

Estimation of gillnet efficiency and selectivity across multiple sampling units: A hierarchical Bayesian analysis using mark-recapture data

Paul J. Askey*, John R. Post, Eric A. Parkinson, Etienne Rivot, Andrew J. Paul, Peter A. Biro

Department of Biological Sciences, University of Calgary, 2500 University Dr. N.W., Calgary, Alta. T2N 1N4, Canada

Received 16 June 2006; received in revised form 9 September 2006; accepted 16 September 2006

Abstract

Fisheries research often involves a repetitive sampling protocol for multiple ecological units (our example is gillnetting of lakes). Estimates of abundance based on catch per unit effort can be erroneous due to lake effects on sampling efficiency. Conversely dividing data into individual lakes may lead to poor inference due to sparse data. Hierarchical Bayesian analysis compromises between these two extreme methods by estimating parameters for an individual lake, but borrowing information from other lakes. We estimated size-selective gillnet efficiency with mark-recapture data across a series of lakes subject to a constant netting effort. Hierarchical Bayesian analysis was able to prevent unrealistic selectivity functions that arose from individual lake analysis. Furthermore, the hierarchical approach was able to derive accurate parameter estimates with very few mark-recaptures in sub-sampled data trials. This paper demonstrates the hierarchical methodology for the estimation of fishery selectivity parameters. The results could be used to derive informative priors for future research which uses the proposed gillnet protocol.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Hierarchical Bayesian; Random effects; DIC; Gillnet; Selectivity; Capture probability; Lake sampling; Gillnet

1. Introduction

Fisheries research and stock assessment often require the estimation of fish abundance across multiple spatial or temporal units. Abundance estimates are primarily influenced by gear selectivity and efficiency parameters, which may or may not be related to the sampling unit. Thus the biologist is faced with the important task of deciding at what level of complexity to model sampling efficiency. In the past this problem has often been avoided by assuming that fishing efficiency is constant across units, and therefore abundance is simply proportional to a pooled CPUE estimate (e.g. Post et al., 1999). However, this assumption should not be taken lightly, as large errors can be present when this is not the case (Fryer, 1991). At the other extreme, gear efficiency in each sampled unit is assumed to be completely independent from the other sampled units, and individual efficiency estimates are required (e.g. Ingólfsson and Jørgensen, 2006). Yet by assuming sampling units are completely unrelated the data may become sparse in some individual units and lead to highly imprecise estimates. More recent work using hierarchi-

cal Bayesian analysis offers a compromise between these two extremes. This approach has proven useful at improving estimates of stock-recruitment parameters (Liermann and Hilborn, 1997; Michielsens and McAllister, 2004), trawler gillnet efficiency (Harley and Myers, 2001), and in mark-recapture data (Rivot and Prevost, 2002).

The advantage of hierarchical Bayesian models (HBM) is that information from the other similar sampling units is incorporated to improve individual estimates. Consider a set of lakes that are sampled with equivalent gillnetting effort, and no information other than the data is available to distinguish one lake from another. An estimate of gillnet efficiency using the hierarchical Bayesian approach does not treat the lakes as completely independent nor equivalent, but rather each lake is an exchangeable random sample. Exchangeability implies that lake specific prior distributions for gillnet efficiency parameters can be considered independent samples from a common population distribution, which is indexed by some hyperparameters (Gelman et al., 2004).

The improvement to parameter estimates in the HBM is not due to increased complexity through the addition of hyperparameters, but rather by the exchange of information implied by the hyperprior(s). This makes it difficult to place HBMs on a complexity scale next to non-hierarchical models, since the

* Corresponding author. Tel.: +1 403 220 5263.

E-mail address: pjaskey@ucalgary.ca (P.J. Askey).

Table 1
Description of study lakes and years sampled

| Lake | Area (ha) | Maximum depth (m) | Years sampled | Sample number (s) |
|----------------------|-----------|-------------------|------------------|-------------------|
| Crater Pothole 4 | 5.4 | 16 | 1993 | 1 |
| Bluey Pothole 2 | 1.4 | 6 | 1993 | 2 |
| Cigar | 3.8 | 9 | 1993 | 3 |
| Crater Pothole 1 | 4.2 | 20 | 1993 | 4 |
| Crater Pothole 3 | 3.3 | 18 | 1993 | 5 |
| Smoke | 2.0 | 8 | 1993 | 6 |
| Big Pantano | 2.1 | 4 | 2003, 2004 | 7, 15 |
| Little Pantano | 1.1 | 3 | 2003, 2004 | 8, 16 |
| Spook | 4.4 | 4 | 2003, 2004, 2005 | 9, 12, 18 |
| Stubby | 6.2 | 8 | 2003, 2004 | 10, 13 |
| Today | 6.5 | 10 | 2003, 2004, 2005 | 11, 14, 19 |
| Wilderness | 12.0 | 12 | 2004 | 17 |
| Pantano ^a | 3.2 | 4 | 2005 | 20 |
| NoFish | 13 | 15 | 2005 | 21 |

^a Note Big and Little Pantano are divided sections of Pantano (Askey et al., in press).

distributional constraints imposed on parameters by the hyperprior reduce the effective dimensionality of the model. However, a measure of complexity would be useful since it is a necessary component to model selection criteria. Spiegelhalter et al. (2002) have proposed a Deviance Information Criterion (DIC), which has similar rationale and interpretation as the Akaike Information Criteria (AIC) (Akaike, 1973) method used under a likelihood approach. The method depends on a measure of the “effective number of parameters”, which may have variable performance depending on the situation (see reviewer discussion in Spiegelhalter et al., 2002). An alternative approach to model selection of hierarchical models is cross-validation with a subset of the data (Stone, 1974; Gelman et al., 2004).

Efficiency and selectivity of fishing gears is a well studied subject in the fisheries literature (Millar and Fryer, 1999; Bromaghin, 2005; Rosenberger and Dunham, 2005; Booth and Potts, 2006; Ingólfsson and Jørgensen, 2006). However, a drawback to results from such studies is their limited applicability to future research. For example gillnet information is often limited to relative size selectivity, which is usually inferred indirectly (Hamley, 1975; Millar and Fryer, 1999). Moreover, the between sample variation is usually ignored (Fryer, 1991). Thus it is difficult to apply this information directly towards predicting probability of capture and abundance at a new experiment or fishery. We feel that studies, which have designed direct methods for estimating sampling efficiency (usually mark-recapture) over multiple sampling units, can provide substantial information for new fishing events when a hierarchical Bayesian approach is used. A hierarchical approach produces posterior predictive distributions that describe the expected outcome for a new sam-

pling unit. This distribution can be combined with future data as an informative prior, or in the absence of mark-recapture data, directly applied as the sample estimate with error equal to the distribution.

This paper investigates the potential of the hierarchical Bayesian approach to analysis of size-selective gillnet efficiency models from mark-recapture experiments. Our experiments were specifically designed to directly quantify gillnet selectivity within and among lakes using a total of 21 lake-years of mark-recapture data. This large data set is used to compare the performance of hierarchically structured models with the pooled and independent model alternatives. We also examine the performance of each approach when data is sparse by reducing the full data set to subsets. Finally, we provide a useful gillnet protocol whereby the posterior predictive parameter estimates can potentially be used as informative priors within future analyses that use the same sampling design.

2. Methods

2.1. Study area and data collection

A total of 21 lake-years of data derived from 13 individual lakes over four sampling years were used in the study (Table 1). The lakes are located in the south-central region of British Columbia, Canada, and generally contain a single fish species, rainbow trout (*Oncorhynchus mykiss*). One lake, Wilderness, also contained a natural population of northern pikeminnow (*Ptychocheilus oregonensis*). Six of the lakes were sampled in 1993 as part of a study on size-structured recruitment dynamics

Table 2
Gillnet types and effort used per hectare and per night

| Net type | Habitat | Nets per hectare | Net depth (m) | Net length (m) | Mesh sizes (mm) |
|----------|----------|------------------|---------------|----------------|----------------------------|
| Lundgren | Littoral | 3 | 1.5 | 11.4 | 13, 16, 20, 25 |
| Sinker | Littoral | 1 | 2.3 | 106.7 | 25, 76, 51, 38, 89, 64, 32 |
| Floater | Pelagic | 0.25 | 6 | 106.7 | 25, 76, 51, 38, 89, 64, 32 |

Total net effort is the summed area over all net types. Mesh sizes are in the order that they occur within the nets and refer to stretched mesh size. Total net effort = 457 m² ha⁻¹ night⁻¹.

(details in Post et al., 1999). The other seven lakes were sampled in 2003–2005 as part of a separate study located north of Kamloops, B.C. at 120°21'54"W 51°9'11"N.

Gillnetting took place in the last week of September and first 2 weeks of October and consisted of three consecutive nights of netting using a constant netting effort (Table 2, Post et al., 1999). Lundgren experimental gillnets consisted of four small mesh sizes (<25 mm) sewn together and were used to sample shallow near-shore habitats (Table 2). The sinking and floating nets consisted of seven panels with differing mesh sizes (25–89 mm) attached at the top and bottom such that a small gap existed between each panel (Table 2). Sinking nets were set at approximately 1–6 m depth in a broad zigzag pattern so that all mesh sizes fished the entire range in depth. Floating panels were set in a straight line or broad curve through the center of the lake. Total netting effort was standardized so that each lake received the same surface area of gillnet per hectare of lake area (Table 2). Mean effort was 457 m² ha⁻¹ night⁻¹, however, there was some variation about the mean because nets come only in integer sizes and could not be split to exactly match fractional sizes of lakes. Nets were set at mid-day and allowed to fish for 18–24 h period. Each net was moved for the following night of fishing to ensure that all areas and habitats were sampled.

Approximately 1 week prior to fall netting a large number (range = 417–726) of marked (fin clipped) rainbow trout ranging in size from 50 to 400 mm fork-length were released into each individual lake. The majority of these fish came from the Fraser

Valley trout hatchery in Abbotsford, B.C. Two age classes (age 0 and age 1) were reared to four separate size classes in order to have marked individuals across the full size range (Fig. 1). In some cases the number of marked fish was augmented with fish captured from within the lakes using beach seines or fyke nets, then fin clipped and released.

2.2. Data analysis

In order to account for the selectivity of gillnets, the fish were divided into 1 cm size bins for data analysis. The data is presented as a simple Petersen mark-recapture where the catch is the sum of three nights of gillnetting effort. Thus the number of marked fish captured (r) in size bin i on lake-year j is modelled by the binomial distribution:

$$r_{ij} \sim \text{bin}(p_{ij}, m_{ij}) \quad (1)$$

where p_{ij} and m_{ij} are the probability of capture and number of marked fish of size i in lake-year j , respectively. We further assume that recapture probability is independent between size bins. Thus, the sampling distribution for lake-year j with data $(m_{ij}, r_{ij})_{i=1, \dots, l}$ is the product of Eq. (1) for all size bins i . The length assigned to all fish within a 1 cm size bin was the midpoint of the size bin.

There are several functions that can be used to model gillnet selectivity (Millar, 1995). We used the logistic function, with an extra parameter to account for an asymptote below one. Thus

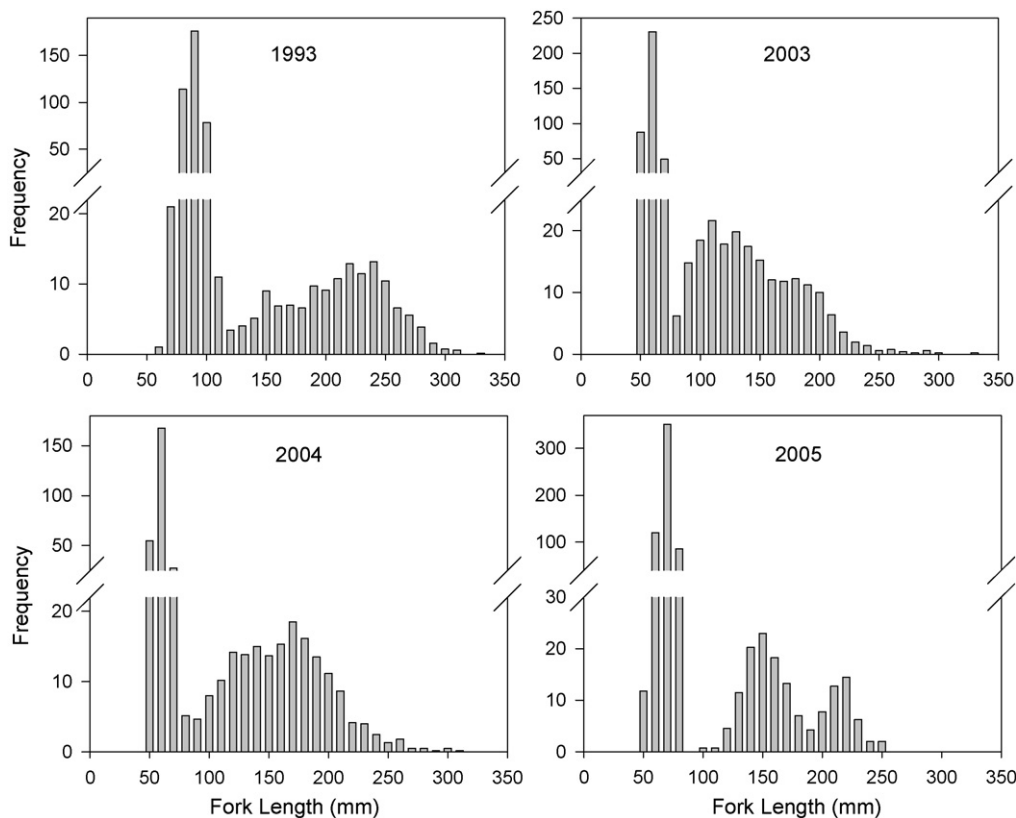


Fig. 1. Size distribution of marked fish stocked approximately 1 week prior sampling in each of the four sampling years. Frequencies and distributions represent the mean number per lake per year.

the probability of capture (p) for a fish in size bin i is

$$p(l) = p_{\max} \left(\frac{e^{\alpha+\beta l}}{1 + e^{\alpha+\beta l}} \right) \quad (2)$$

where p_{\max} is the maximum probability of capture for large fish, α and β are shape parameters and the length l assigned to all fish within a 1 cm size bin i is the midpoint of the size bin. This is the most commonly used fisheries selectivity function (Millar and Fryer, 1999), and fit the data well. For our problem, the parameters α and β model the combined processes of size-specific selectivity and size-specific availability, whereas overall catchability is accounted for by the p_{\max} parameter. Below we denote the full set of three selectivity parameters (α, β, p_{\max}) as θ , and the indices $j = 1, \dots, J$ ($J = 21$) stand for the lake-year units.

The focus of this paper is how to most parsimoniously model the gillnet selectivity function (Eq. (2)) with regards to the J lake-years. The two extreme answers are: (1) a single pooled capture probability model, with a single vector of three selectivity parameters (θ) common for all lake-years, and (2) a lake-year specific model such that each parameter has a subscript j ($3 \times J = 63$ selectivity parameters in total). We may also produce intermediate complexities by removing the lake effect from some parameters and not others. It should also be noted that our complete set of $J = 21$ lake-years of data may not be considered independent, because some lakes were sampled more than once and lakes were sampled in the same year (Table 1). Thus we created models that described lake effects and year effects, and examined their importance in a DIC analysis discussed below. Finally in the case of hierarchical analysis we do not remove the lake effect on selectivity parameters, but rather impose distributional constraints. We modelled gillnet efficiency from both non-hierarchical and hierarchical approaches as well as some intermediates. The descriptions of the three primary modelling approaches are given below, and the set of priors tested for each case is given in Appendix A.

2.3. Model 1: pooled Bayesian model (PBM)

In this case all lake-year effects are ignored and we fit a single size-selective gillnet capture probability model for all lakes and years of data. This model assumes that the constant gillnet protocol should result in a single selectivity curve. We denote $p(\theta)$ the prior probability density function (PDF) for the three-dimensional parameter vector θ . The posterior PDF of the parameters of the pooled model (given all available data) is simply proportional to the product of the prior with the sampling distribution, which is the product of all likelihood terms for each lake-year j . For each lake-year j , the likelihood is obtained from the product over all size bins i of likelihood terms $L(r_{ij}, m_{ij}|\theta)$, each directly obtained from the binomial sampling distributions of the capture-mark-recapture data defined by Eq. (1), were the capture probability p_{ij} is a function of the parameters θ following Eq. (2):

$$p(\theta|\text{data}) \propto p(\theta) \prod_{j=1}^J \prod_{i=1}^I L(r_{ij}, m_{ij}|\theta) \quad (3)$$

We chose diffuse prior PDFs for all three parameters intended to depict no prior information regarding parameter estimates (see Appendix A). As there is no unique definition of non-informative prior PDFs, we performed a sensitivity analysis of our results to the prior specification (see Appendix A).

2.4. Model 2: independent Bayesian model (IBM)

In this case it is assumed that although the gillnet densities and protocol are constant between lake-years, the capture efficiencies of the 21 lake-years are different and mutually independent. It is difficult to imagine reasons for complete independence between lakes, but potentially variation in environmental conditions and differences in lake physical characteristics could change fish behaviour and catchability. As all units (j) are independent, one can write the posterior distribution of the parameters for each lake-year j :

$$p(\theta_j|\text{data}_j) \propto p(\theta_j) \prod_{i=1}^I L(r_{ij}, m_{ij}|\theta_j) \quad (4)$$

The same non-informative prior PDFs as in model 1 were used for all of the $3 \times J = 63$ parameters (see Appendix A).

2.5. Model 3: hierarchical Bayesian model (HBM)

In this model we assume a hierarchical structure on the gillnet efficiency across lake-years. The hierarchical structure allows for lake-year effects on the gillnet selectivity curve, but assumes that lake specific parameter values are sampled from a common population distribution conditional on unknown hyperparameters, ϕ . Thus the lake-year specific parameters θ_j are considered exchangeable, and this is expressed in the joint prior distribution for ϕ and all θ :

$$p(\phi, \theta) = p(\phi) \prod_{j=1}^J p(\theta_j|\phi)$$

The prior is combined with the sampling distribution likelihood, which only depends on the data and not ϕ , thus the full joint posterior distribution for the hierarchical model is simply:

$$p(\phi, \theta|\text{data}) = p(\phi) \prod_{j=1}^J p(\theta_j|\phi) L(r_{ij}, m_{ij}|\theta_j)$$

The hierarchical structure can be visualized with a Directed Acyclic Graph (DAG) (Spiegelhalter et al., 1996), which is presented in Fig. 2 along with the simpler structures of models 1 and 2. Gelman et al. (2004) as well as Rivot and Prevost (2002) provide more detailed introductions into the structuring of hierarchical Bayesian models.

The prior distributions used for the HBM are given in Appendix A. We assessed the sensitivity of our results to the choice of prior and we checked the exchangeability hypothesis (see Appendix A).

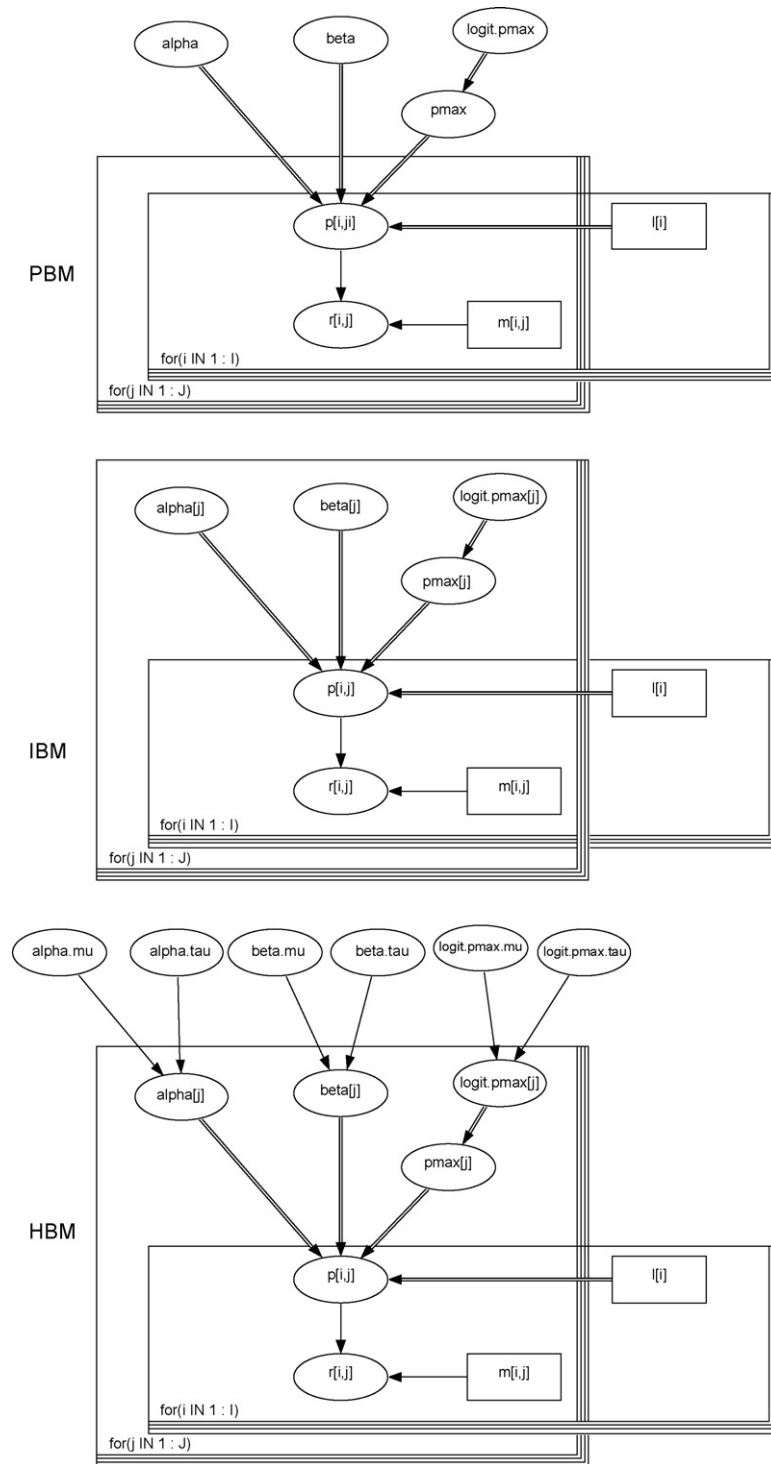


Fig. 2. Directed acyclic graphs (DAGs) give a visual representation of the three model structures: Pooled Bayesian Model (PBM), Independent Bayesian Model (IBM), and Hierarchical Bayesian Model (HBM). Stochastic nodes are represented by ovals, constant nodes are represented by rectangles, the arrows show relationships between nodes and the rectangular plates represent the individual lake and fish size indices. Parameter names conform to model descriptions in text, where mu and tau suffices represent mean and precision parameters of a normal distribution, respectively.

2.6. Model selection

A general approach to choosing between models of differing complexity is information theoretic criteria (e.g. Burnham and Anderson, 2001). In these approaches models are assessed by a balance of their fit to the data (usually deviance) and the model

complexity (the number of parameters) needed to obtain the fit. The most well-known example of this approach is Aikake’s Information Criterion (AIC) which proposes the most parsimonious model is the minimum:

$$AIC = -2LL + 2k$$

where LL is the maximum log-likelihood and k is the number of estimated parameters.

Although the AIC approach can be applied to fixed effects models, such as the PBM and IBM, it is not useful when considering hierarchical structures. This is because the distributional assumptions associated with hierarchical Bayesian structure make it difficult to uniquely define the number of parameters. Recall the lake level parameters within the selectivity function parameter vector, θ :

$$\theta_j = (\alpha_j, \beta_j, p_{max_j})$$

Assuming a normal hyperprior for all three parameters, where p_{max} is logit transformed, then the hyperparameter vector θ is made up of a mean (μ) and variance (σ) parameter for each of the three selectivity parameters in θ . Thus the explicit hyperparameter vector implied above is

$$\phi = (\mu_\alpha, \mu_\beta, \mu_{pmax}, \sigma_\alpha, \sigma_\beta, \sigma_{pmax})$$

Under this formulation it is clear that as the between units variability decreases (σ_θ approaches 0) the 63 lake specific parameters will converge to 3, which results in model 1. Whereas, increasing σ_θ to ∞ would return us to model 2. Therefore, simply counting the parameters is not a measure of model complexity, since complexity will depend on the hyperparameters describing between units variability.

The Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002) has been proposed as an information theoretic approach to compare hierarchical models of arbitrary structure. The rationale and interpretation of DIC is analogous to AIC outlined above. It combines a measure of fit, defined as the posterior mean of the deviance $\overline{D(\theta)}$ (where deviance is -2 times the log-likelihood), with a measure of the model complexity, p_D :

$$DIC = \overline{D(\theta)} + p_D$$

where p_D is the number of “effective parameters”, defined as the posterior mean of the deviance minus the deviance at the posterior means $D(\bar{\theta})$:

$$p_D = \overline{D(\theta)} - D(\bar{\theta})$$

p_D should be approximately equivalent to the actual number of parameters (k) in fixed effects models, but may be less than the apparent number of parameters in hierarchical models for reasons given above. We used posterior medians instead of means for estimates of $D(\bar{\theta})$, as they are better point estimates of parameters if there is any skew in the posterior PDFs.

2.7. Cross-validation

A more pragmatic approach to comparing between model structures is cross-validation by sub-sampling from the full data set (Stone, 1974; Gelman et al., 2004). We assigned indexes to individual fish from each size-class in each lake and randomly sampled a given proportion to the subset for each lake. Thus the subset procedure is equivalent to considering our fall stocked marked population as the total N and the sub-sampling

procedure creates a new mark-recapture from within the known population. Since, we know the destiny (captured or not) of the marked fish, then we have data for the number of fish captured in the sub-sample. Data was sub-sampled to the integer value closest to 10, 20, 30, 40, 50, 60, 70, 80 and 90% of the full marked population in each lake. We first fit the reduced data set with the different model structures and compared posterior estimates of the total marked population with the known true marked population number for all sizes of fish released into the lake. Since the true population size for each lake (N_{true}) was known, the mean squared error (MSE) could be calculated as

$$MSE = \frac{1}{J} \sum_{j=1}^J (N_{est,j} - N_{true,j})^2$$

where $N_{est,j}$ is the total population estimate over all size bins (l_{min} to l_{max}) and calculated as

$$N_{est,j} = \sum_{i=l_{min}}^{l_{max}} \frac{c_{ij} + r_{ij}}{p_{ij}} \quad (5)$$

where c_{ij} is the total catch of fish of size class i from lake j that were not selected to the subset marked population and known to be captured. p_{ij} is the probability of capture estimated from the sub-sampled data for a given model, and r_{ij} represents known recaptures from the subset marked population. Since p_{ij} is a random variable, the MSE and N_{est} estimates are also random variables even though they are logical nodes (as opposed to stochastic) in the model structure. The uncertainty associated with these quantities is purely a result of uncertainty in the probability of capture in gillnets.

2.8. Abundance estimates

The ultimate goal of the capture probability estimation process is to convert catch rates into abundance. As an example, we estimated fall abundance for all lakes in 1993 and 2003. In this case the c_{ij} term in Eq. (5) represents fish that were already in the lake prior to the mark-recapture experiment (previously stocked in spring or from natural recruitment). Thus we make the tacit assumption that the capture probability of the marked fish stocked in the fall does not differ from the fish already present in the lakes. We also present abundance estimates in the absence of new mark-recapture information using the posterior predictive distributions for selectivity parameters ($\tilde{\theta}$) from the HBM given by

$$p(\tilde{\theta}|\text{data}) = \int p(\tilde{\theta}|\phi)p(\phi|\text{data})d\phi$$

The posterior predictive distributions could potentially be used to infer capture probability and estimate abundance based on catch rates alone. Thus, the difference between the posterior predictive estimates and the HBM estimates illustrates the expected gain in precision and accuracy from performing a mark-recapture experiment (given the data already collected).

Table 3
DIC model selection criteria for a set of non-hierarchical (fixed effects) models (HBM DIC given in text)

| Model | k | p_D | $\overline{D(\theta)}$ | $D(\bar{\theta})$ | DIC |
|--|-----|-------|------------------------|-------------------|--------|
| PBM (pooled) | 3 | 3.0 | 1906.8 | 1903.8 | 1909.8 |
| Year effect | 12 | 12.0 | 1812.2 | 1800.2 | 1824.2 |
| Lake-year effect on p_{\max} , but α and β pooled | 23 | 22.7 | 1771.3 | 1748.6 | 1794.0 |
| Lake effect | 42 | 41.6 | 1633.2 | 1591.6 | 1674.8 |
| IBM (lake-year effect) | 63 | 63.8 | 1582.0 | 1518.3 | 1645.8 |

In this subset of models, the IBM is clearly selected as being most parsimonious (minimum DIC). k is the actual number of parameters, p_D the estimated effective number of parameters, $\overline{D(\theta)}$ the average deviance and $D(\bar{\theta})$ is the deviance at the posterior median parameter estimates.

2.9. Bayesian implementation

All cases were modelled from a Bayesian perspective and programmed into WinBUGS[®] (Bayesian Inference using Gibbs Sampling) software, which uses a Monte Carlo Markov Chain (MCMC) algorithm to estimate the posterior PDFs (available at <http://www.mrc-bsu.cam.ac.uk/bugs>). As in all Bayesian analysis, choice of prior distributions may influence posterior estimates. Thus, we tested the sensitivity of prior choice on posteriors, which is included as an Appendix A. There is also a risk that the simulated MCMC chain has not converged to the target distribution. To guard against this we ran two chains starting at randomly seeded initial values. Two chains were run for 30,000 iterations with a 15,000 iteration burn-in (removed from sample) and further thinned to leave 1000 samples from each run. Further we monitored convergence visually by inspecting traces of the MCMC chains as well as using the Gelman–Rubin convergence diagnostics provided in the output of R2WinBUGS software available

at <http://www.stat.columbia.edu/~gelman/bugsR/> (Sturtz et al., 2005).

3. Results

As a first step we compared a set of non-hierarchical (fixed effects) models that range in complexity from the simple three parameter PBM model to the full 63 parameter lake-year specific, IBM model. We compared models with DIC, which is appropriate for a Bayesian framework and is able to incorporate hierarchical models. However, it is worth noting that a frequentist-likelihood based approach with AIC could be used on this subset of non-hierarchical models and will lead to very similar inferences (in terms of model selection/ranking) because non-informative priors were used. Estimates of the effective number of parameters, p_D , were generally quite good at counting parameters when using median parameter estimates (Table 3). The DIC values were quite disparate and clearly selected (minimum DIC) the most complex 63 parameter model (Table 3).

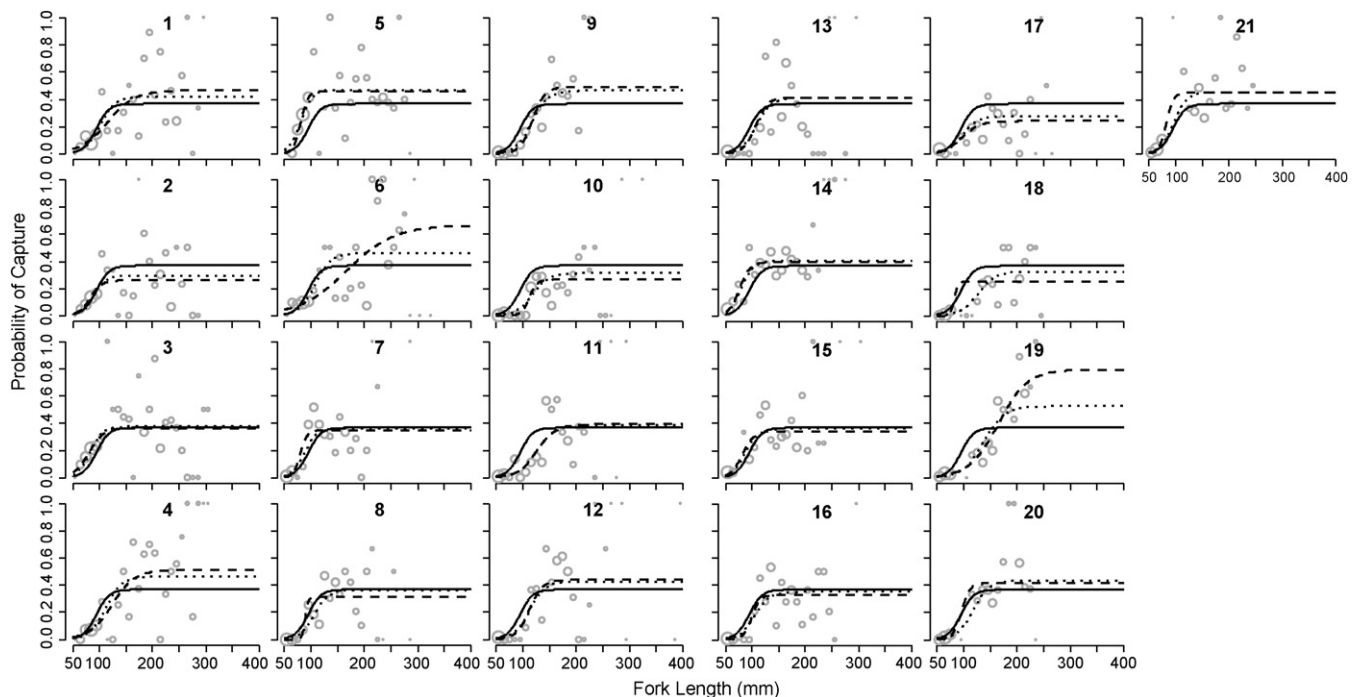


Fig. 3. Three model fits of gillnet efficiency functions to mark-recapture data. Solid line: PBM; long dash: IBM; short dash: HBM. The size of the data points is proportional to the number of fish marked on a log scale. The model are derived using the medians of the posterior distributions for selectivity parameters.

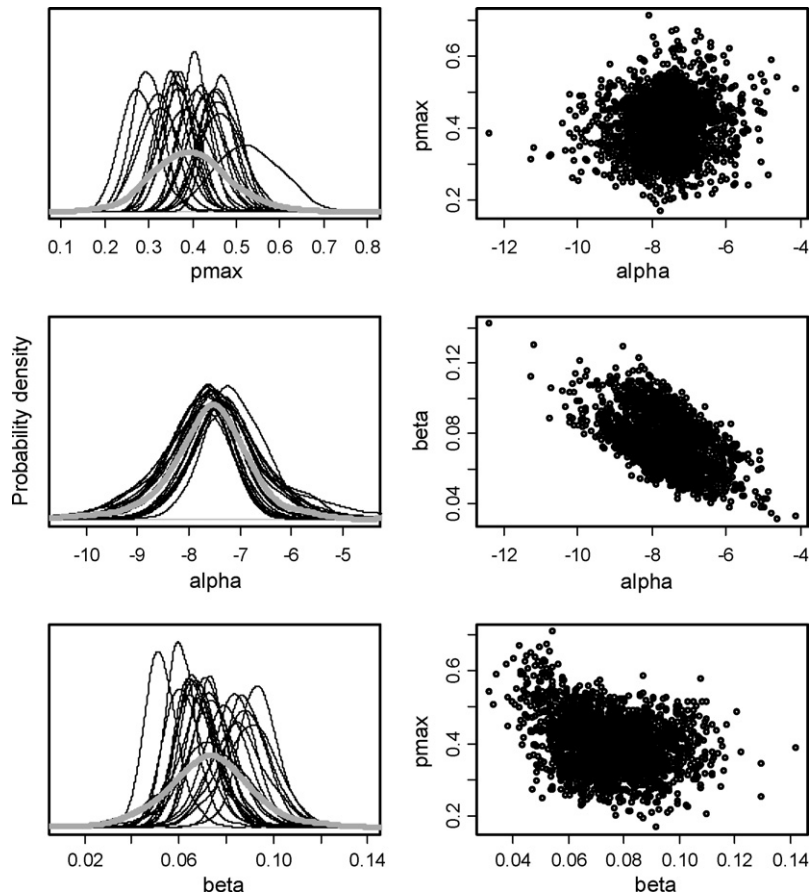


Fig. 4. Left panels: the posterior predictive distributions (thick grey line) and individual lake posterior estimates for each of the three selectivity parameters; p_{max} , α and β when estimated with the HBM. Right panels: correlation plots of the three selectivity parameters. MCMC samples were further reduced to 100 joint draws from each lake-year sample (2100 points per figure) for presentation of scatterplots.

Thus a non-hierarchical approach would indicate that the biologist should estimate abundance with a new selectivity-efficiency function for each lake in each year.

The problem with the individual model (IBM) selected from the non-hierarchical models becomes apparent in the gillnet capture probability plots (Fig. 3). The sparse data for certain size classes in certain years, and low probability of capture for the smallest size classes led to seemingly unrealistic selectivity functions for specific lake-years under an IBM approach (e.g. lakes 8, 18 and 19) (Fig. 3). In these lakes the IBM suggested knife-edge vulnerability, such that a few millimetres in size difference equated to an order of magnitude change in probability of capture or in other cases the IBM predicted unrealistically high p_{max} values, suggesting that approximately 80% of large fish would be removed from a lake in 3 days. Although the pooled model presented a sensible form to the selectivity function, this

function did not fit the data well in some lake-years (e.g. lake-years 5, 10 and 17). The HBM was able to adjust better to the variation between lake-years, and did not produce the unrealistic functions seen in the IBM approach. Thus qualitatively the HBM appeared to produce the best out of sample predictability. The posterior estimates for individual lake-years are shown in Fig. 4, and posterior predictive estimates for the three selectivity parameters are given in Table 4.

In order to compare the model performance more rigorously we sub-sampled the full data set and compared the predictive performance of the three model structures. The relative predictive performance of the three models was dependent on the amount of information available (i.e. the data subset factor). When the data set was reduced the PBM and HBM were able to derive reasonable estimates whereas the IBM produced extremely poor estimates (Fig. 5; top panel). The high MSE for the IBM was

Table 4
Posterior predictive distribution summary statistics for the three selectivity parameters of the hierarchical Bayesian model

| Parameter | Mean | Standard deviation | Median | Quantiles (2.5, 97.5) | Coefficient of variation |
|-----------|--------|--------------------|--------|-----------------------|--------------------------|
| Epmax | 0.399 | 0.0827 | 0.392 | 0.237, 0.566 | 0.207 |
| Ealpha | -7.579 | 0.824 | -7.512 | -9.503, -5.930 | 0.109 |
| Ebeta | 0.0753 | 0.0192 | 0.0753 | 0.0384, 0.112 | 0.255 |

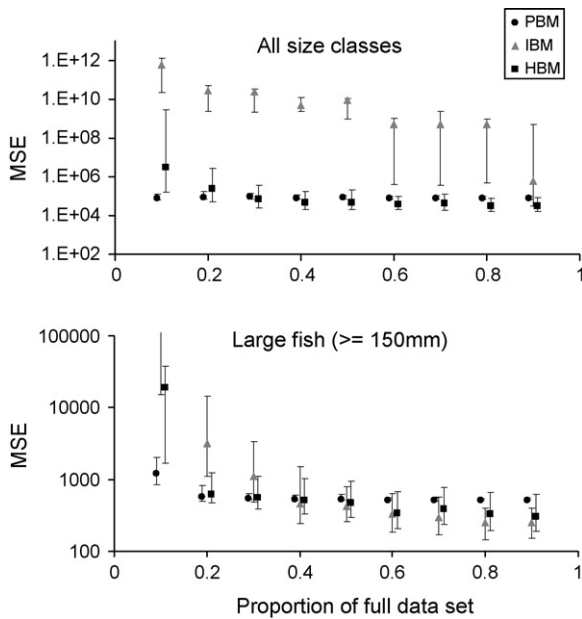


Fig. 5. The mean squared error (MSE) of model abundance estimates (medians with 95% intervals) from the three model structures: PBM, IBM, HBM on reduced data sets. The upper panel shows the MSE for each model structure as a higher proportion of the available data is used to estimate parameters, when all size classes of fish are included in the population. The lower panel shows the trends in MSE if only large fish ≥ 150 mm are considered part of the population. The MSE is on a \log_{10} scale for better visualization, and points have been jiggled to prevent overlap.

due to the low recapture probability for small fish, which led to few or no recaptures in the smallest size bins for some lakes. In these lakes the IBM estimated extremely low capture probabilities, which severely overestimated abundance. If only larger fish are considered (≥ 150 mm) then recapture rates were higher and the IBM was able to produce reasonable estimates with reduced data ($\geq 30\%$ data available). The IBM eventually outperformed the PBM in the large fish only scenario, as it was better able to capture between lake variability (Fig. 5; lower panel).

The subset data illustrated that the hierarchical model is able to maximize the information from data across lakes and elicit realistic estimates of N with very limited data. Overall the HBM was consistently able to better predict N , and thus generally had a lower MSE across strata of sub-sampled data (Fig. 5). The HBM structure was able to efficiently compromise between the PBM and IBM and base parameter estimates on the available data. When data is very limited for individual lakes the HBM parameter are more consistent with the PBM. As more individual level data becomes available, the parameter estimates are pulled towards the IBM estimates (Fig. 6). The analysis reveals the advantage of using a hierarchical approach or similarly the advantage gained in a new study using the joint posterior predictive distribution as an informative prior. Estimation of sampling efficiency in isolation of the information from other lakes can be misleading unless a very high proportion of the population is marked.

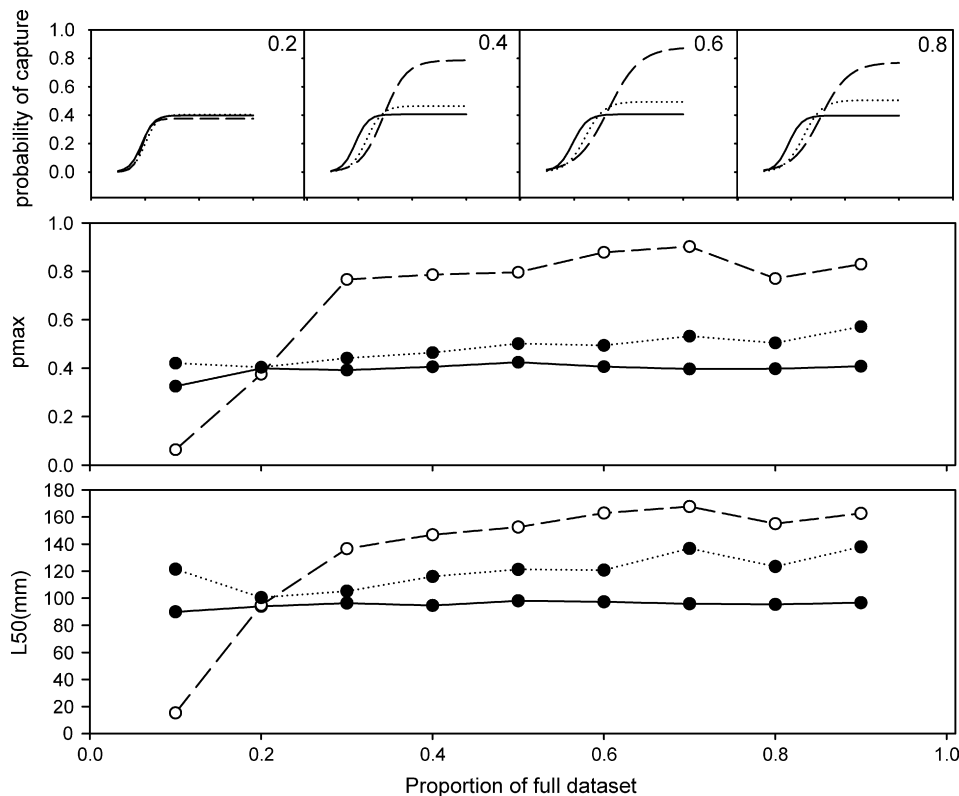


Fig. 6. Upper panel: estimated gillnet selectivity functions for a single lake-year (19) given a data sub-sample equivalent to 20, 40, 60, and 80% of the full data set. The x-axis (fork-length) for each sub-panel is left out to reduce clutter. Lower panels: the posterior medians for p_{max} and $L_{50} (-\alpha/\beta)$ in a single lake-year (19) when the full data set is reduced 10–90% of the full data set. PBM: solid line; IBM: long dash line; HBM: dotted line.

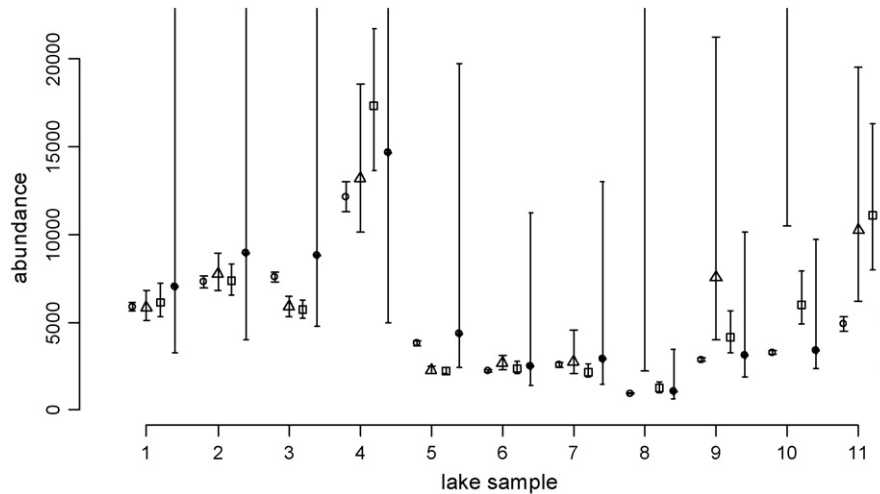


Fig. 7. Posterior median abundance estimates with 80% intervals for all lakes in 1993 and 2003. Open circles: PBM; open triangles: IBM; open squares: HBM, and closed circles are median abundance estimates from the posterior predictive distribution of the HBM (predicted abundance estimate given no new mark-recapture data). The y-axis is truncated at 22,000 for better visualization; the truncated IBM median abundance estimates for lakes 8 and 10 are 39,210 and 51,345, respectively.

Actual fall population estimates for the 1993, 2003 lakes are presented in Fig. 7. The abundance estimates from the posterior predictive parameter distributions had substantial uncertainty, and a long upper tail. The high abundance estimates arose from the potential for very low capture probabilities of small fish, which could not be rejected without data. The uncertainty was drastically improved in the lake-specific HBM estimates, which shows that mark-recapture experiments add substantial information. As in the cross-validation trials, the IBM produced extremely high abundance estimates for lake-samples which had very few recaptures of small fish (lake-years 8 and 10). In the other lakes the IBM produced estimates closer to the HBM, but with higher uncertainty. The PBM model treats all the data as a single sample unit, and thus suggests unrealistically low uncertainty.

We also incorporated the HBM into our DIC analysis, as a further comparison between model structures. The estimated DIC value for the HBM was 1645.7 ($D(\theta) = 1610.5$, $D(\theta) = 1575.3$), with the effective number of parameters estimated at 35.2. Thus the HBM model structure had a lower DIC value, but the difference was trivial at 0.01. This difference is within MCMC error, and p_D error is at least 0.8 (the over estimation of parameters in the IBM), therefore DIC is unable to distinguish the two models. Typically a difference of 5 is necessary to distinguish models with AIC (Burnham and Anderson, 2001; Richards, 2005), and given the extra error associated with p_D , DIC differences should be even greater. Thus we are left with our qualitative observations and quantitative analysis of the subsampled data sets above, which leads us to prefer the HBM model structure.

In order to ensure our inferences were not due to prior specification, we conducted a sensitivity analysis of posteriors to prior specification. We found that posterior inferences were robust to choice of prior (see Appendix A). Further we tested the assumption of exchangeability of lake-years and found it to hold (Appendix A, Fig. A3).

4. Discussion

Fisheries managers who require knowledge of population abundance have two general approaches for estimators: mark-recapture or depletion (CPUE) methods. Mark-recapture requires high sampling effort in order to generate the large sample sizes and high recapture rates necessary to produce accurate estimates. CPUE methods are typically much less demanding, but are associated with a variety of biases (Hilborn and Walters, 1992). Our analysis with the hierarchical Bayesian model suggests that a minimal amount of mark-recapture information can be used to greatly improve the accuracy of a CPUE estimate for individual lake-time strata. The HBM approach provides the means to routinely generate population abundance estimates on multiple lakes that are more precise than simple mark-recapture and more accurate than CPUE estimates.

We have presented an example of a hierarchical Bayesian analysis applied to lakes that were assumed to be exchangeable with regards to gillnet efficiency (for a given fish size). As with any modelling approach, ignorance of covariates can lead to misleading results, and would negate the exchangeability assumption of hierarchical models. However, this does not imply that the hierarchical approach is invalid, but rather that the model structure should be expanded appropriately (Gelman et al., 2004). For example, we could have structured our model to account for fixed effects of some lake characteristics or year effects if appropriate. However, the DIC analysis of the fixed effects model structures indicated that this was unwarranted, as the most parsimonious model considered each lake-year independently. Fisheries managers should be particularly wary of density-dependent catchability, which can lead to hyper-stability of abundance estimates and overharvest of fish stocks (Hilborn and Walters, 1992). Unfortunately, we do not have an independent estimate of fish density near the time of gillnetting to be able to reliably test density dependence in the gillnet efficiency.

The approach we have taken in this manuscript is to compare the structural uncertainty of the models with regards to how to best model lake effects. Conversely, we could have compared the results from distributional assumptions in the likelihood (another form of structural uncertainty). However, this is another topic which is beyond the scope of this manuscript. Further, the expected distribution is binomial and we have shown the distribution to fit the data well in a hierarchical model context.

There are several reasons why the proposed netting protocol and hierarchical analysis has great potential as a research and management tool. First, the application of HBM methods has been limited in the past due to computationally intense nature of estimating the Bayesian posteriors. The advent of accessible MCMC simulation software, such as WinBUGS®, is alleviating this problem with a simple programming code that allows for a specification of a wide variety of Bayesian models. Second the HBM is well suited to stock assessment situations where the stock is divided up amongst many smaller units, such as individual lakes in recreational fisheries. The HBM can derive valid abundance estimates with relatively few tags in each sampling unit. This enables managers to partition resources and focus on a more regional scale as opposed to focusing on reactive management of single, problematic water bodies (Shuter et al., 1998). Finally, the gillnetting method was essentially non-selective over the size range of fish catchable to anglers (≥ 150 mm, Askey et al., in press). This is a convenient property, which facilitates abundance estimation for the catchable population. If only catchable sized individuals are of interest, then the Lundgren nets could be omitted, since these nets select fish < 100 mm (except for a few larger individuals which become tangled). Further, the estimation process for catchables would be reduced to estimation of a single efficiency parameter, which should be equivalent to p_{\max} .

Fisheries applications using DIC as a model selection criterion are at present very rare (we found only one; Michielsens et al., 2006). This is a convenient method for model selection that allows for hierarchical model structures. However, caution should be used in the interpretation of DIC, as it is a fairly recent development, and has several potential sources of error. It is possible to underestimate p_D if posterior distributions are skewed, for example, in our analysis the use of posterior means instead of medians led to large underestimates (by a factor of 10 in the IBM) of p_D . The effective number of parameters will also be reduced if priors are (maybe unintentionally) informative. We widened prior boundaries after finding low p_D estimates for the IBM (see Appendix A). Even if priors are uninformative, different specifications can lead to slightly different DIC values (Spiegelhalter et al., 2002). Thus we recommend that DIC be accompanied by other methods to choose between model structures. Further it seems logical that the difference between DIC values considered sufficient to select models should be greater than AIC differences to accommodate potential error in p_D estimation.

The probability of capture for an individual fish is a combination of two processes: the probability of encountering the fishing gear and the probability of retention given encounter with the fishing gear (Rudstam et al., 1984). It is becoming increasingly

apparent that variation in the first component leads to variation in size-selectivity and capture probability despite a constant fishing effort (Fryer, 1991; Fryer et al., 2003; Trenkel and Skaug, 2005). Thus it is necessary to model the parameters of the size-selective gear efficiency function as stochastic elements that vary about a set of mean function parameters. We have found hierarchical Bayesian analysis to be an efficient approach to this common situation, which appears under-utilized in the literature. Harley and Myers (2001) used the HBM approach in their analysis of research trawl surveys. Unfortunately, they had to fix the shape parameters and only allowed the p_{\max} parameter to vary, which they attributed to limited data. Yet, we found that with mark-recapture data the HBM approach worked well even when data subsets were reduced to very low numbers. This is encouraging as managers may be limited in the resources that can be allotted to tagging individuals.

Appendix A. Sensitivity to prior specification

The priors we used in the Section 3 are presented in Table A1. To investigate the sensitivity of posterior inference to prior specification we compared results with several different priors on the parameters (α, β, p_{\max}) of the size-selectivity function. We present results from three weakly informative priors on the shape of the selectivity function, which is summarized as p_{\max} and L_{50} ($-\alpha/\beta$). The first two priors use the beta distribution, which is a flexible distribution that is commonly used in the fisheries literature (Liermann and Hilborn, 1997; Michielsens and McAllister, 2004). The Beta(1,1) is a uniform distribution over the 0–1 interval, whereas, the Beta(1,2) is skewed to the right. In both cases the distribution were rescaled when necessary to the bounds in Table A1. The third prior was a uniform distribution on the log-scale, also rescaled to match Table A1 on the original scale. Since p_{\max} occurs on the 0–1 interval the prior was on a logit scale (Table A1).

None of the PBM parameters were sensitive to the choice of prior (results not shown). Similarly, the IBM predictions were generally insensitive to prior choice (Fig. A1). Although lake-year 8 had variation in the α and β estimates as well as high

Table A1
Default priors used for each model structure presented in Section 3

| Model | Parameter | Prior | Bounds |
|----------|-------------------|--|-------------------|
| PBM, IBM | alpha | $N(0,1000)$ | (−300,1) |
| PBM, IBM | beta | $N(0,1000)$ | (−0.01,2) |
| PBM, IBM | p_{\max} | $B(1,1)$ | NA |
| HBM | alpha | $N(\mu_{\alpha}, 1/\tau_{\alpha})$ | NA |
| HBM | beta | $N(\mu_{\beta}, 1/\tau_{\beta})$ | NA |
| HBM | p_{\max} | $N(\mu_{p_{\max}}, 1/\tau_{p_{\max}})$ | Note: logit scale |
| HBM | μ_{α} | $B(1,1)$ | Scaled(−300,1) |
| HBM | τ_{α} | $G(0.001, 0.001)$ | NA |
| HBM | μ_{β} | $B(1,1)$ | Scaled(−0.1,2) |
| HBM | τ_{β} | $G(0.001, 0.001)$ | NA |
| HBM | $\mu_{p_{\max}}$ | $N(0, 0.67)$ | Note: logit scale |
| HBM | $\tau_{p_{\max}}$ | $G(0.001, 0.001)$ | NA |

N indicates a normal distribution described by parameters (mean, variance); *B*: beta distribution with parameters (alpha, beta); *G*: gamma distribution with parameters (shape 1, shape 2).

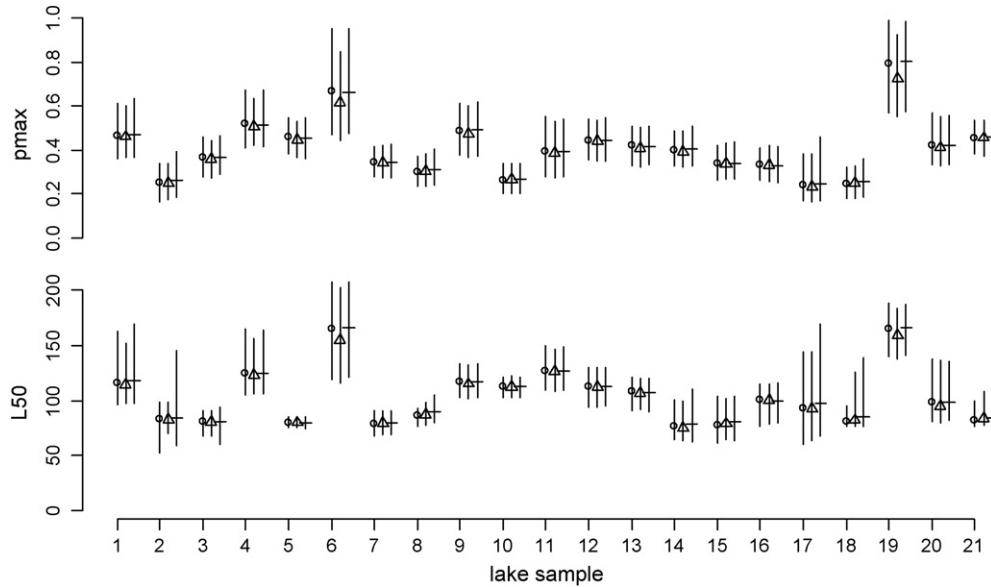


Fig. A1. Posterior medians and 95% intervals for p_{max} (probability of capture for large, fully vulnerable fish) and L_{50} (length at 50% of full vulnerability) for the IBM under three different prior specifications applied simultaneously to all three selectivity parameters. Circles: Beta(1,1); triangles: Beta(1,2); crosses: uniform on log scale.

uncertainty. However, the variation in α and β between prior results in a non-relevant change to the gillnet efficiency function. The p_{max} parameter was insensitive to prior specification for this lake, and all three combinations of α and β result in a similar L_{50} , as shown in Fig. A1. Thus all three priors result in a lake 8 function with an extremely abrupt, knife-edge selectivity that climbs at the same point and saturates at the same level (see Fig. 3).

In the case of the HBM it is possible that the specification of the hyperpriors may influence posterior inference. The variance of the hyperprior may be important as it influences the amount

of shrinkage for individual lakes. There are several potential priors for modelling the precision parameters of the hyperpriors. Typically BUGS models use the gamma distribution to model the precision parameter of a normal distribution (Spiegelhalter et al., 2003). Other choices are a uniform for precision on a log scale (Rivot and Prevost, 2002), or a uniform distribution on the standard deviation (Gelman et al., 2004). However, we found that the individual lake estimates were robust to prior specification of the precision (Fig. A2), as well as for means (results not shown). It is also worth noting the decrease in uncertainty for the HBM parameter estimates as opposed to the IBM estimates in Fig. A2.

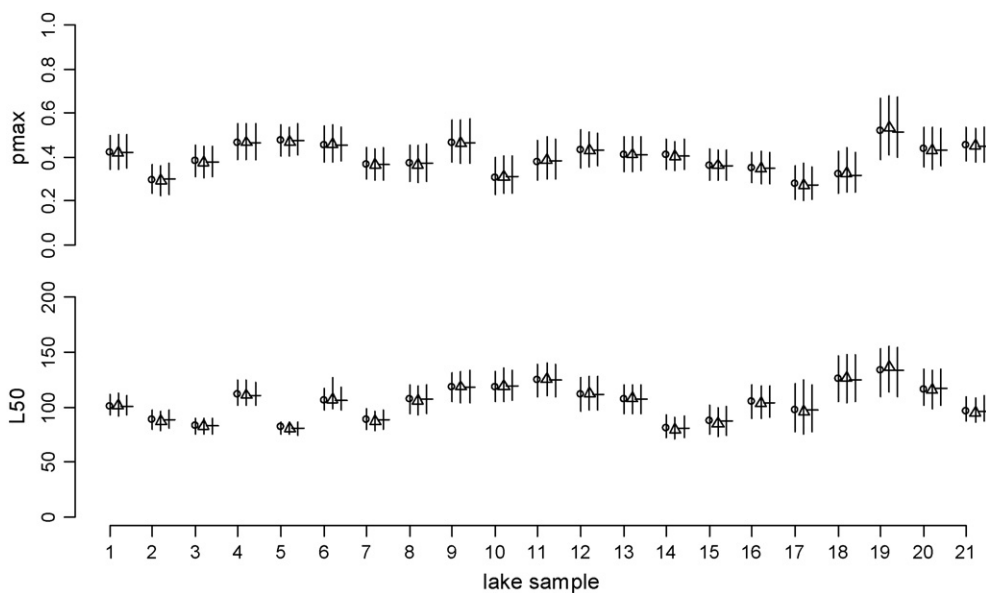


Fig. A2. Posterior medians and 95% intervals for p_{max} (probability of capture for large, fully vulnerable fish) and L_{50} (length at 50% of full vulnerability) for the HBM under three different prior specifications for hyper-variance. Circles: uniform (0,100) on standard deviation; triangles: Gamma(0.001,0.001) on precision; crosses: Log(uniform(-9.2,9.2)) on variance.

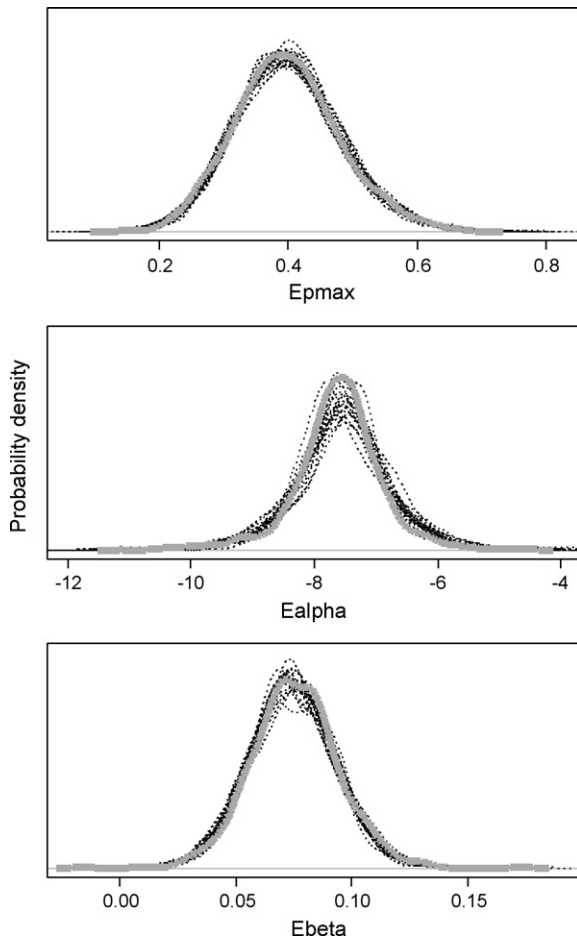


Fig. A3. Posterior predictive estimates of the three gillnet capture probability model parameters when estimated repeatedly with 1 lake-year removed from the data to assess exchangeability (thin dotted lines). Thick dotted lines are the posterior predictive distributions when fit to the full data set.

Lastly posterior predictive estimates of the three logistic parameters were estimated repeatedly with 1 lake-year removed from the data to assess exchangeability. The posterior predictive distributions were not sensitive to these changes (Fig. A3).

References

- Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Petrov, B.N., Csáki, F. (Eds.), Proceedings of the Second International Symposium on Information Theory. Akadémiai Kiadó, Budapest, pp. 267–281.
- Askey, P.J., Richards, S.A., Post, J.R., Parkinson, E.A., in press. Linking angling quality and learning under catch and release regulations. *North Am. J. Fisher. Manage.*
- Booth, A.J., Potts, W.M., 2006. Estimating gill-net selectivity for *Labeo umbratus* (Pisces: Cyprinidae), and an evaluation of using fyke-nets as a non-destructive sampling gear in small reservoirs. *Fish. Res.* 79, 202–209.
- Bromaghin, J.F., 2005. A versatile net selectivity model, with application to Pacific salmon and freshwater species of the Yukon River. *Alaska Fish. Res.* 74, 157–168.

- Burnham, K.P., Anderson, D.R., 2001. Model Selection and Inference: A Practical Information-Theoretic Approach. Springer, New York, 353 pp.
- Fryer, R.J., 1991. A model of between-haul variation in selectivity. *ICES J. Mar. Sci.* 48, 281–290.
- Fryer, R.J., Zuur, A.F., Graham, N., 2003. Using mixed models to combine smooth size-selection and catch-comparison curves over hauls. *Can. J. Fish. Aquat. Sci.* 60, 448–459.
- Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2004. Bayesian Data Analysis. Chapman and Hall/CRC, Boca Raton, FL.
- Hamley, J.M., 1975. Review of gillnet selectivity. *J. Fish. Res. Board Can.* 32, 1943–1969.
- Harley, S.J., Myers, R.A., 2001. Hierarchical Bayesian models of length-specific catchability of research trawl surveys. *Can. J. Fish. Aquat. Sci.* 58, 1569–1584.
- Hilborn, R., Walters, C.J., 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty. Chapman and Hall, New York, 570 pp.
- Ingólfsson, Ó.A., Jørgensen, T., 2006. Escapement of gadoid fish beneath a commercial bottom trawl: relevance to the overall trawl selectivity. *Fish. Res.* 79, 303–312.
- Liermann, M., Hilborn, R., 1997. Depensation in fish stocks: a hierarchic Bayesian meta-analysis. *Can. J. Fish. Aquat. Sci.* 54, 1976–1984.
- Michielsens, C., McAllister, M., 2004. A Bayesian hierarchical analysis of stock-recruit data: quantifying structural and parameter uncertainties. *Can. J. Fish. Aquat. Sci.* 61, 1032–1047.
- Michielsens, C.G.J., McAllister, M.K., Kuikka, S., Pakarinen, T., Karlsson, L., Romakkaniemi, A., Pera, I., Mantyniemi, S., 2006. A Bayesian state-space mark-recapture model to estimate exploitation rates in mixed-stock fisheries. *Can. J. Fish. Aquat. Sci.* 63, 321–334.
- Millar, R.B., 1995. The functional form of hook and gillnet selection curves cannot be determined from comparative catch data alone. *Can. J. Fish. Aquat. Sci.* 52, 883–891.
- Millar, R.B., Fryer, J.F., 1999. Estimating the size-selection curves of towed gears, traps, nets and hooks. *Rev. Fish. Biol. Fisher.* 9, 89–116.
- Post, J.R., Parkinson, E.A., Johnston, N.T., 1999. Density-dependent processes in structured fish populations: interaction strengths in whole-lake experiments. *Ecol. Monogr.* 69, 155–175.
- Richards, S.A., 2005. Testing ecological theory using the information-theoretic approach: examples and cautionary results. *Ecology* 86, 2805–2814.
- Rivot, E., Prevost, E., 2002. Hierarchical Bayesian analysis of capture-mark-recapture data. *Can. J. Fish. Aquat. Sci.* 59, 1768–1784.
- Rosenberger, A.E., Dunham, J.B., 2005. Validation of abundance estimates from mark-recapture and removal techniques for rainbow trout captured by electrofishing in small streams. *North Am. J. Fisher. Manage.* 25, 1395–1410.
- Rudstam, L.G., Magnuson, J.L., Tonn, W.M., 1984. Size selectivity of passive fishing gear: a correction for encounter probability applied to gill nets. *Can. J. Fish. Aquat. Sci.* 41, 1252–1255.
- Shuter, B.J., Jones, M.L., Korver, R.M., Lester, N.P., 1998. A general, life history based model for regional management of fish stocks: the inland lake trout (*Salvelinus namaycush*) fisheries of Ontario. *Can. J. Fish. Aquat. Sci.* 55, 2161–2177.
- Spiegelhalter, D.J., Thomas, A., Best, N.G., 1996. Computation on Graphical Bayesian Models. Oxford University Press, pp. 407–425.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.R., van der Linde, A., 2002. Bayesian measures of model complexity and fit. *J. Roy. Stat. Soc. B* 64, 583–616.
- Spiegelhalter, D.J., Thomas, A., Best, N.G., Lunn, D., 2003. WinBUGS Version 1.4 User Manual. Medical Research Council Biostatistics Unit, Cambridge, available at <http://www.mrc-bsu.cam.ac.uk/bugs>.
- Stone, M., 1974. Cross-validatory choice and assessment of statistical predictions with discussion. *J. Roy. Stat. Soc. B* 36, 111–147.
- Sturtz, S., Ligges, U., Gelman, A., 2005. R2WinBUGS: a package for running WinBUGS from R. *J. Stat. Soft.* 12, 1–16.
- Trenkel, V.M., Skaug, H.J., 2005. Disentangling the effects of capture efficiency and population abundance on catch data using random effects models. *ICES J. Mar. Sci.* 62, 1543–1555.