The Canadian Journal of Statistics
Vol. 36, No. 1, 2008, Pages 5–28
La revue canadienne de statistique

Bayesian analysis of mark-recapture data with travel time-dependent survival probabilities

Saman MUTHUKUMARANA, Carl J. SCHWARZ and Tim B. SWARTZ

Key words and phrases: Bayesian analysis; latent variable; mark-recapture; salmon; WINBUGS.

MSC 2000: Primary 62P10; secondary 62F15.

Abstract: The authors extend the classical Cormack–Jolly–Seber mark-recapture model to account for both temporal and spatial movement through a series of markers (e.g., dams). Survival rates are modeled as a function of (possibly) unobserved travel times. Because of the complex nature of the likelihood, they use a Bayesian approach based on the complete data likelihood, and integrate the posterior through Markov chain Monte Carlo methods. They test the model through simulations and apply it also to actual salmon data arising from the Columbia river system. The methodology was developed for use by the Pacific Ocean Shelf Tracking (POST) project.

Analyse bayésienne de données de capture-recapture à l'aide de probabilités de survie dépendant du temps de déplacement

Résumé : Les auteurs généralisent le modèle de capture-recapture classique de Cormack–Jolly–Seber pour tenir compte de déplacements spatiaux-temporels signalés par des marqueurs (tels que des barrages). Les taux de survie sont modélisés en fonction de temps de déplacement parfois inobservables. Vu la complexité de la vraisemblance, ils optent pour une approche bayésienne fondée sur la vraisemblance des données complètes et intègrent la loi a posteriori par des méthodes de Monte-Carlo à chaîne de Markov. Ils testent le modèle par simulation et l'utilisent pour l'analyse de données sur les saumons du réseau hydrographique de la Columbia. La méthodologie a été développée aux fins du projet POST (Pacific Ocean Shelf Tracking).

1. INTRODUCTION

The Pacific Ocean Shelf Tracking (POST) project (http://www.postcoml.org) is part of the Census of Marine Life study (http://www.coml.org). In the POST project, acoustic transmitters are surgically implanted into fish (e.g., salmon smolt). The fish are then tracked during their migration by a series of listening lines along the ocean shelf. These listening lines record the acoustic-identification of the fish that pass near the receivers and their times of passage. A complicating factor in the analysis of the POST data is that sometimes fish do not pass sufficiently close to the receiver, and hence are not detected. Data are downloaded from the listening lines and are stored in a database that can be queried by researchers.

The POST project is a combination of two types of mark-recapture experiments. In the first type of mark-recapture experiment (Lebreton, Burnham, Clobert & Anderson 1992), animals are initially marked, and then a subset of these animals are recaptured at yearly intervals (for example). In this type of experiment, the interest is in the temporal dimension of survival. For example, one may be interested in the survival rates of a species from year to year. In the POST project, we are also interested in the temporal dimension of survival and this information is captured by measuring the passage of time between listening lines.

In the second type of mark-recapture experiment, marked fish are released, and are detected as they swim past landmarks (Burnham et al. 1987, p. 25). In this type of experiment, the interest is in the spatial dimension of survival. For example, one may be interested in survival rates between particular dams. In the POST project, the listening lines are placed in fixed locations which may correspond to interest in the spatial dimension of survival.

This paper considers methods of combining both the temporal and the spatial dimensions of the problem into a single mark-recapture model. We provide a generalization of the Cormack–

Jolly–Seber model (Cormack 1964; Jolly 1965; Seber 1965). Their model assumes that survival probabilities between listening lines are homogeneous amongst all animals. In our generalization, the travel times of individual animals between listening lines differ amongst animals and survival may be a function of travel time. Cowen & Schwarz (2005) considered a similar problem, but they assumed that survival rates between listening lines are independent of travel time. We do not make this restrictive assumption. We model survival probabilities as a function of travel times. Pollock, Bunck, Winterstein & Chen (1995) considered Kaplan–Meier estimation in a context that allows for differential survival and the possibility of relocation of some animals. We allow for nondetection as in standard mark-recapture models.

A standard likelihood approach is difficult because of the presence of multidimensional integrals. A natural way to approach this problem is via a Bayesian framework through the use of latent (unobservable) variables. A Bayesian approach via simulation avoids the maximization of likelihoods which may be problematic in high-dimensional problems. Latent variables arise in the experiment when marked fish go unobserved at listening lines. A complete data likelihood is easy to construct as it treats the latent variables as though they are observable. We then "integrate" over the complete data likelihood by obtaining a Markov chain Monte Carlo (MCMC) sample from the posterior. Brooks, Catchpole & Morgan (2000) provide a review of Bayesian methods in mark-recapture experiments.

In Section 2, we provide a detailed development of the Bayesian model. The complete data likelihood is derived where survival probabilities depend on travel times. When the distances between listening lines vary greatly, the dependence structure is clearly important. Prior distributions are then defined on the model parameters. Computation is discussed in Section 3. As the posterior distribution is complex and high-dimensional, we obtain posterior summary statistics which describe key features in the study. In particular, posterior expectations are approximated through MCMC methods using WinBUGS software (Spiegelhalter, Thomas & Best 2003). Unlike the Cormack–Jolly–Seber model, it is shown that nonidentifiability is not a problem for the proposed Bayesian model. In Section 4, we discuss the topics of model selection and model adequacy. In Section 5, we provide some examples and demonstrate the reliability of the approach via simulation. We also provide sensitivity analyses with respect to some of the model assumptions. We conclude with a short discussion in Section 6.

2. MODEL DEVELOPMENT

Consider a population of n fish where each fish is implanted with an acoustic transmitter. Without loss of generality, assume that all fish are released at location $j = 0$, and that listening lines are set up at locations $j = 1, \ldots, m$. The observed data for the experiment consist of (ω, T^{obs}) where $\omega = \{\omega_{ij}\}$ is the detection history such that

$$\omega_{ij} = \begin{cases} 1 & \text{if the } i\text{th fish is detected at location } j, \\ 0 & \text{if the } i\text{th fish is not detected at location } j \end{cases}$$

and $\omega_{i0} = 1$. The vector $T^{\text{obs}} = \{T_{ij}\}$ corresponds to observed cumulative travel times such that T_{ij} is the time required for the ith fish to travel from the point of release to location j. When a fish is not detected, then there is no observed cumulative travel time. We refer to the missing or latent cumulative travel times as T^{mis} and let $T = (T^{\text{obs}}, T^{\text{mis}})$. Therefore T represents the complete cumulative travel times. Note that when a fish has died (and is therefore not detected), we still imagine that there is a cumulative travel time associated with the fish. The value is missing but it represents the cumulative travel time that the fish would have taken had it been alive.

Associated with (ω, T) are the quantities (S^{obs}, t) where $S^{\text{obs}} = \{S_{ij}\}$ is a function of the detection history data ω and $t = \{t_{ij}\}$ is a function of the complete cumulative travel times T. The variable S_{ij} denotes the survival status of the ith fish at location j, where $S_{ij} = 1$ (0) indicates that the ith fish is alive (dead) at location j. Since fish are released alive, we have $S_{i0} = 1$.

Note that whereas the entire vector ω is observed, some of the entries S_{ij} are latent. This is due to the fact that an undetected fish may be either alive or dead. As an example, consider the observed data $(\omega_{i0}, \ldots, \omega_{i5}) = (1, 0, 0, 1, 0, 0)$. In this case, $(S_{i0}, \ldots, S_{i3}) = (1, 1, 1, 1)$, but S_{i4} and S_{i5} are latent. We supplement the observed S^{obs} with the missing or latent S^{mis} to give the complete survival history $S = (S^{\text{obs}}, S^{\text{mis}})$. The variable $t_{ij} = T_{ij} - T_{i,j-1}$ denotes the interval travel time for the ith fish from location $j-1$ to j. Because some of the T_{ij} may be missing, this implies that some of the t_{ij} may be missing. In fact, there are at least as many missing t_{ij} as there are missing T_{ij}. As an example, consider $(T_{i0}, T_{i1}, T_{i2}, T_{i3}, T_{i4}, T_{i5}) = (0, x, \text{NA}, \text{NA}, y, z)$ where NA denotes "Not Available". Then $(t_{i0}, t_{i1}, t_{i2}, t_{i3}, t_{i4}, t_{i5}) = (0, x, \text{NA}, \text{NA}, \text{NA}, z-y)$. Therefore, the vector t consists of both observed and latent data. Missing data issues have been previously considered in mark-recapture experiments. For example, Bonner & Schwarz (2006) showed how the classical Cormack–Jolly–Seber model can be extended for time-dependent individual covariates which form a set of missing values when animals are unobserved. Dupuis (1995) used directed graphs for the Bayesian analysis of mark-recapture experiments of the first type (Lebreton, Burnham, Clobert & Anderson 1992).

We now describe the two primary parameters of interest in the model. We let p_j denote the probability of detection at the jth location. As the acoustic transmitters are identical and the fish comprise a sample from an underlying population, one typically assumes that the probability p_j does not depend on fish i. In some instances, it may be reasonable to assume a common probability of detection (i.e., $p_j = p$ for all locations) although the general case causes no additional difficulty. The second parameter of interest concerns survival, where ϕ_{ij} denotes the survival probability of the ith fish when travelling from location $j-1$ to location j given that the fish was alive at location $j-1$. In Cowen & Schwarz (2005), the modelling assumption $\phi_{ij} = \phi_j$ implies that survival probabilities are independent of travel times. In our paper, we consider $\phi_{ij} = f(t_{ij})$ where f is a specified decreasing parametric function. Therefore, the longer that it takes a fish to travel between locations $j-1$ and j, the greater the chance that the fish does not survive. In our datasets, travel times are measured in days, and we define $\phi_{ij} = q_j^{t_{ij}}$ such that q_j denotes the daily survival probability when travelling between locations $j-1$ and j. Our modelling assumption implies that survival is independent across days. Therefore, the proposed framework reduces the primary parameters of interest to (p, q) where $p = \{p_j\}$ and $q = \{q_j\}$. Gimenez et al. (2006) consider a penalized spline approach when modelling survival probabilities in a semi-parametric fashion.

In Cowen & Schwarz (2005), an observed likelihood is obtained based on the observed data (ω, T^{obs}). The observed likelihood is complex as it involves integrals with respect to the latent cumulative travel times T^{mis}. We take an approach based on the complete data likelihood as in van Deusen (2002). The complete data likelihood treats latent variables as though they are available, and is especially well suited to Bayesian analyses (as will be seen). An advantage of the complete data likelihood over the observed likelihood is that it has a much simpler form. In our approach, we develop the complete likelihood based on (ω, S, t).

In obtaining the complete data likelihood, let $[A \mid B]$ generically denote the density function or probability mass function corresponding to A given B. In addition, let $\omega_i = (\omega_{i0}, \ldots, \omega_{im})$, $S_i = (S_{i0}, \ldots, S_{im})$, and $T_i = (T_{i0}, \ldots, T_{im})$. Then the complete data likelihood is given by

$$\begin{aligned}[\omega, S, T] &= \textstyle\prod_{i=1}^{n} [\omega_i, S_i, T_i] \\ &\quad \textstyle\prod_{i=1}^{n} [\omega_i \mid S_i, T_i][S_i, T_i] \\ &\quad \textstyle\prod_{i=1}^{n} [\omega_i \mid S_i, T_i][S_i \mid T_i][T_i]\end{aligned} \tag{1}$$

where the independence of fish is assumed and the expressions in (1) are based on conditional probability. The complete data likelihood for the ith fish is therefore the product of three terms;

the conditional probability mass function of detection history given survival and travel time history, the conditional probability mass function of survival history given travel time history and the travel time density. We now derive expressions for each of the three terms in the product (1). The first term is given by

$$\begin{aligned}[\omega_i \mid S_i, T_i] &= [\omega_i \mid S_i] \\ &= \textstyle\prod_{j=1}^{m} [\omega_{ij} \mid S_{ij}] \\ &= \textstyle\prod_{j=1}^{m} (p_j^{\omega_{ij}} (1-p_j)^{1-\omega_{ij}})^{S_{ij}}\end{aligned} \tag{2}$$

where the key assumption in (2) is that detection at location j does not depend on other locations, and we note that when a fish dies (i.e., $S_{ij} = 0$), then detection is impossible and there is no contribution to the complete data likelihood. Now

$$\begin{aligned}[S_i \mid T_i] &= [S_{im} \mid S_{i0}, \ldots, S_{i,m-1}, T_i][S_{i,m-1} \mid S_{i0}, \ldots, S_{i,m-2}, T_i] \cdots [S_{i1} \mid S_{i0}, T_i] \\ &= [S_{im} \mid S_{i,m-1}, T_i][S_{i,m-1} \mid S_{i,m-2}, T_i] \cdots [S_{i1} \mid S_{i0}, T_i] \\ &= \textstyle\prod_{j=1}^{m} [S_{ij} \mid S_{i,j-1}, T_{i,j-1}, T_{ij}] \\ &= \textstyle\prod_{j=1}^{m} (\phi_{ij}^{S_{ij}} (1-\phi_{ij})^{1-S_{ij}})^{S_{i,j-1}} \\ &= \textstyle\prod_{j=1}^{m} (q_j^{t_{ij} S_{ij}} (1-q_j^{t_{ij}})^{1-S_{ij}})^{S_{i,j-1}}\end{aligned} \tag{3}$$

where $t_{ij} = T_{ij} - T_{i,j-1}$ and there is no survival contribution to the likelihood when a fish has already died (i.e., $S_{i,j-1} = 0$). Putting (1), (2) and (3) together, we have the complete data likelihood

$$[\omega, S, T] = \prod_{i=1}^{n} [T_i] \prod_{j=1}^{m} (p_j^{\omega_{ij}} (1-p_j)^{1-\omega_{ij}})^{S_{ij}} (q_j^{t_{ij} S_{ij}} (1-q_j^{t_{ij}})^{1-S_{ij}})^{S_{i,j-1}}. \tag{4}$$

The last step in the determination of the complete data likelihood (4) is the specification of $[T_i]$. Note that it is preferable to model $[T_i]$ rather than $[t_i] = [t_{i0}, \ldots, t_{im}]$ since there are at least as many missing t_{ij} as missing T_{ij} and therefore we would lose information by modelling $[t_i]$. As the fish arise from the same population and travel times are nonnegative, it may be reasonable to consider a multivariate lognormal distribution. The convenient covariance structure in the multivariate normal distribution is appealing as one might imagine that a fish that is fast (slow) in travelling between two locations may be fast (slow) in travelling between other locations. Specifically, we assume

$$\big(\log(T_{i1}), \ldots, \log(T_{im})\big)^{\top} \sim \mathrm{Normal}_m(\mu, \Sigma), \tag{5}$$

subject to the constraint $0 < T_{i1} \leq \cdots \leq T_{im}$ where a covariance structure is explicitly allowed between the $\log(T_{ij})$. A simpler (but perhaps less realistic) alternative to (5) is $(T_{i1}, \ldots, T_{im}) \sim \mathrm{Normal}_m(\mu, \Sigma)$ subject to the same constraint $0 < T_{i1} \leq \cdots \leq T_{im}$.

In a Bayesian analysis, prior distributions are required for the unknown parameters. Sometimes, strong prior information may be available (e.g., a working knowledge concerning the detection probabilities of listening lines) and it is useful to incorporate this knowledge as can be done in a Bayesian framework. In this paper, we suggest default prior distributions which tend to be diffuse. Diffuse distributions are appealing in that they allow the data to drive the inference. Referring to (4) and (5), we consider the prior density

$$[p, q, \mu, \Sigma] = [p]\,[q]\,[\mu]\,[\Sigma] \tag{6}$$

where prior independence is assumed. As the p and q are probabilities defined on the simplex, it is customary to assume Beta priors. Specifically, we assume independent detection probabilities p_j, where

$$[p_j] \quad \propto \quad p_j^{a_p-1}(1-p_j)^{b_p-1}$$

and independent daily survival probabilities q_j where

$$[q_j] \quad \propto \quad q_j^{a_q-1}(1-q_j)^{b_q-1}.$$

The a and the b may be prespecified based on one's subjective understanding of the listening devices and the daily survival rates. We impose a diffuse $\mathrm{Normal}_m(0, \sigma_\mu I)$ prior for the mean log travel time distribution $[\mu]$ where σ_μ is set large and the normal distribution is constrained according to $\mu_1 \leq \cdots \leq \mu_m$. We set $\Sigma^{-1} \sim \mathrm{Wishart}((1/m)I, m)$. Having specified the complete data likelihood (see (4) and (5)) and the prior (6), the ingredients for a Bayesian analysis have been determined.

3. COMPUTATIONS

We re-express the complete data likelihood $[\omega, S, t]$ in (4) as $\left[X^{\mathrm{obs}}, X^{\mathrm{mis}} \,|\, p, q, \mu, \Sigma\right]$ to emphasize the dependency on the unknown parameters and to emphasize that (ω, S, t) consists of both observed and missing values. The Bayesian paradigm then gives the following expression for the posterior

$$\begin{aligned} \left[p, q, \mu, \Sigma \,|\, X^{\mathrm{obs}}\right] \quad &\propto \quad \left[X^{\mathrm{obs}} \,|\, p, q, \mu, \Sigma\right]\left[p, q, \mu, \Sigma\right] \\ &= \quad \int \left[X^{\mathrm{obs}}, X^{\mathrm{mis}} \,|\, p, q, \mu, \Sigma\right]\left[p, q, \mu, \Sigma\right] dX^{\mathrm{mis}}. \end{aligned} \tag{7}$$

In theory, the posterior density (7) provides a complete description of the uncertainty in the parameters defined in the mark-recapture experiment. However, the dimensionality and the complexity of (7) is such that it is impossible to gain any meaningful insight. Alternatively, we consider the following expression

$$\begin{aligned} \left[p, q, \mu, \Sigma, X^{\mathrm{mis}} \,|\, X^{\mathrm{obs}}\right] \quad &\propto \quad \left[p, q, \mu, \Sigma, X^{\mathrm{obs}}, X^{\mathrm{mis}}\right] \\ &\propto \quad \left[X^{\mathrm{obs}}, X^{\mathrm{mis}} \,|\, p, q, \mu, \Sigma\right]\left[p, q, \mu, \Sigma\right] \end{aligned} \tag{8}$$

where the last expression in (8) is the product of the complete data likelihood and the prior density which are familiar and simple forms.

Therefore, if we are able to sample variates $(p, q, \mu, \Sigma, X^{\mathrm{mis}})$ from (8), then we can use the sampled components (p, q, μ, Σ) as realizations from the posterior distribution. However, sampling directly from (8) is a difficult/impossible task, and instead, a Markov chain is constructed which has the posterior as its stationary distribution. Fortunately, this is easily implemented using WinBUGS software (Spiegelhalter, Thomas & Best 2003). In WinBUGS, the user needs only to specify the form of the complete data likelihood, the prior and the observed data. WinBUGS then produces an appropriate Markov chain. The user may then proceed with the Markov chain output as seen fit. For example, output may be averaged to provide estimates of posterior expectations and marginal posterior densities may be approximated using density estimation techniques. We may even obtain posterior expectations of latent variables. Note that whereas classical methods (e.g., estimation and testing) often rely on asymptotic distributions of statistics, a sample from the posterior is a sample from the distribution of interest. An overview of the use of MCMC methods is provided in the edited text by Gilks, Richardson & Spiegelhalter (1996). Carlin &

Louis (2000) and Gelman, Carlin, Stern & Rubin (2003) provide further information on MCMC and give modern accounts of the Bayesian approach to statistics. A detailed description of the WinBUGS code for the POST project is given in Muthukumarana (2007). There are different versions of WinBUGS and only the full version 1.4.1 was able to handle the complexity of our model. When a model is not too complex, WinBUGS makes use of the Gibbs sampling algorithm to generate a Markov chain. With our model, the full conditional distributions required for Gibbs sampling have nonstandard forms. In this case, WinBUGS makes use of the Metropolis–Hastings algorithm, where Markov chain output contains streams of duplicate values due to the acceptance/rejection step in the algorithm. We note that conditional on S_{ij} and $S_{i,j-1}$, there are two Bernoulli terms in the complete data likelihood. More specifically, we can express (4) as

$$[\omega, S, T] \quad = \quad \prod_{i=1}^{n}[T_i]\prod_{j=1}^{m}(S_{ij}p_j)^{\omega_{ij}}(1-S_{ij}p_j)^{1-\omega_{ij}}(S_{i,j-1}q_j^{t_{ij}})^{S_{ij}}(1-S_{i,j-1}q_j^{t_{ij}})^{1-S_{ij}}.$$

The recognition of this fact enables a simpler expression for the complete data likelihood and dramatically reduces the computational time. We also note that our model contains constrained distributions for the $\log(T_{ij})$ and the μ_j variables. As constraints are not a standard feature of WinBUGS, we overcame this difficulty through the use of indicator variables. Using this approach, it is not necessary to determine the norming constant for the constrained distribution.

WinBUGS coding can initially be difficult; we hope that our code provides a beginning template for future Bayesian analyses in mark-recapture. We also note that some preprocessing was required for the data considered in Section 5.2. It was necessary to extract (ω, S) from T^{obs} prior to running WinBUGS. An R code developed for the preprocessing stage is provided in Muthukumarana (2007).

3.1. Nonidentifiability.

It is well known that final survival and capture rates are confounded in the classical Cormack–Jolly–Seber model as a result of nonidentifiability. To understand the problem at a deeper level, we recall that the observed likelihood in that model is not the same as the complete data likelihood (4). The observed Cormack–Jolly–Seber likelihood differs in that it does not contain the cumulative travel-time distributions $[T_i]$ appearing in (4). Secondly, the individual survival probabilities $\phi_{ij} = q_j^{t_{ij}}$ in (4) are replaced with the simpler probabilities $\phi_{ij} = \phi_j$. Finally, the observed likelihood (which is difficult to write down in the general case) may be derived from the complete data likelihood (4) by summing over cases that are not directly observed. For example, the term in the complete data likelihood corresponding to the unobservable case $(S_{im} = 1, \omega_{im} = 0)$ is added to the term in the complete data likelihood corresponding to the observable case $(S_{im} = 0, \omega_{im} = 0)$. As a result, the terms p_m and ϕ_m only appear in the observed data likelihood as the product $p_m\phi_m$. The implication is that the data only allow us to learn about the product $p_m\phi_m$ and not about the individual parameters p_m and ϕ_m. In this case, we say that p_m and ϕ_m are nonidentifiable.

In general, the typical consequences of nonidentifiability include ridges in the likelihood surface and multimodality which are problematic for estimation. In a Bayesian context, these sorts of problems may not be as problematic as in the classical context. In theory, all that one needs to do in a Bayesian analysis is integrate to obtain the required posterior summaries. However, from a practical perspective, nonidentifiablity still may cause problems in a Bayesian analysis. For example, MCMC algorithms may have difficulty traversing parameter spaces with elongated likelihoods.

Swartz, Haitovsky, Vexler & Yang (2004) have demonstrated that the use of informative (i.e., nondiffuse) priors may be effective in mitigating the effects of nonidentifiability in Bayesian models. In our application, we recommend the use of informative priors particularly for the detection probabilities p_j. There is often good prior knowledge concerning the capabilities of the listening lines and the acoustic transmitters.

In view of the above discussion, there is an appealing by-product of the modelling assumption $\phi_{ij} = q_j^{t_{ij}}$ with respect to nonidentifiablity. In the Cormack–Jolly–Seber model, if one replaces $\phi_{ij} = \phi_j$ with $\phi_{ij} = q_j^{t_{ij}}$, the nonidentifiablity disappears because the product $p_m q_m^{t_{im}}$ appearing in the likelihood differs over the fish $i = 1, \ldots, n$. This result is somewhat paradoxical as the model with $\phi_{ij} = q_j^{t_{ij}}$ is more complex than the traditional Cormack–Jolly–Seber model with $\phi_{ij} = \phi_j$, yet we gain better insight regarding the individual parameters with the more complex model. In the model proposed in this paper (which is a generalization of the Cormack–Jolly–Seber model), it follows that there is no problem with nonidentifiability.

4. MODEL SELECTION AND MODEL ADEQUACY

The topics of model selection and model adequacy are of fundamental importance in applied statistics and these topics are becoming increasingly important with the consideration of more complex models. However, in Bayesian statistics, a myriad of approaches have been proposed for both model selection and assessing model adequacy, and it is fair to say that there is no consensus on the "correct" approach to either of these problems. In this section, we provide some general remarks on model selection and model adequacy, and we provide some concrete suggestions that are relevant to the problem at hand.

4.1. Model selection.

A principled Bayesian approach for comparing a finite number of competing models involves the calculation of the posterior probabilities of the models. When equal prior probabilities are assigned to each of the models, then the posterior comparison of two models (i and j) reduces to a study of the Bayes factor

$$B_{ij} \quad = \quad \frac{\int f_i(x \mid \theta_i)\pi_i(\theta_i)\, d\theta_i}{\int f_j(x \mid \theta_j)\pi_j(\theta_j)\, d\theta_j}, \tag{9}$$

where $f_i(x \mid \theta_i)$ is the likelihood of model i with parameter θ_i, and $\pi_i(\theta_i)$ is the prior density corresponding to parameter θ_i. When the Bayes factor B_{ij} is greater (smaller) than 1, this provides evidence for (against) model i relative to model j. A major practical difficulty with the use of Bayes factors is the calculation of the Bayes factor B_{ij}. The expression (9) can rarely be evaluated analytically and it is typical to attempt to approximate B_{ij}. For example, a method that is often unstable involves the approximation of the numerator by averaging $f_i(x \mid \theta_i^{(k)})$ where $\theta_i^{(k)}$ is the kth iteration of θ_i from the prior distribution. Naturally this approach presupposes a proper prior. Methods of approximation based on output from MCMC simulation have also been proposed. However, even these methods are fraught with difficulties. For example, we have experienced unstable estimation and overflow with the complex models proposed in this paper. Another serious problem with the use of Bayes factors is one of calibration when improper priors are used. An overview of Bayes factors is given by Kass & Raftery (1995).

Due to the practical difficulties with the use of Bayes factors in complex models, a number of alternative diagnostics have been proposed that are often viewed as approximations to Bayes factors. For example, the diagnostics AIC (Akaike 1973), BIC (Schwarz 1978) and DIC (Spiegelhalter, Best, Carlin & van der Linde 2002) have all received prominent attention in the literature.

Amongst the numerous model selection diagnostics, it appears that the DIC is the most widely used in WinBUGS applications. In fact, WinBUGS provides DIC values as an option in its Inference menu. Unfortunately, DIC is unavailable in WinBUGS with our model due to the complexity of the model where some of the stochastic nodes (e.g., survival status S_{ij}) are discrete.

One of the main thrusts of our paper is that complex mark-recapture models can be analyzed fairly easily using WinBUGS software. Therefore we believe that it would be against the spirit

of the paper to require the investigator to fit models using WinBUGS and then carry out model selection using some sophisticated procedure outside of WinBUGS. For example, the method of Laud & Ibrahim (1995) based on posterior predictive discrepancies requires variate generation from marginal distributions outside of WinBUGS. For this reason, we want a model selection approach that can be easily implemented within WinBUGS.

For model selection in our mark-recapture models, we suggest the use of the BIC diagnostic. The BIC diagnostic is a little more sophisticated than AIC as it takes sample size into account. Denote the complete data likelihood (4) by $L(\theta)$ where $\theta = (\omega, S, T)$ and let $\theta^{(i)}$ represent the ith realization of θ from MCMC simulation. Then the BIC is approximated by

$$\text{BIC} \quad = \quad p \log(n^*) - \frac{2}{N} \sum_{i=1}^{N} \log L(\theta^{(i)}), \tag{10}$$

where p is the number of parameters in the model (including missing values), n^* is the number of observed data values, and N is the number of MCMC simulations. A model with a smaller value of (10) is a preferred model. Note that $p \log(n^*)$ may be viewed as a penalty term that takes the dimensionality of the model into account. Note also that the Bayesian formulation of BIC is different than the classical version which evaluates $\log L$ at the maximum likelihood estimate rather than averaging $\log L$ over the posterior. Again, an important feature is that (10) may be evaluated directly in WinBUGS by coding BIC. Some discussion of the use of BIC and other model selection diagnostics in Bayesian applications is given by Aitkin (1991) and DeSantis & Spezzaferri (1997).

4.2. Model adequacy.

As problematic as model selection may be with complex Bayesian models, the assessment of model adequacy in complex Bayesian models is even more problematic. A possible explanation for this is that a posteriori testing of model adequacy is not a Bayesian construct and may be seen as violating the Bayesian paradigm. From the point of view of a subjective Bayesian purist, any uncertainty concerning a model ought to be expressed via prior opinion. For example, if an experimenter is unsure whether the sampling distribution of the data is normal or Student, then the uncertainty might be expressed via a mixture. In theory, if we are able to express uncertainty in a model (and this includes both the sampling model and the parameters given the sampling model), then there is no need to assess model adequacy, as all possible models have been considered and our inferences are subjective. However, from a practical point of view, it is typically difficult or impossible to determine the space of possible sampling models and parameters, and to assign prior opinion to the space.

Therefore, what does the practical Bayesian do in the context of model assessment? An honest answer may be that the assessment of complex Bayesian models is not a routine activity. When Bayesian model assessment is considered, it appears that the prominent modern approaches are based on the posterior predictive distribution (Gelman, Meng & Stern 1996). These approaches rely on sampling future variates y from the posterior predictive density

$$f(y \,|\, x) \quad = \quad \int f(y \,|\, \theta) \pi(\theta \,|\, x) \, d\theta, \tag{11}$$

where x is the observed data, $f(y \,|\, \theta)$ is the sampling density for y and $\pi(\theta \,|\, x)$ is the posterior density. In MCMC simulation, approximate sampling from (11) proceeds by sampling y_i from $f(y \,|\, \theta^{(i)})$, where $\theta^{(i)}$ is the ith realization of θ from the Markov chain. Model assessment then involves a comparison of the future values y_i versus the observed x. One such comparison involves the calculation of posterior predictive P-values (Meng 1994).

A major difficulty with posterior predictive methods concerns a double use of the data. Specifically, the observed data x is used both to fit the model giving rise to the posterior density $\pi(\theta \,|\, x)$ and then is used in the comparison of y_i versus x. For this reason, some authors prefer

a cross-validatory approach (Gelfand, Dey & Chang 1992) where the data $x = (x_1, x_2)$ are split so that x_1 is used for fitting and x_2 is used for validation.

We take the view that in assessing a Bayesian model, the entire model ought to be under consideration, and the entire model consists of both the sampling model of the data and the prior. We also want a methodology that does not suffer from double use of the data. Finally, we want an approach that is not too difficult to implement; as we have argued, our complex mark-recapture models are easily fit using WinBUGS. For the models proposed here, we recommend an approach that is similar to the posterior predictive methods but instead samples "model variates" y from the prior predictive density

$$f(y) \quad = \quad \int f(y \mid \theta)\pi(\theta)\, d\theta, \tag{12}$$

where $\pi(\theta)$ is the prior density. This approach was advocated by Box (1980) before simulation methods were common. Note that generating from (12) presupposes proper priors which are required in WinBUGS.

It is not difficult to write R code to simulate $y_1, \ldots, y_N$ from the prior predictive density in (12). It is then a matter of deciding how to compare the y_i against the observed data x. We advocate simple comparisons that are of direct interest to the application. For example, one might compare the mean observed cumulative travel time $\overline{T}_2 = \sum_{i=1}^{n} T_{i2}/n$ at the second listening line to the histogram formed by the N variates $\overline{T}_2$ obtained from the prior predictive simulation. Naturally, as the priors become more diffuse, it becomes less likely to find evidence of model inadequacy. We investigate the assessment of model adequacy on the Columbia river data in Section 5.2.

5. EXAMPLES

5.1. Simulated data.

Various simulation studies were carried out. We report on one such simulation. A dataset corresponding to $n = 500$ fish with $m = 5$ listening lines was simulated using the R code. Detection probabilities at each listening line were set to $p_j = p = 0.8$, while daily survival probabilities between listening lines were set to $q_j = q = 0.99, j = 1, \ldots, m$. The logarithms of the cumulative travel times between the listening lines were generated from the constrained multivariate normal distribution (5) with $\mu = (1, 2, 3, 4, 5)'$ and $\Sigma = (\sigma_{ij})$ where $\sigma_{ii} = 1.0$ and $\sigma_{ij} = 0.8$ for $i \neq j$. We then generated a survival history S based on $S_{ij} \sim \text{Bernoulli}\,(q_j^{t_{ij}})$ and a detection history ω based on $\omega_{ij} \mid S_{ij} = 1 \sim \text{Bernoulli}\,(p_j)$. Having generated the data as described, we partitioned the data into the observed and missing components of (ω, S, T); this is necessary for the construction of the complete data likelihood (4). In the simulated data, there are $5(500) = 2500$ cumulative travel times T_{ij} of which 858 are missing. Therefore we have considered a challenging test case with a large proportion of missing values. At the fifth listening line, a typical survival probability is $\phi_{i5} = q_5^{T_{i5} - T_{i4}} = (0.99)^{\exp(5) - \exp(4)} \approx 0.39$.

For the analysis of the simulated dataset, we first consider the "full model" which contains all of the parameters described in the paper. Uniform prior distributions for the parameters p_j and q_j were assigned according to $p_j \sim \text{Beta}\,(1.0, 1.0)$ and $q_j \sim \text{Beta}\,(1.0, 1.0), j = 1, \ldots, 5$ with independence across the distributions. The specified priors provide a good test of the robustness of the methods with respect to the priors as the corresponding prior means $\text{E}\,(p) = \text{E}\,(q) = 1/2$ are not close to the preset parameter values. The cumulative travel times are modelled as in (5) and the rest of the prior settings are given as described in Section 2. The model was fit using WinBUGS software where 500 iterations were used for the burn-in period. The posterior estimates in Table 1 were based on 8000 iterations which required approximately 15 minutes of computation on a personal computer. We observe that the posterior means of the primary parameters p_j and q_j are close to the preset values. The posterior means of the secondary parameters μ and Σ also appear in agreement with the preset values.

TABLE 1: Estimates of posterior means and posterior standard deviations in Example 5.1.

Parameter	Mean	SD	Parameter	Mean	SD
p_1	0.77	0.02	Σ_{11}	0.91	0.06
p_2	0.79	0.02	Σ_{12}	0.79	0.05
p_3	0.81	0.02	Σ_{13}	0.77	0.05
p_4	0.82	0.02	Σ_{14}	0.77	0.05
p_5	0.76	0.02	Σ_{15}	0.77	0.06
q_1	0.99	0.00	Σ_{22}	1.01	0.06
q_2	0.99	0.00	Σ_{23}	0.84	0.06
q_3	0.99	0.00	Σ_{24}	0.80	0.06
q_4	0.99	0.00	Σ_{25}	0.85	0.06
q_5	0.99	0.00	Σ_{33}	0.94	0.06
μ_1	1.02	0.04	Σ_{34}	0.82	0.06
μ_2	2.06	0.04	Σ_{35}	0.81	0.06
μ_3	3.06	0.04	Σ_{44}	0.95	0.06
μ_4	4.02	0.04	Σ_{45}	0.83	0.06
μ_5	5.02	0.04	Σ_{55}	0.99	0.07

We now wish to investigate aspects of the model selection diagnostic BIC in (10). With simulated data, we can investigate the diagnostic since we know the true model from which the data were generated. For the full model considered above with the uniform priors, we obtained $\mathrm{BIC} = 9640.3$. Knowing the way that the data were simulated, we also fit the "true model" with $\phi_{ij} = q^{t_{ij}}$ and $p_{ij} = p$ where independent uniform priors were assigned to p and q. For the true model, we obtained $\mathrm{BIC} = 9557.5$. Therefore the BIC diagnostic preferred the true model over the full model.

We next fit an even simpler model with $\phi_{ij} = \phi_j$ and p_j which is analogous to the Cowen and Schwarz (2005) model as it does not consider survival as a function of travel time. In this case, $\mathrm{BIC} = 9725.2$. Therefore the BIC diagnostic rightly suggests that the travel time assumption is important. In fact, the full model is preferred to the model analogous to Cowen and Schwarz (2005).

Finally, we consider the sensitivity of the analysis with respect to the travel time assumption (5). The analysis here is the same as in the full model analysis except that we assume the simpler travel time distribution $(T_{i1}, \ldots, T_{im}) \sim \mathrm{Normal}_m(\mu, \Sigma)$ subject to the constraint $0 < T_{i1} \leq \cdots \leq T_{im}$. Under the simpler assumption, the posterior means of the p_j and q_j are comparable to the posterior means in the analysis of the full model and $\mathrm{BIC} = 9763.3$. This suggests that even though the overall fit of the simpler model is not good (in terms of the BIC), the precise shape of the distribution of the cumulative travel times is not a critical assumption in the estimation of the primary parameters of interest.

5.2. Columbia river data.

The model was then fit to data obtained from the Columbia River system. From April 25, 2001 to May 30, 2001, $n = 324$ radio-tagged chinook salmon were released from the Rock Island Dam. Data were recorded at listening lines established at the $m = 3$ dams downstream at Wanapum, Priest Rapids, and Hanford Reach. For example, corresponding to fish $j = 4$, we have data values $\omega_4 = (1, 1, 0, 0)$ and $(T_{40}, T_{41}, T_{42}, T_{43}) = (0, 6.1, \mathrm{NA}, \mathrm{NA})$. This implies that the fourth fish was released at the Rock Island Dam and then detected at Wanapum but was undetected at both Priest Rapids and Hanford Reach. The fish took 6.1 days to reach Wanapum from the Rock Island Dam. The data gives rise to the survival history $S_4 = (1, 1, \mathrm{NA}, \mathrm{NA})$ since we know

that the fish survived up to Wanupum, but it is unknown whether the fish survived up to Priest Rapids or Hanford Reach. The interdam distances are approximately 37.6 miles, 18.7 miles, and 15.0 miles, respectively. In the Columbia river data, there are $3(324) = 972$ cumulative travel times T_{ij} of which 294 are missing. Cowen & Schwarz (2005) also studied this dataset in the context of radio failure. Here, we ignore radio failure, and therefore survival is a function of both actual survival and radio failure. In this example, we fit a model which allows for the possibility of varying detection probabilities p_j and varying daily survival probabilities q_j at each of the dams. Uniform priors were assigned to the detection and daily survival probabilities. The remaining priors are given as in Section 2. We remark that it is possible to enhance the model by stratifying the salmon according to their release date, although we have not done so.

TABLE 2: Estimates of posterior means and posterior standard deviations in Example 5.2.

Parameter	Mean	SD
p_1	0.96	0.03
p_2	0.97	0.03
p_3	0.98	0.02
q_1	0.99	0.00
q_2	0.91	0.01
q_3	0.77	0.02
μ_1	2.23	0.03
μ_2	2.60	0.03
μ_3	2.78	0.03
Σ_{11}	0.19	0.02
Σ_{12}	0.16	0.02
Σ_{13}	0.15	0.02
Σ_{22}	0.17	0.02
Σ_{23}	0.16	0.02
Σ_{33}	0.18	0.02

Estimates of the posterior means of the parameters are given in Table 2. These are based on a MCMC simulation using WinBUGS with a burn-in period of 500 iterations followed by 4000 iterations. We observe that the detection probabilities p_j are high and are similar across the dams. We note that the daily survival probabilities q_1, q_2, and q_3 decrease and this appears to make biological sense. With respect to the estimated travel-time parameter μ, we refer to Figure 1 which provides density plots of the observed travel times between dams. The average travel times between the three dams are 9.9, 3.6, and 1.9 days, respectively. These values are roughly in agreement with MCMC estimates of the mean interval travel times (e.g. $\mathrm{E}(t_{i1}) \approx \exp(\mu_1 + \Sigma_{11}/2)$) which are 10.3, 4.6, and 2.7 days, respectively. Figure 2 provides an estimate of the posterior density of μ_1 using a kernel smoother from WinBUGS. The plot suggests a nearly symmetric unimodal distribution as might be expected. We observe strong positive correlations in the Σ matrix; this is expected as we are modelling cumulative travel times.

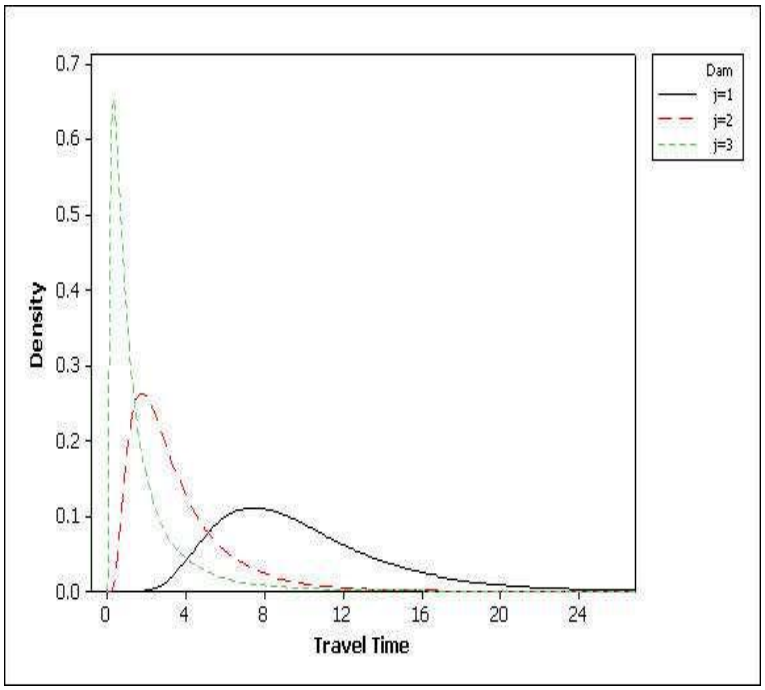

FIGURE 1: Density plots for observed travel times t_{ij} (days) in Example 5.2.

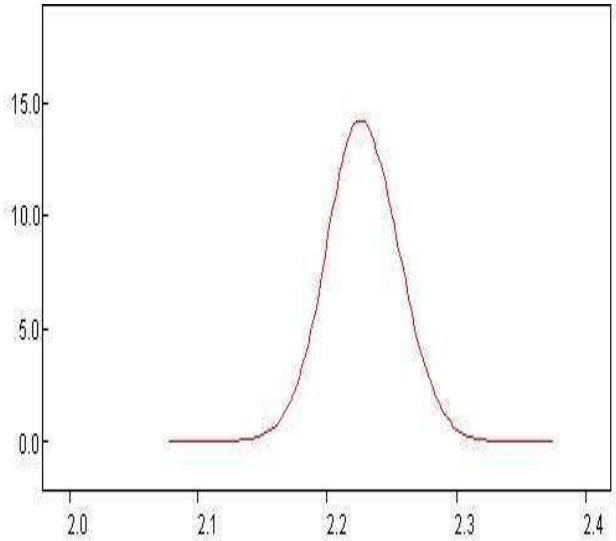

FIGURE 2: Estimate of the posterior density of μ_1 in Example 5.2.

It is instructive to look at some plots related to the MCMC simulation. A trace plot for μ_1 is given in Figure 3. The trace plot appears to stabilize immediately and hence provides no indication of lack of convergence in the Markov chain. In Figure 4, an autocorrelation plot for μ_1 is provided. The autocorrelations appear to dampen quickly. This provides added evidence of the convergence of the Markov chain and also suggests that it may be appropriate to average Markov chain output as though the variates were independent. Similar plots were obtained for all of the parameters in the model. In addition to the diagnostics described, multiple chains were

obtained to provide further assurance of the reliability of the methods. For example, the Brooks–Gelman–Rubin statistic (Brooks & Gelman 1997) gave no indication of a lack of convergence.

As discussed in Section 3.1, nonidentifiability poses neither a theoretical nor a practical obstacle for the model proposed in this paper. To investigate the degree of confounding between the final capture rate p_3 and the final daily survival probability q_3, we calculate the posterior correlation between p_3 and q_3 using the output from the MCMC simulation. The posterior correlation is found to be 0.02 which indicates a lack of confounding between the final capture rate and the final daily survival probability.

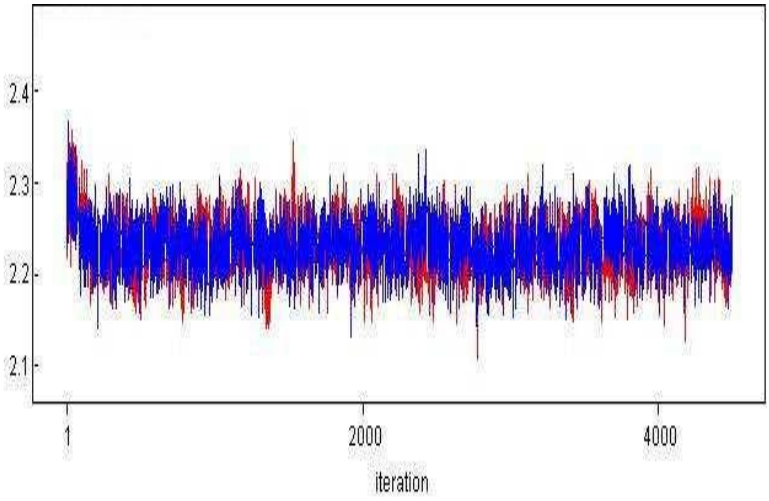

FIGURE 3: Trace plot for μ_1 based on MCMC simulation in Example 5.2.

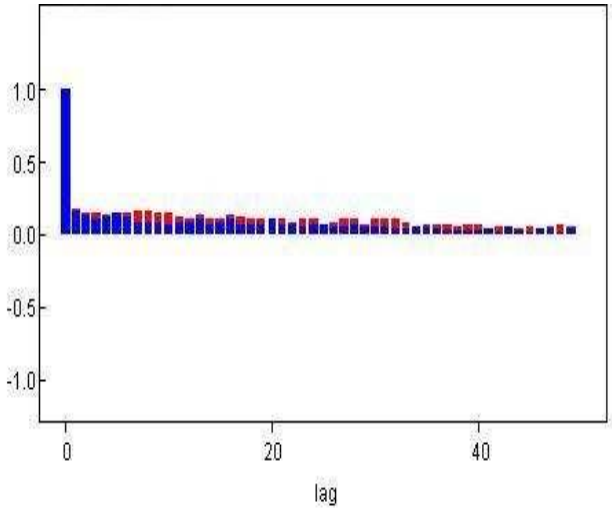

FIGURE 4: Autocorrelation plot for μ_1 based on MCMC simulation in Example 5.2.

In assessing model adequacy, we have mentioned that a model with diffuse priors will almost always be viewed as adequate. The reason for this is that a model with diffuse parameters gives rise to diffuse data, and observed data is unlikely to be seen as inconsistent when compared to diffuse data. To provide a more stringent test, we consider a modification of our model where subjective priors are introduced. Since we know a priori that the listening lines and acoustic devices are of high quality, we assign prior probabilities of detection according to independent

$[p_j] \sim \text{Beta}(19, 1)$. We also know that daily survival probabilities are high and we therefore set independent $[q_j] \sim \text{Beta}(18, 2)$. The prior distribution for the log cumulative travel times is a constrained multivariate normal distribution with mean vector and variance covariance matrix given by the posterior means from the initial analysis. This distribution is roughly consistent with the data and is far less diffuse. With this model, the posterior means of the primary parameters are very close to estimates provided in Table 2. This suggests that the data are informative and dominate the inferences. To assess model fit, we generated 20 datasets from the prior predictive density (12) according to the model described above, and compared the generated data with the observed data. Although there are many features of the data that could be checked, we focus on a study of detection history ω and cumulative travel time T. For ω, Figure 5 provides a histogram of the proportion of fish detected in the simulated datasets. For each simulation, the proportion detected is given by $\sum_{i=1}^{324}\sum_{j=1}^{3}\omega_{ij}/(924)$. The proportion detected for the observed data is 0.70 which appears consistent with the model. For T, there appears to be no limit on the number of features that one may check. For illustration we consider the travel time to the first dam. Figure 6 provides a boxplot for each generated dataset using the 324 total travel times T_{i1}. The boxplot for the observed data is also included and appears to be consistent with the generated data.

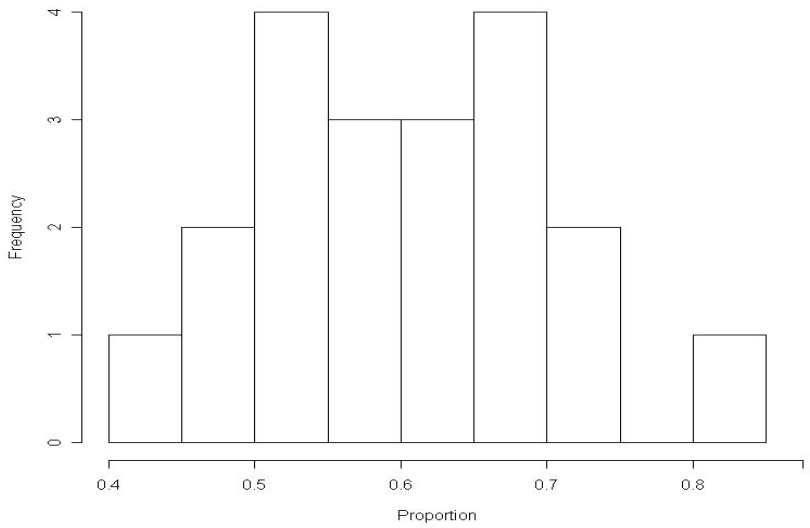

FIGURE 5: Histogram of the proportion of fish detected for the observed data and the 20 generated datasets in Example 5.2.

To check whether the travel-time assumption is relevant to the Columbia river data, we fit a second model that is analogous to the model considered by Cowen & Schwarz (2005). We keep everything the same as in the initial model with the diffuse priors, but let $\phi_{ij} = \phi_j$ rather than $\phi_{ij} = q_j^{t_{ij}}$. For the Cowen and Schwarz (2005) model, we obtain $\text{BIC} = 3988.8$, which is much worse than the initial model with $\text{BIC} = 2521.1$. The extremely bad fit of the Cowen & Schwarz (2005) model may have been anticipated, as Figure 1 suggests large travel time differences between the three dams.

6. DISCUSSION

The Bayesian framework provides a straightforward approach to dealing with the complex observed likelihood which requires integration over the unobservable travel times. Furthermore, the Bayesian approach provides a convenient way of estimating the correlation in travel times

between sampling locations. This addresses an important biological question as to whether some fish are intrinsically faster, or whether travel times are independent random events.

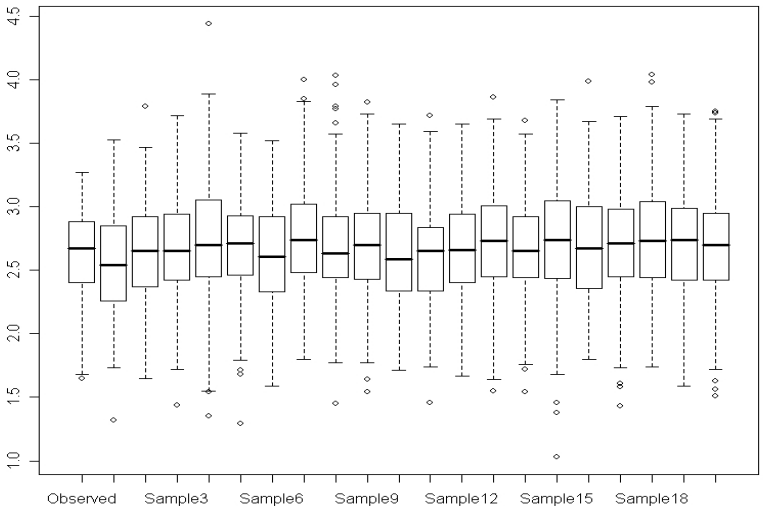

FIGURE 6: Boxplots of the travel time T_{i1} to the first dam for the observed data and the 20 generated datasets in Example 5.2.

An alternative approach might be based on the EM algorithm as used by van Deusen (2002). However, the computation of the expected log-likelihood is not straightforward and numerical methods would likely be needed. In our Bayesian approach, Markov chain Monte Carlo (MCMC) methods avoid the necessity of numerical integration.

Finally, our model is easily extended to allow for individual time-independent covariates such as initial body mass m_{ij} by modelling $q_{ij} = f(m_{ij})$ for some function f. Work is underway to extend our formulation to allow for time-dependent individual covariates and to more fully investigate the choice of f. The revised model must account for missing values both in the travel times and in the individual covariates.

It is not immediately clear how our model might be extended to a two-dimensional spatial setting. For example, listening lines may be set up in the woods to track the movements of animals. In this case, in addition to death and radio tag failure, lack of detection may be due to nonstandard travel paths. We consider this to be an open research problem.

ACKNOWLEDGEMENTS

The authors thank the Editor and two anonymous referees for helpful comments that led to an improvement in the paper. Muthukumarana was partially supported by a MITACS fellowship. Schwarz and Swartz were partially supported by Discovery grants through the Natural Sciences and Engineering Research Council of Canada (NSERC). The WinBUGS code and the R code used in this paper are available from the first author upon request.

REFERENCES

M. Aitken (1991). Posterior Bayes factors. *Journal of the Royal Statistical Society Series B*, 53, 111–114.

H. Akaike (1973). Information theory and an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory* (B. Petrov & F. Csáki, eds.), Akadémiai Kiadó, Budapest, pp. 267–281.

S. J. Bonner & C. J. Schwarz (2006). An extension of the Cormack–Jolly–Seber model for continuous covariates with application to *Microtus pennsylvanicus*. *Biometrics*, 62, 142–149.

G. E. Box (1980). Sampling and Bayes' inference in scientific modelling and robustness (with discussion). *Journal of the Royal Statistical Society Series A*, 143, 383–430.

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

S. P. Brooks & A. Gelman (1997). Alternative methods for monitoring convergence of iterative simulations. *Computational and Graphical Statistics*, 7, 434–455.

K. P. Burnham, D. R. Anderson, G. C. White, C. Brownie & K. H. Pollock (1987). *Design and Analysis Methods for Fish Survival Experiments Based on Release-Recapture*. Monograph 5, American Fisheries Society, Bethesda, Maryland.

B. P. Carlin & T. A. Louis (2000). *Bayes and Empirical Bayes Methods for Data Analysis*, Second edition. Chapman & Hall, London.

R. M. Cormack (1964). Estimates of survival from the sighting of marked animals. *Biometrika*, 51, 429–438.

L. Cowen & C. J. Schwarz (2005). Capture-recapture studies using radio telemetry with premature radio-tag failure. *Biometrics*, 61, 657–664.

F. DeSantis & F. Spezzaferri (1997). Alternative Bayes factors for model selection. *The Canadian Journal of Statistics*, 25, 503–515.

J. A. Dupuis (1995). Bayesian estimation of movement and survival probabilities from capture-recapture data. *Biometrika*, 82, 761–772.

A. E. Gelfand, D. K. Dey & H. Chang (1992). Model determination using predictive distributions with implementation via sampling-based methods (with discussion). In *Baysian Statistics 4* (J. M. Bernardo, J. O. Berger, A. P. Dawid & A. F. M. Smith, eds.), Oxford University Press, Oxford, pp. 147–167.

A. Gelman, J. B. Carlin, H. S. Stern & D. B. Rubin (2003). *Bayesian Data Analysis*, Second edition. Chapman & Hall, London.

A. Gelman, X. L. Meng & H. S. Stern (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica*, 6, 733–807.

W. R. Gilks, S. Richardson & D. J. Spiegelhalter, eds. (1996). *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London.

O. Gimenez, C. Crainiceanu, C. Barbraud, S. Jenouvrier & B. J. T. Morgan (2006). Semiparametric regression in capture-recapture modelling. *Biometrics*, 62, 691–698.

G. M. Jolly (1965). Explicit estimates from capture-recapture data with both death and immigration-stochastic model. *Biometrika*, 52, 225–247.

R. E. Kass & A. E. Raftery (1995). Bayes factors. *Journal of the American Statistical Association*, 90, 773–795.

P. W. Laud & J. G. Ibrahim (1995). Predictive model selection. *Journal of the Royal Statistical Society Series B*, 57, 247–262.

J. Lebreton, K. P. Burnham, J. Clobert & D. R. Anderson (1992). Modelling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecological Monographs*, 62, 67–118.

X.-L. Meng (1994). Posterior predictive p-values. *The Annals of Statistics*, 22, 1142–1160.

S. Muthukumarana (2007). *Bayesian Modelling and Software Development for the POST Project*. MSc project, Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, BC, Canada.

K. H. Pollock, C. M. Bunck, S. R. Winterstein & C. Chen (1995). A capture-recapture survival analysis model for radio-tagged animals. *Journal of Applied Statistics*, 22, 661–672.

G. Schwarz (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6, 461–464.

G. A. F. Seber (1965). A note on the multiple recapture census. *Biometrika*, 52, 249–259.

D. J. Spiegelhalter, N. G. Best, B. P. Carlin & A. van der Linde (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society Series B*, 64, 583–639.

D. J. Spiegelhalter, A. Thomas & N. G. Best (2003). *WinBUGS (Version 1.4) User Manual*. MRC Biostatistics Unit, Cambridge.

T. B. Swartz, Y. Haitovsky, A. Vexler & T. Yang (2004). Bayesian identifiability and misclassification in multinomial data. *The Canadian Journal of Statistics*, 32, 285–302.

P. C. van Deusen (2002). An EM algorithm for capture-recapture estimation. *Environmental and Ecological Statistics*, 9, 151–165.

Received 15 October 2006
Accepted 10 July 2007

Saman MUTHUKUMARANA: muthukumaranasm@yahoo.com
Carl J. SCHWARZ: cschwarz@stat.sfu.ca
Tim B. SWARTZ: tim@stat.sfu.ca
Department of Statistics and Actuarial Science
Simon Fraser University, Burnaby
British Columbia, Canada V5A 1S6

Discussion: Towards a Bayesian analysis template?

Olivier GIMENEZ

1. INTRODUCTION

I would like to congratulate Drs Muthukumarana, Schwarz and Swartz (henceforth MSS) for extending classical mark-recapture models to estimate survival in a complex situation arising from fish monitoring. Because of their model likelihood complexity, MSS used Markov chain Monte Carlo (MCMC) simulations to implement a Bayesian analysis of their data. The Bayesian framework in association with MCMC algorithms is becoming increasingly popular for fitting complex models such as models with latent structures. Two of the main reasons for this are that (i) MCMC methods are well suited to circumvent the issue of high-dimensional integrals involved in these likelihoods and, (ii) fast and powerful computers along with flexible and reliable programs are now available allowing the relatively time-realistic and easy implementation of various MCMC routines. This being said, the possibility to fit complex models comes with methodological issues that should not be overcome. In that sense, MSS have provided an impressive work that deserves to be emphasized. In this discussion, I comment on several technical points and give general considerations that were inspired by the MSS paper.

2. TECHNICAL POINTS

2.1. Model selection.

To determine whether a model specifying survival as a function of travel times was better supported by the data than without, MSS relied on BIC as it was easy to implement within WinBUGS. As acknowledged by MSS, this was one choice among many alternatives, and this is precisely the Achilles' heel of Bayesian analyses. Indeed, many procedures exist and none of them seems to be as consensual as AIC is in the statistical ecology literature (Burnham & Anderson 2002). Basically, there are two groups of methods. One produces a value for each model to be compared among a set of models (e.g., mean square predictive error: Gelfand & Ghosh 1998; DIC: Spiegelhalter, Best, Carlin & van der Linde 2002; BIC: for example, Link & Barker 2007), and the other performs automatic exploration of the model space (e.g., Gibbs variable selection: George & McCulloch 1993; reversible jump MCMC: Green 1995). Most

often, we adopt one option or another because it is convenient to calculate, or because we are familiar with it. MSS went for a method belonging to the first family of methods, in line with recent recommendations (Link & Barker 2007). Interestingly, reversible jump MCMC is now implemented in WinBUGS, and MSS could have used it to compare models. This raises the issue of which method to use. Unfortunately, I'm not aware of any comparison or any review that might give clear guidelines. Having made clear what each method does, the performances of several candidate methods could be assessed by calculating frequencies of ranking the true model as best in Monte Carlo simulations.

2.2. Identifiability and convergence issues.

We are often provided with, even in biological papers, technical details regarding convergence of the MCMC algorithms, while it wouldn't come to one's mind to mention anything about convergence in a classical analysis. This is probably because the procedures in the former case have not yet been implemented in an automatic way, and further work is needed in that direction. MSS have paid careful attention to the identifiability issue, and demonstrated that their model was not parameter redundant. Note that formal methods were developed to assess parameter redundancy of probabilistic models that could be used here too (review in Gimenez et al. 2005). Nevertheless, a cause of poor MCMC convergence is weak identifiability (rather than nonidentifiability) because it leads to large autocorrelations. Calculating the overlap between prior and posterior parameter distributions can help in diagnosing weak identifiability (Garrett & Zeger 2000; Gimenez, Morgan & Brooks 2008). Practical recommendations on checking MCMC convergence are given by experienced statisticians in Kass, Carlin, Gelman & Neal (1998) and comparisons of several available methods can be found in El Adlouni, Favre & Bobée (2006).

2.3. Goodness-of-fit testing.

Goodness–of–fit testing has received little attention in the Bayesian literature and only a few methods are available, which are reviewed by MSS: Bayesian p-values, cross-validation, and another approach developed by Box (1980). Here again, I'm not aware of any evaluation of the frequentist properties (nominal level and power) of these methods. Besides, these procedures tend to be 'omnibus', in that the alternative hypothesis is simply stated as 'the model does not fit the data at hand', without any further indication as to where to go then. Once again, the Bayesian approach may benefit from getting closer to a classical framework. Indeed, goodness-of-fit testing procedures are well developed for single (Lebreton, Burnham, Clobert & Anderson 1992) and multistate (Pradel, Wintrebert & Gimenez 2003) mark-recapture models (review in Pradel, Gimenez & Lebreton 2005). These methods rely on contingency tables that specify well-identified alternative hypotheses (e.g., a trapping effect on recapture probabilities or a memory effect on movement probabilities), which, in case of rejection, are invaluable when it comes to building a model which fits the data better.

2.4. Even more complexity?

I have two further minor suggestions that might help to improve the MSS model. First, MSS made the strong assumption that variation in survival could be fully explained by travel times variation. However, if some extra variation exists, then bias may occur in parameter estimates (Barry, Brooks, Catchpole & Morgan 2003). By using a state-space formulation of their model (Gimenez et al. 2007), MSS could have incorporated individual random effects to cope with (potential) unexplained sources of variability in survival (see Clark et al. 2005; Gimenez et al. 2006; Zheng, Ovaskainen, Saastamoinen & Hanski 2007; Royle 2008). Second, MSS wonder how their model might be extended to a two–dimensional spatial setting. We have recently

extended our nonparametric approach dealing with complex relationships between survival and covariates that was mentioned by MSS (Gimenez et al. 2006; see also Gimenez et al. 2006 in the Additional References) to cope with bivariate smoothing (Gimenez & Barbraud 2008). This methodology has also potential applications in ecology in order to estimate spatial synchrony, as well as in evolutionary biology in order to estimate fitness surfaces made of quantitative phenotypic traits.

3. GENERAL CONSIDERATIONS

3.1. Bayes or not Bayes: is that the question?

As a biostatistician, I have long adopted an eclectic and pragmatic approach and have been using either the Bayesian or the frequentist approach based on a few empirical criterions such as the time it takes to get results, the ease of programming and the nature of the biological question (is there any added value of going for a Bayesian analysis?). From a practitioner's point of view, it is worth repeating that, although it may appear obvious, both approaches are complementary, provided that one is careful in using the terminology. As a nice illustration of this statement, I would like to draw attention to the MCMC procedure recently proposed by Lele, Dennis & Lutscher (2007) which has the appealing feature of producing maximum likelihood estimates. It is fair to say that the Bayes approach is rarely used for what it is intrinsically, but rather as an excuse for implementing the MCMC machinery to cope with complex multidimensional likelihoods. Examples of incorporating prior information are still too few (see, however, Martin, Kuhnert, Mengessen & Possingham 2005; McCarthy & Masters 2005), probably due to our feebleness as referees, while every biologist would agree never to start a new data analysis without prior knowledge of the system.

3.2. Transfer to biologists.

Obviously, in a pragmatic approach, the key question is 'why should I go for a Bayesian analysis'. The answer depends obviously on the analyst and the question, but several steps may be taken to help in deciding whether jumping or not into new territory is worth the price.

- We should think more of teaching Bayesian theory in introductory statistics courses, although some colleagues still hesitate to do so. I've taught both Bayesian and frequentist theories this year in a statistical modelling course for Masters students in ecology and evolutionary biology. The discussions were stimulating, focusing mainly on the incorporation of prior information and on when to use one or the other method. I will repeat this next year as I think that this lecture has not only added an arrow to their bow, but it has also contributed to the development of their critical mind.
- Several excellent textbooks are now available which encourage self–teaching, for applied statistician readers (Gilks, Richardson & Spiegelhalter 1996; Lee 1997; Carlin & Louis 2000; Congdon 2003; Gelman, Carlin, Stern & Rubin 2003; Congdon 2006) as well as for biologist readers (Clark 2007; McCarthy 2007), and many others will surely follow. Attending Bayesian workshops is another very efficient way of learning new material, which has the non negligible advantage of keeping us stuck somewhere (often in exotic places) with limited risk of being disturbed.
- To encourage codes and data sharing, we militate with other colleagues for the creation of a statistical ecology internet platform, with a format similar to Genbank (Benson et al. 2007) in genetics, a database to which nucleotides sequences are submitted prior to publication. This web site would gather material (in particular BUGS codes) that have been used in publications, and would avert the too convenient statement 'the code is available upon request from the authors' which I have used myself too often.

- Related to that, user-friendly and reliable pieces of software are needed. WinBUGS (Spiegelhalter, Thomas & Best 2003) is very flexible (Gimenez et al. 2008), but could gain in conviviality (e.g., by improving its debugging capabilities and implementing in routine several simple analyses). The efforts to build a dialog between R and WinBUGS initiated through the R package R2WinBUGS (Sturtz, Ligges & Gelman 2005) should be continued. An alternative to WinBUGS is AD–Model Builder (Fournier 2001), which seems to be much quicker, but is neither free nor open-source, as is WinBUGS.

Overall, teaching (and research) in ecological statistics largely benefits from collaboration between statisticians and biologists.

4. CONCLUSIONS

Hierarchical analysis of data on marked animals (Clark et al. 2005; Pradel 2005; Gimenez et al. 2007; Zheng, Ovaskainen, Saastamoinen & Hanski 2007; Royle 2008) is experiencing an increasing number of applications in ecology, conservation and evolutionary biology, thanks to the Bayesian framework in conjunction with MCMC methods for its implementation. Note that even though this combination has many advantages, I do not mean to overlook other methods that are valuable to fit models with latent structures, such as particle filtering (Buckland, Newman, Thomas & Koesters 2004; Thomas, Buckland, Newman & Harwood 2005), Kalman filtering (Besbeas, Freeman, Morgan & Catchpole 2002) and Newton-type algorithms (Pradel 2005). Even better, we still need to explore other methods since, although the Bayesian framework is more than three centuries old, we have to confess that its practical implementation using MCMC simulations is not as mature as maximum likelihood analyses using standard optimization methods. As Muthukumarana, Schwarz and Swartz acknowledge, potential issues may arise at various steps of the analysis, such as model identifiability, convergence assessment, model selection and goodness-of-fit testing. We see, however, good signs of a trend towards clear guidelines on how to carry out a Bayesian analysis using MCMC algorithms, the paper by MSS being an important contribution in that direction. In that spirit, and besides the original development of new models for fish monitoring, I consider the paper by MSS as a successful attempt to produce a Bayesian analysis template for future data analyses.

ACKNOWLEDGEMENTS

Many thanks to B. J. T. Morgan who encouraged me to do my first steps in the Bayesian theory while I was his research fellow working at the university of Kent, and to Rachel Borysiewicz who checked my English with patience.

Received 7 November 2007
Accepted 7 November 2007

Olivier GIMENEZ: olivier.gimenez@cefe.cnrs.fr
Centre National de la Recherche Scientifique
Centre d'Ecologie Fonctionnelle et Evolutive –UMR 5175
1919 Route de Mende, 34293 Montpellier Cedex 5, France

Authors' response

We thank Gimenez for his commentary on our article—we agree with him on virtually all of his points.

Swartz was born a Bayesian, Muthukumarana is a novice, and Schwarz is a heretic who converted from likelihood methods. We share many of Gimenez's frustrations with the Bayesian approach in capture-recapture. Perhaps some of us have too many grey hairs and find it difficult to learn new tricks, but the transition from likelihood methods to Bayesian methods has not been without much soul-searching and questioning.

1.MODEL SELECTION

Some of us like the likelihood AIC paradigm – all models are wrong, so if several models seem to fit the data equally well, why not combine inference over these models rather than searching for the best single model. A possibility is to do the same using Bayesian methods although this requires additional complexity and computation.

Like Gimenez, given that Bayes factors can rarely be computed in complex models, we find the Bayesian alphabet soup (AIC, DIC, BIC, etc) somewhat confusing. We are also uneasy about some uses of RJMCMC methods where there are literally millions of potential models (e.g. King et al. 2006). The first rule in using likelihood-AIC methods is NOT to data dredge, but to start with a carefully selected set of candidate models. This advice seems to be discarded in these large scale reversible jump MCMC applications where the model space can't even be listed in advance because it is so large. We wonder about the reliability of inference when millions of MCMC iterations cannot possibly cover the entire model space.

In cases like this, perhaps a better model selection summary is available? Rather than producing the probability of individual models, perhaps obtaining the probability of related groups of models would be more useful? For example, in King, Brooks, Morgan & Coulson (2006), the number and size of age classes was explored. The probability of a model for a given age class structure or all models that are subsets of this age class structure seems appropriate.

2.GOODNESS-OF-FIT

One approach to Bayesian goodness-of-fit involves model comparison between the current model and a very general but well-fitting model. This approach suggests the use of model selection methods rather than trying to mimic likelihood methods via Bayesian p-values, etc. This could be a viable strategy for goodness-of-fit against specified alternatives, but is it possible to determine an adequate and "fully-saturated model" in the Bayesian context?

We are also somewhat at a loss on how to deal with goodness-of-fit for these very complex models. In likelihood methods, the basic goodness-of-fit tests in mark-recapture methods are comparisons between the observed and expected counts of various histories. However, in complex models such as in this paper, every history (a combination of where and when captured) is likely unique, and this approach breaks down. There have been many paper in the literature dealing with sparse multinomial models (e.g., Simonoff, 1985; Eubank, 1997) which assume local smoothness in the cells which would seem like a logical way to proceed, but we are unaware of any such methods being used for capture-recapture.

3.INFLUENCE OF PRIORS

One of the powerful advantages of Bayesian methodology is the ability to include useful prior information. Yet, lip service is often paid to this idea; indeed, "uninformative" priors are often used where the data are supposed to speak for themselves. Some of us feel that there are actually two stages of inference that ought to be practiced. In the first stage, we find it appealing to see what the data from the experiment, and the experiment alone are saying. In this case, non-informative priors would be utilized. In the second stage, a Bayesian analysis using subjective prior information would be carried out. The two-stage procedure permits a comparison of the relative importance of data versus prior opinion. In practice, we find that the second step is rarely done. Often, it is difficult to summarize the state of knowledge about a parameter and researchers want to avoid comments from referees that the prior ignored paper x or paper y.

Another reason for the choice of noninformative priors is the (naive) belief that the results are not influenced by the prior. Schwarz must admit that he has made this mistake in thinking. He finds it unintuitive that a flat beta prior on p in a binomial experiment has the "effect" of adding "1" to the sample size when the posterior is examined. In his naive thinking, Schwarz would like (uninformative) priors to "subtract" 1 from the sample sizes! And how can priors be uninformative when there are sometimes several uninformative priors to choose from in a given problem?

Gimenez's suggestion of looking at the overlap between the prior and posterior distributions (see the additional references) appears to be one way to see how much new information is contributed by the data over the prior. This could be an "automated" approach to looking at the relative contribution of information from priors and data, but another (admittedly) naive approach would be to simply look at the difference between the maximum likelihood estimator (MLE) and the mean of the posterior. Presumably if the MLE cannot be explicitly computed, a stochastic MLE could be found, as outlined in Lele, Dennis & Lutscher (2007).

4.TECHNOLOGY TRANSFER

Gimenez's suggestion of an MCMC bank for code is extremely useful. The 'wiki' paradigm may be more useful where code can be modified and improved. An impediment, unfortunately, is the common problem that the reward system in most academic institutions is completely orthogonal to collaborative approach of this sort.

We also agree with Gimenez's comments about the difficulty in using WINBUGS. Except for simple problems, we find that using this package is not for the faint of heart. We hope that our WINBUGS can serve as a foundation for future applications in capture recapture.

We cannot emphasize too strongly the importance of exercising code using simulated data (with enormous sample sizes) where the answers are known in advance and MCMC simulation error will be negligible. These simulation exercises should try a variety of parameter values and scenarios. Here we grey-haired members of the community actually have an advantage over newcomers as this was common practice when debugging FORTRAN or other codes, but is rarely practiced now.

5.FINAL REMARKS

As Gimenez notes, the Bayesian paradigm together with MCMC methods, allows more complex models to be fit than when using standard likelihood theory. But, at the same time, the fundamental problems of model selection, model averaging and goodness-of-fit are seldom straightforward.

We found that our team worked well on this paper. Often new converts are more zealous than members born to a faith, but our varied team kept each other in check. The novice supplied the energy and drive; the heretic challenged the methods at every step and the cardinal supplied the background and wisdom to ensure that the project did not stray from the proper path.

Saman MUTHUKUMARANA: muthukumaranasm@yahoo.com
Carl J. SCHWARZ: cschwarz@stat.sfu.ca
Tim B. SWARTZ: tim@stat.sfu.ca
Department of Statistics and Actuarial Science
Simon Fraser University, Burnaby
British Columbia, Canada V5A 1S6

ADDITIONAL REFERENCES

S. C. Barry, S. P. Brooks, E. A. Catchpole & B. J. T. Morgan (2003). The analysis of ring-recovery data using random effects. *Biometrics*, 59, 54–65.

D. A. Benson, I. Karsch-Mizrachi, D. J. Lipman, J. Ostell & D. L. Wheeler (2007). Genbank. *Nucleic Acids Research*. 35, D21–D25.

P. Besbeas, S. N. Freeman, B. J. T. Morgan & E. A. Catchpole (2002). Integrating mark-recapture-recovery and census data to estimate animal abundance and demographic parameters. *Biometrics*. 58, 540–547.

G. E. Box (1980). Sampling and Bayes' inference in scientific modelling and robustness (with discus- sion). *Journal of the Royal Statistical Society Series A*, 143, 383–430.

S. T. Buckland, K. B. Newman, L. Thomas & N. B. Koesters (2004). State-space models for the dynamics of wild animal populations. *Ecological Modelling*, 171, 157–175.

K. P. Burnham & D. R. Anderson (2002). *Model Selection and Inference: A Practical Information Theoretic Approach*. Second edition. Springer-Verlag. New York.

J. S. Clark (2007). *Models for Ecological Data: An Introduction*. Princeton University Press.

J. S. Clark, G. A. Ferraz, N. Oguge, H. Hays & J. DiCostanzo (2005). Hierarchical Bayes for structured, variable populations: from recapture data to life-history prediction. *Ecology*, 86, 2232–2244.

P. Congdon (2003). *Applied Bayesian Modelling*. Wiley, Hoboken, New Jersey.

P. Congdon (2006). *Bayesian Statistical Modelling*. Second Edition. Wiley, Hoboken, New Jersey.

R. L. Eubank (1997). Testing goodness-of-fit with multinomial data. *Journal of the American Statistical Association*, 92, 1084–1093.

S. El Adlouni, A.-C. Favre & B. Bobée (2006). Comparison of methodologies to assess the convergence of Markov chain Monte Carlo methods. *Computational Statistics and Data Analysis*, 50, 2685–2701.

D. Fournier (2001). *An introduction to AD MODEL BUILDER Version 6.0.2 for use in nonlinear modeling and statistics*. Available from http://otter-rsch.com/admodel.htm

Otter Research "Integrated Statistical Models", Sidney, British Columbia.

E. S. Garrett & S. L. Zeger (2000). Latent class model diagnosis. *Biometrics*, 56, 1055–1067.

A. E. Gelfand & S. K. Ghosh (1998). Model choice: a minimum posterior predictive loss approach. *Biometrika*, 85, 1–11.

E. I. George & R. E. McCulloch (1993). Variable Selection via Gibbs Sampling. *Journal of the American Statistical Association*, 88, 881–889.

O. Gimenez, & C. Barbraud (2008). The efficient semiparametric regression modeling of capture-recapture data: assessing the impact of climate on survival of two Antarctic seabird species. *Environmental and Ecological Statistics*, in press.

O. Gimenez, S. Bonner, R. King, R. A. Parker, S. P. Brooks, L. E. Jamieson, V. Grosbois, B. J. T. Morgan & L. Thomas (2008). WinBUGS for population ecologists: Bayesian modeling using Markov Chain Monte Carlo methods. *Environmental and Ecological Statistics*, in press.

O. Gimenez, R. Covas, C. R. Brown, M. D. Anderson, M. Bomberger Brown & T. Lenormand (2006). Nonparametric estimation of natural selection on a quantitative trait using mark-recapture data. *Evolution*, 60, 460–466.

O. Gimenez, B. J. T. Morgan & S. P. Brooks (2008). Weak identifiability in models for mark-recapture-recovery data. *Environmental and Ecological Statistics*, in press.

O. Gimenez, V. Rossi, R. Choquet, C. Dehais, B. Doris, H. Varella, J.-P. Vila & R. Pradel (2007). State-space modelling of data on marked individuals. *Ecological Modelling*, 206, 431–438.

O. Gimenez, A. Viallefont, R. Choquet, E. A. Catchpole & B. J. T. Morgan (2005). Methods for investigating parameter redundancy. *Animal Biodiversity and Conservation*, 27, 561–572.

P. Green (1995). Reversible Jump Markov Chain Monte Carlo Computation and Bayesian Model Determination. *Biometrika*, 82, 711–732.

R. E. Kass, B. P. Carlin, A. Gelman & R. M. Neal (1998). Markov chain Monte Carlo in practice: A round-table discussion. *The American Statistician*, 52, 93–100.

R. King, S. P. Brooks, B. J. T. Morgan & T. Coulson (2006). Factors Influencing Soay Sheep Survival: A Bayesian Analysis. *Biometrics*, 62, 211–220.

P. M. Lee (1997). *Bayesian Statistics: An Introduction, Second Edition*. Wiley, New York.

S. R. Lele, B. Dennis & F. Lutscher (2007). Data cloning: easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods. *Ecology Letters*, 10, 551–563.

W. A. Link & R. J. Barker (2007). Model weights and the foundations of multimodel inference. *Ecology*, 87, 2626–2635.

T. G. Martin, P. M. Kuhnert, K. Mengessen & H. P. Possingham (2005). The power of expert opinion in ecological models using Bayesian methods: Impact of grazing on birds. *Ecological Applications*, 15, 266–280.

M. A. McCarthy (2007). *Bayesian Methods for Ecology*. Cambridge University Press.

M. A. McCarthy & P. Masters (2005). Profiting from prior information in Bayesian analyses of ecological data. *Journal of Applied Ecology*, 42, 1012–1019.

R. Pradel (2005). Multievent: An extension of multistate capture-recapture models to uncertain states. *Biometrics*, 61, 442–447.

R. Pradel, O. Gimenez & J.-D. Lebreton (2005). Principles and interest of GOF tests for multistate capture-recapture models. *Animal Biodiversity and Conservation*, 28, 189–204.

R. Pradel, C. M. A. Wintrebert & O. Gimenez (2003). A proposal for a goodness-of-fit test to the Arnason-Schwarz multisite capture-recapture model. *Biometrics*, 59, 43–53.

J. A. Royle (2008). Modeling individual effects in the Cormack-Jolly-Seber model: A state-space formulation. *Biometrics*, in press.

J. S. Simonoff (1985). An Improved Goodness-of-Fit Statistic for Sparse Multinomials . *Journal of the American Statistical Association*, 80, 671–677.

S. Sturtz, U. Ligges & A. Gelman (2005). R2WinBUGS: a package for running WinBUGS from R. *Journal of Statistical Software*, 12, 1–16.

L. Thomas, S. T. Buckland, K. B. Newman & J. Harwood (2005). A unified framework for modelling wildlife population dynamics. *Australian and New Zealand Journal of Statistics*, 47, 19–34.

C. Zheng, O. Ovaskainen, M. Saastamoinen & I. Hanski (2007). Age-dependent survival analyzed with Bayesian models of mark-recapture data. *Ecology*, 88, 1970–1976.