1 Abstract

Literature on abundance estimation models using mark-recapture data often states that Bayesian abundance models outperform classical estimators, particularly when recapture numbers are small. We use simulated data to compare the performance the classical Petersen estimator, a likelihood-based estimator we call the "modeled-p" estimator, and a Bayesian estimator. This study was motivated by a particular mark-recapture study from the Kootenai River in Northwestern Montana, and will focus on the sampling design used in that study.

Contents

2 Introduction

The problem of estimating and monitoring wildlife populations is fundamental to the study of ecology and wildlife biology. Possibly the most common tool utilized in these abundance investigations is the mark-recapture study, of which there are many variations. After collecting mark-recapture data, scientists may choose among a large number of possible methods for modeling and estimating abundance from the mark-recapture data. Literature on the topic often claims that Bayesian abundance models outperform classical estimators, particularly when recapture numbers are small. However, Bayesian models can be more difficult to understand for researchers who do not have a strong background in statistics. Additionally, they can be more time consuming to implement in code. Therefore, we have specified a likelihood based estimator we will call the "modeled-p" estimator, which we believe should show performance similar to a Bayesian estimator in the problematic situation in which recapture numbers are small. The goal of this paper is to use simulated mark-recapture data to assess the relative performance of our modeled-p estimator, our Bayesian estimator, and the classical Petersen estimator.

This simulation study was motivated by mark-recapture data collected at four sites downstream of the Libby Dam on the Kootenai River in Northwestern Montana. Thus, we will use data simulated that reflect the structure of the sampling design used in the Kootenai study.

3　Methods

3.1　Sampling Design

A huge variety of sampling designs for mark-recapture studies exists. The motivation for this paper comes from mark-recapture data collected from 1985 to 2015 at four sites (Dam, Flowerpipe, Rereg, and Troy) downstream of the Libby dam on the Kootenai River in Northwestern Montana. Thus, this paper will focus on the sampling design used to collect these data.

Rainbow trout were the primary species of interest in the Kootenia study, and fish were captured by electrofishing during two sampling periods each year. The two sampling occasions typically occurred approximately one week apart. Not every site was sampled every year; there were 54 unique site/year combinations. Fish were not given individually identifying marks for the majority of the study. Captured fish were measured, weighed, fin-clipped, and returned to the section of river from which they were caught. Electrofishing equipment is known to perform poorly for small fish. Therefore, we will assume the probability of capture varies by length. We will consider 10 different small length classes (< 100mm, $[100,150)$, $[150,175)$, $[175,200)$, $[200,225)$, $[225,250)$, $[250,275)$, $[275,300)$, $[300,325)$, and > 325) and four different large length classes (< 150mm, $[150,250)$, $[250,325)$, > 325). All abundance modeling will be done using the small length classes, but we are interested in making inference for the large length classes. Furthermore, since the motivating study conducted separate sampling efforts at four different sites during various years between 1985 and 2015, we will

also allow the probability of capture to vary across sites and years. Thus, each model will produce an estimate for each $N_{length/site/year}$ (denoted N_{lsy} for the remainder of this paper).

3.2 Data Simulation

In order to assess the relative performance of the various estimators outlined in this paper, we simulated data with a structure similar to that outlined above. To simulate our data set, we made 54 random draws from a NegBinom(size $= 10$, $\mu = 5000$) distribution. These represent our true total trout populations for the 54 unique site/years. For each site/year, we then drew N_{sy} lengths from a Gamma(shape $= 2.2$, rate $= .3$) distribution and multiplied each by 20 (a histogram of which is displayed in figure 1, below). These values represent the lengths of the fish in each site/year population. We then binned the fish from each site/year into our 10 length classes, counting how many fish were in each length/site/year. We then simulated $p_{1_{lsy}}$ (the probability of capture on the first pass) and $p_{2_{lsy}}$ (the probability of capture on the second pass) to resemble those observed in the Kootenai data, allowing for variability in these probabilities across years and sites.

For each N_{lsy} we drew one Bin(N_{lsy}, $p_{1_{lsy}}$), yielding a $n_{1_{lsy}}$ (the number of fish captured in the first pass for a given length/site/year), one Bin(N_{lsy}, $p_{2_{lsy}}$), yielding a $n_{2_{lsy}}$ (the number of fish captured in the second pass for a given length/site/year), and one Hyper($n_{1_{lsy}}$, $N_{lsy} - n_{1_{lsy}}$, $n_{2_{lsy}}$), yielding a $m_{1_{lsy}}$ (the number of recaptured fish for a given length/site/year). The first several lines of the simulated data set, with weight omitted, are displayed below in Table 1. The $\overline{length}$ variable is the median of the length class to which the fish belongs.

20*Gamma(shape = 2.2, rate = 0.3)
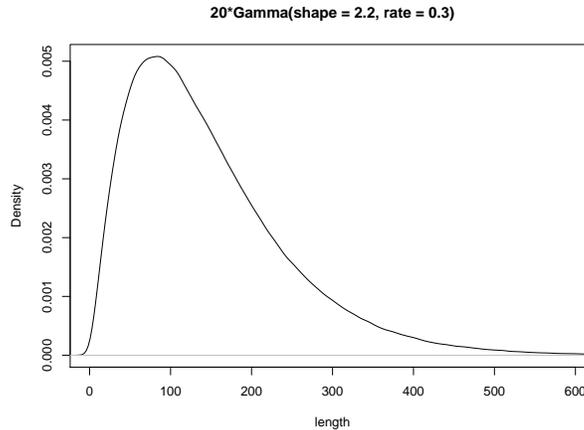
Figure 1: Distribution of Lengths

Table 1: Simulated Data

site	year	length class	n_1	n_2	m	N	p_1	p_2	$\overline{length}$
Dam	1992	[0,100)	36	51	4	662	.058	.079	61.97
Dam	1992	[100,150)	13	24	0	365	.030	.051	123.52
Dam	1992	[150,175)	3	1	0	136	.009	.012	162.01
Dam	1992	[175,200)	0	1	0	117	.014	.008	186.95
Dam	1992	[200,225)	2	0	0	93	.018	.003	211.94
Dam	1992	[225,250)	1	0	0	66	.021	.001	236.85

4 Abundance Estimates

The basic problem when attempting to estimate abundance from such a sampling design is that the p_1, p_2, and N are all unknown and thus must be jointly estimated. To visualize this problem, consider a simplified version in which we assume $p_1 = p_2$ and, thus, n_1 and n_2 are two realizations from a binomial distribution with probability equal to p and total number of trials equal to N (both unknown). Figure 2, plotted below, is the likelihood surface of a binomial distribution, which was generated using simulated data for which the true population size N is 218, the true probability of capture p is 0.1094, $n_1 = 27$, $n_2 = 21$, and $m = 1$. Due to the prominent ridge on the likelihood surface, small errors in estimation

6

Binomial Likelihood: Simulated Data
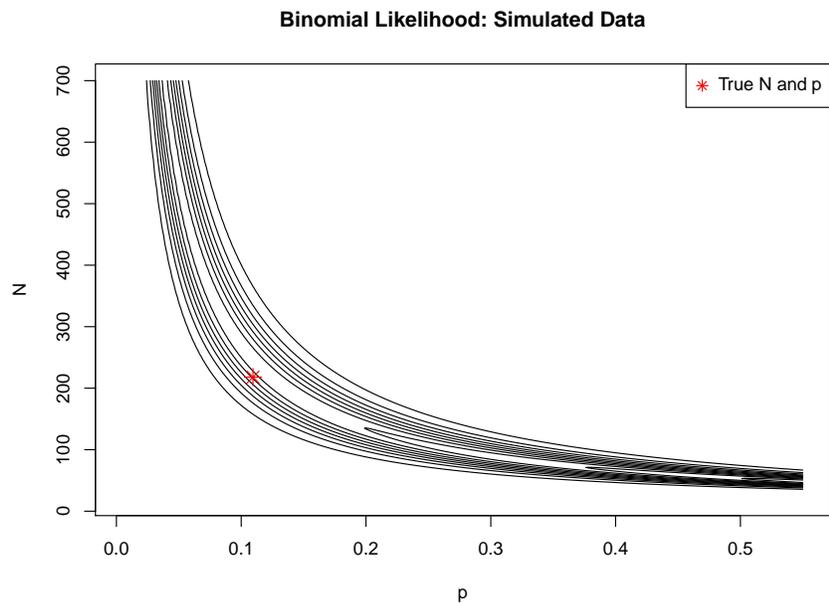
Figure 2: Likelihood Surface

of p can result in drastic errors in the estimation of N.

4.1   Petersen Estimator

Perhaps the simplest mark-recapture estimator available is the Petersen estimator. The Petersen estimator relies on a number of simplifying assumptions (Seber, 1982):

1. The populations is closed, so N is constant from sampling period one to sampling period two (i.e. no animals enter or leave the population from the beginning of the first sampling period to the end of the second).

2. All animals have the same probability of capture.

3. Marking does not affect the catchability of an animal.

4. The second sample is a simple random sample.

7

5. Animals do not lose their marks in the time between the two samples.

6. All marks are reported on recovery in the second sample.

When these assumptions hold, it is reasonable to equate the proportion of marked animals in the second sample to the proportion of captured animals in the first sample. Thus, for each length/site/year combination, we have

$$\frac{m_{lsy}}{n_{2_{lsy}}} = \frac{n_{1_{lsy}}}{\hat{N}_{lsy}} \implies \hat{N}_{lsy} = \frac{n_{1_{lsy}} n_{2_{lsy}}}{m_{lsy}}$$

The Petersen estimator has the benefit of being very simple to calculate, but the drawback of being so simple as to show poor performance in some situations. Indeed, Seber states, "The properties $\hat{N}$ ... have been discussed fully by Chapman [1951]. He shows that although $\hat{N}$ is a best asymptotically normal estimate of N as $N \to$ inf, it is biased, and the bias can be large for small samples," (Seber, 1982). Consider the likelihood surface displayed in figure 2. Above, in figure 3, the same surface is displayed with the Petersen estimator added. For the particular values of n_1, n_2, and m given above, the Petersen estimator is more than two and a half times greater than the true value of N, due in large part to the fact the m is small (equal to 1, in this case), which results in an underestimation of $\hat{p}$.

4.2   Modeled-p Estimator

One of the primary issues with the Petersen estimator is that it performs especially poorly when m is small (and is undefined when $m = 0$). Essentially, this estimator is deriving $\hat{N}$ by dividing the number captured in the first sample by the probability of capture in the second sample, which is estimated by $\frac{m}{n_2}$. If the probability of capture is small (as is typically the case when m is very small), we approach the left hand portion of our likelihood surface,

Binomial Likelihood: Simulated Data
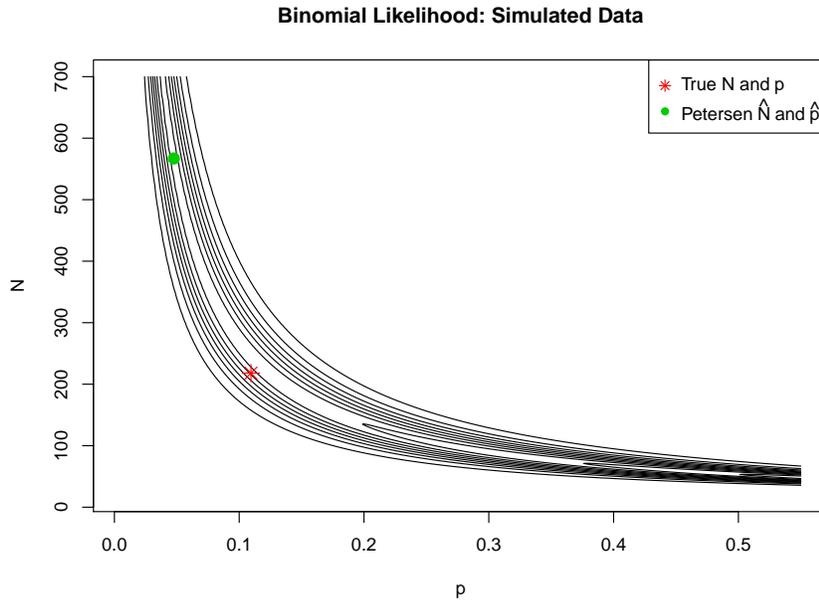
Figure 3: Petersen Estimator

where small shifts in the estimated $\hat{p}$ result in very large shifts in the estimated $\hat{N}$. For mark-recapture data on fish, this is especially problematic for small length classes, where we know the probability of capture to be small, and very large length classes, where there are simply very few fish to capture, which can sometimes result in small sample sizes and thus small m values. Our modeled-p estimator attempts to address this problem by sharing information about the probability of capture for a given length class across sites and years.

When developing the modeled-p estimator for the original data, our aim was to model the probability of capture with length, site, and year effects. To this end, we used a logistic regression model where those fish in the second sample who are recaptures from the first are considered "successes" and the remaining fish in the second sample are considered "failures." We observed a roughly linear trend in the Kootenai data between the log odds of capture
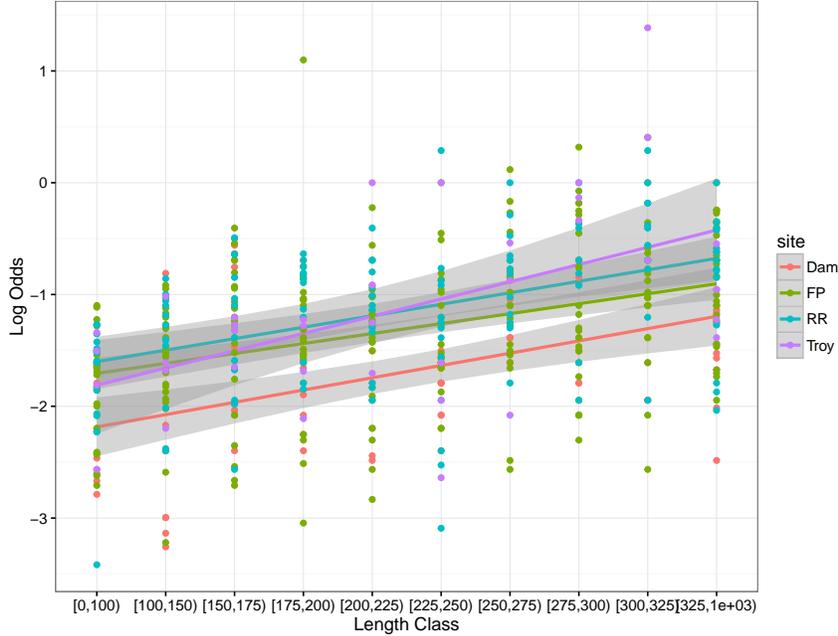
Figure 4: Log Odds and Length Class by Site

in the second sample and length, and thus included in our logistic regression model a fixed effect for length (see figure 4, above). Similarly, we observed in the Kootenai data evidence that different sites had different intercepts, but not different slopes on length. Therefore, we also included a fixed effect for site. Finally, we created a site by year interaction variable, which was included as a random effect. This helps account for some of the heterogeneity in capture probability due to unaccounted for variables. The resulting model is

$$p_{lsy} = logit^{-1}(\beta_0 + \beta_1 length + \beta_2 site_s + \alpha_{sy})$$

This model provides us with a probability distribution that our expected capture probabilities follow. That is, we assume $\boldsymbol{\beta} \sim N(\hat{\boldsymbol{\beta}}, \sigma^2(X^t X)^{-1})$.

Our next step was to simulate random $\hat{p}$ draws from this distribution, giving us a sense of the uncertainty in our $\hat{\boldsymbol{\beta}}$ estimates. We used Gelman's `sim()` function in the `arm` package

(Gelman and Su, 2015). This function works by making use of the fact that, for a random sample from a Normal distribution, $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)}$. For each simulation, the function makes a random draw from a χ^2 distribution, uses this draw to compute a new $\hat{\sigma}^2$, and then generates new $\hat{\beta}$ estimates from the new $\hat{\sigma}^2$. We then calculated a new expected $\hat{p}$ for each length/site/year from each simulated regression curve. Using these new $\hat{p}$'s obtained from our simulated regression curves, we can derive population estimates in the manner of the Petersen estimator. To illustrate this process, suppose we simulate 1000 regression curves using the `sim()` function. We will then predict 1000 $\hat{p}^*$'s for each length/site/year. For each of these $\hat{p}^*$'s, we will calculate an $\hat{N}^* = \frac{n_1}{\hat{p}^*}$, giving us a kind of "distribution" of $\hat{N}$'s. We will take the modeled-p estimator to be the mean of the $\hat{N}^*$'s. Furthermore, we can obtain an uncertainty interval around this estimate.

We believe our modeled-p version of the Petersen estimator should show much better performance than the classical Petersen estimator due to the sharing of information about the probability of capture within each length class across sites and years.

4.3 Bayesian Estimator

Our Bayesian estimator is similar in structure to the modeled p estimator, and the model is fully specified below. Since the Bayesian model specification summarizes the data by capture histories, a brief overview of capture histories follows.

For a study in which two sampling efforts are made each year, possible capture histories are as follows:

- capture history (1,0) corresponds to a fish captured on the first pass, but not captured on the second pass

- capture history (0,1) corresponds to a fish not captured on the first pass, but captured on the second pass

- capture history (1,1) corresponds to a fish captured on both the first and second passes

- capture history (0,0) corresponds to a fish captured on neither the first nor the second pass (this is the only capture history which is unobservable, and hence not observed in the data)

Let x_1 be the number of fish for a given length/site/year that display capture history (1,0), x_2 be the number of fish that display capture history (0,1), and so on. Taken together, x_1, x_2, and x_3 follow a multinomial distribution. That is, $(x_1, x_2, x_3) \sim Multi(p_{10}, p_{01}, p_{11}; N_{cap})$ where

$$f(x_1, x_2, x_3) = \frac{N!}{x_1! x_2! x_3!} p_{10}^{x_1} p_{01}^{x_2} p_{11}^{x_3}$$

Let p_1 equal the probability of captured on the first pass, and p_2 equal the probability of capture on the second pass. Then the parameter $p_{10} = p_1(1 - p_2)$, $p_{01} = (1 - p_1)p_2$, $p_{11} = p_1 p_2$, and $p_{00} = (1 - p_1)(1 - p_2)$. The parameter N_{cap} is equal to the sum of the x_i's. In the Bayesian model specified below, X_{lsy} is the vector of capture histories counts ($x_{1_{lsy}}$, $x_{2_{lsy}}$, $x_{3_{lsy}}$).

$$X_{lsy} \sim Multi(p_{01_{lsy}}, p_{10_{lsy}}, p_{11_{lsy}}; N_{cap_{lsy}})$$

$$N_{cap_{lsy}} = n_{1_{lsy}} + n_{2_{lsy}} - m_{lsy}$$

$$N_{cap_{lsy}} \sim Pois(\lambda_{lsy}), \quad \lambda_{lsy} \sim Gamma(0.1, 0.1)$$

$$N_{lys} = \frac{\lambda_{lsy}}{p_{cap_{lsy}}}, \quad p_{cap_{lsy}} = p_{1_{lsy}} + p_{2_{lsy}} - p_{1_{lsy}}p_{2_{lsy}}$$

$$logit(p_{1_{lsy}}) \sim N(\mu_{ls} + \alpha_{ys}, \sigma_p), \quad \sigma_p \sim Half\ Cauchy$$

$$logit(p_{2_{lsy}}) \sim N(\mu_{ls} + \alpha_{ys}, \sigma_p), \quad \sigma_p \sim Half\ Cauchy$$

$$\mu_{ls} \sim N(\beta_{0s} + \beta_{1s}length, \sigma_s), \quad \sigma_s \sim Half\ Cauchy$$

$$\beta_0 \sim Cauchy(0, 10), \quad \beta_1 \sim Cauchy(0, 2.5)$$

$$\alpha_{ys} \sim N(0, \sigma_{ys}), \quad \sigma_{ys} \sim Half\ Cauchy$$

We assume our three observed capture histories follow a multinomial distribution with unknown capture probabilities ($p_{1_{lsy}}, p_{2_{lsy}}$, and $p_{1_{lsy}}p_{2_{lsy}}$) and given total number captured ($Ncap_{lsy}$). $Ncap_{lsy}$ is assumed to follow a Poisson distribution with rate parameter λ_{lsy}, and λ_{lsy} is assumed to follow a Gamma(0.1, 0.1) distribution. This prior on λ_{lsy} was chosen for two reasons. First, the gamma distribution is a common choice of prior for Poisson rate parameters, and a Gamma(0.1, 0.1) is considered a standard choice of "vague" prior. Second, Raftery suggests $\frac{1}{\lambda}$ as a choice of prior in a similar model (Raftery, 1988). The Gamma(0.1, 0.1) distribution has a similar shape but is proper and thus supported by JAGS. The abundance estimate N_{lsy} is defined to be $\lambda_{lsy}/p_{cap_{lsy}}$, where $p_{cap_{lsy}}$ is the probability of a fish begin captured on one or both passes. Similar to the modeled-p estimator, the logit transformed $p_{1_{lsy}}$'s and $p_{2_{lsy}}$'s are modeled using length/site and site/year effects

$(logit(p_{i_{lsy}}) \sim N(\mu_{ls} + \alpha_{ys}, \sigma_p))$. The length/site effect is assumed to follow a normal distribution with mean equal to $\beta_{0s} + \beta_{1s}length$, where $\beta_0 \sim Cauchy(0, 10), \beta_1 \sim Cauchy(0, 2.5)$, and length is standardized (Gelman et al., 2008). We consider the site/year effect to be random, and thus assign it a normal distribution centered at zero. All variance parameters are given folded Cauchy prior distributions, which are considered vague. This choice of prior is further supported by Gelman's discussion of choice of prior distribution for hierarchical variance parameters in his 2006 paper (Gelman et al., 2006).

5   Simulation Study

In order to compare the performance of the three estimators, we generated $\hat{N}_{lsy}$ estimates from each model using our simulated data. This approach allows us to directly compare estimates to one another and the the true parameter values.

5.1   Obtaining MCMC Draws Using JAGS

The posterior distributions for each N_{lsy} and $p_{cap_{lsy}}$ were obtained using the Gibbs sampler JAGS (Plummer, 2003). We then took the posterior mean of each N_{lsy} to be our Bayesian abundance estimator. Code for our JAGS program can be found in the appendix. We ran three chains for 17000 iterations each and discarded the initial 2000 iterations from each chain, resulting in a total of 45000 draws from each posterior distribution. We assessed convergence using the Gelman-Rubin diagnostic $\hat{R}$, the largest of which was 1.06. Since we have 540 posterior N distributions and 540 posterior p_{cap} distributions, we did not examine

sample path plots for each. However, we did construct sample path plots for parameters with the largest $\hat{R}$'s, to ensure proper mixing and full exploration of the parameter space.

6   Discussion

We can compare the performance of our three estimators by constructing caterpillar plots for each length/site/year combination. The original Kootenia data were broken into four distinct length classes: < 150mm, between 150 and 250mm, between 250 and 325mm, and > 325mm. We will use these larger length classes when constructing our plots, both for the purpose of reducing the number of plots and because these are the length classes about which we were originally interested in making inference. The plots are grouped by length class and site and can be found in the appendix. The plot of abundance estimates for each larger length/site combination is grouped with its associated smaller length/site plots for probability of capture estimates. An example of this grouping is the Dam site for length class <150mm plotted below with its associated $[0, 100)$ and $[100, 150)$ length class estimated probability of capture plots.

Our Bayesian model works by defining $N_{lsy} = \frac{\lambda_{lsy}}{p_{cap_{lsy}}}$, where λ_{lsy} is the expected number captured for a given length/site/year, and $p_{cap_{lsy}}$ is the probability that a fish is ever captured $(p_{1_{lsy}} + p_{2_{lsy}} - p_{1_{lsy}}p_{2_{lsy}})$. We can see intuitively that this estimator should produce accurate estimates of N_{lsy} when the model does a good job of estimating $p_{cap_{lsy}}$. If $\hat{p}_{cap_{lsy}}$ is estimated too low, the resulting $\hat{N}_{lsy}$ will be too large, and vise versa. Similarly, the precision (width of the posterior interval) with which $p_{cap_{lsy}}$ is estimated determines the precision with which

Length < 150mm: Dam
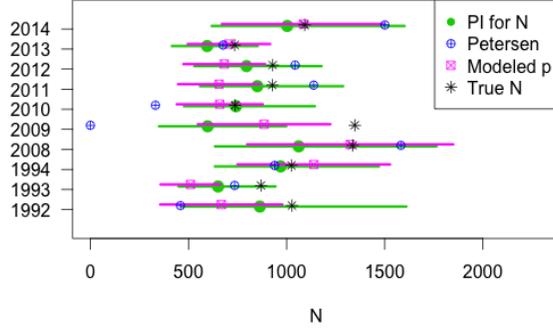
Dam: [0,100)
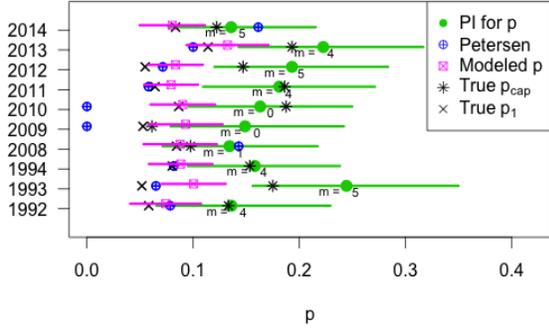
Dam: [100,150)
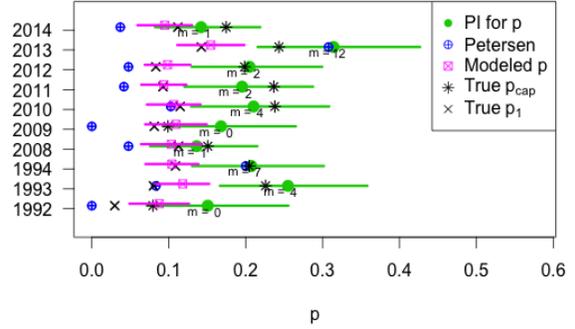
N_{lsy} can be estimated.

The modeled-p estimator works by assuming $p_{1_{lsy}} = \hat{p}_{2_{lsy}}$, where $\hat{p}_{2_{lsy}}$ is the probability of capture predicted from our logistic regression model incorporating length, site, and year effects. We obtain our estimated N_{lsy}'s by computing $\frac{n_1}{\hat{p}_{2_{lsy}}}$. Therefore, we can see intuitively that this estimator should produce accurate estimates of N_{lsy} when $\hat{p}_{2_{lsy}}$ is a good estimator of $p_{1_{lsy}}$.

Our classical Petersen estimator works by assuming $p_{1_{lsy}} = \hat{p}_{2_{lsy}}$, where $\hat{p}_{2_{lsy}} = \frac{m}{n_2}$. We obtain our estimated N_{lsy} by computing $\frac{n_1}{\hat{p}_{2_{lsy}}}$ and therefore, this estimator should produce

accurate estimates of N_{lsy} when $\hat{p}_{2_{lsy}}$ accurately estimates $p_{1_{lsy}}$. It tends to show every poor performance when m is small, and is undefined when $m = 0$. We can see from the plots on the following pages and in the appendix that its overall performance is erratic: for some length/site/years, it estimates $\hat{N}$ almost perfectly, while in other cases it over-estimates $\hat{N}$ by two or three times.

6.1   Uncertainty Intervals

In general, it appears that the confidence intervals associated with our modeled-p estimator are smaller than the posterior intervals associated with our Bayesian estimator. We suspect our choice of vague priors in the Bayesian model contributes to the width of the intervals. Furthermore, the problem appears to be at its worst when recapture counts are small. The variability in the modeled-p estimator, on the other hand, comes entirely from simulating new variance parameters using the identity $s = \hat{\sigma}\sqrt{(n-k)/X^2}$, where X^2 is a random draw from a $\chi^2_{(n-k)}$ distribution. The width of each confidence interval for the predicted $\hat{p}$'s is approximately $4\hat{\sigma}$ (note that $\hat{\sigma}$ is the sum of two sources of variation: the overall variation in the $\hat{p}$'s and the group level variation from the site/year random effect). However, since we obtained N_{lsy} estimates by dividing $n_{1_{lsy}}$ by the $\hat{p}_{lsy}$ predicted from the model, the confidence intervals for the resulting N's come directly from the confidence intervals from the p's. Therefore, when $n_{1_{lsy}}$ is large, the resulting confidence interval will also be large. To more clearly illustrate this idea, suppose the model predicts a 95% confidence interval for a particular length/site/year $\hat{p}$ to be 0.2 to 0.4. Consider the difference between the resulting

intervals on N if the associated n_1 is 3 (7.5 to 15) or if the the associated n_1 is 20 (50 to 100).

6.2 Modeled-p Performance When p_1 and p_2 Differ

Ignoring, for a moment, the associated intervals, let us consider the relative accuracy of the Bayesian posterior mean and the modeled-p estimator. Since the modeled-p estimator assumes $p_{1_{lsy}} = \hat{p}_{2_{lsy}}$, it stands to reason that it will perform poorly if capture probability varies between the first and second passes. For our simulated data, the largest differences between first and second pass capture probabilities occurred in the Flowerpipe section in 1986, for which all length classes showed differences greater than 0.11. However, for each length class, the 1986 modeled-p estimator for Flowerpipe showed good performance, as we can see in the plots on the following page. This seems to indicate that differences between $p_{1_{lsy}} = \hat{p}_{2_{lsy}}$ need to be fairly large (i.e. greater than 0.11) for this issue alone to result in poor performance of the modeled-p estimator.

6.3 Estimator Comparison for Zero Recapture Count Events

As noted earlier, the primary motivation for using a Bayesian model for abundance estimation is that Bayesian estimators are supposed to show better performance when recapture numbers are small. Our motivation for developing the modeled-p estimator is that it may show performance as good or almost as good as the Bayesian estimator, but with the added benefit of being easier to understand and explain to researcher without a strong background

18

Length < 150mm: FP

Length Between 150 and 250mm: FP

Length Between 250 and 325mm: FP

Length > 325mm: FP
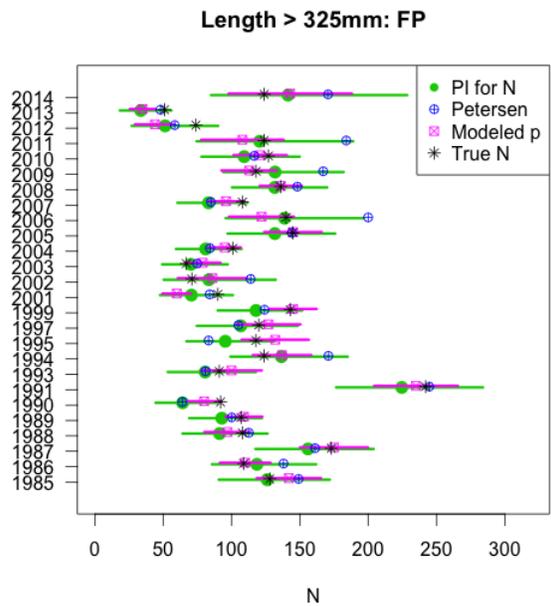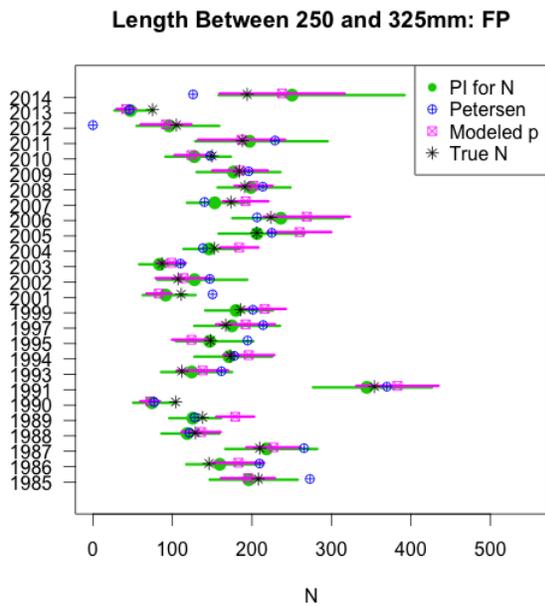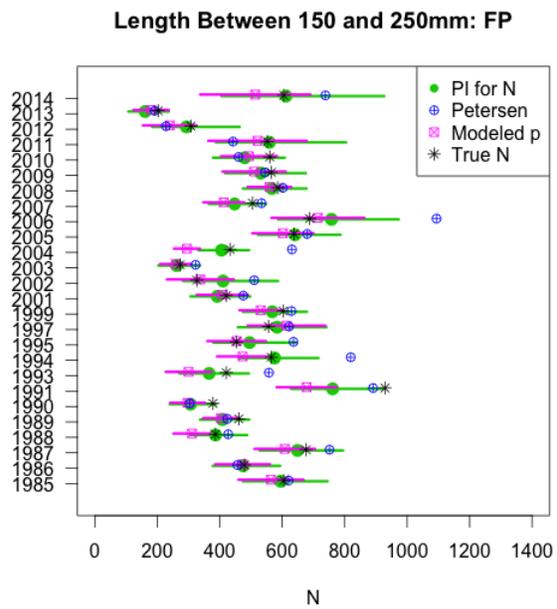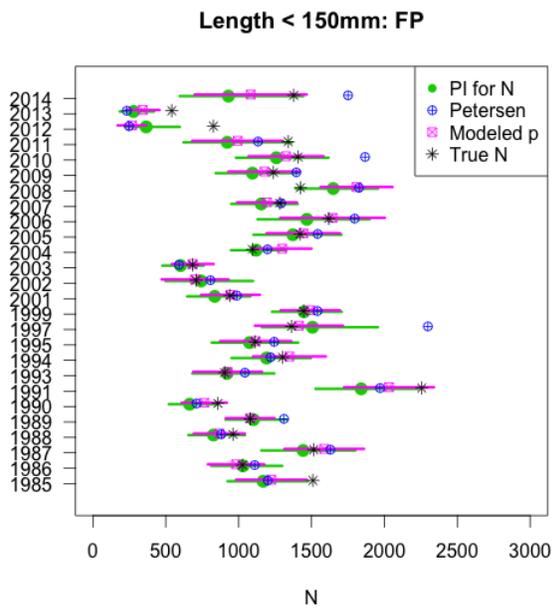
in statistics. In the interest of exploring the relative performance of these two estimators, we will examine in depth the Dam 1992 site/year, for which eight out of the 10 small length classes had recapture numbers of zero in the simulated data. Caterpillar plots for the Dam section are displayed below.

Clearly, both the Bayesian and modeled-p estimators showed poor performance in 1992, particularly in the middle two length classes. In the < 150 mm length class, one out of two of the smaller associated length classes had a recapture count of 0, in the $[150,250)$ and $[250,325)$ mm length classes, all of the associated smaller length classes had recapture counts of zero, and in the >325 length class, none of the associated smaller length classes had recapture counts of zero. Therefore, it appears that when all sub-length classes have recapture counts of zero, both models struggle to produce good abundance estimates. To further examine why these models are performing poorly, we will consider the associated probability plots, also plotted below.

Length < 150mm: Dam

Length Between 150 and 250mm: Dam

Length Between 250 and 325mm: Dam

Length > 325mm: Dam
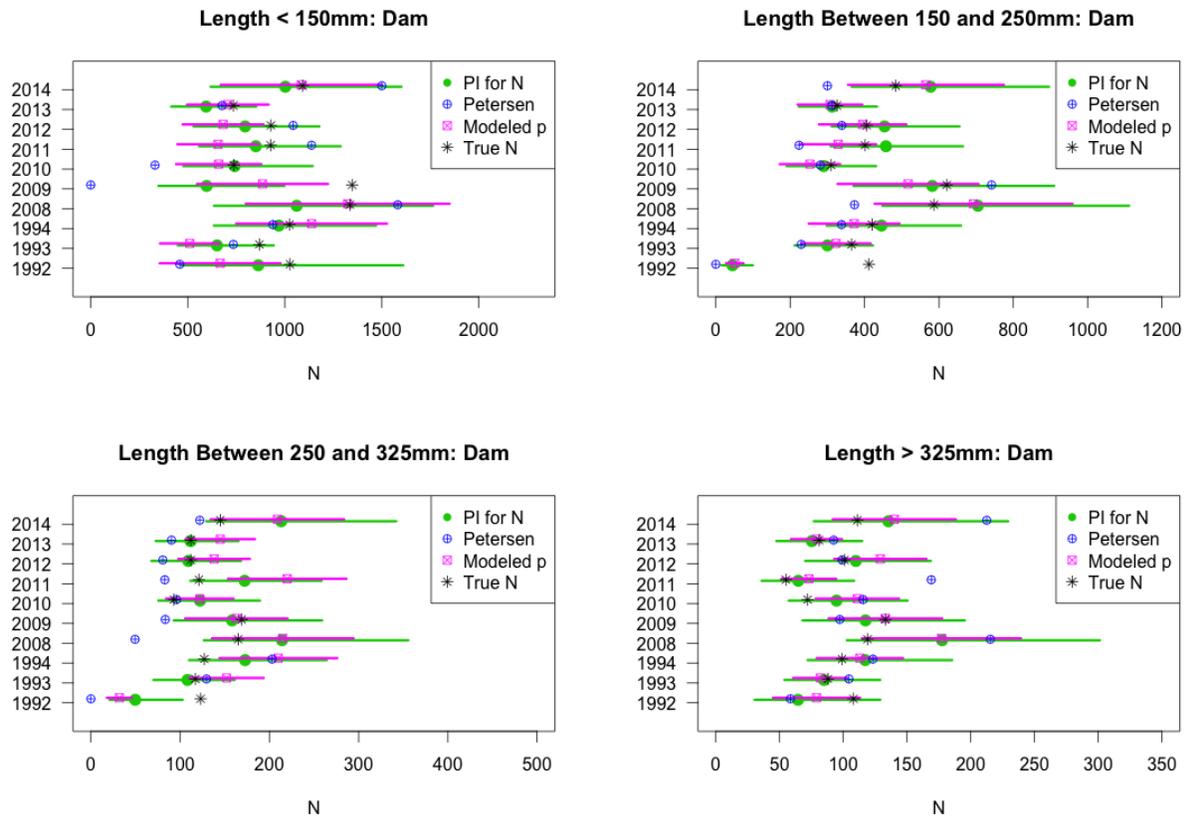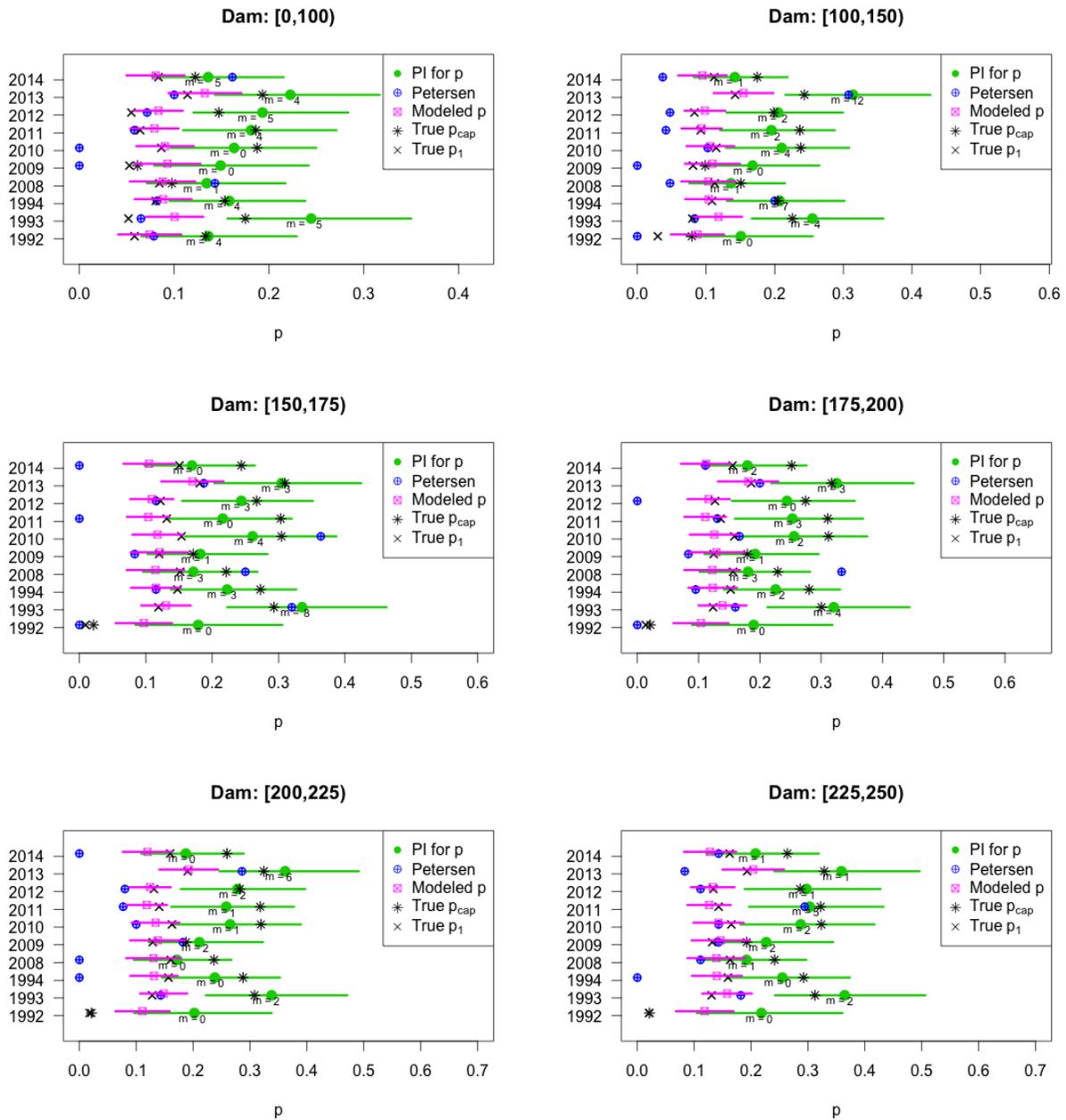
Recall that the Bayesian model is attempting to model p_{cap}, the probability of ever being captured, while the modeled-p model is attempting to model the probability of capture on the first pass. From these probability plots, we make two notable observations. First when $m = 0$, both models are far over-estimating their respective targets. Second, for the length classes in which the models are showing the poorest estimates of p_{cap} and p_1, the true capture

Dam: [250,275) Dam: [275,300) Dam: [300,325) Dam: >325
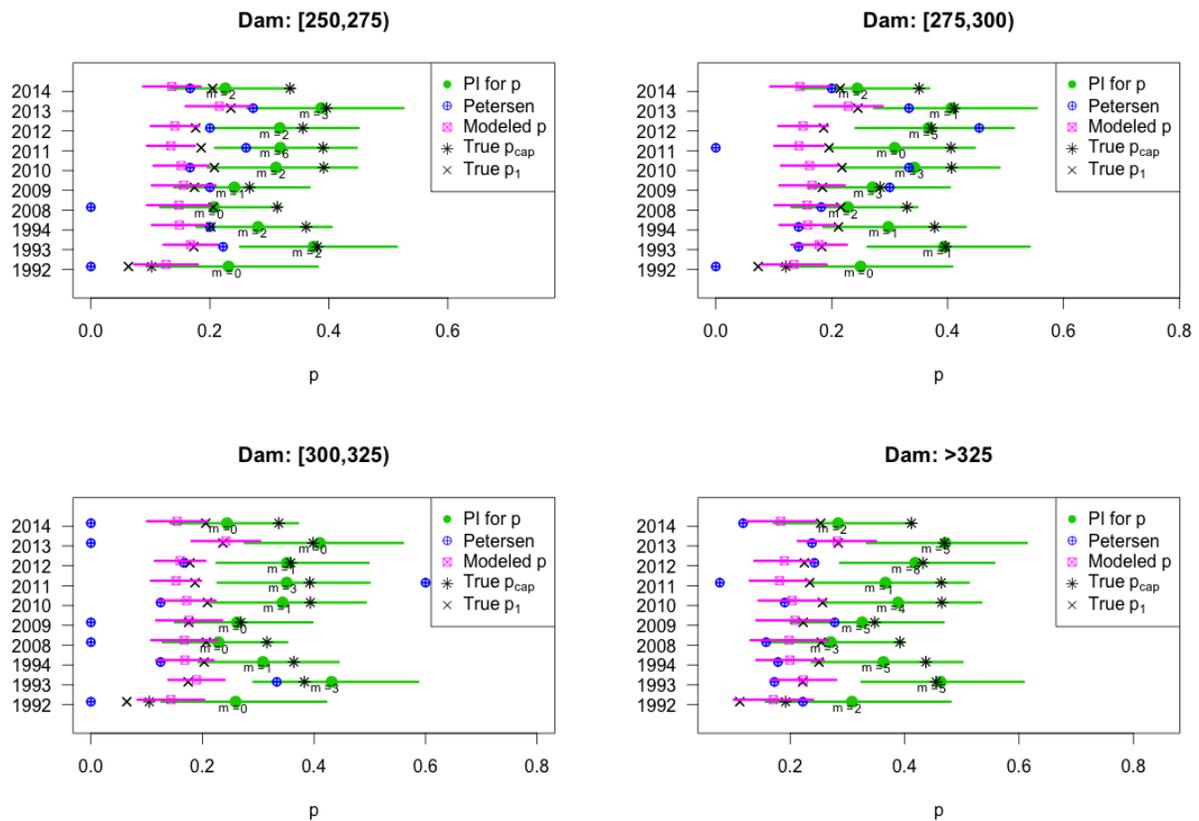
probabilities are much smaller than in the other years for the Dam site. This is interesting because it appears that the models are actually performing exactly as we had hoped: when a recapture count of zero is observed, the models are using information pooled from other years to estimate a reasonable probability of capture, instead of naively assuming it to be zero, as the classical Petersen estimator does. However, this strategy does not work well in this case, because the 1992 probabilities of capture in the Dam site are so unlike the other years for that site. However, we see other length/site/years for which $m = 0$ and the models do a fine job of estimating the probabilities of capture (such as [175,200)/Dam/2012). This seems to happen when these capture probabilities are more similar to other capture probabilities for the site in question. This suggests that it is not so much the recapture counts of zero that are the problem, but the unusually small probabilities of capture.

22

We believe the poor performance of both the modeled-p and Bayesian estimators in site/years for which the probabilities of capture are unusually low is notable because it suggests these models probably perform best when probabilities of capture are relatively similar across years, and may perform quite poorly if probabilities of capture are very different across years. However, when working with real data, there is no way of knowing which situation is the truth. The best way to avoid the latter scenario is to carefully design a study in which the sampling effort is as uniform as possible from year to year, while also attempting to control for other variables that may affect the probability of capture.

7 Further Work

There are a variety of ways in which we could extend this work in the future. First, the problems associated with the Petersen estimator are well known. Therefore, its comparatively poor performance is not surprising. It would be interesting to compare the modeled-p and Bayesian estimators to a classical estimator with better properties than the Petersen estimator.

Second, there are a variety of ways in which we could adapt our Bayesian model. For instance, we chose to use only vaguely informative priors for our Bayesian estimator. However, it seems reasonable to assume that researchers may be able to provide some amount of prior information for at least some of the parameters in the model. Exploring the effect of informative priors, particularly on the precision of the Bayesian estimator, would be worthwhile.

We could also consider adjusting the structure of the Bayesian model slightly. Specifically, our model specifies

$$logit(p_{1_{lsy}}) \sim N(\mu_{ls} + \alpha_{ys}, \sigma_p), \quad \sigma_p \sim Half\ Cauchy$$

$$logit(p_{2_{lsy}}) \sim N(\mu_{ls} + \alpha_{ys}, \sigma_p), \quad \sigma_p \sim Half\ Cauchy$$

$$\mu_{ls} \sim N(\beta_{0s} + \beta_{1s}length, \sigma_s), \quad \sigma_s \sim Half\ Cauchy$$

$$\alpha_{ys} \sim N(0, \sigma_{ys}), \quad \sigma_{ys} \sim Half\ Cauchy$$

However, it may be interesting to explore the effect of either of the following modifications. First we could define the logit transformed probabilities of capture to be equal to $(\mu_{ls} + \alpha_{ys}$ and putting a normal distribution only on μ_{ls} and α_{sy}. Second, we could define μ_{ls} to be equal to $\beta_{0s} + \beta_{1s}length$ and put normal distributions only on the logit transformed probabilities and α_{sy}. We believe either of these options would likely result in narrower posterior intervals. The former would more closely resemble the modeled-p model.

Agresti and Caffo explore the effect of adding one success and one failure when computing confidence intervals for proportions (Agresti and Caffo, 2000). It may be worthwhile to explore the impact a similar approach would have on the modeled-p and Petersen estimates.

We produced abundance estimates for our three chosen estimators using only one simulated data set. It could be interesting to generate hundreds or thousands of data sets under the same assumptions, and then produce estimates for each method and each data set. This would give us a sense of how the performance of our estimators may vary due to natural

sampling variability.

Finally, it could be worthwhile to explore the effect of changing some of the assumptions under which we generated the data. For instance, we could induce more variability in the capture probabilities both between sampling events within a site/year and across sites and years. We could also examine the effect of smaller or larger true population sizes or different distributions of lengths. Since the Petersen estimator is known to have larger bias when samples are small, and both our modeled-p and Bayesian estimators are based on the Petersen estimator, exploring the bias in these estimators as it relates to sample size may be enlightening.

8   Conclusions

Going through the process of simulating data and fitting different models provided a lot of insight into this problem. In most length/site/year combinations, both the Bayesian and modeled-p estimators performed fairly well. Additionally, we identified conditions under which we believe these estimators may perform poorly. We examined the difference between the uncertainty intervals associated with the Bayesian and modeled-p estimators, noting that the posterior intervals are much more intuitive, while the uncertainty intervals for the modeled-p estimator are difficult to interpret. Provided that it is reasonable to assume the probability of capture is relatively uniform across years within a site, both estimators appear to perform very well. Under this condition, the modeled-p estimator is a reasonable choice

for producing accurate point estimates from an easily understandable model.

It is worth noting the manner in which the data were simulated was in agreement with the manner in which these two models were specified. Therefore, it is difficult, if not impossible, to assess whether the model fits the data well or the data fit the models well. The assumptions under which we simulated the data were largely driven by the sampling design, but there is always the possibility that factors exist which affect the probability of capture or some other aspect of the problem that we are unaware of and therefore not accounting for.

References

A. Agresti and B. Caffo. Simple and effective confidence intervals for proportions and differences of proportions result from adding two successes and two failures. *The American Statistician*, 54(4):280–288, 2000.

A. Chao. An overview of closed capture-recapture models. *Journal of Agricultural, Biological, and Environmental Statistics*, 6(2):158–175, 2001.

A. Gelman and J. Hill. *Data Analysis Using Regression and Multilevel/Hierarchical Models*, volume Analytical methods for social research. Cambridge University Press, New York, 2007.

A. Gelman and Y.-S. Su. *arm: Data Analysis Using Regression and Multilevel/Hierarchical Models*, 2015. URL `https://CRAN.R-project.org/package=arm`. R package version 1.8-5.

A. Gelman, A. Jakulin, M. G. Pittau, and Y.-S. Su. A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, pages 1360–1383, 2008.

A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian data analysis*, volume 2. Taylor & Francis, 2014.

A. Gelman et al. Prior distributions for variance parameters in hierarchical models. *Bayesian analysis*, 1(3):515–534, 2006.

M. Kéry and J. Royle. *Applied hierarchical modeling in ecology: analysis of distribution, abundance and species richness using R and BUGS, vol 1. Prelude and static models.* Academic Press, San Diego, 2016.

S. Mäntyniemi and A. Romakkaniemi. Bayesian mark recapture estimation with an application to a salmonid smolt population. *Canadian Journal of Fisheries and Aquatic Sciences*, 59(11):1748–1758, 2002.

M. Plummer. *JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling*, 2003.

K. H. Pollock. The use of auxiliary variables in capture-recapture modelling: an overview. *Journal of Applied Statistics*, 29(1-4):85–102, 2002.

A. E. Raftery. Inference for the binomial n parameter: A hierarchical bayes approach. *Biometrika*, 75(2):223–228, 1988.

G. A. Seber and C. J. Schwarz. Capture-recapture: before and after euring 2000. *Journal of Applied Statistics*, 29(1-4):5–18, 2002.

G. A. F. Seber. *The Estimation of Animal Abundance and Related Parameters. Second Editon.* The Blackburn Press, Caldwell, New Jersey, 1982.

H. Wickham and R. Francois. *dplyr: A Grammar of Data Manipulation*, 2015. URL `https://CRAN.R-project.org/package=dplyr`. R package version 0.4.1.

E. Zwane and P. Van Der Heijden. Semiparametric models for capture–recapture studies with covariates. *Computational Statistics & Data Analysis*, 47(4):729–743, 2004.

9 Acknowledgments

10   Appendix A

10.1   Dam Section Plots

Length < 150mm: Dam
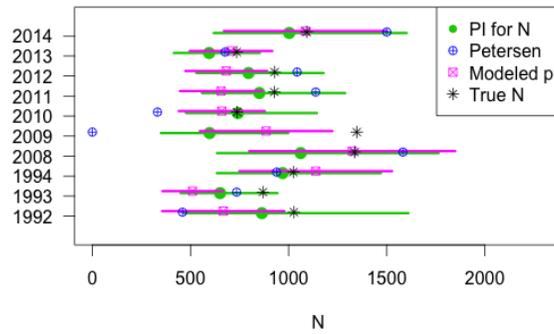
Dam: [0,100)
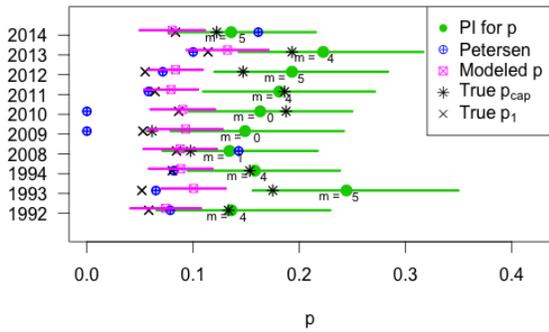
Dam: [100,150)
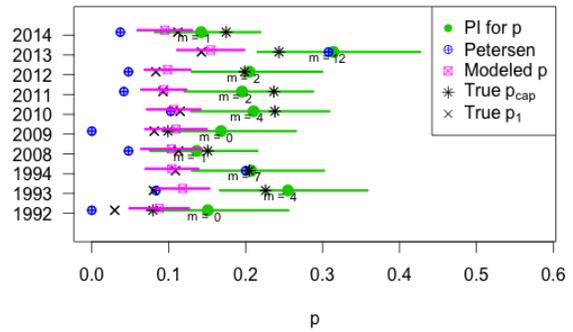
Length Between 150 and 250mm: Dam
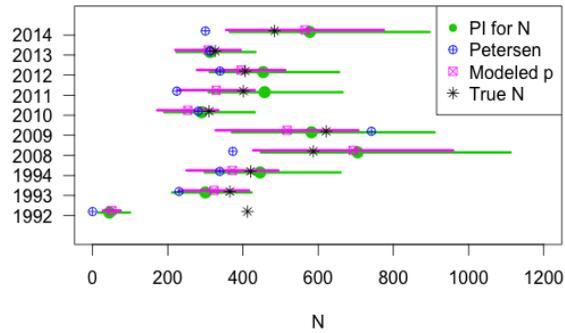
Dam: [150,175)
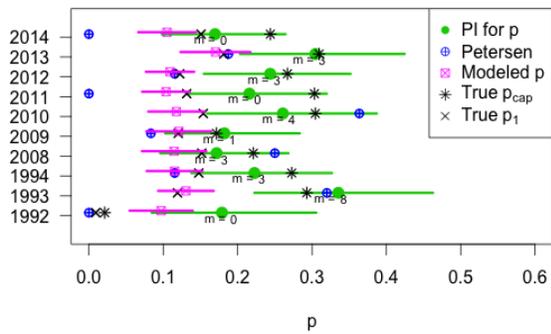
Dam: [175,200)
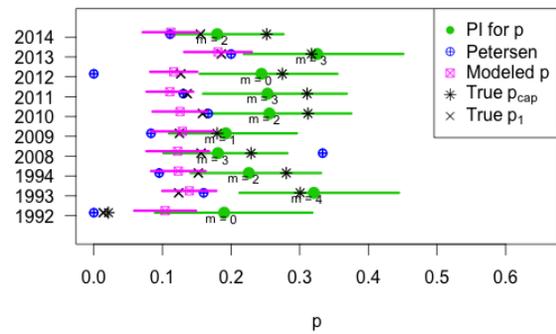
Dam: [200,225)
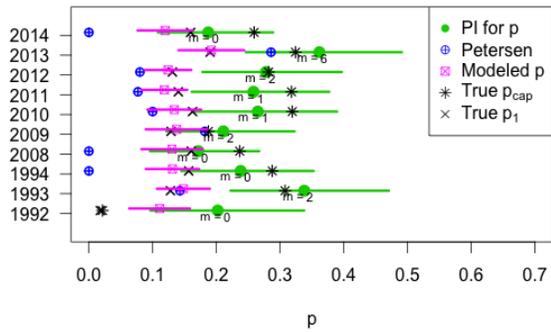
Dam: [225,250)
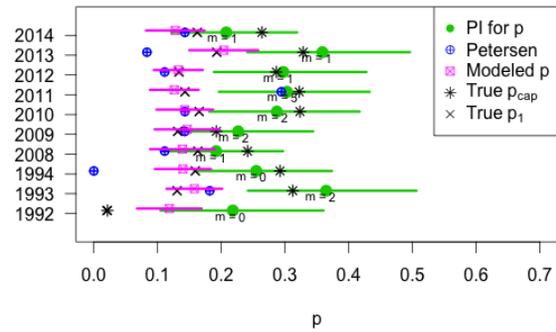
Length Between 250 and 325mm: Dam
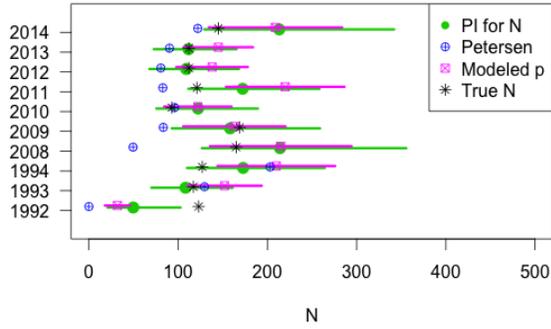
Dam: [250,275)
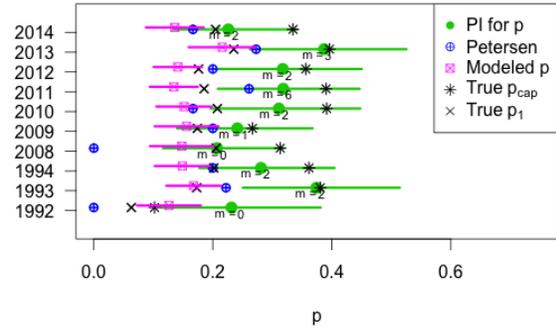
Dam: [275,300)
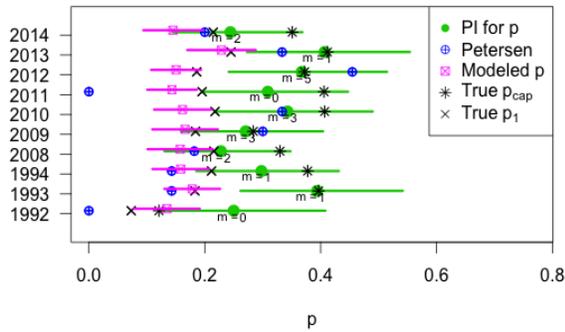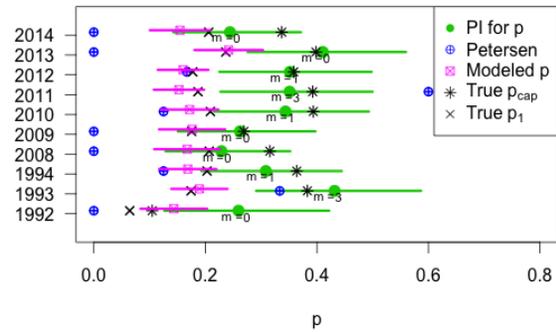
Dam: [300,325)

Length > 325mm: Dam

Dam: >325

32
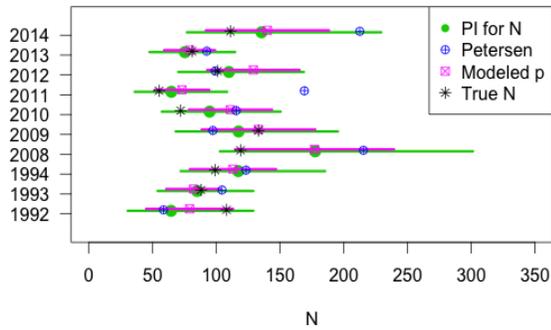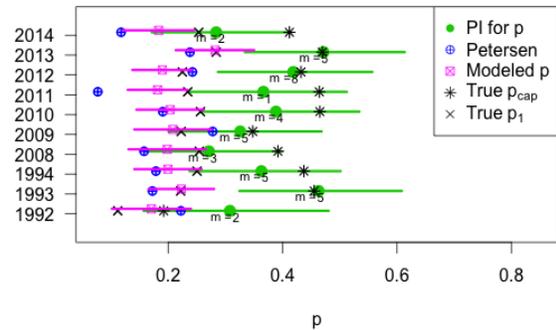
10.2 Flowerpipe Section Plots

Length < 150mm: FP

FP: [0,100)

FP: [100,150)

33
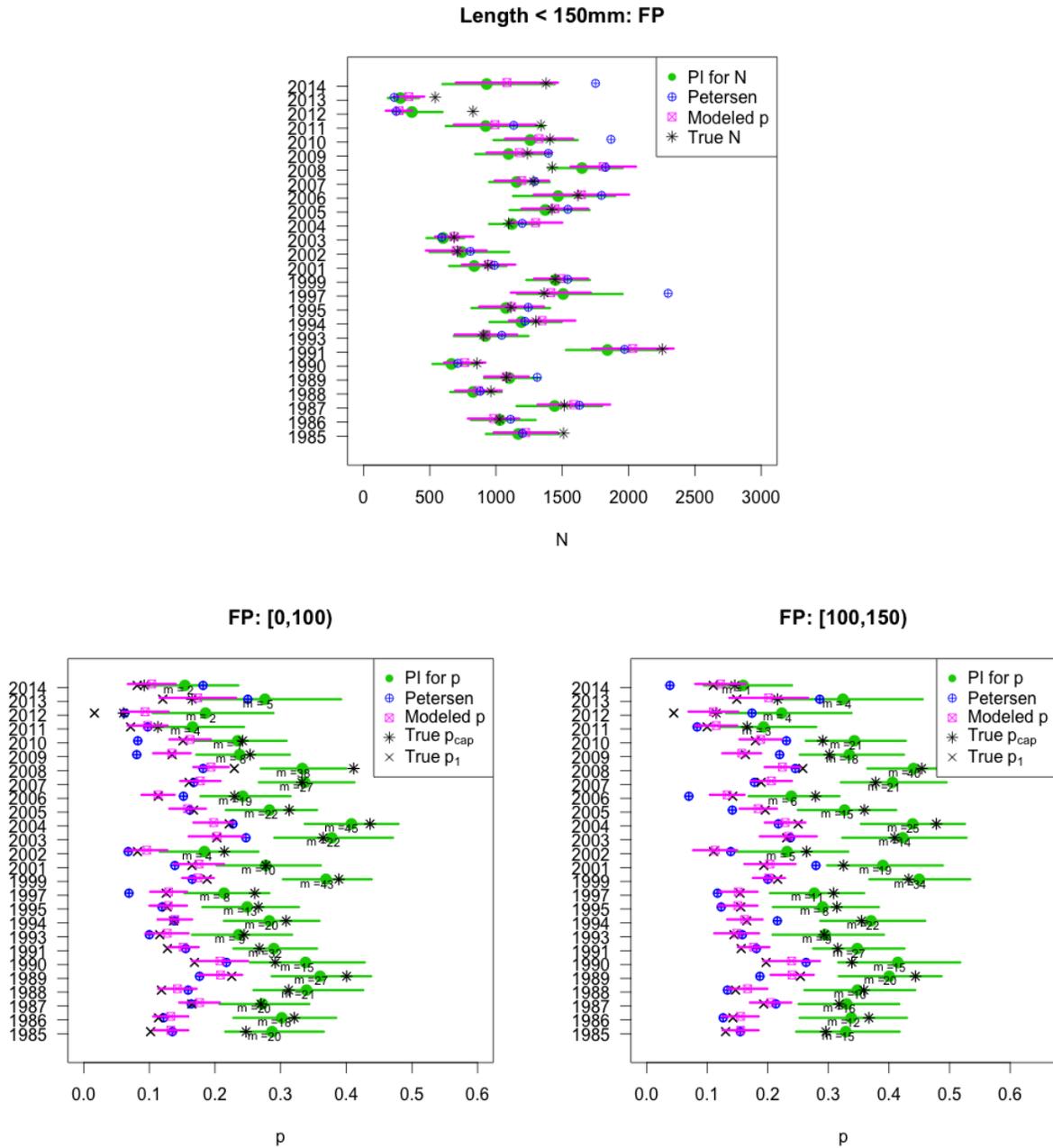
Length Between 150 and 250mm: FP
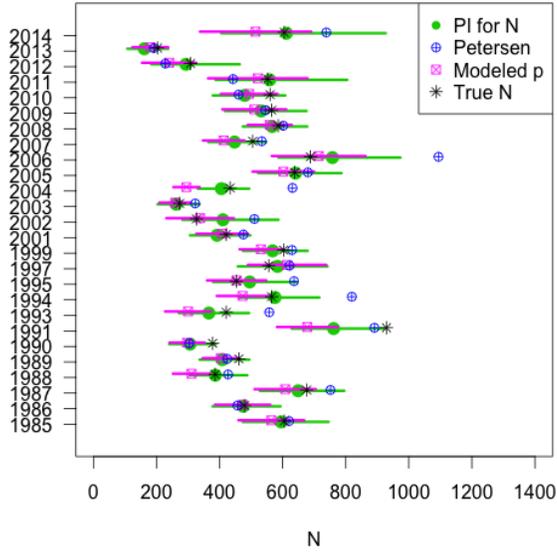
FP: [150,175)
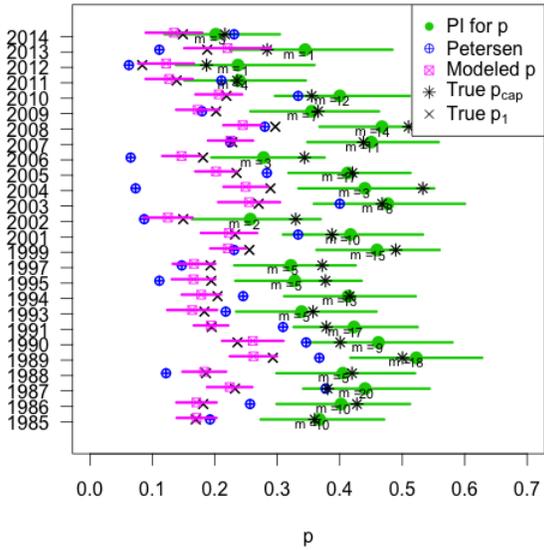
FP: [175,200)
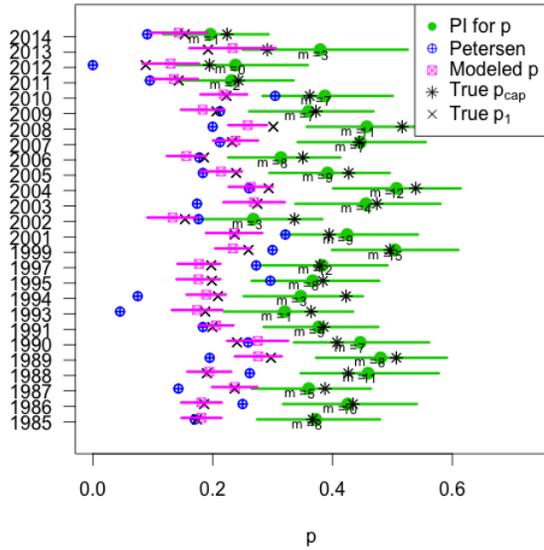
FP: [200,225)
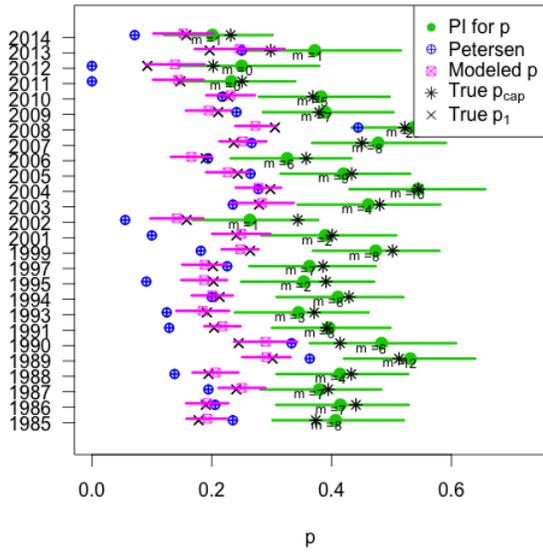
FP: [225,250)
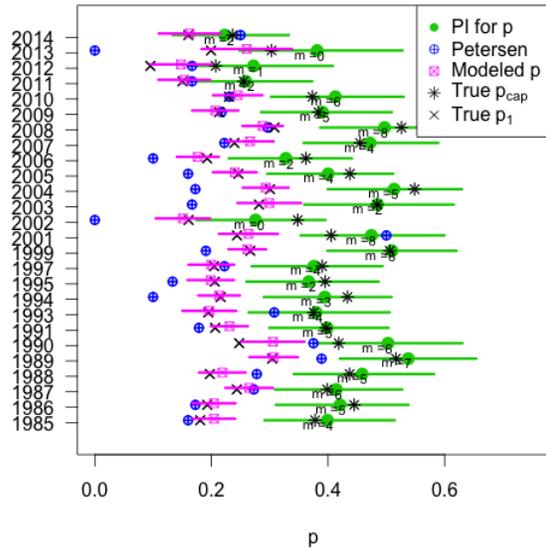
Length Between 250 and 325mm: FP
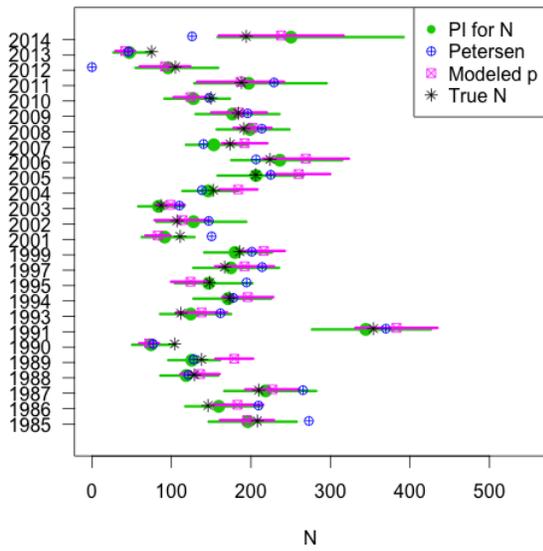
FP: [250,275)
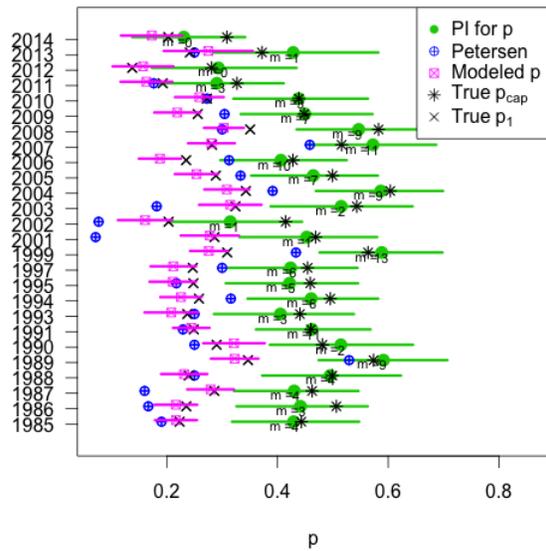
FP: [275,300)
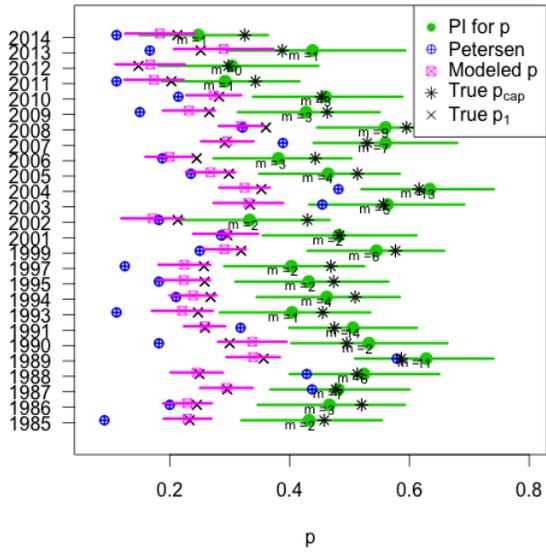
FP: [300,325)
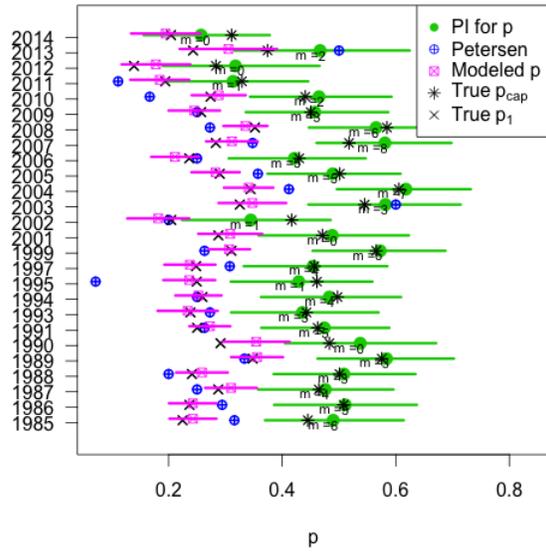
Length > 325mm: FP
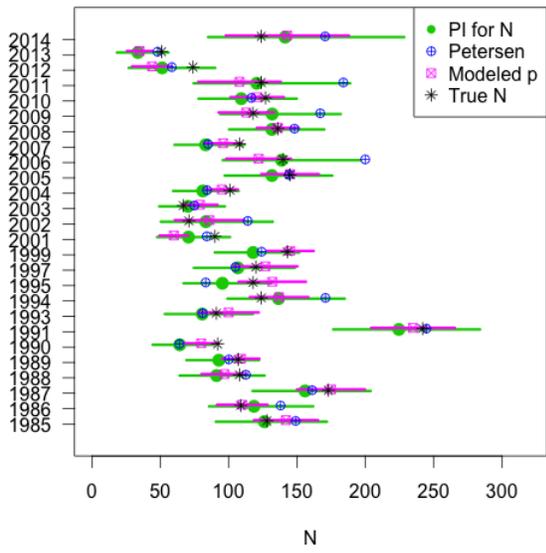
FP: >325
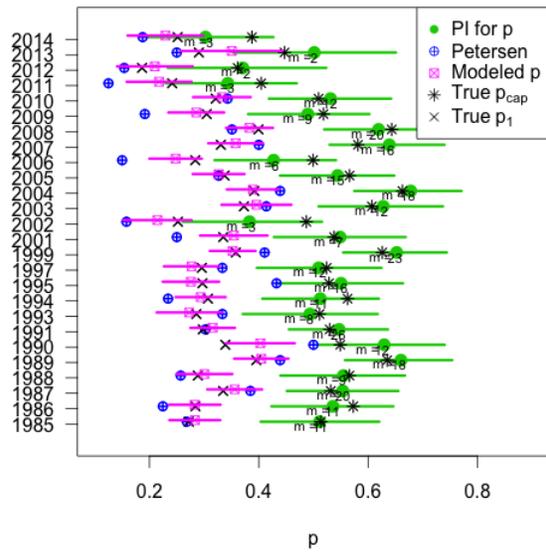
10.3 Rereg Section Plots

Length < 150mm: RR
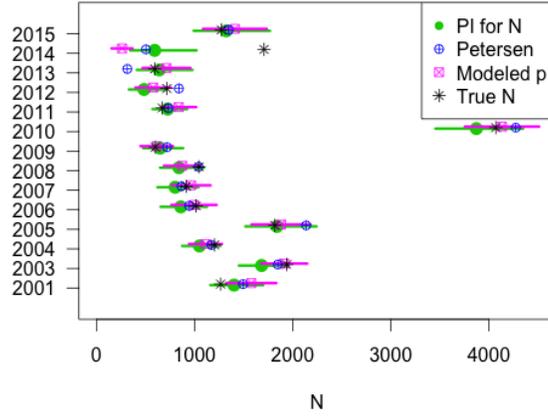
RR: [0,100)
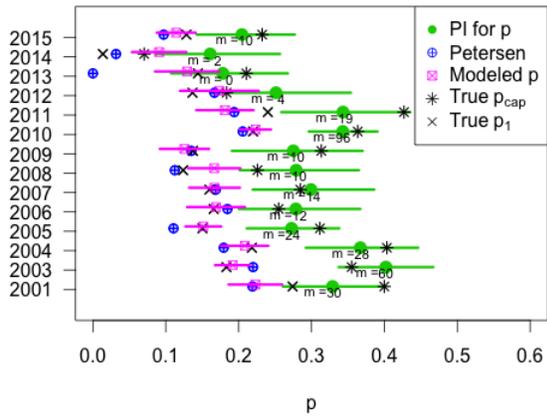
RR: [100,150)
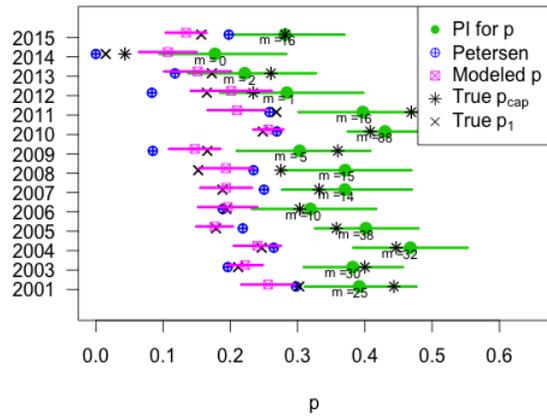
Length Between 150 and 250mm: RR
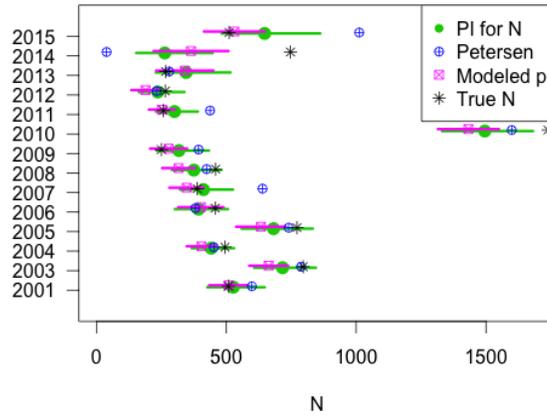
RR: [150,175)
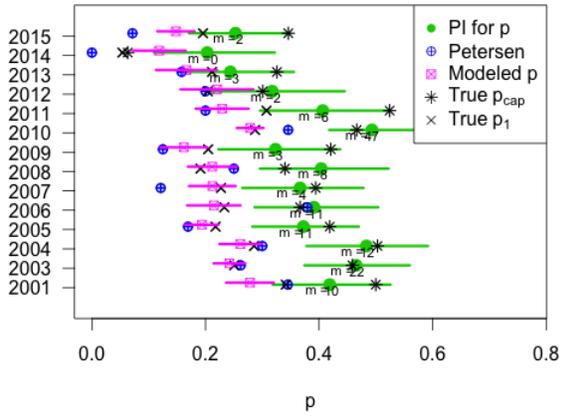
RR: [175,200)
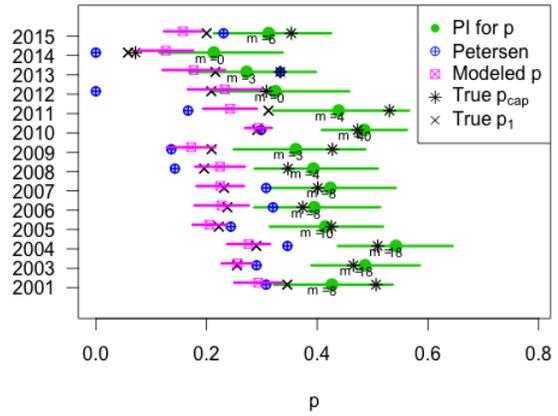
RR: [200,225)
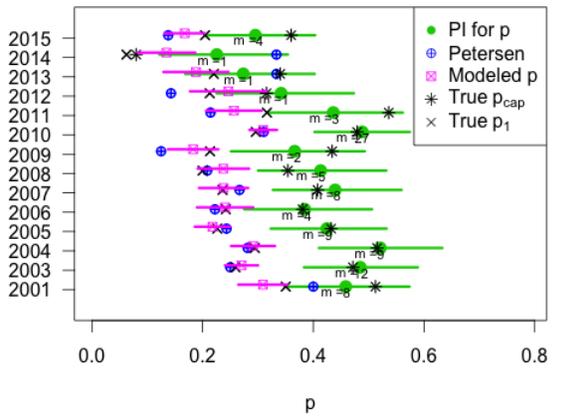

RR: [225,250)
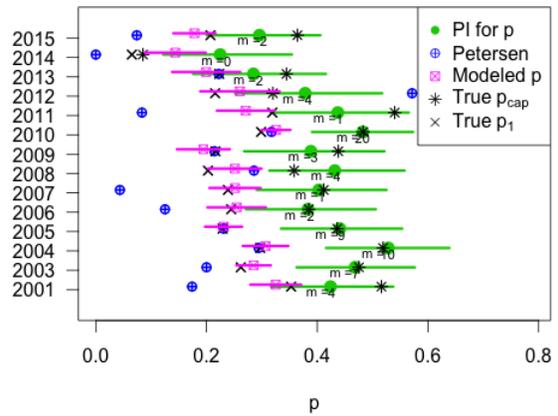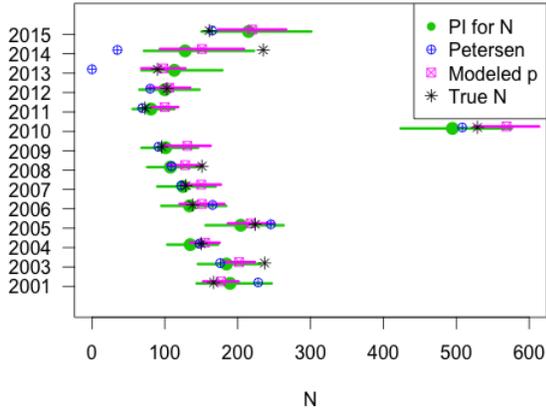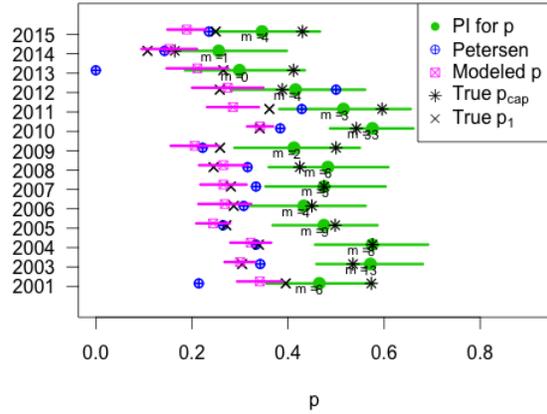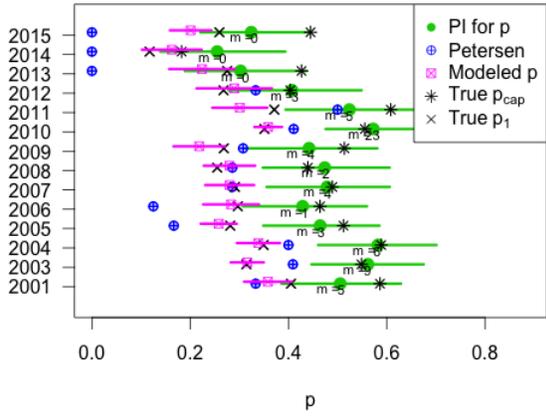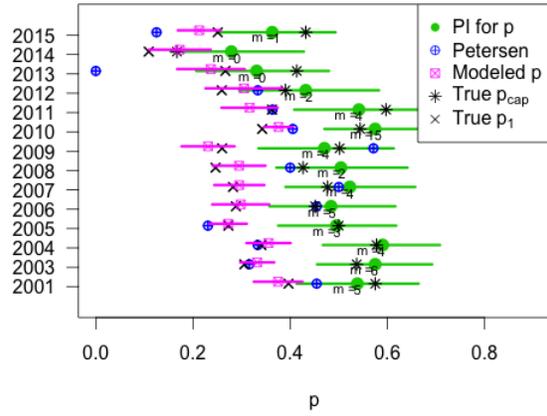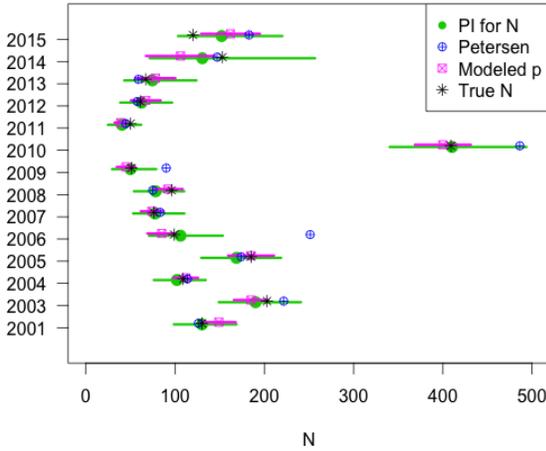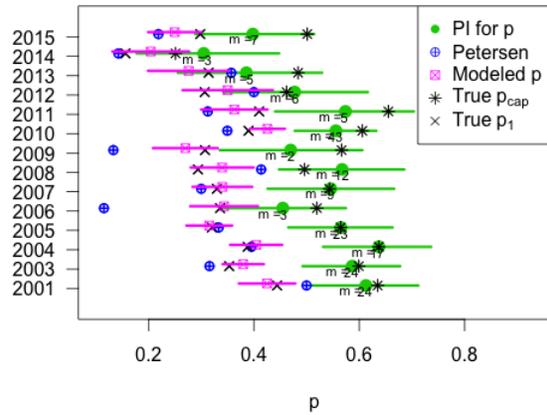
10.4   Troy Section Plots

Length < 150mm: Troy
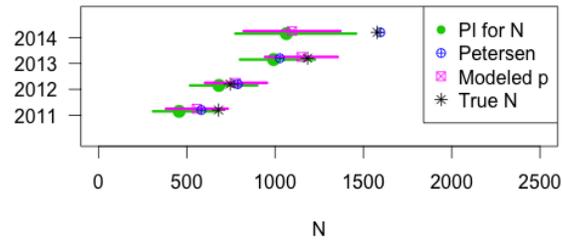
Troy: [0,100)
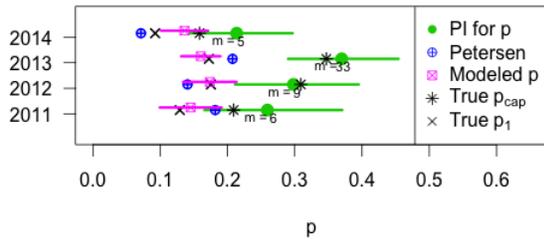
Length Between 150 and 250mm: Troy
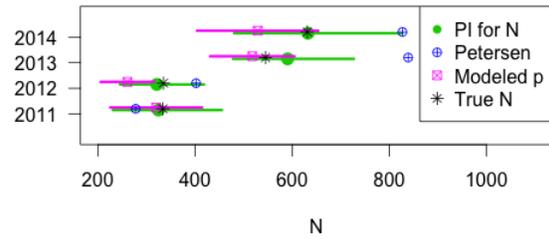
**Length Between 150 and 250mm: Troy**

Troy: [150,175)

Troy: [175,200)

Troy: [200,225)

Troy: [225,250)

Length Between 250 and 325mm: Troy

Troy: [250,275)

Troy: [275,300)

Troy: [300,325)

Length > 325mm: Troy

Troy: >325
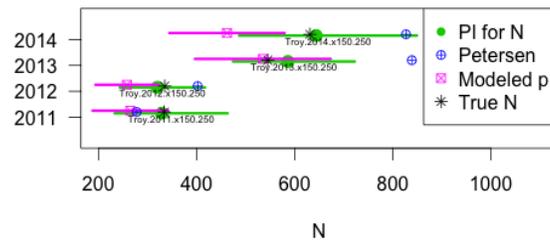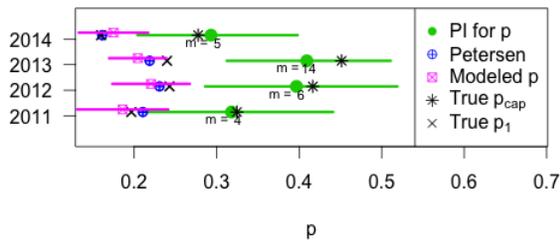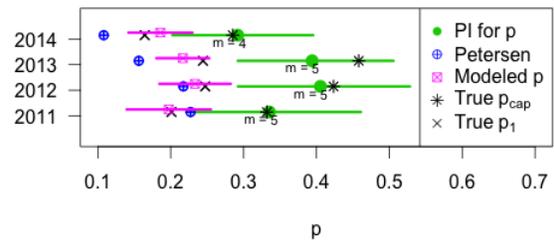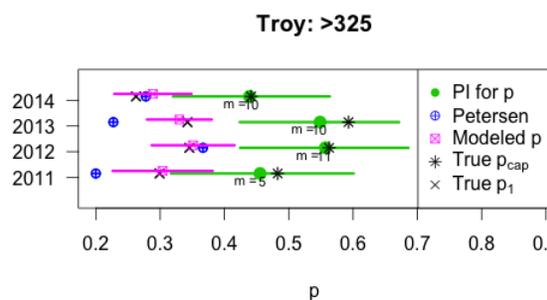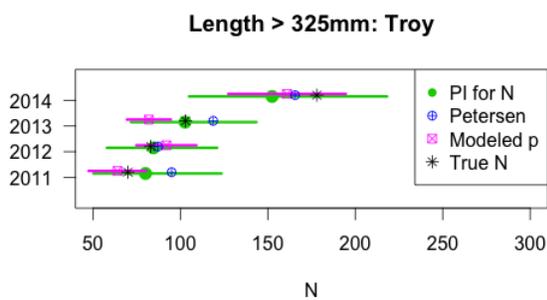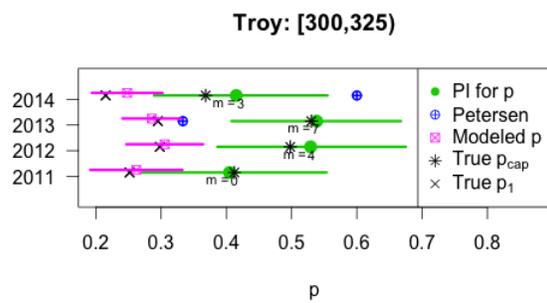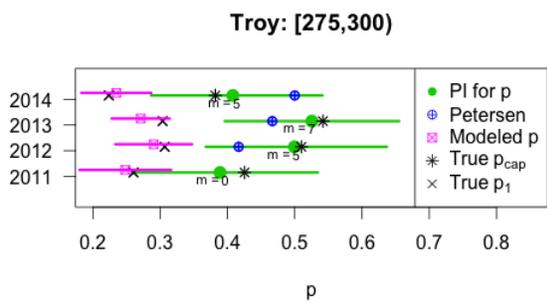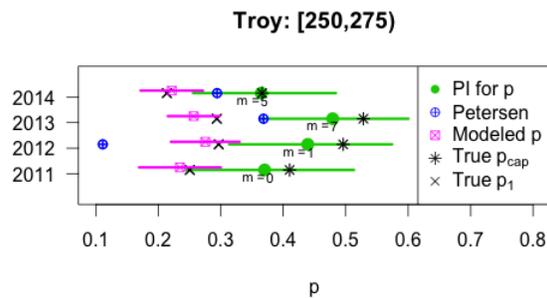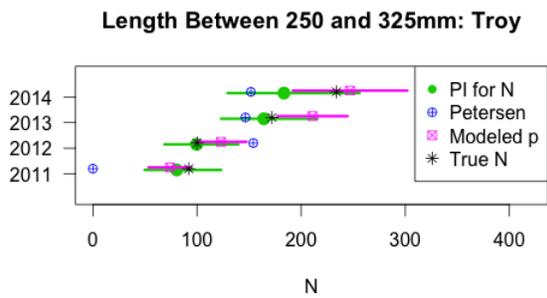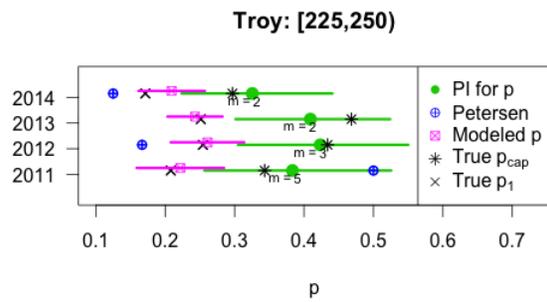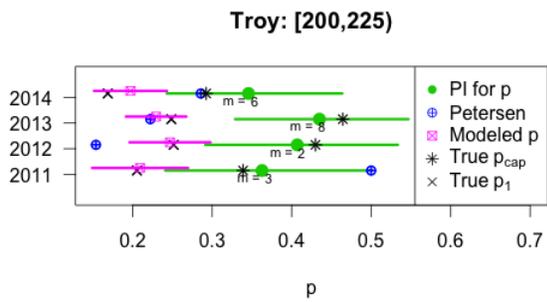
11   Appendix B

11.1   R Code

```
cat("model{
```

```
for(len in 1:10){ ##10 unique lengths
  for(ste in 1:nSites){ ##4 unique sites
    ## each length class has its own effect on Prob across all years
    mu.p[len, ste] <- beta0.p[ste] + beta1.p[ste] * (nLen[len])
    len.eff.p[len,ste] ~ dnorm(mu.p[len, ste], tau.len.p)
    ##len.eff.p[len] <- len.eff.raw.p[len, ste] -
      mean(len.eff.raw.p[]) + mu.p[len]
  }
}


for(ste in 1:nSites){
  beta1.p[ste] ~ dt(0, 10, 1) ## cauchy(0,10) on intercept (Gelman paper)
  beta0.p[ste] ~ dt(0, 2.5, 1)   ## cauchy(0,2.5) on slope (Gelman paper)
}


tau.len.p <- pow(sigma.len.p, -2)   # tau = 1/sigma^2
sigma.len.p ~ dt(0, 1, 1)I(0,)


for(t in 1:nYears){ ##28 unique years
  for(ste in 1:nSites){
    year.eff.raw.p[ste, t] ~ dnorm(0, tau.year.p)  ## year effects  on prob's
    year.eff.p[ste,t] <-  year.eff.raw.p[ste,t] -
      mean(year.eff.raw.p[,])   ## adj
  }
}
tau.year.p <- pow(sigma.year.p, -2)  # tau = 1/sigma^2
sigma.year.p ~ dt(0, 1, 1)I(0,)      # informative prior on sigma


for(ndx in 1:nRows ){
##nRows is (number unique lengths) * (number unique site.years)
##############################################################
### probabilities  #########################################
##############################################################
  mu.x[ndx] <- len.eff.p[length[ndx], site[ndx]] + year.eff.p[site[ndx], year[ndx]]
  x1[ndx] ~ dnorm(mu.x[ndx], tau.p)
  x2[ndx] ~ dnorm(mu.x[ndx], tau.p)
  p1[ndx]  <- ilogit(x1[ndx])
  p2[ndx]  <- ilogit(x2[ndx])
  ## p1 and p2 are functions of length and year effects.
  pCap[ndx] <-  p1[ndx] + p2[ndx] - p1[ndx] * p2[ndx]
  ## prob of getting captured in one or both passes
  pNeverCap[ndx] <- 1 - (p1[ndx] + p2[ndx] - p1[ndx] * p2[ndx])
##############################################################
### population effects  ####################################
##############################################################
```

```
        totalCaps[ndx]  ~  dpois(lambda[ndx])
        lambda[ndx] ~ dgamma(.1, .1)
        Npop[ndx] <-  lambda[ndx]/pCap[ndx]  # + epsilon[ndx] ##from Raftery paper

        p.multi[ndx,1] <- p1[ndx] * (1 - p2[ndx]) ## prob capture history = (1,0)
        p.multi[ndx,2] <- (1 - p1[ndx]) * p2[ndx] ## prob capture history = (0,1)
        p.multi[ndx,3] <- p1[ndx] * p2[ndx] ## prob capture history = (1,1)

        counts[ndx,1:3] ~ dmulti(p.multi[ndx,1:3], totalCaps[ndx])
    }

    tau.p <- pow(sigma.p, -2) # tau = 1/sigma^2
    sigma.p ~ dt(0, 1, 1)I(0,)      # informative prior on sigma
    }", file="betterModel.txt")


############################################################################
######################### Run for simulated data  ##########################
############################################################################
set.seed(7951)
cmr.sim.data <- with(sim.data, list(nYears = 28,
                            totalCaps = n1 + n2 - m,
                            nRows = 540,
                            counts = cbind(n1-m, n2-m, m),
                            nLen = (unique(mean.len) -
                            mean(unique(mean.len)))/(2*sd(unique(mean.len))),
                            ##scaled to have sd of .5 (Gelman paper),
                            nSites = 4,
                            length = length.cat3,
                            site = site,
                            year = year2))
warmup.sim.cmr <- jags.model("betterModel.txt", data=cmr.sim.data,
  n.chains=3, n.adapt=2000, quiet=TRUE)

params <- c("Npop", "pCap")
samples.sim <- coda.samples(warmup.sim.cmr, params, n.iter=15000, quiet=TRUE)
```