

# Smoothing Population Size Estimates for Time-Stratified Mark-Recapture Experiments using Bayesian P-Splines

Simon J Bonner<sup>1\*</sup> and Carl J Schwarz<sup>2\*</sup>

<sup>1</sup>Department of Statistics, University of Kentucky, Lexington, KY 40506, USA

<sup>2</sup>Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada

\**email*: simon.bonner@uky.edu

\*\**email*: cschwarz@stat.sfu.ca

**SUMMARY:** Petersen-type mark-recapture experiments are often used to estimate the number of fish or other animals in a population moving along a set migration route. A first sample of individuals is captured at one location, marked, and returned to the population. A second sample is then captured farther along the route, and inferences are derived from the numbers of marked and unmarked fish found in this second sample. Data from such experiments are often stratified by time (day or week) to allow for possible changes in the capture probabilities, and previous methods of analysis fail to take advantage of the temporal relationships in the stratified data. We present a Bayesian, semi-parametric method that explicitly models the expected number of fish in each stratum as a smooth function of time. Results from the analysis of historic data from the migration of young Atlantic salmon (*Salmo salar*) along the Conne River, Newfoundland, and from a simulation study indicate that the new method provides more precise estimates of the population size and more accurate estimates of uncertainty than the currently available methods.

**KEY WORDS:** Bayesian inference; Hierarchical model; Mark-recapture; OpenBUGS; Penalized spline; Atlantic salmon; Stratified-Petersen

## 1. Introduction

Single-recapture, Petersen-type mark-recapture experiments are commonly used to monitor populations of Atlantic and Pacific salmon (*Salmo salar* and *Onchorhynchus sp.*) as they migrate between their freshwater spawning grounds and the oceans where they spend their adult lives. In these experiments, samples of fish are captured at one location, marked, and returned to the population, and new samples containing both marked and unmarked fish are captured from a second location farther along the migration route. Estimates of abundance are then derived by modelling the numbers of marked and unmarked fish captured at the second location. If all fish have the same probability of being captured at the second location then the simple Petersen estimator computed from the overall numbers of marked and unmarked fish captured at the second location provides a valid estimate of population size. However, salmon migrations can last for weeks or months and the probability of capture can change considerably over this time. As a result, simple Petersen estimates of the population size, and the degree of uncertainty in these estimates, may be severely biased (see Seber, 2002, pg. 85).

One way to allow for such variations in catchability is to stratify the population by time – essentially, to divide the experiment into shorter time strata (days or weeks) and then to estimate the number of individuals passing the second location separately for each stratum. Darroch (1961) provided the first rigorous treatment of the stratified-Petersen model and derived maximum likelihood estimators by conditioning on the numbers of individuals marked and released in each stratum at the first location. Similar methods were developed by Macdonald and Smith (1980) for the modified experiment in which trapping occurs at only one location and marked fish are introduced by transporting them upstream to be released. Plante et al. (1998) developed unconditional likelihood methods by modelling the number of individuals at the first location, and further allowed the capture probabilities to depend

26 on stratum specific covariates like rate of flow. Estimators of abundance from the stratified-  
27 Petersen model have also been obtained by the methods of moments (Chapman and Junge,  
28 1956) and by least squares (Banneheka et al., 1997).

29 Although stratification reduces the bias of the abundance estimates, it also increases the  
30 number of parameters which leads to a bias/variance trade-off. In general, marked individuals  
31 may pass the second location in any stratum so that it is necessary to model the movements  
32 of these fish between the trapping locations. The number of parameters increases as the  
33 product of the number of strata at the two locations, and if few marked fish are recaptured  
34 then the movements are difficult to model and the resulting estimates of population size will  
35 be imprecise.

36 The number of parameters can be reduced by partially pooling the data, but there are  
37 several pitfalls. Proposed methods for pooling strata essentially entail testing for differences  
38 in the capture probabilities between neighbouring strata and then combining these strata if  
39 the null hypothesis is not rejected (see Darroch, 1961; Schwarz and Taylor, 1998; Bjorkstedt,  
40 2000). However, such tests will have low power, because the number of fish marked in each  
41 strata is small, and important differences are likely to be overlooked. Guidelines as to how far  
42 the data should be pooled are not available, and, moreover, standard errors and confidence  
43 intervals computed from these methods do not account for the pooling decisions and will  
44 underestimate the true variability in the estimated population size (Steinhorst et al., 2004).

45 Schwarz and Dempson (1994) introduced another method to reduce the number of pa-  
46 rameters by assigning parametric distributions to the travel times of the marked individuals.  
47 Their particular formulation assumes that travel times between the trapping locations follow  
48 independent log-normal distributions with mean and variance parameters that depend only  
49 on the release strata. Probabilities governing the movement of the marked fish can then be  
50 computed from standard normal calculations, and the number of parameters in the model can

51 be greatly reduced relative to the model of Darroch (1961) which allows complete flexibility in  
52 the movements between strata. Mantyniemi and Romakkaniemi (2002) extended this model  
53 to the hierarchical Bayesian framework by modelling the capture probabilities, the means and  
54 variances of the travel times, and the numbers of unmarked fish in each stratum as random  
55 draws from hyper-priors with unknown parameters. This allows for sharing of information  
56 between the strata through estimation of the hyper-parameters and further decreases the  
57 uncertainty in total abundance.

58 One shortcoming of all of these methods is that they fail to account for the natural,  
59 temporal ordering of the data. Most salmon runs follow a fairly predictable pattern with  
60 small numbers of fish passing the traps early in the migration followed by a steady increase  
61 to one or two peaks near the middle of the migration and then a slow decrease toward  
62 the end. Accounting for these patterns has the potential to improve estimates of the total  
63 population size, and our primary contribution is the implementation of a Bayesian, semi-  
64 parametric method that takes advantage of the temporal ordering by explicitly modelling  
65 the expected number of unmarked fish in each stratum as a smooth function of time. Further  
66 to this, we develop a hierarchical, non-parametric model for the movements of the marked  
67 fish that reduces the number of parameters by building on similarities in the movements  
68 between strata rather than by assuming a specific parametric form. In the following sections  
69 we develop the general model, describe methods for assessing goodness-of-fit and comparing  
70 candidate models, and provide results from an application to historical data and a simulation  
71 study that illustrate the primary advantages of our method.

## 72 **2. Methods**

73 We describe the implementation of the Bayesian P-spline model with specific application  
74 to programs monitoring the productivity of salmon populations, though the methods are  
75 generally applicable to temporally stratified data. Most species of salmon are anadromous,

76 meaning that they are born in freshwater lakes and rivers but spend the majority of their  
 77 adult lives in the ocean before returning to freshwater to spawn. Every year, the young fish  
 78 born one or two years earlier (called smolts) leave the spawning grounds, and the productivity  
 79 of the population can be quantified by the total number of out-migrating smolts. To estimate  
 80 the size of the migration (also called the run), fish are trapped at one location, marked,  
 81 and released back into the population. Samples containing both marked and unmarked  
 82 individuals are then trapped at a second location farther along the migration route, and  
 83 the size of the population is estimated by modelling the capture of marked and unmarked  
 84 individuals at the second trap. As discussed in Section 1, these runs may last for weeks or  
 85 months and the data are often stratified to account for changes in the capture probabilities  
 86 over this time. We assume that the data are stratified by day for notational convenience, but  
 87 the methods apply equally if the data are grouped in other ways.

88 Data obtained from these experiments consist of the daily numbers of marked fish released  
 89 at the first trap, and the daily numbers of both marked and unmarked fish captured at  
 90 the second trap. The two traps need not be operated for the same length of time, and it is  
 91 common to continue trapping at the second location for several days after the last marked  
 92 fish are released to allow time for these individuals to move past the second trap. We denote  
 93 the number of days of capture at the first and second traps by  $s$  and  $t$ , the number of marked  
 94 fish released on day  $i$  at the first trap by  $n_i$ , the number of these fish recaptured on day  $j$   
 95 by  $m_{i,j}$ , and the number of unmarked fish captured at the second trap on day  $j$  by  $u_j$ . The  
 96 total number of unmarked fish that pass the second trap on day  $j$  is denoted by  $U_j$ , the sum  
 97 over all days by  $U_{Tot}$ , and the total population size by  $N_{Tot} = \sum_{i=1}^s n_i + U_{Tot}$ . Vectors and  
 98 matrices are used to simplify the notation where possible such that  $\mathbf{n}$  represents the vector  
 99  $(n_1, \dots, n_s)'$ ,  $\mathbf{u}$  the vector  $(u_1, \dots, u_t)'$ , and  $\mathbf{M}$  the matrix with  $ij$  entry  $m_{i,j}$ .

100 Common assumptions for modelling such data essentially imply that the samples of marked

101 and unmarked fish captured at the second trap constitute simple, random samples from the  
 102 populations available on each day. Specifically, it is assumed that:

- 103 (1) all fish passing the second trap on day  $j$  have the same probability of capture,  $p_j$ , and  
 104 (2) whether or not any one fish is captured at the second trap is independent of whether or  
 105 not any other fish is captured.

106 Further to this, it is necessary to model the movement of the marked fish between the two  
 107 traps. Standard assumptions concerning the movements of marked fish are that:

- 108 (1) they do not emigrate or die between the two traps,  
 109 (2) marks are not lost or overlooked when fish are recaptured,  
 110 (3) all marked fish released at the first trap on day  $i$  have the same probability of passing  
 111 the second trap on day  $j$ ,  $\theta_{i,j}$ , and  
 112 (4) the movements of marked fish are independent of one and other.

113 Given these assumptions, the recaptures of fish marked on day  $i$  will follow the multinomial  
 114 distribution:

$$\mathbf{m}_i \sim \text{MN} \left\{ n_i, \left( \theta_{i,1}p_1, \dots, \theta_{i,t}p_t, 1 - \sum_{j=i}^t \theta_{i,j}p_j \right) \right\} \quad (1)$$

115 where  $\mathbf{m}_i = (m_{i,1}, \dots, m_{i,t}, n_i - m_{i,\cdot})$  and  $m_{i,\cdot} = \sum_{j=i}^t m_{i,j}$ . Conditioning on  $\mathbf{U}$ , as in the  
 116 original formulation of Darroch (1961), the number of unmarked fish captured at the second  
 117 trap on day  $j$  will be distributed according to the binomial model:

$$u_j \sim \text{Binomial}(U_j, p_j). \quad (2)$$

118 The likelihood function is computed by multiplying the contributions from (1) and (2) over  
 119 all days and is proportional to:

$$L(\mathbf{U}, \mathbf{p}, \Theta | \mathbf{n}, \mathbf{M}, \mathbf{u}) = \prod_{i=1}^s \left\{ \prod_{j=1}^t (\theta_{i,j}p_j)^{m_{i,j}} \cdot \left( 1 - \sum_{j=1}^t \theta_{i,j}p_j \right)^{n_i - m_{i,\cdot}} \right\} \cdot \prod_{j=1}^t \binom{U_j}{u_j} p_j^{u_j} (1 - p_j)^{U_j - u_j}. \quad (3)$$

120 A complete Bayesian formulation is then obtained by specifying prior distributions for the  
 121 parameters  $\mathbf{U}$ ,  $\boldsymbol{\rho}$ , and  $\Theta$ .

122 In their formulation, Mantyniemi and Romakkaniemi (2002) (hereafter MR02) define the  
 123 prior for  $\mathbf{U}$  by first assigning  $U_{Tot}$  a vague prior distribution and then modelling  $\mathbf{U}$  conditional  
 124 on  $U_{Tot}$ . Specifically, they define the improper Jeffrey's prior  $\pi(U_{Tot}) \propto 1/U_{Tot}$  and model  
 125 the individual elements of  $\mathbf{U}$  by the multinomial:

$$\mathbf{U}|U_{Tot}, \boldsymbol{\rho} \sim \text{MN}(U_{Tot}, (\rho_1, \dots, \rho_t)).$$

126 The cell probabilities,  $\boldsymbol{\rho}$ , are then assigned a Dirichlet prior with fixed parameter  $\boldsymbol{\alpha}$ .

127 In theory, individual elements of  $\boldsymbol{\alpha}$  could be chosen to encode prior information about  
 128 the elements of  $\mathbf{U}$ , but this would require the user to provide explicit details regarding the  
 129 exact shape of the run. The simplest prior, with  $\alpha_j = \alpha$  for all  $j = 1, \dots, t$ , implies that the  
 130 marginal prior distribution of  $U_j$  is the same for all  $j$  and ignores the ordering of the strata  
 131 in that the joint prior distribution of  $(U_{j_1}, U_{j_2})$  is the same for all  $j_1$  and  $j_2$ . If the run was  
 132 expected to peak on day  $k$  then one could increase the values of  $\alpha_j$  in some neighbourhood  $k$ ,  
 133 but this would require specific prior knowledge about when the peak was expected and how  
 134 high it would be relative to the other days. Instead, we sought a prior that would incorporate  
 135 the natural ordering of the data without the need for such specific prior information.

136 To accomplish this, we construct an alternative prior for  $\mathbf{U}$  which directly models the  
 137 number of unmarked individuals passing the second trap each day. In particular, we model  
 138 the expected log daily run size,  $E(\log(U_j))$ , as a smooth function of  $j$  using the Bayesian  
 139 penalized spline or P-spline model of Lang and Brezger (2004). Briefly, splines are semi-  
 140 parametric regression functions formed as linear combinations of sets of basis functions  
 141 associated with a set of knot points. The shape of a spline is very flexible and to avoid  
 142 overfitting the data the complexity of the spline is controlled by two factors: the number  
 143 of knot points and their locations, and the relationship amongst the regression coefficients.

144 The Bayesian P-spline algorithm uses the cubic B-spline basis functions (often chosen for  
 145 their computational simplicity) and addresses overfitting by fixing a relatively large number  
 146 of knots but specifying a prior which favours small changes in the coefficients.

147 Letting  $B_1(\cdot), \dots, B_K(\cdot)$  denote the cubic B-spline functions for the chosen set of knots,  
 148 we model  $\mathbf{U}$  as:

$$\log(U_j) = \sum_{k=1}^K b_k B_k(j) + \epsilon_j \quad (4)$$

149 where the daily deviations from the spline,  $\epsilon_j$ , are assumed to be independent, mean zero  
 150 normal random variables with constant variance,  $\tau_\epsilon^2$ . Following the recommendations of Lang  
 151 and Brezger (2004), we space the knot points evenly across the data at intervals of 4 days  
 152 (or as close to this as  $t$  allows) and model the regression coefficients by the second difference  
 153 prior:

$$b_{k+1}|b_1, \dots, b_k \sim N(b_k + (b_k - b_{k-1}), \tau_b^2) \quad (5)$$

154 for  $k = 3, \dots, K$  with  $b_1$  and  $b_2$  assigned the improper flat prior,  $U(-\infty, \infty)$ . The parameter  
 155  $\tau_b^2$  controls the smoothness of the spline and is assigned an inverse gamma prior distribution  
 156 with parameters  $\alpha_b$  and  $\beta_b$  chosen so that  $E(\tau_b^2)$  is small, favouring small differences in the  
 157  $b_k$  and hence a smooth fit, but  $\text{Var}(\tau_b^2)$  is large. Lang and Brezger (2004) recommend setting  
 158  $\alpha_b = 1$  and choosing a small value for  $\beta_b$ , but note that there is no common choice for all  
 159 models. We found that the values  $\alpha_b = 1$  and  $\beta_b = .05$  worked well in our application.

160 Concerning the capture probabilities, the core model of MR02 defines a hierarchical prior  
 161 for  $\mathbf{p}$  which ignores possible temporal associations, and we have adopted exactly the same  
 162 prior. Specifically, the prior distribution for  $\mathbf{p}$  assumes that:

$$\text{logit}(p_j) \sim N(\eta_p, \tau_p^2) \quad (6)$$

163 independent but with common mean,  $\eta_p$ , and variance,  $\tau_p^2$ , assigned diffuse normal and inverse  
 164 gamma priors respectively. As an extension, MR02 do allow for the effect of covariates such

165 that  $E(\text{logit}(p_j)) = X'_j\beta$ . The same approach can easily be incorporated in our model, but  
 166 we have ignored such effects to keep the models simple.

167 Finally, we consider two possible priors for the matrix of movement probabilities. The  
 168 first is exactly that of MR02 which extends the log-normal model of Schwarz and Dempson  
 169 (1994) (hereafter SD94) to the hierarchical Bayesian framework and the second provides a  
 170 non-parametric alternative. SD94 proposed to model the times that marked fish take to move  
 171 between the two traps as independent log-normal random variables with stratum specific  
 172 parameters. Denoting the mean and variance of the log travel times for fish marked on day  
 173  $i$  by  $\mu_i^{(LN)}$  and  $\sigma_i^{(LN)^2}$ , the probability that a fish marked on day  $i$  passes the second trap on  
 174 day  $i + k$  is:

$$\theta_{i,i+k} = \begin{cases} 0 & k < 0 \\ \Phi\left(\frac{\log(1+d)-\mu_i^{(LN)}}{\sigma_i^{(LN)}}\right) & k = 0 \\ \Phi\left(\frac{\log(k+1+d)-\mu_i^{(LN)}}{\sigma_i^{(LN)}}\right) - \Phi\left(\frac{\log(k+d)-\mu_i^{(LN)}}{\sigma_i^{(LN)}}\right) & k > 0 \end{cases} \quad (7)$$

175 where  $\Phi(\cdot)$  is the standard-normal cumulative distribution function and  $d$  corrects for the  
 176 continuous release of marked fish (see pg. 100 of SD94 for details). MR02 extend this to the  
 177 hierarchical Bayesian framework by assigning  $\mu_i^{(LN)}$  and  $\sigma_i^{(LN)^2}$  common normal and inverse  
 178 gamma priors for all  $i$ . We call this the log-normal (LN) prior for  $\Theta$ .

179 To provide more flexibility in modelling the movement probabilities, we introduce an  
 180 alternative prior based on the continuation ratio for ordinal response variables (see e.g.  
 181 Dobson, 2002, sec. 8.4.4). Assuming the existence of a known cutoff value,  $\Delta_{\max}$ , such that  
 182  $\theta_{i,i+k} = 0$  when  $k > \Delta_{\max}$  for all  $i = 1, \dots, s$ , we model  $\theta_{i,i+k}$  by:

$$\log\left(\frac{\theta_{i,i+k-1}}{\sum_{l=k}^{\Delta_{\max}} \theta_{i,i+l}}\right) \sim N(\mu_k^{(NP)}, \sigma^{(NP)^2}) \quad (8)$$

183 for  $k = 1, \dots, \Delta_{\max}$  and set  $\theta_{i,i+\Delta_{\max}} = 1 - \sum_{k=1}^{\Delta_{\max}-1} \theta_{i,i+k}$ . The hyper-parameters  $\mu_k^{(NP)}$  and  
 184  $\sigma^{(NP)^2}$  are then assigned independent priors:

$$\mu_k^{(NP)} \sim N(0, 1.5), \quad k = 1, \dots, \Delta_{\max} \text{ and } \sigma^{(NP)^2} \sim \Gamma^{-1}(.1, .1).$$

185 Drawing on the hierarchical framework, this allows information about the movement prob-  
 186 abilities in one strata to influence the posterior distribution of the movement probabilities  
 187 in all other strata. In Section 3 we chose  $\Delta_{\max} = \max\{k \text{ s.t. } m_{i,i+k} > 0 \text{ for any } i\}$ . Cases in  
 188 which  $t < s + \Delta_{\max}$  can easily be accommodated by adding  $s - t + \Delta_{\max}$  columns of 0s to  
 189  $\mathbf{M}$  and fixing the capture probabilities for these strata to 0. We call this the non-parametric  
 190 (NP) model of  $\Theta$ .

191 In Section 3, we fit a variety of models to the sample data by combining the alternative  
 192 priors for  $\mathbf{U}$ ,  $\mathbf{p}$ , and  $\Theta$ . Along with the Bayesian P-spline model of  $\mathbf{U}$  (P-spline), we also  
 193 considered a cubic model for  $E(\log(U_j))$  (Cubic), a hierarchical model that assigned all  
 194  $\log(U_j)$  normal prior with common but unknown mean and variance (Hier), and a simple  
 195 Bayesian model that assigned all  $\log(U_j)$  identical normal priors with fixed mean and variance  
 196 (Simple). Most of the selected models made use of the hierarchical prior for  $\mathbf{p}$  in equation (6)  
 197 (Hier), but we did consider the combination of the simple prior for  $\mathbf{U}$  with a prior assuming  
 198  $p_1 = \dots = p_t$  (Pooled). Each of these choices of priors for  $\mathbf{U}$  and  $\mathbf{p}$  was then combined with  
 199 both the log-normal (LN) and non-parametric (NP) models of the movement probabilities to  
 200 create a total of 10 candidate models. In Sections 3 and 4, the models are labelled by triplets  
 201 identifying the priors for  $\mathbf{U}$ ,  $\mathbf{p}$  and  $\Theta$  in order using the abbreviations given in parantheses  
 202 above.

203 The models Simple/Hier/LN and Hier/Hier/NP are almost equivalent to the model of  
 204 MR02 with different choices of  $\alpha$  and provided very similar results. However, the Markov  
 205 chain Monte Carlo (MCMC) algorithms needed to compute summaries of the posterior  
 206 distribution converged much more quickly for the new models. The models with constant  
 207 capture probability (Simple/Pooled/LN and Simple/Pooled/NP) are Bayesian equivalents of  
 208 the classical pooled-Petersen estimators which ignore changes in  $\mathbf{p}$  and which are commonly  
 209 used when recapture probabilities are low and data are sparse (Seber, 2002, pg. 60).

210 The complete structure of model P-spline/Hier/NP (the best fitting model in Section 3) is  
 211 illustrated by a directed acyclic graph in Figure 1 and DAGs for the remaining models are  
 212 provided in Web Figures 9 – 18. Prior distributions at the highest level of each model were  
 213 chosen to be weakly informative about the respective parameters. For example, the prior  
 214 assigned to  $\eta_p$  in equation (6) was chosen such that the prior median was close to .1 but the  
 215 5<sup>th</sup> and 95<sup>th</sup> percentiles covered a broad range from .0 to .5.

216 As with most complex Bayesian models, the posterior distributions are not tractable and  
 217 inferences were obtained via MCMC simulation implemented in OpenBUGS (Thomas et al.,  
 218 2006). Five parallel chains starting from diffuse initial values were run for each model and  
 219 convergence was monitored via the Gelman-Rubin-Brooks diagnostics for both the mean and  
 220 upper 97.5% quantile of each parameter (Brooks and Gelman, 1998). Each chain was run for  
 221 a total of 500,000 iterations, the first 100,000 iterations were discarded, and the remaining  
 222 iterations were thinned by a factor of 50 to save storage-space and reduce auto-correlations,  
 223 producing a final sample size of 8,000 values from each of the 5 chains.

224 Comparisons between the candidate models were assessed with the deviance information  
 225 criterion (DIC) (Spiegelhalter et al., 2002). Denoting the deviance by:

$$D(\mathbf{U}, \mathbf{p}, \Theta | \mathbf{n}, \mathbf{M}, \mathbf{u}) = -2 \log(L(\mathbf{U}, \mathbf{p}, \Theta | \mathbf{n}, \mathbf{M}, \mathbf{u}))$$

226 the DIC for any model is:

$$\text{DIC} = D(\hat{\mathbf{U}}, \hat{\mathbf{p}}, \hat{\Theta} | \mathbf{n}, \mathbf{M}, \mathbf{u}) + 2p_D$$

227 where  $\hat{\mathbf{U}}$ ,  $\hat{\mathbf{p}}$ , and  $\hat{\Theta}$  are point estimates (taken to be the posterior means) and  $p_D$  is a measure  
 228 of the effective number of parameters computed as the mean value of the deviance over all  
 229 MCMC iterations minus the value of the deviance at the posterior means (see Spiegelhalter  
 230 et al. (2002, pg. 591) for further details). In essence, the DIC provides a Bayesian analog of  
 231 the AIC that is easily computed for hierarchical models for which the number of parameters  
 232 is poorly defined.

233 Fit of the models was assessed by computing posterior predictive p-values (also called  
 234 Bayesian p-values) (Gelman and Meng, 1996; Gelman et al., 2003, pg. 157–177). For a given  
 235 discrepancy measure,  $DM(\boldsymbol{\Omega}, \mathbf{X})$  with  $\boldsymbol{\Omega}$  representing the complete set of parameters and  
 236  $\mathbf{X}$  the observed data, the Bayesian p-value is the probability that new data,  $\mathbf{X}'$ , generated  
 237 from the posterior predictive distribution produces a greater discrepancy than the observed  
 238 data ( $p = P(DM(\boldsymbol{\Omega}, \mathbf{X}') > DM(\boldsymbol{\Omega}, \mathbf{X}) | \mathbf{X})$ ). Values of  $p$  close to 0 or 1 indicate that the  
 239 observed data are unlikely under the assumed model. Different discrepancy measures can  
 240 be chosen to assess different components of the model, and we have chosen three specific  
 241 measures:

$$DM_1(\boldsymbol{\Omega}, \mathbf{X}) = -2 \cdot L(\mathbf{U}, \mathbf{p}, \boldsymbol{\Theta} | \mathbf{n}, \mathbf{M}, \mathbf{u})$$

242 providing an overall assessment of the model:

$$DM_2(\boldsymbol{\Omega}, \mathbf{X}) = \sum_{j=1}^t \left( \sqrt{\sum_{i=1}^s m_{i,j}} - \sqrt{\sum_{i=1}^s n_i \theta_{i,j} p_j} \right)^2$$

243 comparing the observed and expected number of marked fish recaptured in each strata, and:

$$DM_3(\boldsymbol{\Omega}, \mathbf{X}) = \sum_{i=1}^s \sum_{j=1}^t \left( \sqrt{m_{i,j}} - \sqrt{n_i \theta_{i,j} p_j} \right)^2$$

244 assessing the number fish in each cell of  $\mathbf{M}$ . The latter two are based on the Freeman-Tukey  
 245 statistic as recommended for mark-recapture data by Brooks et al. (2000).

246 An R package called BTSPAS (Bayesian Time-Stratified Population Analysis System)  
 247 which implements the Bayesian P-spline model of  $\mathbf{U}$  and both models of  $\boldsymbol{\Theta}$  along with  
 248 computations of the DIC and Bayesian p-values is available from the Comprehensive R  
 249 Archive Network (<http://cran.r-project.org/index.html>). This packages also includes  
 250 a demonstration based on the analysis of the following section and implements extensions  
 251 of this model to the case of diagonal  $\mathbf{M}$ , separation of stocks (e.g., hatchery and wild fish),  
 252 and inclusion of covariates for  $\mathbf{p}$ .

253 [Figure 1 about here.]

### 3. Application

Here we apply the models described in the previous section to analyse modified data from the study of Atlantic salmon smolts migrating along the Conne River, Newfoundland, in 1987. In this study, smolts were trapped at two sites along the river for 46 days between April 26 and June 10. A total of 4975 smolts were marked at the first trap of which 998 (20%) were recaptured at the second trap along with 13,363 unmarked smolts.

SD94 and MR02 both used this data set to illustrate their methods. SD94 excluded data from the first 3 and final 17 days of the study because small numbers fish were marked in these periods and estimated the total size of the run in the remaining period to be 75 thousand fish with 95% CI (68,82) thousand. MR02 were able to analyse the entire data set and also estimated the total run size at 75 thousand with 95% CI (69,81) thousand.

Although SD94 found it necessary to exclude some strata, data sparsity is not a severe issue. Our focus is to improve estimation for sparse data sets, and so we simulated new data with smaller capture probabilities by artificially reducing the numbers of fish captured at the second trap. To do this, we generated new values of  $m_{i,j}$  and  $u_j$  as:

$$m_{i,j} \sim \text{Binomial}(m_{i,j}^*, .2) \text{ and } u_j \sim \text{Binomial}(u_j^*, .2)$$

where  $m_{i,j}^*$  and  $u_j^*$  are the original data values. This effectively reduced the capture probability on each day of the study by a factor of 5, but maintained the other relationships in the data. The new data set comprised the same number of marked fish released, but the total numbers of marked and unmarked smolts captured at the second site were reduced to 183 and 2697 respectively. A complete listing of the modified data set is provided in Web Table 1.

Tables 1 summarizes inferences about the total population size from the 10 candidate models, and Table 2 presents the DIC and Bayesian p-values. Summaries of the posterior distributions of  $\mathbf{U}$  and  $\mathbf{p}$  for the 5 models using the NP prior of  $\Theta$  are depicted in Figures

277 2 and 3. Summaries of  $\mathbf{U}$  and  $\mathbf{p}$  for the models using the LN model were indistinguishable  
 278 from these at the daily level and so have been omitted.

279 One clear result is that the NP model of  $\Theta$  was preferred to the LN model. Regardless  
 280 of the priors chosen for  $\mathbf{U}$  and  $\mathbf{p}$ , switching from the LN to NP model of  $\Theta$  decreased the  
 281 DIC by more than 15 units (Table 2). In all cases this decrease reflected reductions in both  
 282  $p_D$  and the deviance, implying that the NP model produced better fit to the data with a  
 283 smaller effective number of parameters. Switching from the LN to NP model did move the  
 284 first and second Bayesian p-values closer to 0, but not by enough to question the goodness-  
 285 of-fit (Table 2). Despite the strong evidence in favour of the NP prior, the choice of model  
 286 for  $\Theta$  had only a small effect on the posterior distribution of  $N_{Tot}$ . For any combination of  
 287 priors for  $\mathbf{U}$  and  $\mathbf{p}$ , the posterior mean of  $N_{Tot}$  was about 3 thousand higher under the NP  
 288 model of  $\Theta$ , but the 95% CIs from the paired models overlapped by roughly 90% (Table 1).

289 Comparing alternative models of  $\mathbf{p}$  given the NP model of  $\Theta$ , the results clearly indicate  
 290 that the hierarchical model was preferable to the pooled model. Considering only the posterior  
 291 distribution of  $N_{Tot}$ , one might be tempted to select model Simple/Pooled/NP as it provides  
 292 a point estimate that is similar to those from the other models and the narrowest interval  
 293 estimates (Table 1). However, the DIC for this model was higher than the DIC for all other  
 294 models with the NP prior for  $\Theta$  and the Bayesian p-values provided strong evidence that  
 295 constant  $p_j$  is not appropriate for this data. Results from the simulation study in Section  
 296 4 also indicated that the interval estimates from the pooled model would have coverage far  
 297 below the nominal value when the capture probabilities vary between strata.

298 Considering the remaining four models which combined the NP model of  $\Theta$  and the  
 299 hierarchical model of  $\mathbf{p}$ , the P-spline model of  $\mathbf{U}$  was preferred both for its fit to the data  
 300 and precision in estimating  $N_{Tot}$ . Figure 1 shows that the model Simple/Hier/NP provided  
 301 relatively precise estimates at the peak of the run, but the posterior of  $U_j$  was highly variable

302 for  $j$  near 1 or  $t$ . Model Hier/Hier/NP decreased the variability of these values, as well as  $N_{Tot}$ ,  
303 but increased the DIC (Table 2) indicating that the hierarchical prior with constant mean is  
304 also not appropriate for this data. In contrast, models Cubic/Hier/NP and P-spline/Hier/NP  
305 lowered the posterior variability of the individual  $U_j$  and of  $N_{Tot}$  (Figure 2 and Table 1),  
306 and also decreased both  $p_D$  and the DIC (Table 2). Compared to each other, the curves fit  
307 to  $E(\log(U_j))$  were difficult to distinguish by eye (Figure 2) and the estimates of  $N_{Tot}$  were  
308 almost coincidental, but the DIC for model P-spline/Hier/NP was lower by 6 units (Table  
309 1). Bayesian p-values for these models presented no evidence to question their fit to the data  
310 (Table 2).

311 Convergence diagnostics indicated that the Markov chains for all models converged quickly  
312 to stable distributions and that the samples obtained represented the posterior distributions  
313 well. Trace plots and GRB diagnostics comparing the 5 parallel chains for each model  
314 indicated that the burn-in period was more than sufficient. Web Appendix A provides details  
315 for model P-spline/Hier/NP (the model with the smallest DIC) and results were similar for  
316 the remaining models.

317 Effects of the choice of prior for the P-spline parameters were assessed through a sensitivity  
318 analysis. We refit model P-spline/Hier/NP a total of 25 times for all combinations of 5 inverse  
319 gamma priors for  $\tau_b^2$  and  $\tau_\epsilon^2$  that varied according to their location and relative spread. Details  
320 of the selected priors and complete results are provided in Web Appendix B. In summary, the  
321 sensitivity analysis indicated that the shape of the fitted spline and the posterior distribution  
322 of  $N_{Tot}$  were robust to the choice of priors. However, the posterior distributions of the  
323 individual  $U_j$  were affected by the choice of prior for  $\tau_\epsilon^2$ , and so we recommend that any  
324 analysis be repeated using a range of prior parameterizations – particularly if one is interested  
325 in local characteristics of the run (e.g., peak day).

[Table 2 about here.]

[Figure 2 about here.]

[Figure 3 about here.]

#### 4. Simulation Study

To examine the P-spline model further, we conducted a simulation study based on the results of the analysis of Section 3. In short, 300 data sets were simulated from the model described by equations (1) and (2) with  $\log(U_j)$  generated from a smooth curve similar to the spline fit in Section 3,  $\theta_{i,j}$  generated from the LN model but truncated so that  $\theta_{i,i+k} = 0$  for  $k > 6$ , and  $p_j$  generated from a distribution approximating the posterior obtained in Section 3. The 10 models described previously were fit to each data set and compared based on 1) the resulting bias and mean-squared error (MSE) of the posterior mean of  $N_{Tot}$  as a point estimate of the true population size and 2) the width and coverage of the associated 95% credible intervals. Performance of the DIC for model selection was also assessed. Full details of the simulation study are presented in Web Appendix C; here, we summarize the key results.

Overall, the Bayesian P-spline models produced estimates of  $N_{Tot}$  that were nearly unbiased and had small MSE relative to the other models, and also gave 95% credible intervals with exactly the nominal coverage. The only models which performed similarly were those fitting a cubic polynomial to  $E(\log(U_j))$ . These models actually had smaller bias and MSE, though coverage of their credible intervals was slightly below the nominal value. The models with the simple or hierarchical fit to  $E(\log(U_j))$  performed poorly in that they had large bias and MSE, and although point estimates from the models with pooled capture probabilities were nearly unbiased, the coverage of the credible intervals was only  $\frac{2}{3}$  of the nominal value.

In accordance with these results, the DIC selected the Bayesian P-spline model of  $E(\log(U_j))$  for 66% of the simulated data sets (26% and 40% in combination with the LN and NP models

351 of  $\Theta$  respectively). The cubic model was selected for 30% of data sets (11% and 19% in  
 352 combination with the LN and NP models of  $\Theta$ ), and of the remaining 4% of the simulated  
 353 data sets, model Simple/Pooled/NP was selected 8 times, Simple/Hier/NP 3 times, and  
 354 Hier/Hier/NP once. The fact that the cubic model was selected for almost  $\frac{1}{3}$  of the data  
 355 sets was not surprising given that a cubic polynomial fits the true  $E(\log(U_j))$  very closely  
 356 and requires fewer parameters than the spline. Despite the similar performance of the two  
 357 models, we prefer the Bayesian P-spline because of its abilities to fit a wide range of curves  
 358 and to properly account for model uncertainty. These issues are discussed further in Section  
 359 5. Similar DIC values were also to be expected from the alternative priors for  $\Theta$  as the true  
 360  $\theta_{ij}$  were truncated, as in the NP model, but were well approximated by the LN model. These  
 361 results confirm that the differences in DIC seen in Section 3 provide strong evidence for  
 362 selecting between models and further support our conclusion that P-Spline/Hier/NP is the  
 363 best of the candidate models for the subsampled Atlantic salmon data.

## 364 5. Conclusion

365 The results of Sections 3 and 4 clearly demonstrate the advantages of the P-spline model of  $\mathbf{U}$   
 366 for estimating the size of a population from temporally-stratified mark-recapture data. The  
 367 P-spline accounts for the natural ordering of the data by explicitly modelling the expected  
 368 (log) population size per strata as a smooth function of time. The result is a nearly unbiased  
 369 estimate of  $N_{Tot}$  that is as or more precise than those from the other models.

370 The model which produced results most similar to those of the P-spline model was the  
 371 Cubic model of  $E(\log(U_j))$ . The cubic and P-spline curves fit to  $E(\log(U_j))$  in Section 3  
 372 were almost identical and these models performed almost equally well in the simulation  
 373 study. However, it is easy to imagine data for which the cubic would perform poorly, for  
 374 example if the run size had two distinct peaks. Several strategies might be used to fit higher  
 375 degree polynomials to such data, but each has drawbacks. One might test polynomials with

376 increasing degree until an adequate model was found, but this would constitute data-snooping  
377 and variance estimates would not account for the trial-and-error process. Alternatively, a high  
378 degree polynomial could be fit to all data, but this would lead to overfitting, and reversible  
379 jump MCMC could be used to select the degree of the polynomial (Green, 1995), but this  
380 requires more complex computation. The Bayesian P-spline solves the problem of unknown  
381 model complexity by allowing flexibility in the degree of the curve while penalizing overly  
382 complex models. Moreover, by including  $\tau_b^2$  as an unknown parameter in the model the  
383 variance estimates properly account for uncertainty in the complexity.

384 This is the first time, to our knowledge, that splines have been applied to model changes in  
385 population size over time from mark-recapture data, though similar methods have been used  
386 to model other aspects of population dynamics. Methods using splines to allow flexibility  
387 in the relationship between covariates and survival probabilities have been presented by  
388 Gimenez et al. (2006), Bonner et al. (2009), Gimenez and Barbraud (2009), and Gimenez  
389 et al. (2009). While the nature of these models are similar, and all apply Bayesian inference,  
390 the specific spline implementations vary. For example, Gimenez et al. (2006) applied a  
391 penalized spline approach using the truncated polynomial basis functions, whereas Bonner  
392 et al. (2009) applied a free-knot method in which the knot locations were chosen as part  
393 of the inference procedure. Gimenez and Barbraud (2009) introduced a two-stage inference  
394 procedure to reduce computation by approximating the likelihood function, and Gimenez  
395 et al. (2009) developed fitness surfaces that simultaneously model the effects of two or  
396 more variables on survival with multivariate splines. Each method has advantages and  
397 disadvantages, and there is a large body of literature on splines (see for example Ruppert  
398 et al., 2003). While it is not clear if one method is particularly suited to modelling mark-  
399 recapture data, the Bayesian P-spline model of Lang and Brezger (2004) has worked well in  
400 the examples we have tried thus far.

401 As mentioned previously, there is a variant of the two sample experiment which requires  
402 only one trap. When only one trap is used, a subset of the individuals captured in each  
403 strata are marked and released back into the population upstream so that they pass the  
404 trap a second time. The only difference in modelling is that the marked fish do not need  
405 to be recounted when they pass the trap a second time and so the total population size  
406 satisfies  $N_{Tot} = U_{Tot}$ . An example of the one trap experiment is provided in (Bonner, 2008).  
407 In some studies it may also be known that  $\theta_{i,j} = 1(i = j)$  and so it is not necessary to model  
408 the movements of the marked fish. Such data are called diagonal because  $\mathbf{M}$  is a diagonal  
409 matrix and examples of fitting the P-spline model to such data are provided by Schwarz  
410 et al. (2009).

411 The P-spline model also provides a way to deal with some common data anomalies. Fish-  
412 eries traps often cannot be operated continuously because of adverse weather conditions, and  
413 the data from such studies may contain missing values in some strata. Classical estimators  
414 ignore these strata completely, while standard hierarchical Bayesian models can produce  
415 inference by drawing on the data from other days but treat all other strata equally rather  
416 than weighting information from the neighbouring strata more heavily. The P-spline model  
417 can easily be fit to data with missing days and essentially smooths the run size across the  
418 strata with unobserved data, allowing for uncertainty by addition of the error term. Of course,  
419 this relies on the untestable assumption that the patterns in the observed data continue over  
420 the missing strata.

421 One drawback of the P-spline approach is that MCMC simulations can be time consuming  
422 and sometimes misleading. While the algorithms implemented in OpenBUGS generally  
423 performed well, we did encounter one data set in the simulation study in which the single  
424 chain became trapped in a local maximum of the posterior density. The error was fixed simply  
425 by choosing a new seed value, but this emphasized the importance of running multiple chains

426 to assess convergence, as we did in Section 3. Gains from the model will also be small if the  
427 run size has a very simple underlying shape or if the number of strata is small so that  
428 the smoothness cannot be well estimated. The model may also produce misleading results  
429 if  $E(\log(U_j))$  is not smooth – for example, if there are jumps in  $U_j$  associated with sudden  
430 outmigration or the release of hatchery fish. If the timing of such events are known then they  
431 can be accommodated by incorporating breaks in the P-spline, and an example is provided  
432 in Bonner (2008).

433 Another concern is that the model treats the expected number of unmarked fish in each  
434 stratum,  $E(\log(U_j))$ , as the parameter of interest rather than the total number of fish,  
435  $E(\log(N_j))$ . The difficulty is that  $N_j$  does not appear in the likelihood in equation (3) which  
436 conditions on  $\mathbf{n}$ . One could adopt a likelihood which models captures at both locations, as  
437 in Plante et al. (1998), and apply the P-spline to the numbers of fish passing the first trap  
438 in each strata, but this likelihood effectively doubles the number of parameters. Instead,  
439 we have considered applying the Bayesian melding algorithm of Poole and Raftery (2000)  
440 which allows specification of prior information for derived parameters like  $N_j$  or  $N_{Tot}$ . Note,  
441 however, that if relatively few fish are marked then  $N_j \approx U_j$  and whether the prior is defined  
442 for  $N_j$  or  $U_j$  will have little effect on inference.

443 Another minor issue is that the coefficients of the P-spline do not have direct interpre-  
444 tations. However, inference about quantities like run timing can easily be generated from  
445 the MCMC output. For example, inference about the day by which 50% of the population  
446 has passed the traps can be generated simply by computing this quantity for each iteration  
447 and then computing posterior summary statistics. Inference about many other biological  
448 quantities, like the peak day of the run, can be obtained in exactly the same way.

449 Although we have focused on smoothing abundance, the Bayesian P-spline could also be  
450 applied to smooth  $\mathbf{p}$  or the elements of  $\Theta$ . If  $p_j$  was expected to change smoothly with

451 time then it would be sensible to model  $E(\text{logit}(p_j))$  in a similar manner. The Bayesian  
452 P-spline model could also be applied to multiple sets of parameters simultaneously, though  
453 we found that such models did not perform well in application to the Conne River data set.  
454 Specifically, models which smoothed both  $\mathbf{U}$  and  $\mathbf{p}$  produced poor estimates of abundance  
455 and DIC values that were much higher than those of the other models. Developing techniques  
456 to smooth multiple parameters simultaneously is a continuing aspect of our research.

### 457 **Supplementary Materials**

458 Web Appendices, Figures, and Tables referenced in Sections 3 and 4 are available under the  
459 Paper Information link at the Biometrics website <http://www.biometrics.tibs.org>.

### 460 **Acknowledgements**

461 This research was supported by the National Science and Engineering Research Council  
462 of Canada, the Pacific Institute for the Mathematical Sciences, and the National Science  
463 Foundation (Grant No. 0814194). Further funding and data for model development were  
464 provided by the Trinity River Restoration Program. We would also like to thank the editor  
465 and two anonymous referees for their constructive comments on earlier versions of the  
466 manuscript.

### 467 **References**

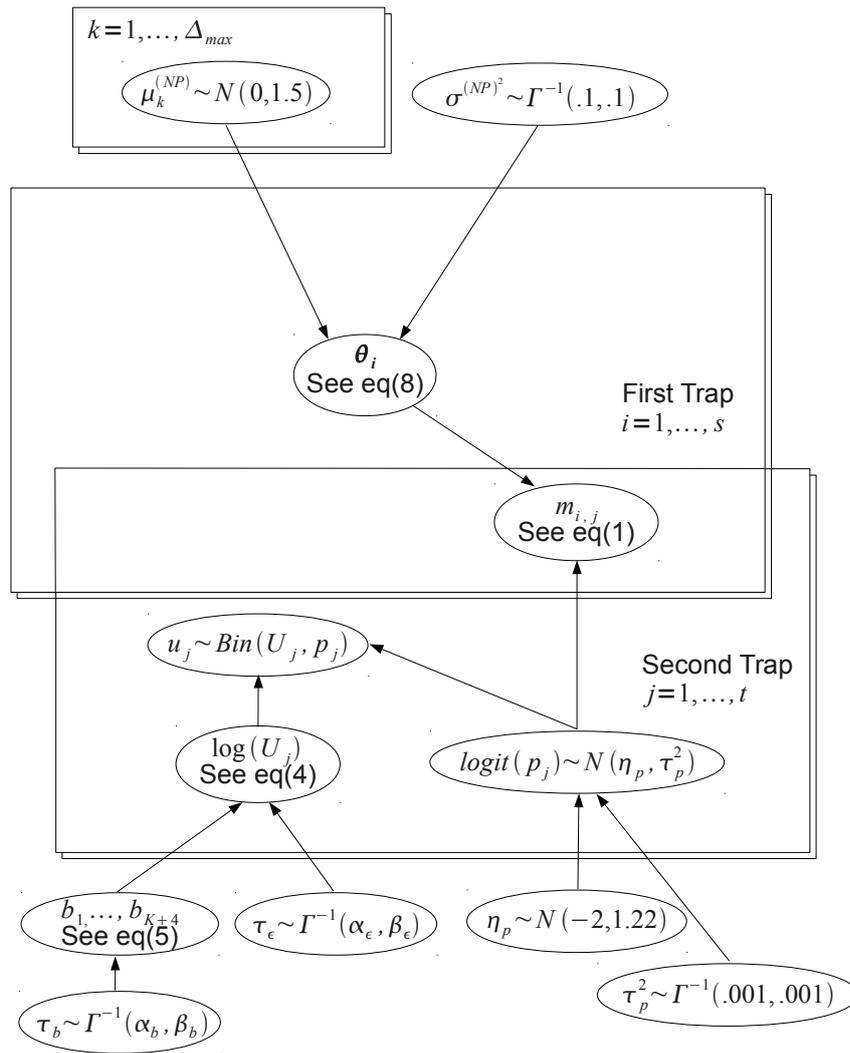
- 468 Banneheka, S. G., Routledge, R. D., and Schwarz, C. J. (1997). Stratified two-sample tag-  
469 recovery census of closed populations. *Biometrics* **53**, 1212–1224.
- 470 Bjorkstedt, E. P. (2000). DARR (Darroch analysis with rank-reduction): A method for  
471 analysis of stratified mark-recapture data from small populations, with application to  
472 estimating abundance of smolts from outmigrant trap data. Technical report, National  
473 Marine Fisheries Service, Southwest Fisheries Science Center.

- 474 Bonner, S. J. (2008). *Heterogeneity in capture-recapture: Bayesian methods to balance realism*  
475 *and model complexity*. PhD thesis, Simon Fraser University, Burnaby, BC, Canada.
- 476 Bonner, S. J., Thomson, D., and Schwarz, C. J. (2009). Time-varying covariates and semi-  
477 parametric regression in capture-recapture: an adaptive spline approach. In Thomson,  
478 D. L., Cooch, E. G., and Conroy, M. J., editors, *Modeling Demographic Processes in*  
479 *Marked Populations*, Environmental and Ecological Statistics 3, pages 657–676. Springer.
- 480 Brooks, S. P., Catchpole, E. A., and Morgan, B. J. T. (2000). Bayesian animal survival  
481 estimation. *Statistical Science* **15**, 357–376.
- 482 Brooks, S. P. and Gelman, A. (1998). General methods for monitoring convergence of  
483 iterative simulations. *Journal of Computational and Graphical Statistics* **7**, 434–455.
- 484 Chapman, D. G. and Junge, Jr, C. O. (1956). The estimation of the size of a stratified  
485 animal population. *The Annals of Mathematical Statistics* **27**, 375–389.
- 486 Darroch, J. N. (1961). The two-sample capture-recapture census when tagging and sampling  
487 are stratified. *Biometrika* **48**, 241–260.
- 488 Dobson, A. J. (2002). *An Introduction to Generalized Linear Models*. Chapman and  
489 Hall/CRC Press, Boca Raton, Fl., 2nd edition.
- 490 Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2003). *Bayesian Data Analysis*.  
491 Chapman and Hall, 2nd edition.
- 492 Gelman, A. and Meng, X.-L. (1996). Model checking and model improvement. In Gilks, W.,  
493 Richardson, S., and Spiegelhalter, D., editors, *Markov chain Monte Carlo in practice*,  
494 pages 189–201. Chapman and Hall, London.
- 495 Gimenez, O. and Barbraud, C. (2009). The efficient semiparametric regression modeling  
496 of capture-recapture data: Assessing the impact of climate on survival of two Antarctic  
497 seabird species. In Thomson, D. L., Cooch, E. G., and Conroy, M. J., editors, *Modeling*  
498 *Demographic Processes in Marked Populations*, Environmental and Ecological Statistics

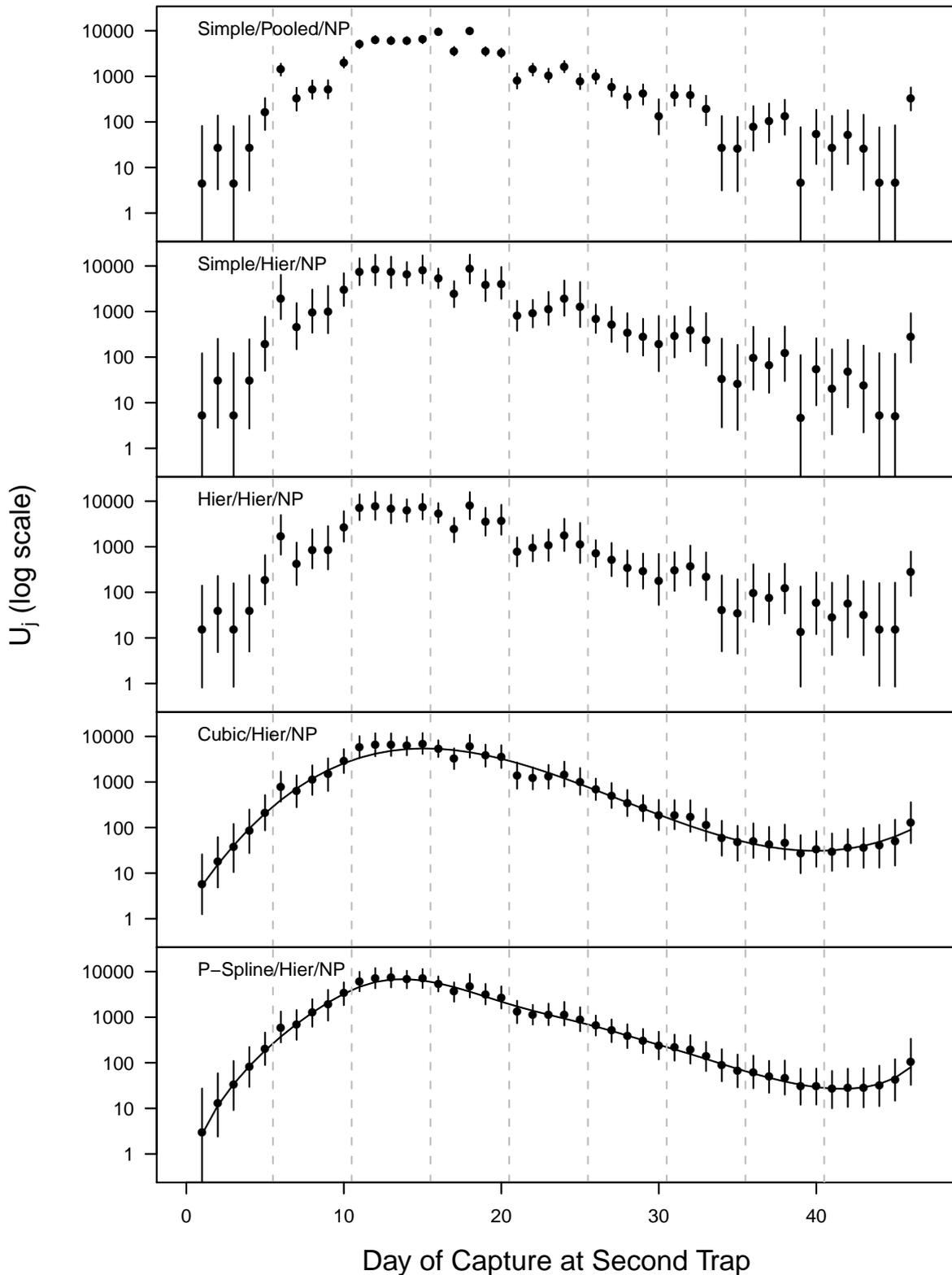
- 3, pages 43–58. Springer.
- Gimenez, O., Crainiceanu, C., Barbraud, C., Jenouvrier, S., and Morgan, B. J. T. (2006). Semiparametric regression in capture-recapture modelling. *Biometrics* **62**, 691–698.
- Gimenez, O., Grégoire, A., and Lenormand, T. (2009). Estimating and visualizing fitness surfaces using mark-recapture data. *Evolution* **63**, 3097–3105.
- Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82**, 711–7332.
- Lang, S. and Brezger, A. (2004). Bayesian P-splines. *Journal of Computational and Graphical Statistics* **13**, 183–212.
- Macdonald, P. D. and Smith, H. D. (1980). Mark-recapture estimation of salmon smolt runs. *Biometrics* **36**, 401–417.
- Mantyniemi, S. and Romakkaniemi, A. (2002). Bayesian mark-recapture estimation with an application to a salmonid smolt population. *Canadian Journal of Fisheries and Aquatic Science* **59**, 1748–58.
- Plante, N., Rivest, L.-P., and Tremblay, G. (1998). Stratified capture-recapture estimation of the size of a closed population. *Biometrics* **54**, 47–60.
- Poole, D. and Raftery, A. E. (2000). Inference for deterministic simulation models: The Bayesian melding approach. *Journal of the American Statistical Association* **95**, 1244–1255.
- Ruppert, D., Wand, M. P., and Carroll, R. J. (2003). *Semiparametric Regression*. Cambridge University Press, Cambridge, UK.
- Schwarz, C. J. and Dempson, J. B. (1994). Mark-recapture estimation of a salmon smolt population. *Biometrics* **50**, 98–108.
- Schwarz, C. J., Pickard, D., Marine, K., and Bonner, S. J. (2009). Juvenile salmonid outmigrant monitoring evaluation. Technical report, Prepared for the Trinity River

- 524 Restoration Program, Weaverville, CA.
- 525 Schwarz, C. J. and Taylor, C. G. (1998). Use of the stratified-Petersen estimator in fisheries  
526 management: estimating the number of pink salmon (*Oncorhynchus gorbuscha*) spawners  
527 in the Fraser River. *Canadian Journal of Fisheries and Aquatic Science* **55**, 281–296.
- 528 Seber, G. A. F. (2002). *The Estimation of Animal Abundance and Related Parameters*. The  
529 Blackburn Press, New Jersey, USA, 2nd edition.
- 530 Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. (2002). Bayesian  
531 measures of model complexity and fit (with discussion). *Journal of the Royal Statistical  
532 Society B* **64**, 583–640.
- 533 Steinhorst, K., Wu, Y., Dennis, B., and Kline, P. (2004). Confidence intervals for fish out-  
534 migration estimates using stratified trap efficiency methods. *Journal of Agricultural,  
535 Biological and Environmental Statistics* **9**, 284–299.
- 536 Thomas, A., O’Hara, B., Ligges, U., and Stutz, S. (2006). Making BUGS open. *R News* **6**,  
537 12–17.

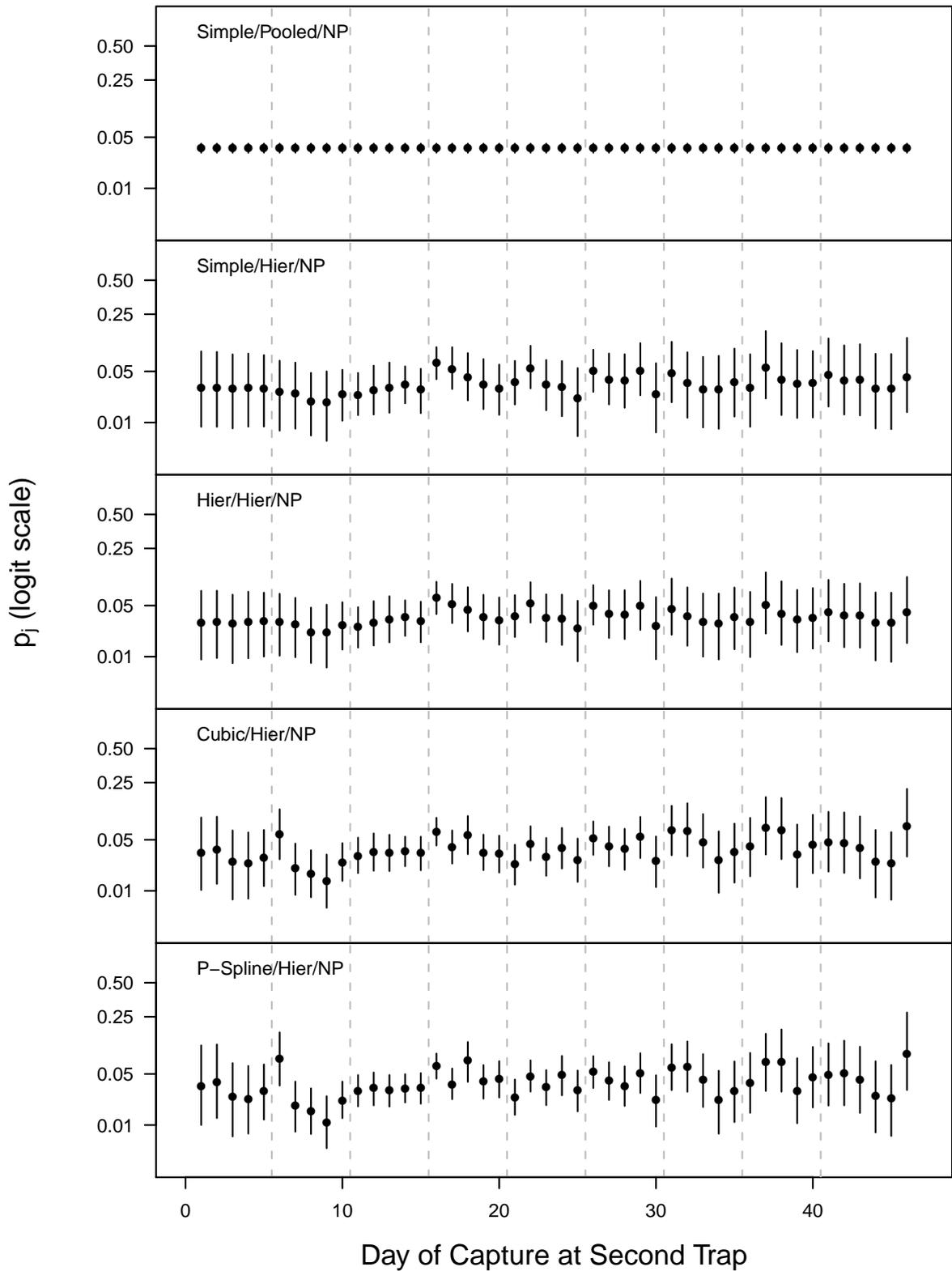
1 March 2011



**Figure 1.** Directed acyclic graph (DAG) illustrating the structure of the Bayesian P-spline model with non-parametric travel times (P-Spline/Hier/NP), the best fitting model for the subsampled Conne River data. Ellipses represent nodes that are stochastic and diamonds nodes that are deterministic conditional on their parents. The rectangular frames represent structures that are repeated over the days of trapping at either the first or second trap locations. Nodes outside of these frames do not repeat.



**Figure 2.** Estimated daily abundance of unmarked fish for the subsampled Conne River Atlantic Salmon data. Each panel summarizes the posterior distribution of  $\log(U_j)$ ,  $j = 1, \dots, 46$ , resulting from one of the 5 base models combined with the non-parametric model of the travel time probabilities, as indicated. Points represent the daily posterior medians and the error bars represent the 95% credible intervals. The curves in the final two panels represent the functions fit to  $E(\log(U_j))$ .



**Figure 3.** Estimated daily capture probabilities for the subsampled Conne River Atlantic Salmon data. Each panel summarizes the posterior distribution of  $\text{logit}(p_j)$ ,  $j = 1, \dots, 46$ , resulting from one of the 5 base models combined with the non-parametric model of the travel time probabilities, as indicated. Points represent the daily posterior medians and the error bars represent the 95% credible intervals.

**Table 1**

Summaries of the posterior distributions for  $N_{Tot}$  obtained from the 10 models fit to the subsampled Conne River Atlantic salmon data. The first two columns provide the posterior mean ( $\hat{N}_{Tot}$ ) and 95% highest posterior density credible interval (95% CI) in thousands of fish. The final two columns give the width of this credible interval absolutely (CI Width) and as a percentage of the posterior mean (% CI Width). Each model is labelled by a triplet identifying the components of the prior as described in the text.

Model	$\hat{N}_{Tot}$	95% CI	CI Width	% CI Width
Simple/Pooled/LN	77	(67,89)	22	29
Simple/Pooled/NP	80	(69,92)	22	28
Simple/Hier/LN	87	(10,107)	37	42
Simple/Hier/NP	90	(73,111)	37	42
Hier/Hier/LN	81	(67,100)	33	40
Hier/Hier/NP	85	(70,103)	32	38
Cubic/Hier/LN	76	(64,90)	26	35
Cubic/Hier/NP	79	(67,92)	26	33
P-Spline/Hier/LN	76	(64,89)	25	33
P-Spline/Hier/NP	78	(66,92)	26	33

**Table 2**

*Model selection and goodness-of-fit results for the 10 Bayesian models fit to the subsampled Conne River Atlantic Salmon smolt data. The left half of the table presents the values of  $p_D$  and DIC for each model, and the right half presents the Bayesian  $p$ -values for each discrepancy measure. The models are labelled by triplets identifying the components of the prior as described in the text.*

Model	Model Selection		Bayesian P-values		
	$p_D$	DIC	1	2	3
Simple/Pooled/LN	83.4	691	0.26	0.01	0.02
Simple/Pooled/NP	80.7	671	0.05	0.00	0.04
Simple/Hier/LN	94.4	672	0.41	0.29	0.32
Simple/Hier/NP	89.5	656	0.19	0.19	0.29
Hier/Hier/LN	93.0	676	0.47	0.20	0.24
Hier/Hier/NP	88.0	659	0.17	0.17	0.23
Cubic/Hier/LN	84.6	670	0.58	0.25	0.31
Cubic/Hier/NP	79.8	650	0.28	0.21	0.32
P-Spline/Hier/LN	85.1	666	0.55	0.26	0.33
P-Spline/Hier/NP	79.7	644	0.28	0.24	0.36