

Web-based Supplementary Materials for “Partial Stratification in Two-Sample Capture-Recapture Experiments”

W. A. Lasantha Premarathna^{1,*}, Carl James Schwarz^{1,} and Thomas S. Jones^{2,***}**

¹Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, Canada.

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

^{*}*email:* wpremara@sfu.ca

^{**}*email:* cschwarz@stat.sfu.ca

^{***}*email:* tom.jones@state.mn.us

Appendix A. MLE Approach to Partial Stratification in Two-Sample Capture-Recapture Experiment

A.1 Sampling Protocol

Consider an animal population that can be divided into non-overlapping categories where the stratification variable has been determined. At the first sample occasion a random sample of size n_1 is captured. Then a sub-sample of size n_1^* is selected from n_1 and the stratum is determined for all animals in the sub-sample. All captured animals are marked, 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 contains animals captured and marked at the first occasion (some of them might be stratified and some of them might not be stratified) as well as the animals not captured at the first occasion. From animals not previously captured, a sub-sample of size n_2^* is selected and the stratum determined. A pictorial view of the sampling protocol is given in Figure A.1.

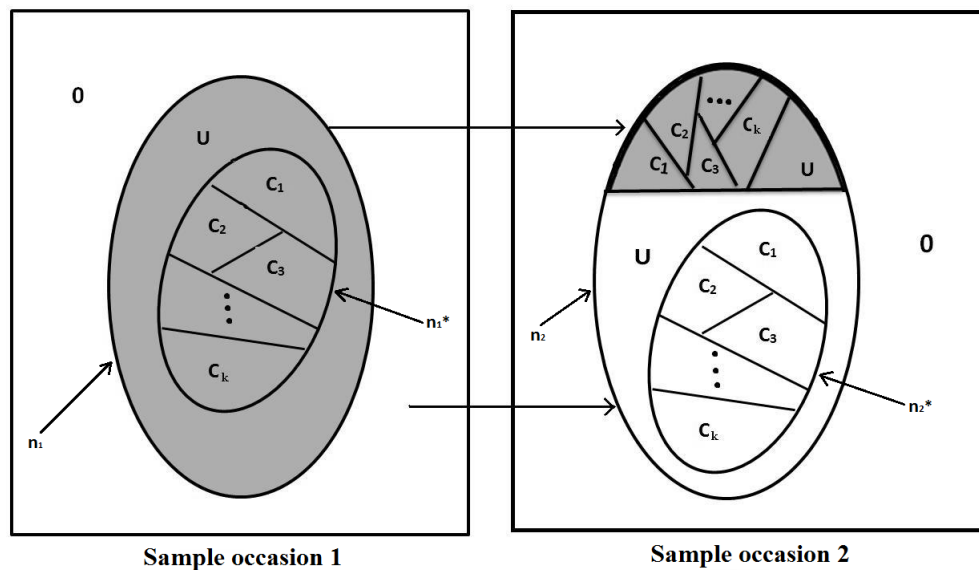


Figure A.1: Sampling protocol: Partial stratification in two-sample capture-recapture experiments. The shaded areas represent marked animals. some of the marked animals in the sample occasion 1 are stratified into strata C_1, \dots, C_k ; some are left unstratified (U). A similar representation is used at sample occasion 2.

A.2 Identify Poor Model Fit and Violation of Assumptions Through Residual Plots

The assumptions of partial stratification in two-sample capture-recapture experiments might be violated in an experiment. In such cases it is important to know whether the model assessment procedures can identify model violations through standardized residual plots and goodness of fit plots. We considered the following four cases for identifying model violation of assumptions through standardized residual plots.

- (a) Heterogeneity in catchability among the animals within each category
- (b) Failure of non-death assumption
- (c) Entering of new animals to the sampled population between the first and second time period
- (d) Fit a model forcing the capture probabilities to be equal when actually they vary between categories and/or occasion

Each of these cases was tested using simulated data. The violation of assumptions given in the first three cases cannot be identified by residual plot or goodness of fit plots because the generated capture histories do not have any information attached regarding those failures of assumptions. However the residual plots revealed that the violation given by the last case can be well identified.

To test the last case given above, we simulated data where the capture probabilities varied between categories and sample occasions, and fitted a model where they are forced to be equal. We considered the following information in Table A.1 for data generation when there are two categories in the population, M and F .

Table A.1: Data generation information to test model fit

Parameter	Value
p_{1M}	0.07
p_{1F}	0.05
p_{2M}	0.08
p_{2F}	0.10
λ_M	0.60
λ_F	0.40
θ_1	0.60
θ_2	0.50
sample size = n =	4000

The model $\{p(\cdot) \theta(t) \lambda(c)\}$ was fitted to the generated data assuming that capture probabilities do not vary by time and by category. The standardized residual plot for this model clearly shows that the fitted model does not fit the generated data because some residuals fall outside ± 1.96 (Figure A.2). In general, we can find poor model fitting of a given data set through residual plots.

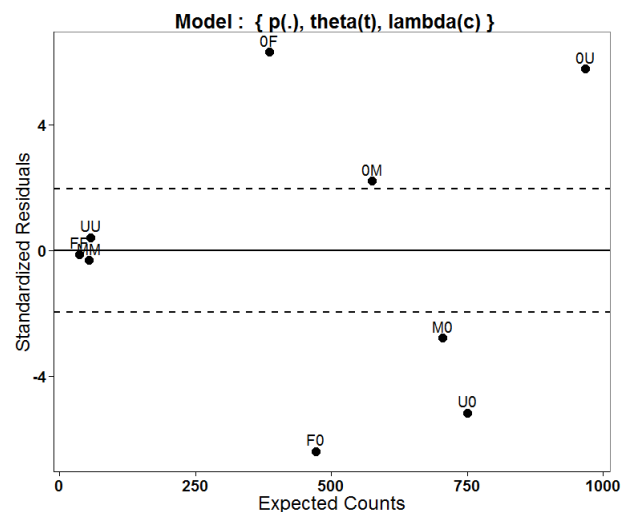


Figure A.2: Identify violation of assumptions for generated data using standardized residual plot. Data is generated from the model $\{p(c * t) \theta(t) \lambda(c)\}$. The fitted model is $\{p(\cdot) \theta(t) \lambda(c)\}$. Dashed lines are drawn at ± 1.96 to represent the region where approximately 95% of residuals should lie if model assumptions were met completely.

A.3 Power Analysis

Power analysis is carried out with generated data with two categories (M and F) in the population to detect the changes of capture probabilities, category proportions and sample sizes. The method based on expected values (Devinue et al., 2006) is considered for power analysis and then we used simulation to verify the power. In each situation, we set $\alpha = 0.05$.

For power analysis to detect changes of capture probabilities, we generated data where the capture probabilities vary between categories and sample occasions and fit a model where capture probabilities vary only by sample occasions. We then calculated the power with the change (Δ_p) of the capture probabilities for category F . We considered the values for Δ_p from 0 to 0.1. Parameter values used for data generation are given in Table A.2.

Table A.2: Data generation information for power analysis with capture probabilities

Parameter	Value
p_{1M}	0.08
p_{1F}	$0.08 + \Delta_p$
p_{2M}	0.04
p_{2F}	$0.04 + \Delta_p$
λ_M	0.60
λ_F	0.40
θ_1	0.80
θ_2	0.50
sample size = n =	2000

We considered the null model H_0 and the alternative model H_a as follows

$$H_0 : \{ p(t) \theta(t) \lambda(c) \}$$

$$H_a : \{ p(c * t) \theta(t) \lambda(c) \}$$

Figure A.3 shows the power analysis using the method based on expected values (Devinue et al., 2006) and verifies the power using a simulation study for capture probability with $\alpha = 0.05$. According to Figure A.3, 80% power was achieved with $\Delta_p = 0.040$ for a sample size 2000.

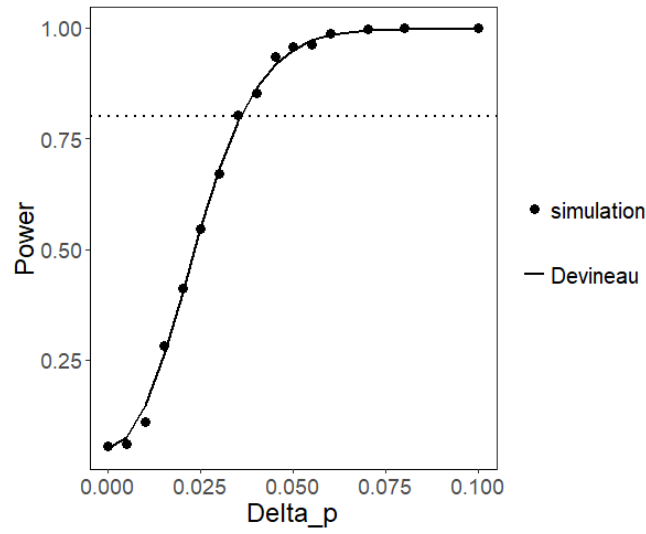


Figure A.3: Power analysis with capture probability using Devineau method and verified power using simulation study with sample size = 2000 at $\alpha = 0.05$. Δ_p (denoted in the graph as Delta_p) is the difference between the capture probability of two categories (F - M capture probabilities) at each sample occasion. Models under H_0 and H_a are $H_0 : \{ p(t) \theta(t) \lambda(c) \}$ and $H_a : \{ p(c * t) \theta(t) \lambda(c) \}$. Horizontal dotted line indicates 80% power.

The power analysis was repeated for sample sizes ranging from 1,000 to 4,000 (Figure A.4). Parameter values used for data generation are given in Table A.2. Larger samples allowed us to detect small differences between the capture probabilities of two categories at each occasion. For example, with a sample size of 4,000, 80% power was achieved when $\Delta_p = 0.025$.

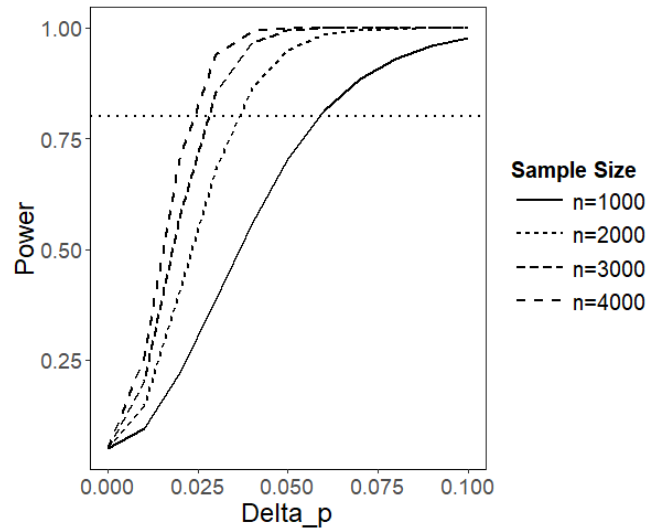


Figure A.4: Power analysis with different sample sizes at $\alpha = 0.05$. Delta_p is the difference between the capture probability of two categories (F - M capture probabilities) at each sample occasion. Models under H_0 and H_a are $H_0 : \{ p(t) \theta(t) \lambda(c) \}$ and $H_a : \{ p(c * t) \theta(t) \lambda(c) \}$. Horizontal dotted line indicates 80% power.

For the power analysis to detect changes of category proportions, we used generated data such that the capture probabilities varied between categories and sample occasions, and category proportions were different. Then a model was fitted that forced fixed category proportions. Parameter values used for data generation are given in Table A.3.

Table A.3: Data generation information for power analysis with category proportions

Parameter	Value
p_{1M}	0.05
p_{1F}	0.08
p_{2M}	0.10
p_{2F}	0.12
λ_M	0.4
λ_F	0.6
θ_1	0.80
θ_2	0.50
sample size = n =	2000

We considered the the Null model H_0 and the alternative model H_a as follows.

$$H_0 : \{ p(c * t) \theta(t) \lambda(0.4 + \Delta_M) \}$$

$$H_a : \{ p(c * t) \theta(t) \lambda(c) \}$$

We considered different values for Δ_M as 0, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, and 0.4. Then the corresponding values of λ_M under H_0 are 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, and 0.8. We used these values for power analysis. Finally we verified the power using a simulation study.

Figure A.5 shows the power analysis for λ_M with $\alpha = 0.05$ with sample size 2000. At $n = 2000$, 80% power was achieved with $\Delta_M = 0.2$.

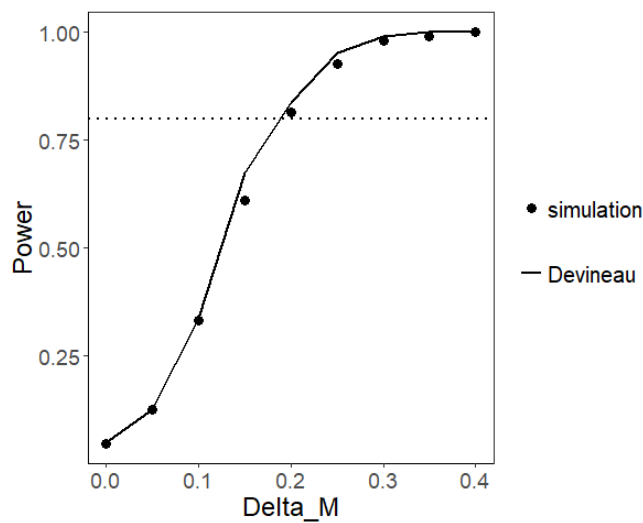


Figure A.5: Power analysis with category proportion (λ) using Devineau method and verified power using simulation study with sample size = 2000 and $\alpha = 0.05$. Delta_M (Δ_M) is the increment of λ_M from 0.4. Models under H_0 and H_a are $H_0 : \{ p(c * t) \theta(t) \lambda(0.4 + \Delta_M) \}$ and $H_a : \{ p(c * t) \theta(t) \lambda(c) \}$. Horizontal dotted line indicates 80% power.

Appendix B. Analysis of Mille Lacs Lake Walleye Data - MLE Approach

B.1 Standardized Residual Plots

Standardized residual plots for the top four models in Table 2 in the paper are given in Figure B.1 .

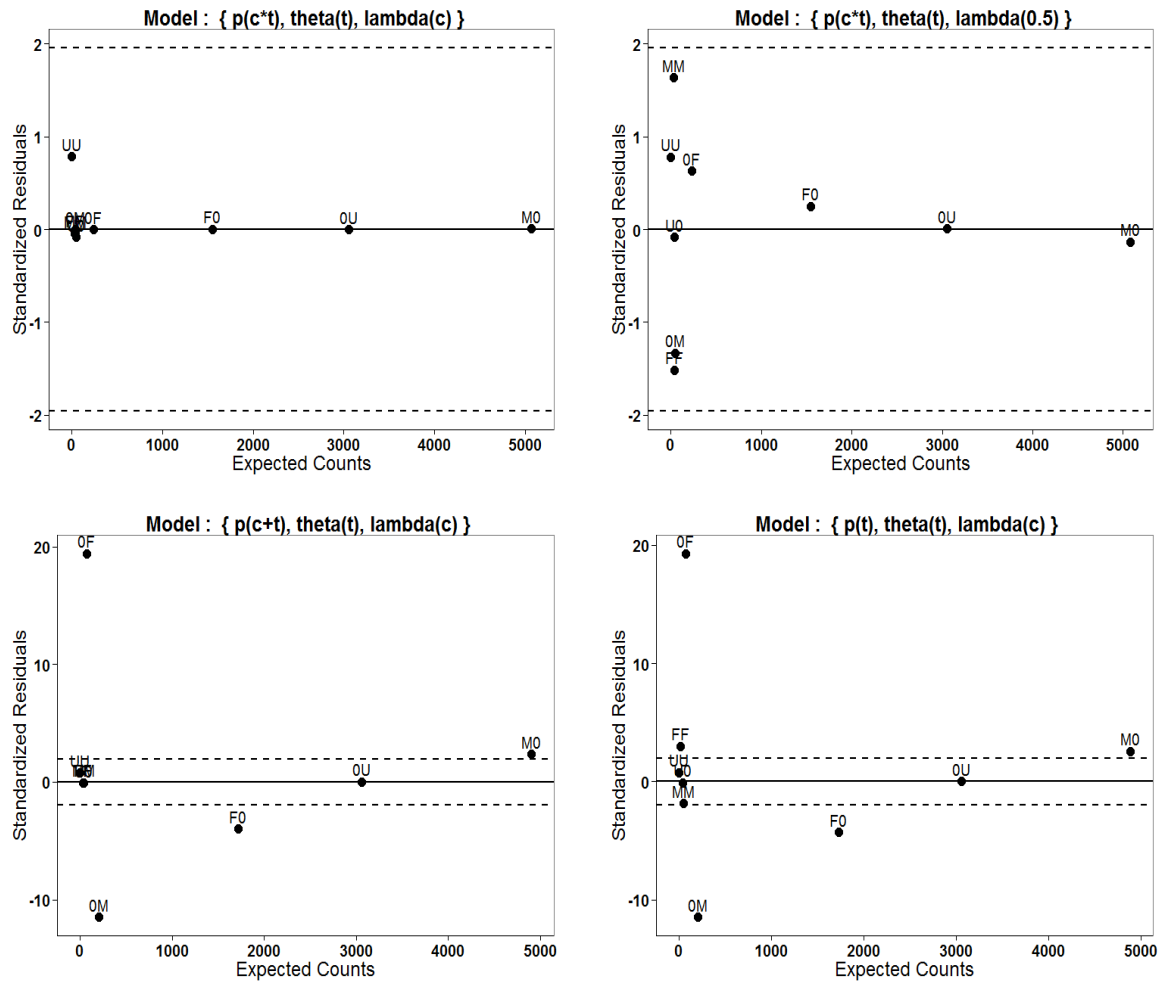


Figure B.1: Standardized residual plots for the best four models according to AICc values in Table 2. Dashed lines are drawn at ± 1.96 to represent the region where approximately 95% of residuals should lie if model assumptions were met completely. Good fit in top row; poor fit in bottom row.

B.2 Model Assessment using Parametric Bootstrap Method

Figure B.2 and B.3 show the parametric bootstrap goodness of fit plots for the best four models according to the AICc criteria using the two discrepancy statistics; the deviance statistic and the Tukey statistic for 1000 bootstrap samples. Plot (a) in both of these figures refers to the model $\{p(c * t) \theta(t) \lambda(c)\}$ which has the smallest AICc value. The observed deviance and the observed Tukey statistic and the corresponding p-values suggest that the model $\{p(c * t) \theta(t) \lambda(c)\}$ fits to the walleye data.

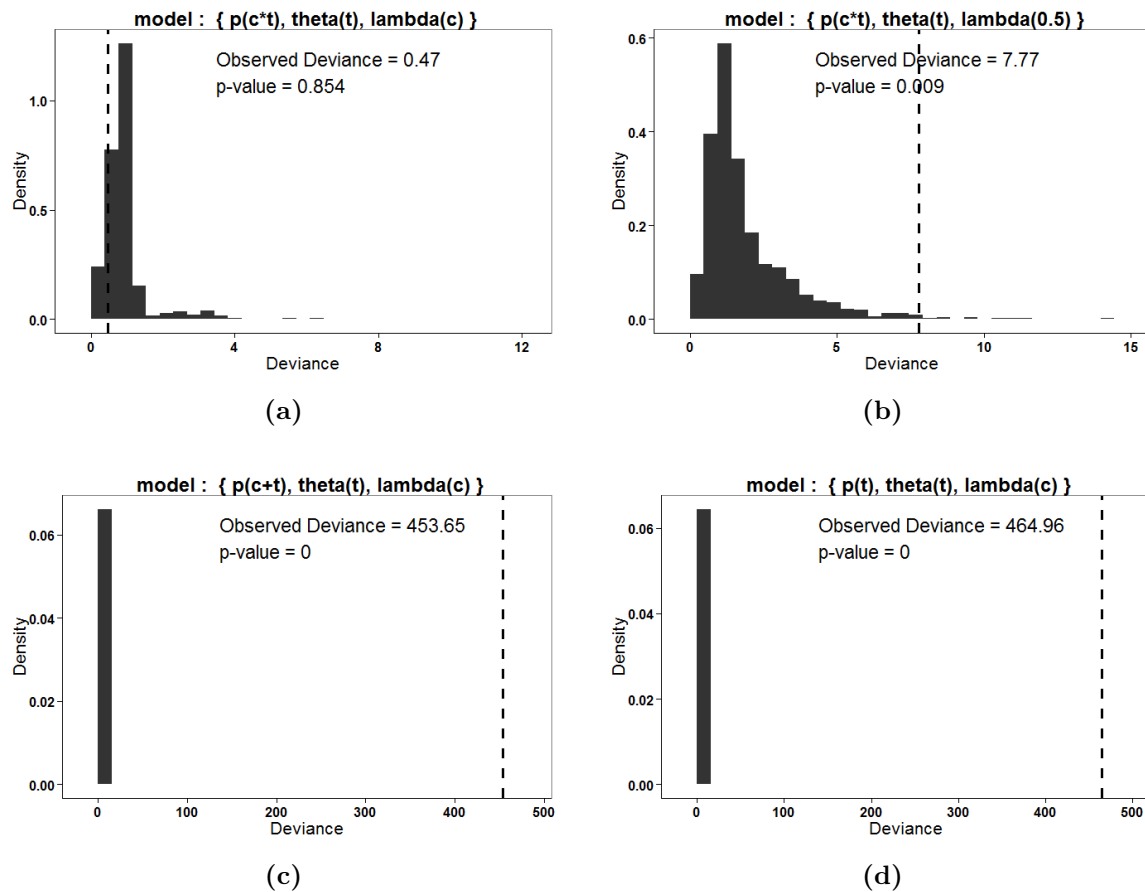


Figure B.2: Parametric bootstrap plots using discrepancy statistic, deviance for 1000 bootstrap samples. The vertical dashed line indicates the deviance for the observed walleye data set.

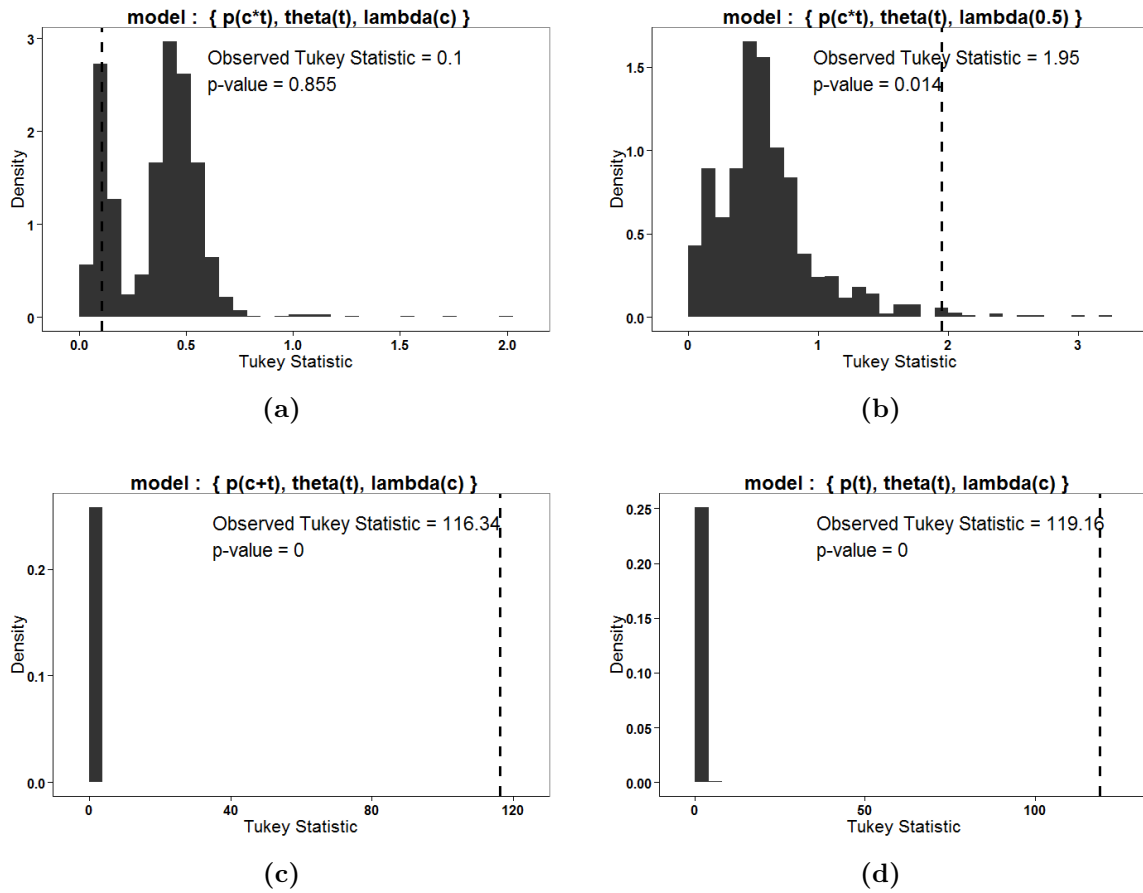


Figure B.3: Parametric bootstrap plots using the Tukey statistics for 1000 bootstrap samples. The vertical dashed line indicates the Tukey statistic for the observed walleye data set.

B.3 Approximate Closed Form Estimates

A closed form solution for the parameter estimates is derived with the use of a moment type estimation method (Davidson and Solomon, 1974). We consider two-sample capture-recapture experiments with two categories (M and F) in the population. The two capture histories UU and $U0$ are combined as U^* , and counts n_{UU} and n_{U0} for the combined histories are combined to n^* .

The moment type estimation method gives

$$n_i = E[n_i] = N \times P_i$$

where $n_i = n^*$, n_{M0} , n_{MM} , n_{F0} , n_{FF} , n_{0M} , n_{0F} , and n_{0U} and $P_i = P_{U^*}$, P_{M0} , P_{MM} , P_{F0} , P_{FF} , P_{0M} , P_{0F} , and P_{0U} respectively.

Then the resulting moment equations are

$$\begin{aligned}
n^* &= N [\lambda_M p_{1M} (1 - \theta_1) + (1 - \lambda_M) p_{1F} (1 - \theta_1)] \\
n_{M0} &= N \lambda_M p_{1M} \theta_1 (1 - p_{2M}) \\
n_{MM} &= N \lambda_M p_{1M} \theta_1 p_{2M} \\
n_{F0} &= N (1 - \lambda_M) p_{1F} \theta_1 (1 - p_{2F}) \\
n_{FF} &= N (1 - \lambda_M) p_{1F} \theta_1 p_{2F} \\
n_{0M} &= N \lambda_M (1 - p_{1M}) p_{2M} \theta_2 \\
n_{0F} &= N (1 - \lambda_M) (1 - p_{1F}) p_{2F} \theta_2 \\
n_{0U} &= N [\lambda_M (1 - p_{1M}) p_{2M} (1 - \theta_2) + (1 - \lambda_M) (1 - p_{1F}) p_{2F} (1 - \theta_2)]
\end{aligned}$$

Closed form solutions are obtained by solving these equations to give:

$$\begin{aligned}
\theta_1 &= \frac{1}{1 + \frac{n^*}{n_{0M} + n_{MM} + n_{0F} + n_{FF}}} \\
\theta_2 &= \frac{1}{1 + \frac{n_{0U}}{n_{0M} + n_{0F}}} \\
p_{1M} &= \frac{1}{1 + \frac{n_{0M}(n_{0M} + n_{MM} + n_{0F} + n_{FF})(n_{0M} + n_{0F} + n_{0U})}{n_{MM}(n_{0M} + n_{0F})(n^* + n_{0M} + n_{MM} + n_{0F} + n_{FF})}} \\
p_{1F} &= \frac{1}{1 + \frac{n_{0F}(n_{0M} + n_{MM} + n_{0F} + n_{FF})(n_{0M} + n_{0F} + n_{0U})}{n_{FF}(n_{0M} + n_{0F})(n^* + n_{0M} + n_{MM} + n_{0F} + n_{FF})}} \\
p_{2M} &= \frac{1}{1 + \frac{n_{M0}}{n_{MM}}} \\
p_{2F} &= \frac{1}{1 + \frac{n_{F0}}{n_{FF}}} \\
\lambda_M &= \frac{1}{1 + \frac{n_{F0}p_{1M}(1 - p_{2M})}{n_{M0}p_{1F}(1 - p_{2F})}} \\
\lambda_F &= 1 - \lambda_M \\
N &= \frac{n_{M0} + n_{MM}}{\lambda_M p_{1M} \theta_1}
\end{aligned}$$

The closed form solutions for the parameter estimates using the model $\{p(c*t)\theta(t)\lambda(c)\}$

for the walleye example are given in Table B.1. These estimates are very similar to the numerical MLEs in Table 3 in Section 3.2.

Table B.1: Closed form MLEs using the model $\{p(c * t) \theta(t) \lambda(c)\}$.

Parameter	Moment estimate
p_{1M}	0.074
p_{1F}	0.011
p_{2M}	0.008
p_{2F}	0.020
λ_M	0.323
λ_F	0.677
θ_1	0.994
θ_2	0.083
N	215,113
N_M	69,385
N_F	145,728

B.4 Optimal Allocation of Sampling Effort

Optimal allocation of the sample sizes for a given cost for partial stratification in two-sample capture-recapture experiments with two categories in the population is discussed in Section 3.3 using Mille Lacs walleye Data 2013.

Optimal allocation of sample sizes and sub-sample sizes produced by the numerical methods for the given costs are $n_1 = 8929$, $n_1^* = 8908$, $n_2 = 8359$ and $n_2^* = 1412$. At these optimal values $SE(\hat{N})$ is 13,657

A conditional contour plot for standard error of \hat{N} is given in Figure B.4 when the n_1^* and n_2^* are fixed at the optimal values. Figure B.5 shows the conditional contour plot for standard error of \hat{N} when the n_1 and n_2 are fixed at the optimal values. These contour plots show that many solutions are possible for optimal allocation.

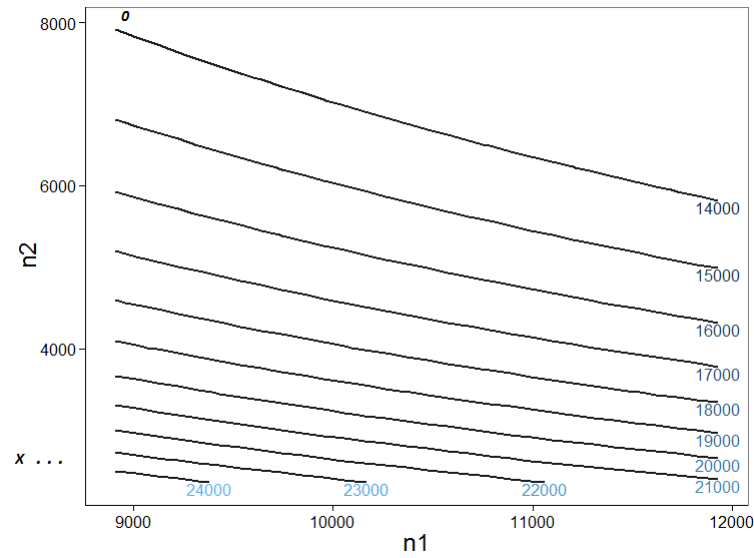


Figure B.4: Conditional contour plot for the standard error of \hat{N} when n_1^* and n_2^* are fixed at the optimal values. ‘ \times ’ represents the current allocation ($n_1 = 6741$ and $n_2 = 3409$) and ‘ o ’ represents the optimal allocation ($n_1 = 8929$ and $n_2 = 8908$). Note that the current allocation falls outside the graph on the left side.

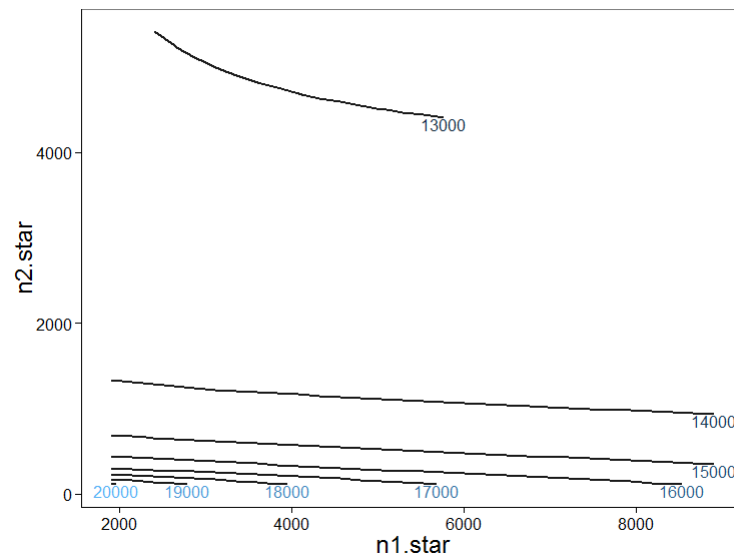


Figure B.5: Conditional contour plot for the standard error of \hat{N} when n_1 and n_2 are fixed at the optimal values.

B.5 Precision of the Estimates when Additional Information is Available

We compare the standard errors (Table B.2) of the estimates for the parameters under two models using the Mille Lacs Lake walleye data to see how well the precision of the estimate of population abundance can be improved in the presence of additional information. Let the first model be $\{p(c * t) \theta(t) \lambda(c)\}$ and the second model be $\{p(c * t) \theta(t) \lambda(MLE)\}$. The first model is the best fitted model to the walleye data under the MLE method 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.

Table B.2: Comparison of the precision of the estimates of two models; the first model is the best model $\{p(c * t) \theta(t) \lambda(c)\}$ and the second model $\{p(c * t) \theta(t) \lambda(MLE)\}$ where the sex ratios are fixed at the MLE values obtained from the best model

Parameter	SE - best model $\{p(c * t) \theta(t) \lambda(c)\}$	SE - second model $\{p(c * t) \theta(t) \lambda(MLE)\}$
p_{1M}	0.015	0.010
p_{1F}	0.002	0.001
p_{2M}	0.001	0.001
p_{2F}	0.004	0.003
λ_M	0.060	-
λ_F	0.060	-
θ_1	0.001	0.001
θ_2	0.005	0.005
N	27,032	26,103
N_M	13,335	8,431
N_F	24,771	17,672

Appendix C. Bayesian Analysis

C.1 Bayesian Analysis Model Comparison Table

Table C.1 shows the DIC values and related information for six different models using two methods as described in Section 4.1.2.

C.2 Bayesian p-value Scatter Plots

Bayesian p-value plots are given in Figure C.1 using discrepancy functions: deviance statistic and Freeman Tukey statistic.

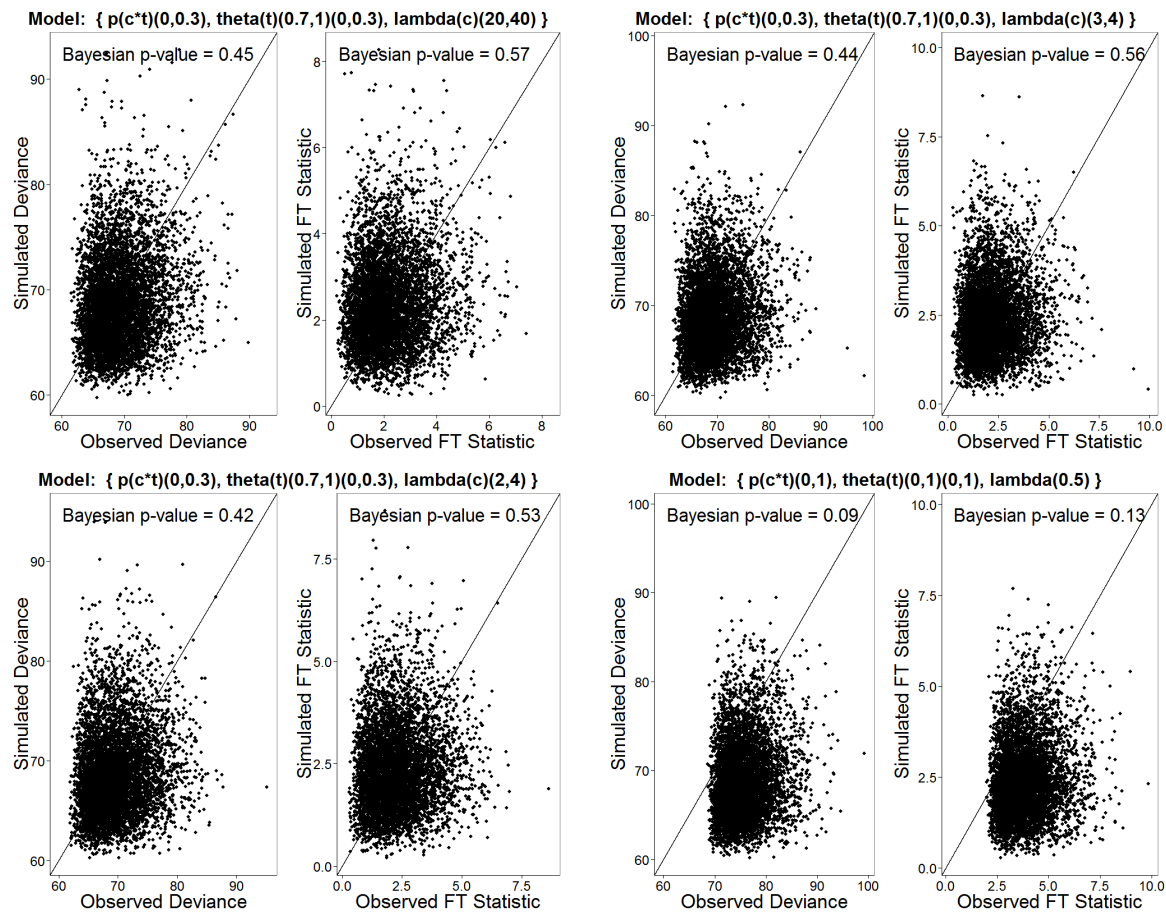


Figure C.1: Bayesian p-value scatter plots for the first four models according to DIC criteria as in Table C.1 using the discrepancy functions (a) Deviance Statistic and (b) Freeman Tukey Statistic (FT). A sample of 6000 (3 chains each with 2000) observations considered from the posterior distribution for each of the models used.

Table C.1: Bayesian analysis model comparison table.

Model	np	Using pD				Using pv			
		pD	DIC	Δ DIC	DIC Weight	pD	DIC	Δ DIC	DIC Weight
$M1 \equiv \{p(c*t)(0, 0.3) \theta(t)(0.7, 1)(0, 0.3) \lambda(c)(20, 40)\}$	8	7.25	76.57	0.05	0.27	8.23	77.55	0.00	0.34
$M2 \equiv \{p(c*t)(0, 0.3) \theta(t)(0.7, 1)(0, 0.3) \lambda(c)(3, 4)\}$	8	7.01	76.52	0.00	0.27	8.93	78.44	0.89	0.22
$M3 \equiv \{p(c*t)(0, 0.3) \theta(t)(0.7, 1)(0, 0.3) \lambda(c)(2, 4)\}$	8	7.23	76.94	0.42	0.22	8.74	78.45	0.90	0.22
$M4 \equiv \{p(c*t)(0, 1) \theta(t)(0, 1)(0, 1) \lambda(c)(7, 3)\}$	8	7.25	76.87	0.35	0.23	8.98	78.60	1.05	0.20
$M5 \equiv \{p(c*t)(0, 1) \theta(t)(0, 1)(0, 1) \lambda(0.5)\}$	7	6.91	82.63	6.11	0.01	7.69	83.40	5.85	0.02
$M6 \equiv \{p(c)(0, 1) \theta(t)(0.7, 1)(0, 0.3) \lambda(c)(2, 4)\}$	6	5.82	1,658	1,581	0.00	6.83	1,659	1,581	0.00

np = number of parameters

C.3 Trace Plots for the Beta Parameters

Consider the best fitted model $\{p(c * t)(0, 0.3) \theta(t)(0.7, 1)(0, 0.3) \lambda(c)(20, 40)\}$. Figure C.2 shows the trace plots for the beta parameters for post burn-in samples.

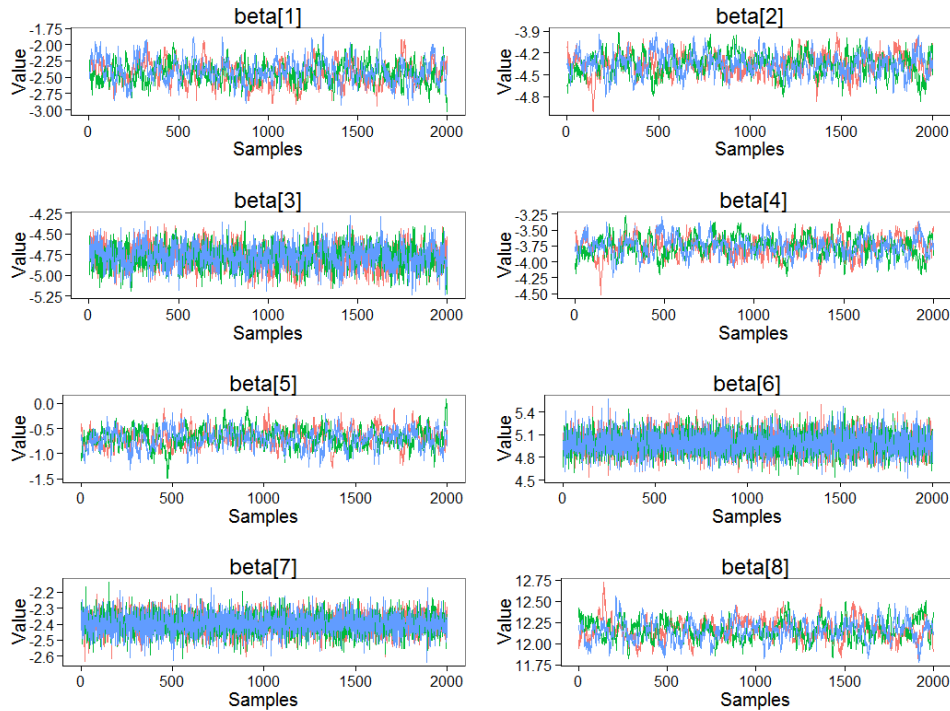


Figure C.2: Trace plots for all the beta parameters using a sample of 6000 (3 chains each with 2000 observations were plotted as overlays in 3 different colours) observations from the posterior distribution for the best model $\{p(c * t)(0, 0.3) \theta(t)(0.7, 1)(0, 0.3) \lambda(c)(20, 40)\}$. Parameters beta[1] to beta[4] correspond to capture probabilities (p_{1M} , p_{1F} , p_{2M} and p_{2F}), beta[5] corresponds to category proportion (λ_M), beta[6] and beta[7] correspond to sub-sample proportions (θ_1 and θ_2), and beta[8] corresponds to population abundance (N). Values in the y-axis are in logit form except for the beta[8], which is in log form.

C.4 Posterior Summary Values

Results produced by the MCMC Metropolis-Hasting 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.

Table C.2: Numerical output of the posterior summary values for the best fitted model $\{p(c * t)(0, 0.3) \theta(t)(0.7, 1)(0, 0.3) \lambda(c)(20, 40)\}$. n.eff is a rough measure of the effective sample size for each parameter. Rhat is the potential scale reduction factor.

Parameter	mean	sd	2.5%	97.5%	Rhat	n.eff
beta[1]	-2.428	0.169	-2.752	-2.077	1.0	170
beta[2]	-4.358	0.147	-4.650	-4.072	1.0	6000
beta[3]	-4.773	0.139	-5.047	-4.505	1.0	630
beta[4]	-3.761	0.150	-4.061	-3.467	1.0	2100
beta[5]	-0.688	0.188	-1.055	-0.326	1.0	350
beta[6]	4.988	0.148	4.709	5.286	1.0	6000
beta[7]	-2.396	0.063	-2.522	-2.274	1.0	6000
beta[8]	12.159	0.117	11.931	12.394	1.0	530
p_{1M}	0.082	0.013	0.059	0.111	1.0	170
p_{1F}	0.013	0.002	0.009	0.016	1.0	6000
p_{2M}	0.008	0.001	0.006	0.011	1.0	630
p_{2F}	0.023	0.003	0.017	0.030	1.0	2100
λ_M	0.336	0.042	0.258	0.419	1.0	340
λ_F	0.664	0.042	0.580	0.742	1.0	350
θ_1	0.993	0.001	0.991	0.995	1.0	6000
θ_2	0.083	0.005	0.074	0.093	1.0	6000
N	192,229	22,923	151,969	241,445	1.0	530
N_M	64,327	9,990	46,250	85,524	1.0	170
N_F	127,902	18,631	96,027	167,791	1.0	6000

Parameters beta[1] to beta[7] are in logit scale and beta[8] is in log scale. Parameters beta[1] to beta[4] correspond to capture probabilities (p_{1M} , p_{1F} , p_{2M} and p_{2F}), beta[5] corresponds to category proportion (λ_M), beta[6] and beta[7] correspond to θ_1 and θ_2 , and beta[8] corresponds to population abundance (N).

C.5 Posterior Distributions for Capture Probabilities, Category Proportions and Sub-Sample Proportions

Figure C.3 gives the posterior distribution of the capture probabilities p_{1M} , p_{1F} , p_{2M} and p_{2F} . Figure C.4 and Figure C.5 give the posterior distributions of the category proportions and the sub-sample proportions.

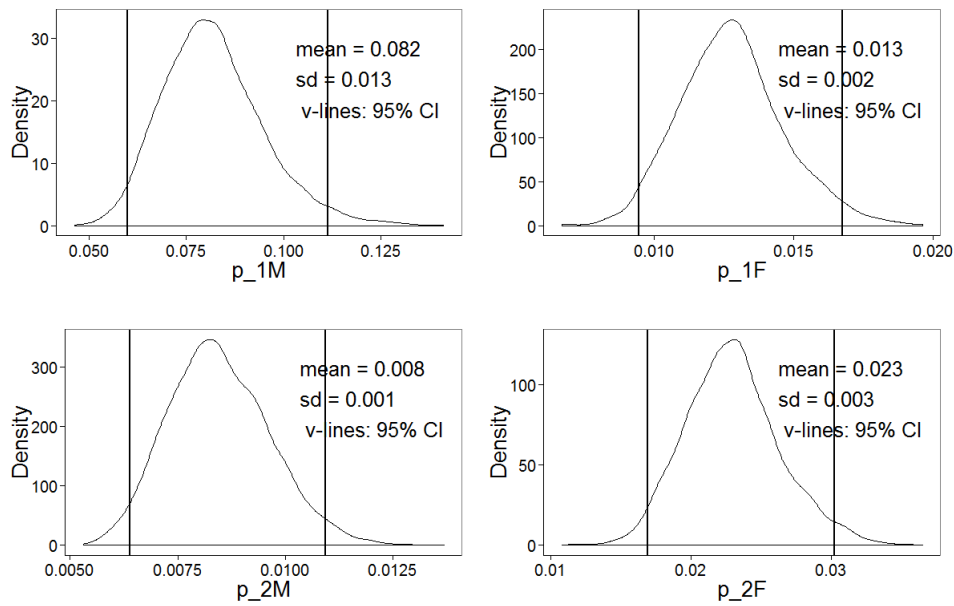


Figure C.3: Posterior distributions for capture probabilities using a sample of 6000 (3 chains each with 2000) observations from the posterior distribution for the best model $\{p(c * t)(0, 0.3) \theta(t)(0.7, 1)(0, 0.3) \lambda(c)(20, 40)\}$. Vertical bars are the 95% credible interval.

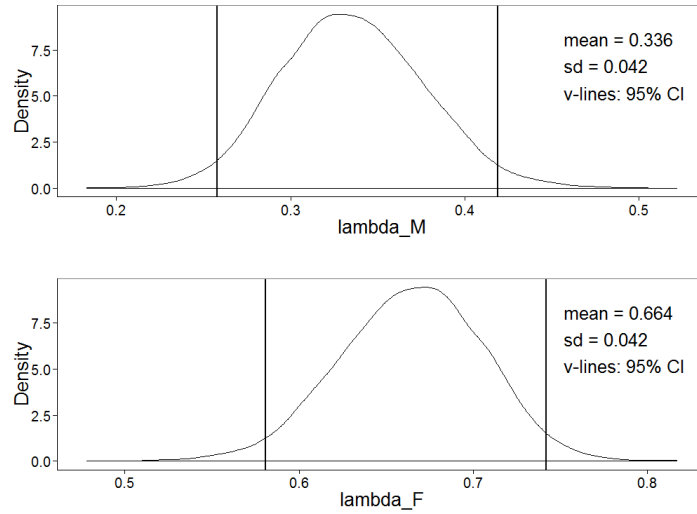


Figure C.4: Posterior distributions for category proportions (λ_M and λ_F) using a sample of 6000 (3 chains each with 2000) observations from the posterior distribution for the best model $\{p(c * t)(0, 0.3) \theta(t)(0.7, 1)(0, 0.3) \lambda(c)(20, 40)\}$. Vertical bars are the 95% credible interval.

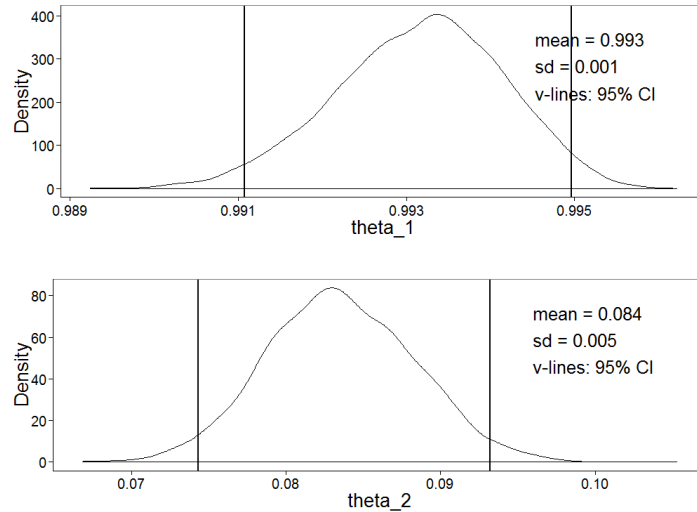


Figure C.5: Posterior distributions for sub-sample proportions (θ_1 and θ_2) using a sample of 6000 (3 chains each with 2000) observations from the posterior distribution for the best model $\{p(c * t)(0, 0.3) \theta(t)(0.7, 1)(0, 0.3) \lambda(c)(20, 40)\}$. Vertical bars are the 95% credible interval.

Appendix D. Analysis with Individual Covariates

D.1 Summary of Distribution of Lengths of Walleyes

The summary distribution of length of Mille Lacs Lake walleyes captured in both sample occasions is given in Figure D.1. More males were captured than females at the first sample occasion and more females were captured than males at the second sample occasion. Sub-sample size was much larger at the first sample occasion compared to the second occasion. Females tended to be larger than males.

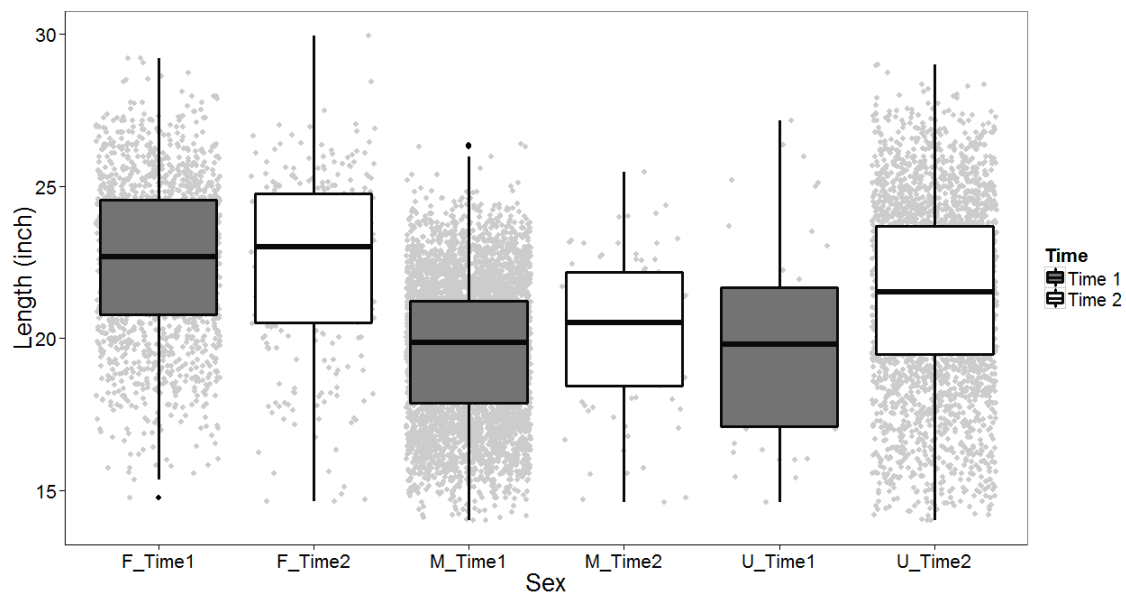


Figure D.1: Summary of distribution of lengths of walleyes in the sample. M, F and U in the x-axis denote male, female and not identify the sex along with they captured in time 1 (first sample occasion) or time 2 (second sample occasion). Dots spread though each box plot represent the actual data points.

D.2 Model Selection

Seven different models were fitted to walleye data (Table D.1). Note that the last two models do not have the individual covariate *length*. These two are the same models given in Table 2 of Section 3. According to the AICc criteria, the best model for the walleye data is the model $\{p(\text{length} * \text{category} * \text{time} + \text{length}^2 * \text{category} * \text{time}), \theta(t), \lambda(c)\}$.

Model specification with the individual covariate *length* is similar to models specified in the MLE method described in Section 2 except for the capture probabilities. Because we

Table D.1: Model comparison table for walleye example with individual covariate ‘length’.

Model	np	\hat{N} ‘000s	$s.e.(\hat{N})$ ‘000s	AICc	$\Delta AICc$	AICc Weights
$\{p(\text{length} * \text{category} * \text{time} + \text{lengthsq} * \text{category} * \text{time}), \theta(t), \lambda(c)\}$	15	279.2	57	21230	0	1.0
$\{p(\text{length} * \text{category} * \text{time} + \text{lengthsq} * \text{category}), \theta(t), \lambda(0.32)\}$	12	326.9	75	21251	21	0.0
$\{p(\text{length} * \text{category} * \text{time} + \text{lengthsq}), \theta(t), \lambda(c)\}$	12	393.3	72	21474	244	0.0
$\{p(\text{length} * \text{category} * \text{time}), \theta(t), \lambda(c)\}$	11	752.3	393	21501	271	0.0
$\{p(\text{lengthsq} * \text{category} * \text{time}), \theta(t), \lambda(c)\}$	11	217.3	28	23261	2031	0.0
$\{p(\text{category} * \text{time}), \theta(t), \lambda(c)\}$	7	209.1	27	23602	2372	0.0
$\{p(\text{category} + \text{time}), \theta(t), \lambda(c)\}$	6	357.1	49	24053	2823	0.0

np = number of parameters

lengthsq represents the length^2 term in the models

Used standardized *length* to reduced the correlation between *length* and length^2

consider a logistic regression model for the capture probabilities that depends on a capture formula, we specify the capture formula in the model specification. Capture formula in the best fitted model was given as $p(\text{length} * \text{category} * \text{time} + \text{length}^2 * \text{category} * \text{time})$. Then the capture probabilities for each individual i for each category C at time t can be found as follows.

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

D.3 Parameter Estimation

The estimates for the parameters and their standard errors using the best model $\{p(\text{length} * \text{category} * \text{time} + \text{length}^2 * \text{category} * \text{time}), \theta(t), \lambda(c)\}$ for the walleye data are shown in Table D.2. Estimates for $\beta_0, \beta_1, \dots, \beta_{11}$ are in logit scale. All the other estimates are in regular scale. We use Equation 1 in Appendix D to calculate the estimated capture probability for each individual.

We use the values 1 and 0 for category ‘M’ and ‘F’ (for males and females) and the values 0 and 1 for ‘time 1’ and ‘time 2’ respectively with the standardized *length* to reduce the correlation between *length* and *length*² in Equation 1 in Appendix D when calculating the estimated individual capture probabilities (Table D.2).

The last two models in Table D.2 do not have the individual covariate *length*. For these models, estimated capture probabilities for all the individuals for a certain category at a certain time are the same. These two models are the same models given in Table 2 of Section 3. The estimates for the parameters for these two models under the analysis with covariates gave the same results that we found in Section 3. However the methods used were different. Having similar results for the same models using different methods shows that the methods used in the analysis with individual covariates are reliable.

Table D.2: MLEs using the model $\{p(\text{length} * \text{category} * \text{time} + \text{length}^2 * \text{category} * \text{time}), \theta(t), \lambda(c)\}$.

Parameter	MLE	SE
$\beta_0 \equiv p : (\text{Intercept})$	-4.660	0.2532
$\beta_1 \equiv p : \text{length}$	0.307	0.0564
$\beta_2 \equiv p : \text{category}_M$	2.608	0.3510
$\beta_3 \equiv p : \text{time}_T2$	0.659	0.0583
$\beta_4 \equiv p : \text{length}^2$	-0.037	0.0135
$\beta_5 \equiv p : \text{length} : \text{category}_M$	-0.519	0.0160
$\beta_6 \equiv p : \text{length} : \text{time}_T2$	-0.136	0.0174
$\beta_7 \equiv p : \text{category}_M : \text{time}_T2$	-3.118	0.2610
$\beta_8 \equiv p : \text{category}_M : \text{length}^2$	-0.059	0.0046
$\beta_9 \equiv p : \text{time}_T2 : \text{length}^2$	0.013	0.0039
$\beta_{10} \equiv p : \text{length} : \text{category}_M : \text{time}_T2$	0.310	0.0619
$\beta_{11} \equiv p : \text{category}_M : \text{time}_T2 : \text{length}^2$	0.031	0.0185
λ_M	0.293	0.068
λ_F	0.707	0.068
θ_1	0.993	0.001
θ_2	0.083	0.005
N_M	81,833	25,398
N_F	197,397	44,625
N	279,230	57,034

Values related to β 's are in *logit* scale and all the other values in regular scale

References

- [1] Davidson, R. R., and Solomon, D. L. (1974). Moment-type estimation in the exponential family. *Communication in Statistics* **3**, 1001–1008.
- [2] Devineau, O., Choquet, R., and Lebreton, J. (2006). Planing capture–recapture studies: Strightforward precision, bias and power calculations. *Wildlife Society Bulletin*, 34, 1028–1035.