Ministry for Primary Industries
Manatū Ahu Matua

Comparison of bycatch estimation for fish species using a ratio estimator and model-based method

New Zealand Aquatic Environment and Biodiversity Report No. 154

C. T. T. Edwards, I. Doonan and O. F. Anderson

ISSN 1179-6480 (online)
ISBN 978-0-908334-40-7 (online)

June 2015

New Zealand Government

Growing and Protecting New Zealand

Requests for further copies should be directed to:

Publications Logistics Officer
Ministry for Primary Industries
PO Box 2526
WELLINGTON 6140

Email: brand@mpi.govt.nz
Telephone: 0800 00 83 33
Facsimile: 04-894 0300

This publication is also available on the Ministry for Primary Industries websites at:
http://www.mpi.govt.nz/news-resources/publications.aspx
and http://fs.fish.govt.nz (go to Document library/Research reports)

This report was produced using a LaTeX template developed
by NIWA Ltd., which is available from the authors on request.

Contents

EXECUTIVE SUMMARY

Edwards, C.T.T.; Doonan, I.; Anderson, O.F. (2015) Comparison of bycatch estimation for fish species using a ratio estimator and model-based method.

New Zealand Aquatic Environment and Biodiversity Report No. 154. 30 p.

Commercial fisheries in New Zealand often catch species that are not necessarily targeted by the fishing operation (bycatch). Since this bycatch is not fully recorded it is necessary to estimate the absolute quantity using observer data, which is collected from a subset of vessels every year. The observer data can then be used to extrapolate unobserved catches across the remainder of the fishing fleet. This extrapolation can be based on a simple ratio-based estimate of the catch rate. However this procedure is prone to bias and is shown here to produce results with a spuriously high level of accuracy. In this study we compare estimates from a ratio-based approach with a model-based estimator using data from the ling longline fishery, and scampi and squid trawl fisheries. A simulation study, parameterised using the empirical data, was first used to examine performance of the estimators; which were subsequently applied to the empirical data itself. The simulation results illustrated that a model-based approach is preferable, and reliability of the empirical results are discussed in this context. Although not intended to provide accurate bycatch estimates, the results of this study are used to argue that a model-based estimator should be the preferred approach for future work of this kind.

Project code: ENV2013-01

1. Introduction

Two approaches have been used to estimate total catches for bycatch species in New Zealand's commercial fisheries. For fin-fish bycatch, ratio-based methods have typically been used (e.g. Anderson, 2013, Griggs & Baird, 2013). The data consist of an assumed known total effort for the fishery, and catch per effort unit for a subset of observed fishing events occuring on a subset of vessels fishing in that year. These data can often be stratified in a manner intended to minimise variance of the ratio estimate within each strata, depending on the distribution of observations across vessels and the spatial locations of fishing.

If we consider the random variables y_i and $t_i^{[S]}$, which refer to the observed bycatch and associated effort value respectively, for a given fishing event i, then the ratio r is calculated as the *ratio of averages*:

$$r = \frac{\sum y_i}{\sum t_i^{[S]}}$$

which is equivalent to the *average ratio* when $t_i^{[S]} = 1$, but in other situations is generally preferred over this alternative, since it has lower bias and an approximate normal distribution at large sample sizes (Barnett, 1991). For a particular strata, the total bycatch N is then estimated assuming a constant r value across unobserved effort. Denoting the ratio-based estimate of the bycatch as $N^{[R]}$, and the total effort for each fishing event (observed and unobserved) as $t_i^{[T]}$, then this can be written as::

$$N^{[R]} = \sum y_i + \left(\sum t_i^{[T]} - \sum t_i^{[S]}\right).r = \sum t_i^{[T]}.r$$

A variance around this estimate can be obtained either analytically or more usually via a bootstrap procedure. Due to data availability, stratification usually takes place at the level of the area, meaning that the ratio and bycatch are estimated for each FMA and then summed across the fishery. Although this method is easy to apply, it has a number of drawbacks that will be further described and examined in this study. For particularly important species therefore, a more intensive model-based approach has been adopted, for example when estimating the total catches of marine mammals (Abraham, 2008, Smith & Baird, 2007a,b, Thompson et al., 2013) or birds (Abraham & Thompson, 2011).

For model-based methods a statistical model is used to estimate the expected catch for unobserved fishing events. Covariate data associated with a particular observation are denoted by the design matrix $\mathbf{x}$, and consist of spatial and temporal information as well as other covariates such as depth of fishing or type of fishing gear. For this information to be useful, it must be assumed to exist for unobserved fishing events also. In general we can write the expected observation as a function of the design matrix and vector of coefficients (β) associated with each covariate. This function is denoted $g(.)$, and is known as the *inverse link function*. The effort term t_i is treated as an offset and therefore not included in the function arguments. Using this notation we can write our estimate of the total bycatch as:

$$\mathbb{E}[y_i] = g(\mathbf{x}_i'\beta).t_i^{[S]}$$

$$N^{[M]} = \sum g(\mathbf{x}_i'\beta).t_i^{[T]}$$

with the summation taken across all fishing effort to give the model-based estimate of the bycatch $N^{[M]}$. Note that the recorded bycatch (and associated effort $t_i^{[S]}$) is used only to parameterise the model, and not in the final summation when estimating the total $N^{[M]}$. This is because, for a given fishing event, only a fraction of the actual gear deployments are ever actually observed. The bycatch associated with a particular effort $t_i^{[T]}$ is therefore usually unknown and must be estimated, even if that fishing event was observed.

2. Modelling approaches for over-dispersed data

Catch data can be represented as a Poisson process when it is assumed that catches occur randomly in space or time at a constant rate. The Poisson distribution has a single rate parameter λ, which is equal to both the expectation (mean) and variance. For a given fishing event i, consisting of a single effort unit, the distribution of observed catch values can therefore be written as:

$$y_i \sim Poisson(\lambda).$$

This distribution can be used to model catch data using a log-link function: $\mathbb{E}[y_i] = \exp(\mathbf{x}_i'\beta)$, where $\mathbf{x}$ is the design matrix and β is the vector of regression coefficients. However empirical data rarely conform to a Poisson distribution because the variance in the response exceeds the expectation. If the variance $\mathbb{V}[y_i] > \lambda$, the data are *over-dispersed* and this model is no longer appropriate. Two alternative models are capable of representing over-dispersed catch data, differentiated by the assumed distribution behind the over-dispersion process.

2.1. The Poisson-Gamma model

The simplest way to model over-dispersed data is to estimate an additional term that scales the mean of the distribution to the variance, so that $\mathbb{V}[y_i] = \lambda.\phi$, with the expectation unchanged. This can be extended by assuming that the multiplicative scalar is a *random effect* designed to account for unobserved heterogeneity in the data, which is equivalent to the assumption that each observation is drawn from an independent Poisson distribution. In other words, the expected value for a particular data record i is unique: $\mathbb{E}[y_i|\lambda_i] = \lambda_i$; with the expectation referring specifically to a Poisson distribution with rate λ_i. If λ_i is assumed to follow a gamma distribution across fishing events with a shared shape parameter ζ and an expected value of μ_i, then this can be written as:

$$y_i|\lambda_i \sim Poisson(\lambda_i)$$
$$\lambda_i|\zeta,\mu_i \sim Gamma(\zeta,\zeta/\mu_i).$$

Similarly we could write $\lambda_i = \mu_i.\varepsilon_i$, in which case ε_i would be considered a random effect with distribution $\varepsilon_i \sim Gamma(\zeta,\zeta)$ and expected value of one. The expected value of λ_i is denoted $\mathbb{E}[\lambda_i|\mu_i] = \mu_i = \exp(\mathbf{x}_i'\beta)$, and is an expectation across the over-dispersion process. Since $\mathbb{E}[y_i|\lambda_i] = \exp(\mathbf{x}_i'\beta).\varepsilon_i$, using the law of total expectation we obtain: $\mathbb{E}[y_i] = \exp(\mathbf{x}_i'\beta)\mathbb{E}[\varepsilon_i] = \exp(\mathbf{x}_i'\beta)$. Overall, the unconditional distribution of y_i follows a Negative Binomial distribution with parameters (μ_i,ζ):

$$p(y_i) = \frac{\Gamma(\zeta+y_i)}{\Gamma(\zeta).y_i!}\left(\frac{\mu_i}{\zeta+\mu_i}\right)^{y_i}\left(\frac{\zeta}{\zeta+\mu_i}\right)^{\zeta}$$

which has an expectation $\mathbb{E}[y_i] = \mu_i$ and variance $\mathbb{V}[y_i] = \mu_i + \mu_i^2/\zeta$.

2.2. The Poisson-LN model

An alternative approach is to assume that λ_i follows a log-normal (LN) distribution:

$$y_i|\lambda_i \sim Poisson(\lambda_i)$$
$$\lambda_i|\mu_i,\sigma \sim LN(\mu_i,\sigma^2)$$

again with $\mu_i = \exp(\mathbf{x}_i'\beta)$. Following a similar argument to that given above, the over-dispersion term ε_i is now log-normal and it is convenient to write $\mathbb{E}[y_i|\lambda_i] = \exp(\mathbf{x}_i'\beta + e_i)$, where $e_i \sim Normal(0, \sigma^2)$. There is no analytical form for the unconditional distribution of y_i, but it is useful nevertheless to report the expectation: $\mathbb{E}[y_i] = \exp(\mathbf{x}_i'\beta + \sigma^2/2)$.

2.3. Including random effects

Fixed and random effects within a statistical model are interpreted differently, and it is on this basis that they can be distinguished. For fixed effects, the estimated coefficient is of direct interest, whereas for random effects we are concerned primarily with the distribution of these coefficients. During inference we often need to incorporate the variation associated with unsampled components of the population, and this variation is captured by a random effects component to the model. Because of this distinction fixed and random effects are usually noted separately in the model definition:

$$\mu_i = \exp(\mathbf{x}_i'\beta + \mathbf{z}_i'\gamma)$$

where $\mathbf{z}$ is the random effects design matrix and γ is the vector of random effect coefficients. A model of this type, that includes both fixed and random effects, is known as a general linear mixed model (GLMM).

3. Model-based estimation and prediction

In this study we assumed that the catch data follow a Poisson-LN probability:

$$y_i|\lambda_i \sim Poisson(\lambda_i)$$
$$\lambda_i|\mu_i, \sigma \sim LN(\mu_i, \sigma^2)$$

with $\mu_i = \exp(\mathbf{x}_i'\beta + \ln(t_i^{[S]}))$. Because the over-dispersion term is log-normal we can write $\mathbb{E}[y_i|\lambda_i] = \exp(\mathbf{x}_i'\beta + \ln(t_i^{[S]}) + e_i)$, where $e_i \sim Normal(0, \sigma^2)$. The unconditional expectation is therefore: $\mathbb{E}[y_i] = \exp(\mathbf{x}_i'\beta + \ln(t_i^{[S]}) + \sigma^2/2)$.

Including random effects the full log-linear predictor is written as:

$$\ln(\mu_i) = \mathbf{x}_i'\beta + \mathbf{z}_i'\gamma + \ln(t_i^{[S]})$$

with $\ln(t_i^{[S]})$ treated as an offset term. Parameterisation then involves the estimation of β and γ, as well as the variance terms $\sigma^2_{[\gamma]}$ and σ^2 that describe the (log-normal) distributions of γ and e_i respectively.

3.1. Available estimation procedures

The strength and difficulty inherent to any GLMM is the need to distinguish between variance of the random effect terms and the over-dispersion. This is still an active area of statistical research and a variety of approaches exist. The R-packages `lmer4` and `hglm` can be used to fit Poisson-LN and Poisson-Gamma models, using maximum likelihood. However because of the approximations otherwise required to quantify the error, a Bayesian approach is preferable. A negative binomial GLMM can be fitted using Bayesian methods with `glmmADMB` or coded directly, usually as a

Poisson-Gamma process, in `bugs` or `stan`. Alternatively, a Poisson-LN can be fitted by MCMC using `MCMCglmm`.

In practice, the Poisson-Gamma and Poisson-LN distributions are similarly capable of modelling over-dispersed data, and although previous work in New Zealand has tended to assume a Poisson-Gamma process, coded in `winbugs`, for this study the Poisson-LN was used instead. This choice was based largely on availability of the R-package `MCMCglmm`, which proved to be much faster than an equivalent model coded in either `winbugs` or `rstan`. Validation of the package was performed by comparing fitted estimates from simulated data with those obtained using `lme4::glmer`, and both were found to produce equivalent results.

The R-package `MCMCglmm` executes an efficient Bayesian estimation procedure for Poisson-LN models. A Bayesian approach is desirable since it provides a better representation of the uncertainty during prediction, without the need for asymptotic approximations of the type required for a maximum-likelihood based approach.

3.2. Posterior prediction

Following a Bayesian fit to the data the parameterised model can then be used to predict the bycatch occurring at unobserved fishing events. In this context posterior predictive simulation can be used to produce a distribution of estimated bycatch values that represent uncertainty around the true bycatch value. This is most easily described by referring back to the original model formulation, with a subscript i referring to a data record (fishing event) and an additional subscript p used to denote a particular sample from the posterior:

$$\tilde{y}_{ip} \sim Poisson(\lambda_{ip})$$
$$\lambda_{ip} \sim LN(\mu_{ip}, \sigma_p^2)$$

where $\tilde{y}_i$ represents a simulated observation generated by: i) sampling from the posterior distributions of $\hat{\beta}$, $\hat{\gamma}$, $\sigma^2_{[\gamma]}$ and $\hat{\sigma}^2$; ii) calculating a value for μ_{ip} using the log-linear predictor; iii) randomly generating a value for λ_{ip} from the log-normal distribution parameterised by μ_{ip} and σ_p^2; and, iv) randomly generating $\tilde{y}_{ip}$ from a Poisson distribution with mean λ_{ip}. This procedure captures uncertainty in both the estimation process, the random effect and the random nature of over-dispersion. However posterior prediction was not performed in this study, since our main intention was to compare estimators. When calculating confidence limits we therefore used the percentiles of the posterior probability (the credibility intervals), which can be compared more directly to the confidence intervals produced by the ratio-based estimator.

4. Data and methods

Three fisheries were selected for analysis in this study, namely the ling longline (Anderson, 2014), scampi trawl (Anderson, 2012) and arrow squid trawl (Anderson, 2013) fisheries. The analysis for each fishery consisted of two parts. First, we simulated bycatch and observational data from a Poisson-LN process assuming a distribution of fishing effort and observer sampling across time, space and vessels that matched the empirical data. To this data we applied both ratio and model-based estimators and compared their predictive performance. Second, we applied both methods to the

Table 1: Parameter specifications for simulation of catch rate data. Catch rates shown are an average for the specified area across all empirical data, in kilograms per unit of fishing effort. These values were used to create the initial catch rate vector $\mathbf{I}_0$.

	Ling		Squid		Scampi
Area	Catch rate (kg/set)	Area	Catch rate (kg/trawl)	Area	Catch rate (kg/trawl)
BNTY	321.57	AUCK	413.96	AUCK	74.35
CAMP	132.15	CHAL	594.50	BANK	279.48
COOK	240.84	CHAT	820.69	CHAT	466.25
LIN1	54.49	NRTH	442.79	NRTH	150.38
LIN2	572.34	PUYS	589.83	PUYS	396.78
LIN3	139.25	WAIR	435.03	SNAR	54.62
LIN4	83.44	WCSI	194.00	SUBA	76.96
LIN7	78.31				
PUYS	151.10				
Vessel random effect ($\sigma^2_{[\gamma]}$)					
Ling	0.78	Squid	0.20	Scampi	1.16

empirical data itself. Comparison of these empirical results could then be informed by the simulation study.

In the ling, scampi and squid fisheries being examined, bycatch estimates typically consider groups of species, classified as those within the Quota Management System (QMS), those that are not, and invertebrates. For this study we consider only a single group, namely the non-QMS species. This group is the primary focus of Anderson (2012, 2013, 2014) and was considered sufficient for current purposes, since our intention was to compare alternative estimators, rather than provide actual estimates of the bycatch for these fisheries. Effort was measured as the number of sets (for ling), or the number of trawls (for scampi and squid). Therefore t_i refers to the number of sets/trawls for data record i. For ling, Anderson (2014) actually used the number of hooks rather than the number of sets, and so we expect our ratio-based estimates of the bycatch to be different.

4.1. Data simulation

The simulation of bycatch and associated observational data required a representation of the underlying catch rate dynamics over time and space. For each fishery the initial catch rate vector $\mathbf{I}_0$, with length equal to the number of areas (as defined by Anderson, 2012, 2013, 2014), was set equal to the average catch rate for that area across all the empirical data (Table 1). The catch rate dynamics by year t then proceeded according to a simple random walk process:

$$\mathbf{I}_{(t+1)} = \mathbf{T}\mathbf{I}_t$$

with $\mathbf{T}$ representing a diagonal matrix with elements sampled from a multi-variate log-normal distribution with arbitrarily chosen parameters $\sigma = 0.1$ and $\rho = 0.8$ (Code Listing A1). This procedure was used to generate correlated dynamics over time and space.

Given the average catch rate vector for a particular year $\mathbf{I}_t$, a Poisson-LN process was then used to simulate the bycatch assuming a vessel random effect (Code Listing A2). The vessel random effect variance was calculated from the empirical distribution of mean catch rates across all sampled vessels (Table 1). The bycatch data were then sampled at random according to the distribution of observer sampling effort. No attempt was made to simulate the effect of other covariates such as depth or season, which are known to influence the bycatch rate (e.g., Anderson, 2013). However the vessel random effect can be considered as an integration over these other covariate effects, without them being explicitly included. Although the model-based method could have made use of additional covariate information, it was considered sufficient for the model to estimate the vessel random effects directly. This greatly simplified the experimental design and analysis. We nevertheless note that covariate data could improve the performance of model based methods, although this will not always be the case, being dependent on the realised quality of covariate data available.

In summary, our simulation procedure was able to produce correlated bycatch and observational data, in a manner that was consistent with the actual relative distribution of observational effort across the fisheries. It is the relative distribution of fishing and sampling effort that is likely to determine the bias of ratio-based estimators (see Appendix I), and because the data will be consistent across iterations, this simulation design should allow bias to be detected from the residual difference between simulated and estimated bycatch values. Two hundred simulations were performed, with the same simulated data used to compare the two estimation methods.

4.2. Estimation procedures

For ratio-based estimation, the ratio r was calculated directly from the empirical data for each year j and area k, giving the estimated bycatch as:

$$\hat{N}_j^{[R]} = \sum_k r_{jk}.t_{jk}^{[T]}$$

Confidence intervals were obtained via a similar procedure to that described by Anderson (2012, 2013, 2014). This calculation involved a nested bootstrap by vessel, meaning that for each combination of j and k, vessels were sampled, and then observations within each vessel, yielding r_{jkb} and $N_{jb}^{[R]}$, where b refers to a particular bootstrap iteration. Confidence intervals were obtained as:

$$\left(\hat{N}_j^{[R]} - \overline{\delta}, \hat{N}_j^{[R]} - \underline{\delta}\right)$$

where $\overline{\delta}$ and $\underline{\delta}$ are the upper and lower 95th percentiles respectively of the distribution of $N_{jb}^{[R]} - \hat{N}_j^{[R]}$ (Rice, 1995). This procedure differs slightly from that executed by Anderson (2012, 2013, 2014), but the results produced were shown to be almost identical.

In many cases there were insufficient data to calculate r_{jk} directly, which was considered the case if there were less than 25 data records. As an alternative, the ratio calculated using all the data for that year (i.e. including all areas) was substituted for r_{jk}, provided there were more than 50 records available, otherwise the ratio was calculated using all the data for that area (i.e. including all years).

The model-based estimate of bycatch was calculated by first fitting a log-linear model to the observed data using `MCMCglmm` (Code Listing A3):

$$\ln(y_i) = \mathbf{x}_i'\beta + \mathbf{z}_i'\gamma + \ln(t_i^{[S]}) + e_i$$

with area and time fixed effects β, random effects $\gamma \sim Normal(0, \sigma_{[\gamma]}^2)$ and over-dispersion $e \sim Normal(0, \sigma^2)$. The bycatch N was calculated as the marginal expectation (Code Listing A4):

$$\mathbb{E}[N_i] = \exp\left(\mathbf{x}_i'\beta + \ln(t_i^{[T]}) + \sigma_{[\gamma]}^2/2 + \sigma^2/2\right)$$

noting that in this case the design matrix **x** refers to the complete commercial data, rather than the observer data only. For each fishing year the total bycatch was then calculated as:

$$N_{jp} = \sum_i N_{ip}$$

where the summation is across data records in year j, and the subscript p refers to a particular sample from the posterior distribution. This yields a posterior distribution of N_j values for each year, from which we obtained $\hat{N}_j^{[M]}$ as the posterior median, and credibility intervals as the 95th percentiles.

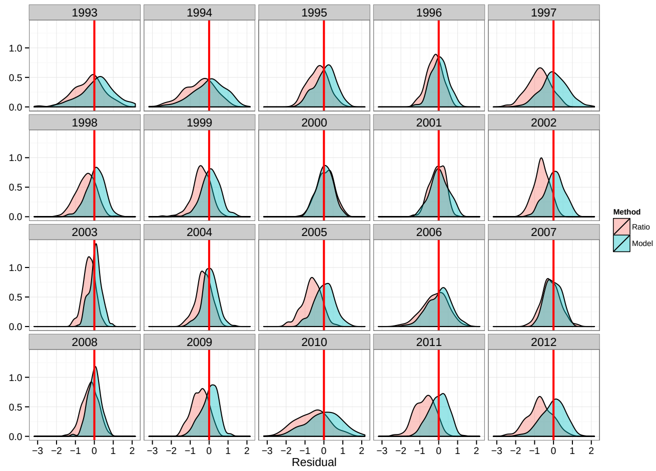
5. Simulation Results

The simulation results are summarised by the residual distributions in Figure 1. The most notable feature of these results is that the ratio-based estimate is often biased, whereas the central tendency of the model-based estimates is close to the simulated value. The simulation study was designed so that the simulated data share important features of the empirical data, namely a consistent area-specific and realistic effect size across iterations, a realistic vessel random effect variance, and a relative observer sampling distribution that matched the empirical data. The simulation results demonstrate that under these conditions the ratio-based estimates have low accuracy. This can be most clearly seen for squid (Figure 1c), for which the ratio-based results are often underestimates. A similar feature is observed for ling (Figure 1a), but is less obvious for scampi (Figure 1b), particularly in recent years. In general, the model-based estimates appear to have a lower tendency towards bias, more often giving a result close to the simulated value.

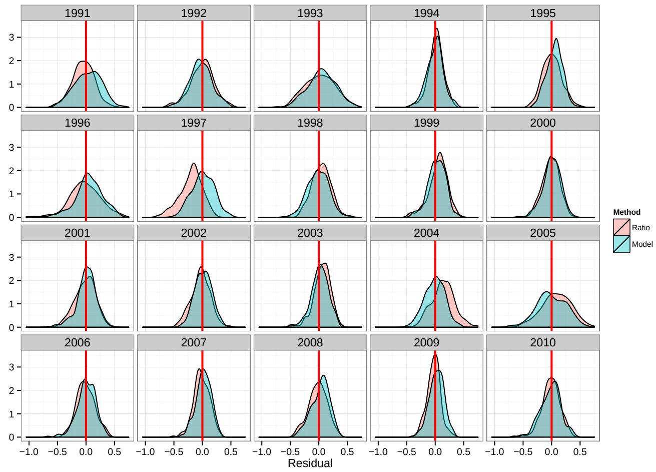
Alongside accuracy, an important consideration when evaluating estimator performance is quantification of the uncertainty. This was measured from the proportion of times that the estimated confidence bounds enclose the true simulated value. This probability is referred to as a p-value, since it is the probability of a Type 1 statistical error (i.e. the probability of incorrectly concluding that the estimated value is different from the simulated value). Therefore smaller p-values generally indicate better performance. More specifically, we would expect p-values of approximately 5% if the confidence intervals are being estimated accurately, since the intervals should enclose the true value 95% of the time. The results of this calculation are given in Table 2. In every case the p-value is smaller for the model-based estimator, indicating that this method is more likely to produce confidence bounds that enclose the true value. This is rarely the case for the ratio-based estimator, indicating that the confidence intervals produced by this method are likely to be too narrow, and for this reason the model-based estimator is to be preferred.

6. Empirical Results

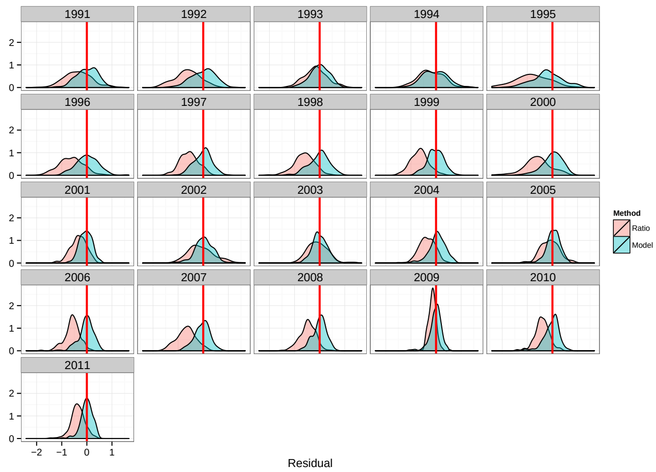
The estimators of bycatch for each method, for each fishery are illustrated in Figure 2 and Tables 3, 4 and 5. For ling, the ratio-based estimates are often an order of magnitude less than the model-based estimates and have much tighter confidence intervals. From the simulation results in Figure 1a, it appears that the ratio-based estimator has a negative bias, giving values that tend to be underestimates. From Table 2a it is also clear that the confidence intervals are likely to be too narrow. For both these reasons, the model-based estimates are preferable.



(a) Ling



(b) Scampi



(c) Squid

Figure 1: Distribution of log-residuals $\ln(\hat{N}/N)$ for ratio and model-based estimators of bycatch applied to simulated data, where N and $\hat{N}$ refer to the simulated and estimated bycatch respectively. The residual distribution should be close to zero (indicated by the vertical red line) if the estimator is performing well.

Table 2: Comparative performance of ratio and model-based estimators of bycatch, using simulated data. The p-value refers to the proportion of times that the 95% confidence or credibility intervals for an estimate do not include the simulated value. We would expect a p-value of approximately 5% if the uncertainty intervals are being estimated correctly: if $p > 0.05$ the intervals are too narrow; if $p < 0.05$ the intervals are too wide.

(a) Ling

Fishing year	p-value Ratio	p-value Model
1993	0.945	0.160
1994	0.945	0.100
1995	0.885	0.200
1996	0.695	0.110
1997	0.920	0.160
1998	0.830	0.100
1999	0.715	0.060
2000	0.675	0.045
2001	0.750	0.235
2002	0.885	0.190
2003	0.765	0.060
2004	0.815	0.090
2005	0.890	0.110
2006	0.855	0.200
2007	0.785	0.120
2008	0.850	0.030
2009	0.885	0.070
2010	0.950	0.175
2011	0.850	0.045
2012	0.950	0.030
Average	0.842	0.114

(b) Scampi

Fishing year	p-value Ratio	p-value Model
1991	0.395	0.135
1992	0.360	0.180
1993	0.675	0.135
1994	0.220	0.085
1995	0.590	0.010
1996	0.685	0.130
1997	0.605	0.025
1998	0.245	0.070
1999	0.345	0.045
2000	0.380	0.030
2001	0.505	0.040
2002	0.465	0.045
2003	0.320	0.045
2004	0.485	0.065
2005	0.590	0.045
2006	0.325	0.025
2007	0.235	0.010
2008	0.260	0.010
2009	0.230	0.005
2010	0.480	0.020
Average	0.420	0.058

(c) Squid

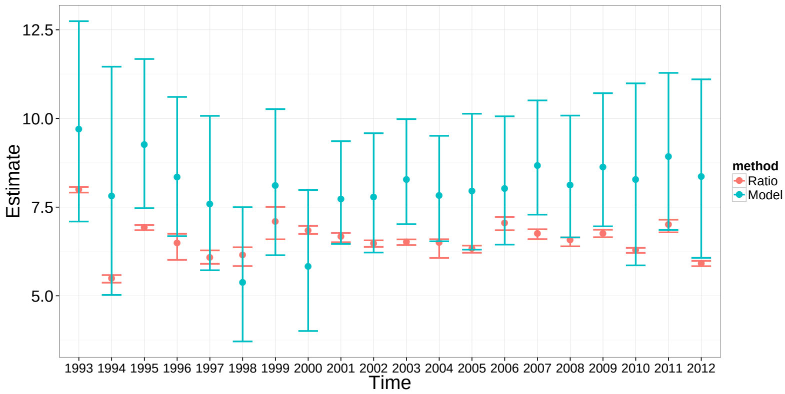
Fishing year	p-value Ratio	p-value Model
1991	0.560	0.100
1992	0.750	0.150
1993	0.480	0.235
1994	0.665	0.220
1995	0.760	0.255
1996	0.795	0.160
1997	0.765	0.235
1998	0.765	0.220
1999	0.840	0.180
2000	0.810	0.210
2001	0.540	0.155
2002	0.690	0.215
2003	0.470	0.185
2004	0.600	0.225
2005	0.470	0.215
2006	0.835	0.190
2007	0.840	0.200
2008	0.735	0.130
2009	0.180	0.115
2010	0.725	0.135
2011	0.685	0.100
Average	0.665	0.182

For scampi, the bias expected from application of the ratio-based method is less obvious (Figure 1b), and Table 4 shows that both methods produce comparable estimates of the total bycatch (see also Figure 2b). There is a clear difference however in the confidence intervals, which are much wider for the model-based estimator. From the simulation study (Table 2b) we know that the confidence intervals for the ratio-based method are likely to be conservative. For squid, the confidence intervals for both methods are more similar. However the estimates of bycatch are slightly different (Table 5 and Figure 2c) and we suspect that the ratio-based results are underestimates, based on the bias observed in Figure 1c.

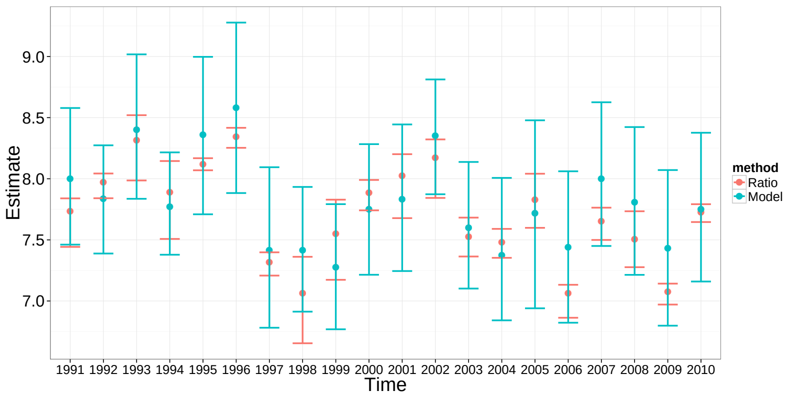
7. Conclusions

This study has compared the performance of ratio and model-based estimators of bycatch using data from three example fisheries. A simulation study was used to indicate likely performance of each method, followed by application to the empirical data. In all cases it was shown that the model-based estimator should be preferred, since it appears to be less prone to bias and capable of producing a more reliable representation of uncertainty.

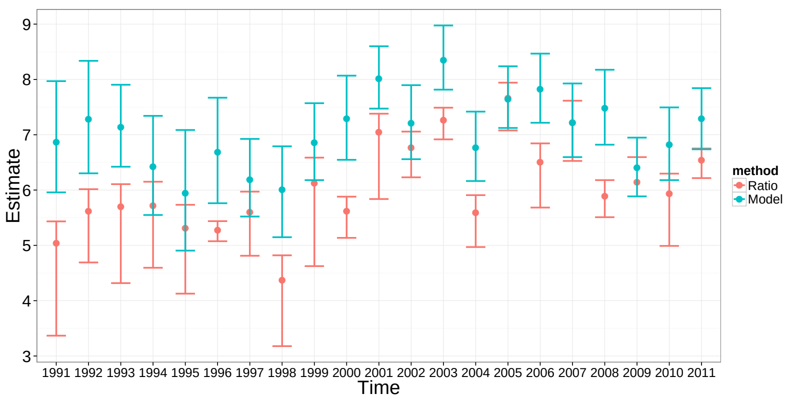
The results are not intended to supplant previous estimates of the bycatch made by Anderson (2012,



(a) Ling



(b) Scampi



(c) Squid

Figure 2: Estimates of non-QMS fish bycatch from empirical data using ratio and model-based estimators. Values are given on a log scale in units of tonnes, with 95% confidence intervals shown.

Table 3: Comparative estimates of total non-QMS bycatch in tonnes from the ling longline fishery using ratio and model-based estimators.

Fishing	Total bycatch $\hat{N}$ and 95% confidence intervals	
year	Ratio	Model
1993	2965 [2727;3199]	16438 [1205;342109]
1994	243 [215;266]	2480 [152;94772]
1995	1024 [944;1091]	10551 [1755;117800]
1996	661 [409;853]	4213 [797;40446]
1997	439 [366;534]	1974 [305;23713]
1998	469 [343;583]	218 [41;1805]
1999	1203 [730;1822]	3307 [466;28681]
2000	933 [848;1067]	342 [55;2928]
2001	794 [672;873]	2274 [642;11599]
2002	652 [590;709]	2393 [503;14516]
2003	679 [620;728]	3917 [1118;21635]
2004	666 [431;731]	2501 [690;13512]
2005	566 [500;612]	2861 [546;25180]
2006	1153 [943;1366]	3077 [629;23386]
2007	863 [733;969]	5841 [1464;36617]
2008	719 [599;769]	3365 [770;23889]
2009	866 [774;957]	5628 [1052;44929]
2010	538 [498;574]	3923 [350;59224]
2011	1104 [888;1269]	7516 [949;79620]
2012	369 [342;398]	4289 [433;66357]

2013, 2014), simply to provide a comparative representation of performance. Our ratio-based estimate of the confidence intervals was slightly different from that applied previously, specifically sampling vessels rather than vessel trips during the bootstrap. This was necessary to allow the same estimator to be applied to both simulated and empirical data, and would have led to slightly different estimates of the confidence bounds. Discrepancies were observed between our ratio-based estimates of the bycatch and those reported by Anderson (2012, 2013, 2014). For ling, this was because we used the number of sets rather than the number of hooks to quantify effort. For scampi and squid the reasons are unclear, but the discrepancies themselves are small.

Based on the results presented here and in Appendix I we would argue that the model-based approach could be usefully implemented to provide actual estimates of the bycatch, particularly given the unbalanced sampling design associated with observer data collected from real fisheries. The model-based approach could be extended through the inclusion of other covariates, which will improve performance if the covariate data are of good quality. This is an aspect of the model-based approach that has not been considered in the current study, and whether these covariate data are reliable enough for inclusion in a model-based application will have to be assessed on a case-specific basis. The current study provides an argument for further work in developing this approach.

8. Acknowledgements

This work was funded by the Ministry for Primary Industries (Wellington, New Zealand), under project code ENV2013-01, and received a helpful review from Ian Tuck (NIWA, New Zealand).

Table 4: Comparative estimates of total non-QMS bycatch in tonnes from the scampi fishery using ratio and model-based estimators.

Fishing	Total bycatch $\hat{N}$ and 95% confidence intervals	
year	Ratio	Model
1991	2288 [1707;2539]	2984 [1739;5318]
1992	2902 [2542;3112]	2527 [1617;3919]
1993	4080 [2940;5015]	4449 [2530;8249]
1994	2674 [1822;3446]	2372 [1601;3699]
1995	3351 [3195;3527]	4275 [2229;8082]
1996	4197 [3839;4522]	5338 [2652;10697]
1997	1503 [1350;1633]	1661 [881;3276]
1998	1169 [776;1574]	1661 [1005;2788]
1999	1898 [1304;2512]	1443 [870;2422]
2000	2655 [2302;2952]	2326 [1359;3957]
2001	3062 [2160;3645]	2525 [1401;4649]
2002	3544 [2548;4113]	4246 [2626;6718]
2003	1857 [1578;2169]	1999 [1214;3420]
2004	1773 [1561;1978]	1592 [936;3002]
2005	2508 [1995;3106]	2252 [1033;4809]
2006	1167 [956;1252]	1705 [918;3169]
2007	2104 [1807;2353]	2987 [1720;5574]
2008	1820 [1446;2285]	2465 [1358;4550]
2009	1183 [1065;1264]	1688 [896;3201]
2010	2263 [2092;2421]	2325 [1286;4343]

Table 5: Comparative estimates of total non-QMS bycatch in tonnes from the squid fishery using ratio and model-based estimators.

Fishing	Total bycatch $\hat{N}$ and 95% confidence intervals	
year	Ratio	Model
1991	154 [29;229]	962 [387;2891]
1992	275 [109;410]	1449 [546;4172]
1993	298 [75;449]	1252 [615;2708]
1994	304 [99;469]	617 [257;1543]
1995	203 [62;309]	380 [135;1195]
1996	196 [160;230]	798 [318;2142]
1997	271 [123;392]	487 [250;1017]
1998	79 [24;124]	407 [172;890]
1999	455 [102;725]	952 [483;1940]
2000	275 [170;358]	1460 [697;3192]
2001	1143 [343;1604]	3033 [1759;5433]
2002	866 [508;1161]	1349 [705;2688]
2003	1431 [1009;1788]	4208 [2479;7908]
2004	268 [144;368]	865 [475;1666]
2005	2114 [1186;2812]	2077 [1240;3782]
2006	668 [294;938]	2496 [1362;4754]
2007	1365 [683;2029]	1366 [732;2771]
2008	360 [247;483]	1768 [916;3551]
2009	463 [-54;732]	603 [360;1041]
2010	379 [147;543]	919 [483;1798]
2011	690 [501;855]	1469 [843;2548]

References

Abraham, E.R. (2008). Evaluating methods for estimating the incidental capture of New Zealand sea lions. *New Zealand Aquatic Environment and Biodiversity Report* 15: 25p.

Abraham, E.R.; Thompson, F.N. (2011). Estimated capture of seabirds in New Zealand trawl and longline fisheries, 2002-03 to 2008-09. *New Zealand Aquatic Environment and Biodiversity Report* 79: 74p.

Anderson, O.F. (2012). Fish and invertebrate bycatch and discards in New Zealand scampi fisheries from 1990-91 until 2009-10. *New Zealand Aquatic Environment and Biodiversity Report* 100: 65p.

Anderson, O.F. (2013). Fish and invertebrate bycatch and discards in New Zealand arrow squid fisheries from 1990-91 until 2010-11. *New Zealand Aquatic Environment and Biodiversity Report* 112: 62p.

Anderson, O.F. (2014). Fish and invertebrate bycatch and discards in New Zealand ling longline fisheries from 1992-93 until 2011-12. *New Zealand Aquatic Environment and Biodiversity Report* 138: 66p.

Barnett, V. (1991). Estimating a ratio. In: Sample Survey: Principles and Methods, chap. 4.1, 78–84. Edward Arnold.

Griggs, L.H.; Baird, S.J. (2013). Fish bycatch in New Zealand tuna longline fisheries 2006-07 to 2009-10. *New Zealand Fisheries Assessment Report* 2013/13: 73p.

Kendall, M.; Stuart, A. (1977). Advanced Theory of Statistics. Charles Griffin & Company.

Rice, J.A. (1995). Mathematical Statistics and Data Analysis. Duxbury Press.

Smith, M.H.; Baird, S.J. (2007a). Estimation of incidental captures of New Zealand sea lions (*Phocarctos hookeri*) in New Zealand fisheries in 2003-04, with particular reference to the SQU 6T squid fishery. *New Zealand Fisheries Assessment Report* 2007/7: 32p.

Smith, M.H.; Baird, S.J. (2007b). Estimation of incidental captures of New Zealand sea lions (*Phocarctos hookeri*) in New Zealand fisheries in 2004-05, with particular reference to the SQU 6T squid trawl fishery. *New Zealand Aquatic Environment and Biodiversity Report* 12: 31p.

Thompson, F.N.; Abraham, E.R.; Berkenbusch, K. (2013). Common dolphin (*Delphinus delphis*) bycatch in New Zealand commercial trawl fisheries. *PloS ONE* 8 (5): e64438.

Appendix I: Justification for a model-based estimator of fish bycatch

Introduction

This appendix aims to compare the performance of ratio and model-based estimates of bycatch within a simulation framework. To design an appropriate study it is first worthwhile to describe some properties of ratio-based estimators, when they are likely to perform well, and when they are likely to perform poorly. Alternative modelling approaches will then be described, followed by the simulation study itself.

Accuracy of ratio-based estimators

The accuracy of an estimator can be broken down into the *precision*, synonymous with the variance of the estimator, and *bias*, which is a measure of the difference between the central tendency of the estimator and the true population value. The properties of ratio-based estimators can suggest the conditions under which accuracy is likely to be compromised, and we investigate these properties here. To simplify this outline, we assume that y_i and t_i are independent and define T and S as the total fishing effort, and the total observed fishing effort respectively, with corollaries T_k and S_k that are specific to strata k.

We first note that since the variance of the effort term is zero (i.e. measured without error), the ratio is an unbiased estimate with variance $\mathbb{V}[r] \approx \mathbb{V}[y.]/S$ (Kendall & Stuart, 1977). Since $N^{[R]} = T.r$ we can write:

$$\mathbb{V}[N^{[R]}] \approx \frac{T^2}{S}.\mathbb{V}[y.]$$

This indicates quite simply that variance in the estimate of $N^{[R]}$ is inversely related to the observed fishing effort. In other words, the more observations the more precise the estimate.

We next define an analytical formula that describes the expected bias B, which can be introduced by ratio-based methods when applied to data that have not been appropriately stratified, usually as a result of only partial observer coverage. Consistent with current practice in New Zealand fisheries, we assume that the ratio estimator is stratified at the level of the area and investigate how an unbalanced sample allocation across areas creates a bias in the bycatch estimate.

We start with an assumption that observed catch is described perfectly by a series of coefficients β:

$$y_{ik} = g(\mathbf{x}_i'\beta).t_{ik}^{[S]}$$

$$N_k = T.\sum g(\mathbf{x}_i'\beta).f_{ik}^{[T]}$$

where $f_{ik}^{[T]} = t_{ik}^{[T]}/T$ is the proportion of the total effort T that correspond to event i in area k. For the ratio-based method, we assume an area-level stratification and obtain the bycatch estimate as:

$$N_k^{[R]} = T_k.r_k = T_k.\frac{\sum y_{ik}}{\sum t_{ik}^{[S]}} = T_k.\frac{\sum g(\mathbf{x}_i'\beta).t_{ik}^{[S]}}{\sum t_{ik}^{[S]}} = T_k.\frac{S}{S_k}\sum g(\mathbf{x}_i'\beta).f_{ik}^{[S]}$$

where $f_{ik}^{[S]} = t_{ik}^{[S]}/S$. We are now in a position to write down the bias as:

$$\begin{aligned} B_k &= N_k^{[R]} - N_k \\ &= T_k \frac{S}{S_k} \sum g(\mathbf{x}_i'\beta).f_{ik}^{[S]} - T.\sum g(\mathbf{x}_i'\beta).f_{ik}^{[T]} \end{aligned}$$

where the summations on the left and right refer to the observed and total fishing effort respectively. However if we set $f_{ik}^{[S]} = 0$ for unobserved tows and define:

$$h_k = \frac{T_k}{T}\frac{S}{S_k}$$

then:

$$\begin{aligned} B_k &= T.h_k.\sum g(\mathbf{x}_i'\beta)f_{ik}^{[S]} - T.\sum g(\mathbf{x}_i'\beta).f_{ik}^{[T]} \\ &= T.\sum g(\mathbf{x}_i'\beta).(h_k.f_{ik}^{[S]} - f_{ik}^{[T]}) \\ &= T.\sum g(\mathbf{x}_i'\beta).\left(\frac{T_k}{T}\frac{S}{S_k}.\frac{t_{ik}^{[S]}}{S} - \frac{t_{ik}^{[T]}}{T}\right) \end{aligned}$$

which simplifies to:

$$B_k = T_k.\sum g(\mathbf{x}_i'\beta).\left(\frac{t_{ik}^{[S]}}{S_k} - \frac{t_{ik}^{[T]}}{T_k}\right)$$

The predicted bias is the sum across areas $B = \sum B_k$ and is therefore a function of the difference between the proportion of effort observed and the proportion of the total fishing effort, for each fishing event, summed over the fishing events in each area.

From the above derivation we can see that for bias to occur $t_{ik}^{[S]}/S_k \neq t_{ik}^{[T]}/T_k$. Intuitively, this equates to a difference between the distribution of actual and observed tows across strata. As this difference increases, we expect the bias to also increase.

The magnitude of the bias will also be partly determined by the correlation between the sign of $t_{ik}^{[S]}/S_k - t_{ik}^{[T]}/T_k$ and the sign of $g(\mathbf{x}_i'\beta)$. If they are not correlated then there will be a partial cancelling out of bias effects since the overall sign of the bias at the *ik* level is random. However, should the signs of these terms be highly correlated, then they will have a reinforcing effect, since most will have the same sign (either positive or negative).

From these analytical results we can obtain two important insights concerning the reliability of ratio-based estimators of bycatch. First, as the sample size gets smaller, the precision will decrease. Second, as the distribution of sampled fishing events becomes increasingly unrepresentative, the bias will likely also increase. When comparing the performance of ratio-based methods against model-based alternatives, we will therefore include scenarios of low data quantity and an unbalanced sampling design.

Simulation Testing

A simulation experiment was carried out to compare performance of model-based bycatch prediction with simple ratio estimators. The study consisted of data simulation under a Poisson-LN process (Code Listing A5) and estimation of the fixed effects and associated bycatch using each of the methods. These estimates could then be compared to the simulated values. Iteration of this process allowed visualisation of the precision and bias of each estimator under a variety of data scenarios. The *precision* is defined as variation in the estimated value, whereas the *bias* concerns tendency of the estimator to converge on a value that is different from the true value as more data are collected.

The data scenarios themselves were primarily based on the quantity of data and distribution of sampled vessels between areas, both of which are expected to affect performance of the estimators (see "Accuracy of ratio-based estimators" above). Specifically, performance was expected to decrease if few vessels are sampled and if those vessels are clustered into a particular area. However we further specified whether or not coefficient values were held constant across the data iterations. This is because when the bias and effect-size are uncorrelated, they will tend to cancel out over multiple iterations. By fixing the input values, the iterated simulations can be more easily used to ascertain the bias associated with estimation, since error will be in a consistent direction. For randomised coefficient values, an impression of precision can be obtained.

We made no attempt to include the effects of additional covariates in the simulation, such as depth or the number of hooks (e.g. Anderson, 2014). This would have markedly complicated the analysis, without necessarily improving the quality of our conclusions.

Experimental design

The experimental design explored all combinations of different levels of data sampling (two levels) and data quantity (two levels), and whether the true area coefficient values used to simulate the data were fixed or randomly generated. For each of these eight combinations, 200 simulations were executed; and for each simulated dataset, both ratio and model-based methods were applied. Specific details regarding this experimental design are given here.

Dimensions

The simulation study assumed two spatial areas: A_1, A_2; represented by the fixed effect coefficients (β_1, β_2) and a random effect for 20 vessels $(\gamma_1, \dots, \gamma_{20})$, with each vessel assigned twenty fishing events. Each fishing event consisted of one effort unit (i.e. $t_i = 1$ for all i). The inclusion of a random vessel effect captures the influence of other covariates that may determine the catch rate, without needing to represent these covariates explicitly. This means that performance of the model-based estimator will be weaker than otherwise had covariate data been explicitly included, but the overall analysis is simplified greatly.

Data quantity

According to the above dimensions, fishing under 40 area/vessel combinations were simulated and sampled at different intensities:

- High data: 38 area/vessel combinations were sampled, giving 18 vessels sampled fishing across both areas;
- Low data: only 22 of these combinations were sampled, giving two vessels sampled across both areas.

For each area/vessel combination sampled, it was assumed that all fishing events were included.

Data sampling

The sampled area/vessel combinations were distributed according to the following designs:

- Unbalanced: all the vessels in A_2 were sampled, but only a subset of vessels in A_1 (18 and two for the high and low data scenarios respectively);
- Null: the sampled combinations were distributed evenly between A_1 and A_2.

Coefficients

Area specific input coefficients were either fixed across iterations, or randomly sampled. Under random sampling we assumed that $\{\beta_1, \beta_2\} \sim Uniform(1,3)$. When fixed, we assumed that $\beta_1 = 2.5$ and $\beta_2 = 1.5$. Random vessel effects were sampled from a normal distribution: $\{\gamma_1, \ldots, \gamma_{20}\} \sim Normal(0, \sigma^2_{[\gamma]})$; with $\sigma^2_{[\gamma]} = 1$. When area-specific coefficients were fixed, random effects were also fixed, being sampled once and held constant across iterations.

Iterations

Bycatch data were generated for all 40 area/vessel combinations using the simulator in Code Listing A5, iterated 200 times, once with `fixed.linear.predictor` set to `FALSE` and once with it set to `TRUE`; this determined whether fixed or randomly generated β values were used. These iterations were then sampled according to the data quantity and sampling designs given above, meaning that each scenario used a consistent set of simulated data.

Estimation methods

For each iterated data sample generated, bycatch (N) and fixed effect coefficients (β_1, β_2) were estimated using ratio and model-based methods. For the ratio-based method, the estimated bycatch was calculated as the average catch rate per area multiplied by the summed fishing effort:

$$\mathbb{E}[y_k] = \bar{y}_k$$

$$\hat{N}^{[R]} = \sum_k \left(\mathbb{E}[y_k] . \sum_{ij} t_{ijk} \right)$$

where i, j and k refer to the fishing event (equivalent to a single data record), vessel and area respectively. The area specific coefficient, for comparison to the model-based estimate, was calculated as $\beta_k = \ln(\bar{y}_k) - \sigma_k^2/2$, where σ_k^2 is the residual variance for that area: $\mathbb{V}[\ln(N^{[R]}_{ijk}/\hat{N}^{[R]}_{ijk})]$.

The model-based estimate of bycatch was calculated by first fitting a log-linear model to the observed data using `MCMCglmm` (Code Listing A6):

$$\ln(y_i) = \mathbf{x}_i'\beta + \mathbf{z}_i'\gamma + e_i$$

The catch rate was then calculated as the marginal expectation across a log-normal over-dispersion process (Code Listing A6):

$$\mathbb{E}[y_{jk}] = \exp\left(\beta_k + \gamma_j + \sigma^2/2\right)$$

which is specific to the particular area/vessel combination, so that the bycatch is estimated as:

$$\hat{N}^{[M]} = \sum_k \sum_j \left(\mathbb{E}[y_{jk}] . \sum_i t_{ijk} \right)$$

This estimator specifically makes use of the estimated random effect coefficient γ_j. Given the experimental design it could equally have been considered a fixed effect since every vessel is sampled at least once. However it is treated as a random effect here to maintain consistency with a more realistic scenario in which only a subset of vessels are sampled. In this latter case an estimate of the variance across vessels is important for bounding the uncertainty during prediction, and therefore treatment of the vessel effect as random is more appropriate.

Performance measures

The accuracy of ratio and model-based approaches were compared with respect to their precision and tendency towards bias. Precision was measured using the mean residual error:

$$MRE_\theta = \sqrt{\frac{1}{n}\sum(\theta - \mathbb{E}[\theta])^2}$$

with residuals measured on a log-scale, so that θ is one of $\ln(N/\hat{N})$, $\ln(\beta_1/\hat{\beta}_1)$ or $\ln(\beta_2/\hat{\beta}_2)$, and the expectation taken as an average across iterated residual values. Note that this statistic is equivalent to a coefficient of variation. For estimates of bias, we are concerned with detecting a consistent directional tendency in the distribution of estimated values. A suitable statistic for this purpose is the mean prediction error:

$$MPE_\theta = \left|\frac{\sum(\theta - \mathbb{E}[\theta])}{\sum\theta}\right|$$

where θ is one of N, β_1 or β_2 and the corresponding $\mathbb{E}[\theta]$ is $\hat{N}$, $\hat{\beta}_1$ or $\hat{\beta}_2$.

Simulation results

The evaluation results do not separate into clear categories, but for ease of presentation and discussion we make a distinction between simulations conducted with randomly generated β coefficients, and those conducted with fixed β coefficients. The former category is arguably more instructive for questions concerning precision, whereas the latter is more pertinent to the bias. To illustrate comparative precision, we therefore focus on data generated using randomly sampled input coefficients under conditions of high and low data quantity, and with performance summarised using the *MRE* diagnostic. For bias, the focus is on data generated using fixed input coefficients under Unbalanced and balanced (Null) sampling distributions, and compared using the *MPE* diagnostic.

Precision

Comparative precision of the two methods is illustrated in Figure A1, which shows the correlation between simulated and estimated values of the bycatch N, and the residual distribution, both plotted on the log-scale. Results have been combined over the unbalanced and null sampling designs but are shown separately for the high and low data scenarios. Overall there is a close correlation between simulated and estimated values, but from the residual distribution it can be clearly seen that the relative precision of each method is dependent on the quantity of data. As expected, the ratio-based method becomes less precise when there are fewer data. This observation is reflected in the *MRE* statistic (Figure A2a). For the high data scenario, the ratio-based method performs well. Because vessel effects are sampled from the same log-normal distribution, their effects average to zero on

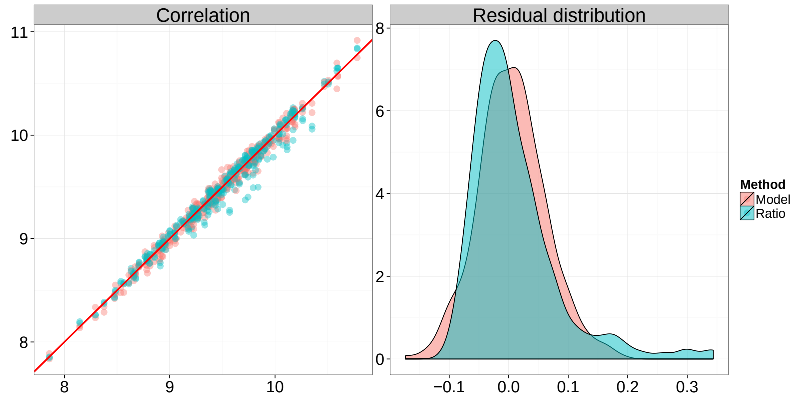
Table A1: Complete list of simulation testing results, showing the mean residual error (MRE) and mean prediction error (MPE) diagnostics, for estimated area specific coefficients (β_1 and β_2) and total bycatch values (N). Smaller MRE values indicate better precision, and smaller MPE values indicate a smaller bias.

(a) Randomly generated input coefficients

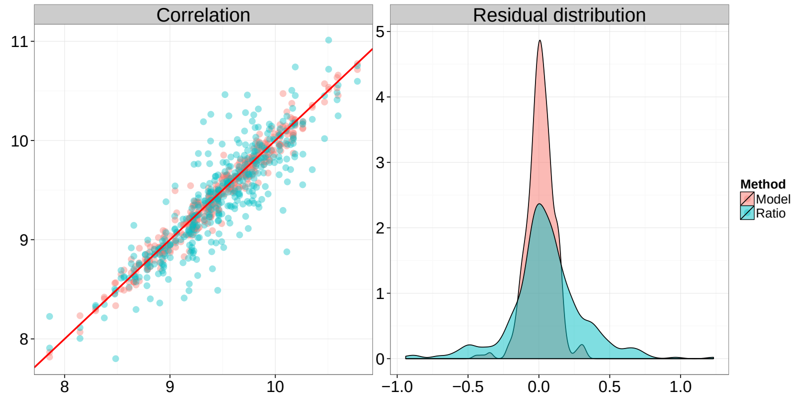
Parameter	Sampling	Data	MRE		MPE	
	design	quantity	Model	Ratio	Model	Ratio
β_1	Null	High	0.108	0.108	0.008	0.015
	Null	Low	0.121	0.132	0.009	0.014
	Unbalanced	High	0.111	0.114	0.008	0.010
	Unbalanced	Low	0.147	0.328	0.011	0.057
β_2	Null	High	0.109	0.104	0.003	0.016
	Null	Low	0.118	0.126	0.004	0.004
	Unbalanced	High	0.109	0.100	0.000	0.018
	Unbalanced	Low	0.108	0.100	0.003	0.018
N	Null	High	0.044	0.050	0.010	0.006
	Null	Low	0.058	0.089	0.017	0.019
	Unbalanced	High	0.045	0.056	0.003	0.013
	Unbalanced	Low	0.088	0.267	0.002	0.041

(b) Fixed input coefficients

Parameter	Sampling	Data	MRE		MPE	
	design	quantity	Model	Ratio	Model	Ratio
β_1	Null	High	0.024	0.015	0.048	0.048
	Null	Low	0.033	0.018 7	0.059	0.120
	Unbalanced	High	0.025	0.016	0.050	0.028
	Unbalanced	Low	0.061	0.033	0.047	0.215
β_2	Null	High	0.043	0.023	0.078	0.090
	Null	Low	0.061	0.030	0.060	0.009
	Unbalanced	High	0.040	0.021	0.079	0.113
	Unbalanced	Low	0.044	0.021	0.074	0.113
N	Null	High	0.046	0.004	0.005	0.049
	Null	Low	0.059	0.025	0.005	0.108
	Unbalanced	High	0.047	0.011	0.005	0.050
	Unbalanced	Low	0.111	0.085	0.003	0.167



(a) High data scenario



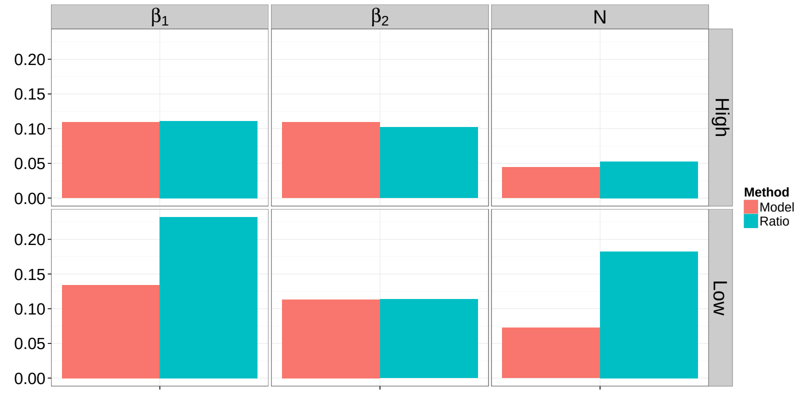
(b) Low data scenario

Figure A1: Predicted bycatch diagnostics for simulated data iterations, combined across sampling designs and assuming high and low data scenarios with randomly generated input coefficients. The correlation shows the relationship between $\ln(N)$ and $\ln(\hat{N})$, where N and $\hat{N}$ refer to the simulated and estimated bycatch respectively. Points close to the line indicate that simulated and estimated values are similar. The residual distribution shows a probability density plot of $\ln(N/\hat{N})$, which should be close to zero if the estimator is performing well.

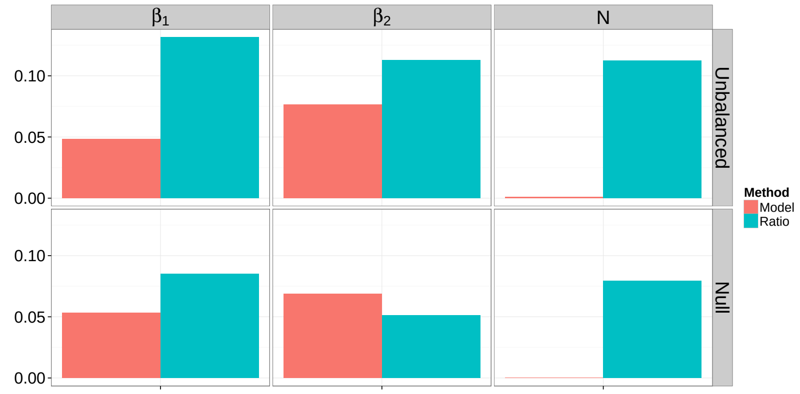
the log-scale; and so provided enough vessels are sampled in each area the ratio-based method will be accurate. However if the sampling intensity is decreased, then precision of the estimator also decreases, and in the low data scenario precision of the model-based estimator is superior.

The patterns observed in Figure A1 and Figure A2a can be examined in more detail in Table A1a. For the β coefficients, MRE values are around 10%. For estimates of N, MRE values are around 4% for the High data scenarios and slightly higher for the Low data scenarios. The model-based estimator has the lower coefficient of variation (MRE) for N, although for the High data scenarios this difference is only about 1%. For the Low data scenarios there is a particularly obvious difference between estimators when the sampling design is unbalanced, with $MRE^{[M]}$ much less than $MRE^{[R]}$. A superior performance of the model-based method at a Low data quantity (Figure A2a) is therefore largely due to the fact that it can outperform the ratio-based method when the design is Unbalanced. When the sampling design is balanced (Null), and there are a lot of data, precision of the ratio-based method is very similar.

From Table A1a it can be summarised that both methods lose precision with less data, but the effect is further compounded depending on whether or not sampling is balanced. For an unbalanced sample



(a) Comparative precision for high and low data scenarios measured using the mean residual error (MRE). Results are integrated over unbalanced and null sampling designs and assume randomly generated coefficients for the simulated data. Smaller MRE values indicate better precision.



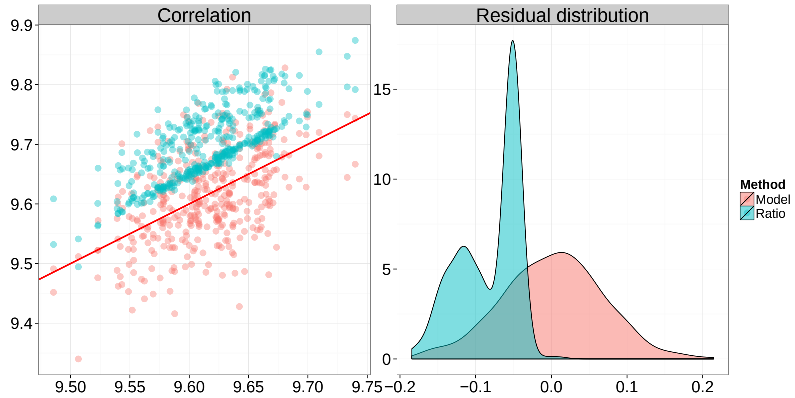
(b) Comparative bias for unbalanced and balanced (null) sampling designs measured using the mean prediction error (MPE). Results are integrated over high and low data scenarios and assume fixed input coefficients for the simulated data. Smaller MPE values indicate a smaller bias.

Figure A2: Comparative precision and bias of ratio and model-based estimators of the area specific coefficients (β_1 and β_2) and total bycatch values (N).

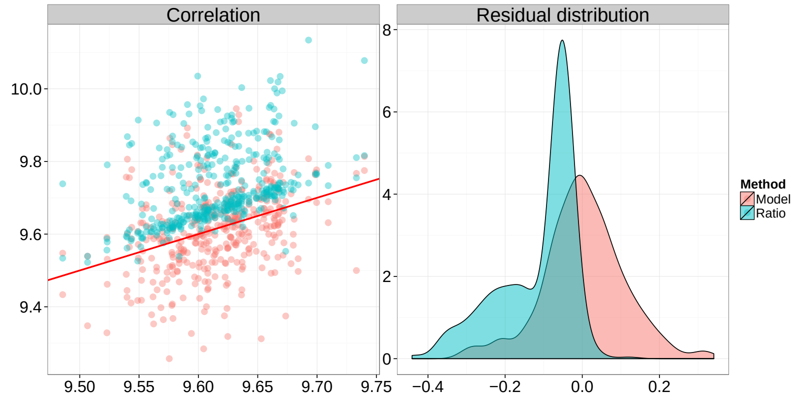
design there is a much more pronounced loss of accuracy for the ratio-based approach. This is responsible for the large differences in precision observed in Figure A2a. If only a small fraction of the vessels in one particular area are sampled, then they are unlikely to be representative of the overall bycatch rates in that area. In this example, very few vessels were sampled from A_1, resulting in an inaccurate estimate for the β_1 coefficient, in turn leading to an inaccurate estimate of the bycatch. For the model-based method however, information on the vessel effects can be shared across areas if available, making it less susceptible to the sampling design.

Bias

By keeping the fixed effects constant, directional error is more easily observed, and this can be seen in Figure A3. Even though the effect size is small, the ratio-based method is clearly biased, since the residual distribution has a central tendancy that has shifted away from zero. Figure A3 illustrates that an Unbalanced sampling design increases bias for the ratio based method, and the bimodal residual distribution demonstrates the componding effect of data quantity, with less data leading to an



(a) Balanced (null) sampling design



(b) Unbalanced sampling design

Figure A3: Predicted bycatch diagnostics for simulated data, combined across high and low data scenarios and assuming balanced and unbalanced sampling designs with constant input coefficients. The correlation shows the relationship between $\ln(N)$ and $\ln(\hat{N})$, where N and $\hat{N}$ refer to the simulated and estimated bycatch respectively. Points close to the line indicate that simulated and estimated values are similar. The residual distribution shows a probability density plot of $\ln(N/\hat{N})$, which should be close to zero if the estimator is performing well. Note the difference in scale compared to Figure A1.

increased bias. The influence of data quantity on bias can be clearly seen in the *MPE* diagnostics reported in Table A1b. In general, the model-based method has only a small bias in estimates of N, being less than 1% for all scenarios. In contrast, the ratio-based method has a bias in the region of 5% to 15% (Table A1b), with the larger values associated with the Low data scenarios.

The *MPE* values in Figure A2b summarise the results in Table A1b. The ratio-based method produces a biased estimate of β_1, which is translated into a biased estimate of the bycatch. In contrast, the model-based approach is not only less biased, but also more resilient to the sampling design, giving similar results under the Unbalanced and Null sampling scenarios. Figure A3 shows how accuracy of the method is reasonably consistent across sampling scenarios, which is an attractive property of the model-based approach.

The simulations conducted so far have only assumed either fixed or random coefficients, and so to further investigate the different propensities towards bias we conducted a secondary analysis, again with fixed area-specific coefficients, but fixed across a range of values. This allowed us to examine the bias associated with each estimator as the difference in effect size between areas increased. Since we

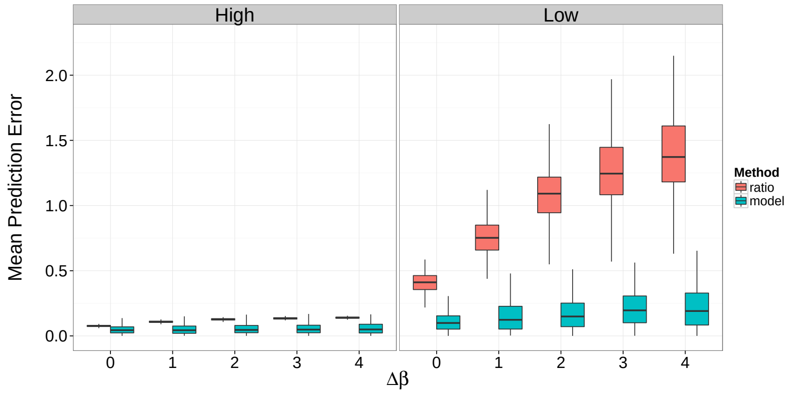


Figure A4: The bias of ratio and model-based estimators plotted against different combinations of area-specific simulation coefficients, summarised as the difference $\Delta\beta$. Both High and Low data scenarios are shown.

expected *a priori* that estimation bias would be more apparent for an unbalanced sampling design, an expectation that has been confirmed by the results already presented, this secondary simulation was conducted for the Unbalanced scenario only. The following area-specific coefficients were explored:

β_1	1.0	1.5	2.0	2.5	3.0
β_2	1.0	0.5	0.0	-0.5	-1.0
$\Delta\beta$	0	1	2	3	4

where $\Delta\beta$ is the difference between coefficient values. The results of this experiment are summarised in Figure A4. The bias for both methods increases with $\Delta\beta$ but it can be clearly seen that the ratio-based method is more susceptible.

Conclusions

The results presented here were to a large extent expected given the experimental design and the way analyses were conducted. Regarding performance of the model-based method, the model was applied to simulated data generated from the same model. Hence, the model-based estimator should

Table A2: Summary table illustrating comparative performance of ratio and model-based estimators of the bycatch (N), referred to using the $[R]$ and $[M]$ superscripts respectively. The estimators are compared using diagnostic ratios of the mean residual error (MRE, a measure of precision) and mean prediction error (MPE, a measure of bias).

Input coefficients	Sampling design	Data quantity	MRE diagnostic ratio $MRE_{N^{[R]}}/MRE_{N^{[M]}}$	MPE diagnostic ratio $MPE_{N^{[R]}}/MPE_{N^{[M]}}$
Random	Null	High	1.124	0.570
	Null	Low	1.548	1.125
	Unbalanced	High	1.247	5.161
	Unbalanced	Low	3.037	25.465
Fixed	Null	High	0.096	9.993
	Null	Low	0.423	21.214
	Unbalanced	High	0.229	9.874
	Unbalanced	Low	0.769	56.994

be largely unbiased, as our results show. However the parameters are estimated with error and so it has low precision for low data quantity. This imprecision is worse for an unbalanced sampling design, which is again what one would expect.

For the ratio estimator, the incomplete sampling design guarantees that bias will occur. Because the ratio-based method cannot accomodate the random vessel effects (which are estimated in the model-based method) the estimator will be biased unless all vessels are sampled in all areas, which was never true for our experiment. However the ratio-based method does not require estimation of covariates for prediction, and can produce results of comparable (and potentially better) precision with adequate sampling. Therefore if covariate effects are weak or absent the ratio based method may perform well, provided the sampling coverage is adequate.

In Table A2 we have calculated diagnostic ratios for each of the performance statistics to summarise our results. It can be seen quite clearly that precision can be higher or lower for the ratio-based estimator but that the bias associated with these estimates is often substantial (i.e. $MPE_N^{[R]}$ is much greater than $MPE_N^{[M]}$). It is not clear why the MRE diagnostic should favour the model-based estimator when input coefficients are Random, and not when they are Fixed, but it may be due to the particular choice of vessel specific coeffcients generated for the Fixed input scenario. Nevertheless overall, given that covariate effects, an unbalanced sampling design, and incomplete data sampling would be the norm for a real fishery, we can conclude that the model-based estimator is likely to be superior when applied to real data.

This appendix has illustrated some of the basic shortcomings inherent to ratio-based estimators of fisheries bycatch. Although limited in scope, it is sufficient to highlight the importance of a balanced sampling design (with respect to the level of stratification) for the method to work well. If it is not balanced, with sampling concentrated in a few potentially unrepresentative vessels or covariate realisations, then ratio-based methods do not perform well, being prone to bias. Furthermore, biased estimates can still be associated with a high level of precision (low variance) giving an unwarranted and misleading impression of accuracy. Model-based methods on the other hand can be less accurate, but we have shown that they are far more robust, being more likely to give accurate results under an unbalanced sampling design even when there are relatively few data available.

Appendix II: Code listings

Code Listing A1: Catch rate simulation function

```
# CATCH-RATE SIMULATION FUNCTION
catch.rate <- function(initial.catch.rate,ntime,niter) {

   require(MASS)

   # number of areas
   narea <- length(initial.catch.rate)

   # biomass array
   x <- array(NA,c(narea,ntime,niter))

   # initial values
   x0 <- initial.catch.rate

   # transition matrix
   T <- diag(narea)

   # covariance matrix
   sd <- 0.1
   rho <- 0.8

   sigma <- matrix(rho*sd*sd,narea,narea)
   diag(sigma) <- sd^2

   # generate array of correlated
   # log-normal random variates
   rnd <- array(NA,c(narea,ntime,niter))

   for(t in 1:ntime) {
      rnd.normal <- mvrnorm(niter,rep(0,narea),sigma)
      rnd.lognormal <- exp(rnd.normal)

      rnd[,t,] <- t(rnd.lognormal)
   }

   # projection
   x[,1,] <- x0
   if(ntime>1) {
      for(i in 1:niter) {
         for(t in 2:ntime) {
            diag(T) <- rnd[,t-1,i]
            x[,t,i] <- T%*%x[,t-1,i]
         }
      }
   }

   return(x)
}
```

Code Listing A2: Simulation function for catch data from a Poisson-LN process, for a single time step only, with design matrices `x.mat` and `z.mat` for fixed and random effects respectively, and associated model data frame `model.dfr`. The output `summary.data` provides the input for the model-based estimator in Code Listing A4.

```
# DATA SIMULATION FUNCTION
sim <- function(mean.catch.rate,fish.effort,obs.sampling,model.dfr)
{
   # design matrices
   x.mat <- model.matrix(~area-1,model.dfr)
   z.mat <- model.matrix(~vessel-1,model.dfr)

   # over-dispersion
   sigma2e <- 0.5

   # random effects
   sigma2r <- sigma2.vessel
   gmma  <- rnorm(ncol(z.mat),0,sqrt(sigma2r))

   # fixed effects
   beta <- log(mean.catch.rate) - sigma2r/2 - sigma2e/2

   # log-linear predictor
   log.mu <- x.mat %*% beta + z.mat %*% gmma

   # bycatch per tow with log-normal over-dispersion
   # and Poisson catch process
   bycatch.list <- apply(data.frame(log.mu,fish.effort),1,
                    function(x) {
                       lambda <- exp(x[1] + rnorm(x[2],0,sqrt(sigma2e)))
                       rpois(x[2],lambda)
                    })

   # sample tows
   y.list <- list()
   for(i in 1:length(bycatch.list)) {
      if(obs.sampling[i]>0)
         y.list[[i]] <- sample(bycatch.list[[i]],
                           obs.sampling[i],
                           replace = ifelse(obs.sampling[i] <
                              length(bycatch.list[[i]]), F, T)
                           )
      else y.list[[i]] <- NA
   }

   # summary data
   model.dfr$y   <- unlist(lapply(y.list,sum))
   model.dfr$bycatch <- unlist(lapply(bycatch.list,sum))

   # return
   list(beta = beta,
        bycatch = bycatch.list,
        observer.samples = y.list,
        summary.data = model.dfr)
}
```

Code Listing A3: Function for fitting a Poisson-LN model to observational bycatch data using `MCMCglmm` and assuming a sampling effort offset term

```
# MODEL FITTING FUNCTION
fitmodel <- function(dat) {

   require(MCMCglmm)

   # construct normal prior distributions
   # for fixed effects with informative unit
   # prior for log(sampling) offset term
   x.mat <- model.matrix(~area+time-1,dat)
   prior <- list(B = list(mu = matrix(0,dim(x.mat)[2]+1),V =
      diag(dim(x.mat)[2]+1)*(10^4)))
   prior$B$mu[dim(x.mat)[2]+1]<-1
   diag(prior$B$V)[dim(x.mat)[2]+1]<-1e-9

   # fit model
   model.formula <- y ~ -1 + area + time + log(sampling)
   m <- MCMCglmm(model.formula,random =
      ~idh(1):vessel,prior=prior,family="poisson",data=dat,pr=TRUE)

   return(m)
}
```

Code Listing A4: Bycatch estimation function for model-based estimator using the model fitting routine in Code Listing A3.

```
# MODEL-ESTIMATOR FUNCTION
fitfun.mcmc <- function(dat) {

   require(plyr)

   # assign model data.frame
   model.dfr <- dat

   # remove missing values for
   # fitting
   dat <- subset(dat,sampling>0)

   # fit model
   m <- fitmodel(dat)

   # random effect
   sigma2r <- m$VCV[,1,drop=FALSE]

   # over-dispersion term
   sigma2e <- m$VCV[,2,drop=FALSE]

   # extract bycatch.hat values
   # as the marginal expectation
   x.mat <- model.matrix(~area+time-1,model.dfr)

   beta <- apply(m$Sol[,1:(dim(m$X)[2]-1)],1,function(x) x.mat %*% x)
```

```
    log.mu <- sweep(beta,2,sigma2r/2 + sigma2e/2,'+')

    bycatch.hat <- exp(log.mu + log(model.dfr$effort))

    # convert to data.frame
    bycatch.dfr <- melt(bycatch.hat,varnames=c('time','iter'))
    bycatch.dfr$time <- model.dfr$time

    # summarize by time
    bycatch.dfr <- ddply(bycatch.dfr,.(time,iter),summarize,
                         value=sum(value))
    bycatch.dfr <- ddply(bycatch.dfr,.(time),summarize,
                         bycatch.hat=posterior.mode(mcmc(value)),
                         med=quantile(value,0.5),
                         low=quantile(value,0.025),
                         upp=quantile(value,0.975))

    # return
    list(y = y.dfr, bycatch = bycatch.dfr)
}
```

Code Listing A5: Simulation function for Poisson-LN process, with design matrices `x.mat` and `z.mat` for fixed and random effects respectively, and associated model data frame `model.dfr`.

```
# SIMULATION FUNCTION
sim <- function(sampling=1,fixed.linear.predictor)
{
   sigma2r <- 1
   sigma2e <- 0.5

    if(fixed.linear.predictor) {
       beta <- beta.fixed
      gmma <- gmma.fixed
    } else {
       beta  <- runif(ncol(x.mat),1,3)
       gmma   <- rnorm(ncol(z.mat),0,sqrt(sigma2r))
    }

    # log-linear predictor
    log.mu <- x.mat %*% beta + z.mat %*% gamma

    # log-normal over-dispersion
    e <- rnorm(length(log.mu),0,sqrt(sigma2e))

    # Poisson rate parameter
    lambda <- exp(log.mu + e)

    # Poisson catch process
    bycatch <- rpois(length(lambda),lambda)

    # observation
    y <- sampling * bycatch

    # return
```

```
    list(beta=beta,gmma=gmma,sigma2=sigma2r,
                    data=data.frame(id=1:length(y),
                          model.dfr,
                          effort=1,
                          sampling,bycatch,y))

}
```

Code Listing A6: Function for model based estimation of the bycatch using observed values y.

```
# GLMM FITTING FUNCTION
fit.mcmc <- function(dat) {

   require(MCMCglmm)

   # fit mixed model
   m <- MCMCglmm(y ~ -1 + area,random =
       ~idh(1):vessel,family="poisson",data=dat,pr=TRUE,pl=TRUE)

   # extract fixed and random effects
   theta <- apply(m$Sol,2,function(x) posterior.mode(mcmc(x)))
   # extract variance terms
   sigma2 <- apply(m$VCV,2,function(x) posterior.mode(mcmc(x)))
   # extract latent variables
   fitted <- apply(m$Liab,2,function(x) posterior.mode(mcmc(x)))

   # construct design matrix for fixed
   # and random effects
   W <- cBind(m$X,m$Z)

   # calculate marginal expectation of
   # bycatch rate
   expectation <- exp(W%*%theta + sigma2[2]/2)

   # calculate bycatch per record
   bycatch <- dat$effort * expectation

   # calculate total bycatch
   bycatch.total = sum(bycatch)

   # return
   list(fixef = theta[1:dim(m$X)[2]],reff =
       sigma2[1],odisp=sigma2[2],yhat=fitted,bycatch=bycatch.total)
}
```