with higher precision.

When additional information such as a continuous covariate length is available for each captured sample, then the model development with individual covariates should be expected to produce better estimate for the population

abundance compared with the methods developed without individual covariates. It is important to define an appropriate logistic regression model (e.g., quadratic relationship of catchability versus length for the walleye data) for capture probabilities to fit a good model for the given data with individual continuous covariates. However, the precision of the population abundance produced by a highly parameterized model that uses continuous covariates can be worse than the precision given by models that were developed without individual covariates because there is always a compromise between accuracy and precision.

Bayesian models produced estimates with improved precision for population abundance and category totals over the MLE method for walleye data. Simulation studies show that performances under large population sizes and small population sizes are similar for both MLE and Bayesian model development methods. When the individual lengths are considered for the walleye data, the conditional likelihood method produced estimates for category proportions and subsample proportions that were quite similar to the MLE and Bayesian methods. However, the estimates of population abundance under the conditional likelihood method produced somewhat different values with less precision compared with the previous two methods because the best fitted covariate model was highly parameterized.

We need to consider the issue of nonidentifiability in model fitting. All the parameters can be estimated in two-sample capture–recapture studies similar to Lincoln–Peterson (Williams et al., 2002) model fitting. However, we would not be able to estimate some of the parameters, for example, if no females were observed. There is no nonidentifiable issue on model fitting with walleye data because both males and females were observed. However, there can be a nonidentifiability issue with the population abundance parameter when modeling with individual heterogeneity, as described by Link (2003).

ORCID

W. A. Lasantha Premarathna http://orcid.org/0000-0003-4118-9287

REFERENCES

Alho, J. M. (1990). Logistic regression in capture-recapture models. *Biometrics*, 46(3), 623–635.

Brooks, S. P., Catchpole, E. A., & Morgan, B. J. T. (2000). Bayesian animal survival estimation. *Statistical Science*, 15(4), 357–376.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33(2), 261–304.

Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5), 1190–1208.

Davidson, R. R., & Solomon, D. L. (1974). Moment-type estimation in the exponential family. *Communications in Statistics*, 3(11), 1101–1108.

Devineau, O., Choquet, R., & Lebreton, J-D. (2006). Planing capture–recapture studies: straightforward precision, bias and power calculations. *Wildlife Society Bulletin*, 34(4), 1028–1035.

Dupuis, J. A., & Schwarz, C. J. (2007). A Bayesian approach to the multistate Jolly-Seber capture-recapture model. *Biometrics*, 63(4), 1015–1022.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis* (2nd ed.). Boca Raton, FL: Chapman and Hall/CRC.

Ghalanos, A., & Theussl, S. (2015). Rsolnp: general non-linear optimization using augmented Lagrange multiplier method. R package version 1.16.

Huggins, R. M. (1989). On the statistical analysis of capture experiments. *Biometrika*, 76(1), 133–140.

Huggins, R. M. (1991). Some practical aspects of a conditional likelihood approach to capture experiments. *Biometrics*, 47(2), 725–732.

King, R., Morgan, B. J. T., Gimenez, O., & Brooks, S. P. (2010). *Bayesian analysis for population ecology*. Boca Raton, FL: Chapman and Hall/CRC.

Link, W. A. (2003). Nonidentifiability of population size from capture–recapture data with heterogeneous detection probabilities. *Biometrics*, 59(4), 1123–1130.

Pollock, K. H., Nichols, J. D., Brownie, C., & Hines, J. E. (1990). Statistical inference for capture-recapture experiments. *Wildlife Monographs*, 107, 3–97.

R Development Core Team (2008). R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Schwarz, C. J., & Seber, G. A. F. (1999). Estimating animal abundance: review III. *Statistical Science*, *14*(4), 427–456.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *64*(4), 583–639.

White, G. C., & Burnham, K. P. (1999). Program MARK: survival estimation from populations of marked animals. *Bird Study*, *46*, S120–S139.

Williams, B. K., Nichols, J. D., & Conroy, M. J. (2002). *Analysis and management of animal populations*. San Diego, CA: Academic Press.

Xiang, Y., Gubian, S., Suomela, B., & Hoeng, J. (2013). Generalized simulated annealing for efficient global optimization: the GenSA package for R. *The R Journal*, *5*, 13–28.

Ver Hoef, J. M. (2012). Who invented the delta method? *The American Statistician*, *66*(2), 124–127.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Premarathna WAL, Schwarz CJ, Jones TS. Partial stratification in two-sample capture–recapture experiments. *Environmetrics*. 2018;29:e2498. https://doi.org/10.1002/env.2498