

Prior versus data: A new Bayesian method for fishery stock assessments

Y. Lei^a, S. Zhou^b and N. Ye^a

^a*School of Mathematics and Physics, The University of Queensland, St Lucia, QLD 4072, Australia*

^b*CSIRO Oceans and Atmosphere, 306 Carmody Road, Saint Lucia, Brisbane, QLD 4072, Australia*

Email: yeming.lei@uq.edu.au

Abstract: Bayesian method has become an indispensable technique in fisheries stock assessments, due to its flexibility and effectiveness in handling the various types of uncertainties in stock assessments, such as a lack of knowledge about the biological process and the presence of process and measurement errors. In particular, in the data-limited or data-moderate case, while frequentist methods may struggle to provide satisfactory results, the Bayesian approach can obtain more accurate estimates with its ability to incorporate the existing knowledge in the form of prior distributions on the models.

The effectiveness of Bayesian methods stems from its ability to combine the prior and observation in a principled way using the Bayes theorem. However, specifying a good prior can be a difficult task, and a poor prior may lead to a poor parameter estimate, even though the data are of high-quality for a frequentist-based approach, such as maximum likelihood estimation, to obtain accurate parameter estimates.

We developed a novel Bayesian method called *multi-pass Bayesian estimation* (MBE) that allows us to adjust the relative importance of the prior and the data, and thus offers more flexibility in combining the prior and the data. The MBE enables such flexibility to control the relative importance of the prior and the data by performing multiple Bayesian updates using the given dataset. By controlling the relative importance of the prior and the data, we obtain a spectrum of different interpolations of them, including only relying on the prior which completely ignores the data, the standard Bayesian posterior which combines the prior and the data using the Bayes rule once, and distributions that essentially rely just on the data.

We applied the MBE to a Bayesian surplus production model based on the popular Schaefer population dynamics model. Results in simulation studies show that the MBE provides more accurate parameter estimates than the standard Bayesian approach, in various settings with different levels of uncertainties in the data. The method is applicable to any Bayesian stock assessment models, and our results suggest that further investigation of the method's performance is promising.

Keywords: *Bayesian method, Multi-pass Bayesian estimation, maximum likelihood estimate, fishery stock assessments*

1 INTRODUCTION

Fishery stock assessments form a quintessential component in sustainable fisheries management: it provides information on the past and current status of a fishery stock, by assessing the size of the stock, whether the stock is growing or declining, and the effect of fishery management interventions (Cooper 2006, Methot & Wetzel 2013).

A key challenge in fishery stock assessments is the presence of various sources of uncertainties, including a lack of knowledge about the biological process and measurement errors (Cooper 2006). As a consequence, it is not possible to perfectly determine key stock parameters using data alone. In addition, we often want to know the uncertainties on the estimated parameters.

The Bayesian approach provides a principle and effective way to reason with the uncertainties present in fishery stock assessments. It can be used to combine prior domain knowledge and data using the Bayes rule; at the same time, it estimates uncertainties in the stock assessment results. Bayesian methods have been developed for the stock assessment using different types of data, including catch data (Meyer & Millar 1999), catch-age data (McAllister & Ianelli 1997), and length-data (Breen et al. 2003). An informative prior is often constructed using data from similar stocks via meta-analysis or empirical Bayes methods, or using expert judgement (e.g., see (McAllister & Kirkwood 1998) for an early extensive review, and see (Winker et al. 2018) for a discussion on some recent works). One of the most popular Bayesian stock assessment models is the Bayesian surplus production model (Punt & Hilborn 1997, McAllister et al. 2001, Winker et al. 2018), which will be used in our simulation studies in this paper.

The use of Bayesian methods allows us to exploit informative priors to obtain more reliable estimates when data are limited or noisy. However, specifying a good prior itself can be difficult. In that case, it is preferable to allow the data to play a more important role in the inference process. It is thus desirable to have the flexibility to control the relative importance of the prior and the data.

In this paper, we propose a novel Bayesian method that allows us to control the relative importance of the prior and the data. Our method allows us to navigate a spectrum of different degrees of relative importance on the prior and the data: at one end of the spectrum, we ignore the data and simply rely on the prior; at the other end of the spectrum, we depend mainly on the data with the prior is essentially ignored; the standard Bayesian method is one point in this spectrum – it is closer to the “prior” endpoint than the “data” endpoint as compared to most other points in the spectrum. Figure 1 provides a schematic illustration on the spectrum, where the parameter J (defined in Section 3) controls the relative importance of the prior and the data, with $J = 0$ indicating using the prior only, and $J = \infty$ indicating that the prior is essentially ignored.



Figure 1. The MBE spectrum. The Bayesian posterior is located in between the prior and data while the MBE can move from one end to the other.

We compared the MBE and the standard Bayesian method in simulation studies using the Schaefer form of the Bayesian surplus production model. We focused on estimating the carrying capacity, and our results suggest that the MBE can be used to obtain more accurate parameter estimates than the standard Bayesian method. In addition, the MBE also appears to be able to produce more accurate estimates of the population biomass trajectories.

In the remainder of this paper, we first briefly describe Bayesian stock assessment with an emphasis on the Schaefer model that we used in our simulation studies in Section 2, then we present our MBE method in Section 3. We compare the standard Bayesian method and the MBE in simple simulation studies in Section 4, and finally discuss extensions of our work and concludes this paper in Section 5.

2 BAYESIAN STOCK ASSESSMENT

The surplus production model (SPM) often uses a time series of catch per unit efforts (CPUEs) $\{I_t\}$, and a time series of catches $\{C_t\}$ to estimate key stock parameters (Prager 1994). The population dynamics model describes how a given population grows and shrinks over time by aggregating spawning, growth, and natural mortality (Cooper 2006). The SPM is commonly used because it is among the simplest and least

data demanding ones that can be used to produce estimates of the Maximum Sustainable Yield and a few associated reference points used in making fishery management decisions (Winker et al. 2018). A surplus production model can be expressed as a discrete-time model with the form

$$B_{t+1} = B_t + SP_t - C_t, \quad (1)$$

where the B_t , SP_t and C_t are the biomass, surplus production and catch in year t respectively. The surplus production term SP_t is used to aggregate the effects of recruitment, growth, and natural mortality. A general form SPM with three parameters is the Pella-Tomlinson model (Pella & Tomlinson 1969):

$$SP_t = \frac{r}{m-1} B_t \left(1 - \left(\frac{B_t}{K} \right)^{m-1} \right), \quad (2)$$

where r is the intrinsic rate of population increase, K is the carrying capacity which can be interpreted as the maximum biomass that can be sustained by the environment, and m is a shape parameter. The case of $m = 2$ is known as the classic two-parameter Schaefer model (Schaefer 1954), which is assumed in this paper.

To take process and measurement errors into account, we consider the following stochastic model which relates fishing efforts and biomasses to the catch and CPUE time series:

$$B_{t+1} = \left[B_t + rB_t \left(1 - \frac{B_t}{K} \right) - C_t \right] e^{\varepsilon_B}, \quad (3)$$

$$C_t = qE_t B_t e^{\varepsilon_C}, \quad (4)$$

$$I_t = \frac{C_t}{E_t} e^{\varepsilon_I} = qB_t e^{\varepsilon_C + \varepsilon_I}, \quad (5)$$

where E_t is the effort at time t , q is the catchability constant, and $\varepsilon_B \sim N(0, \sigma_B^2)$, $\varepsilon_C \sim N(0, \sigma_C^2)$, and $\varepsilon_I \sim N(0, \sigma_I^2)$ are random Gaussian noises used to generate simulated data. The key parameters that need to be estimated are r , K , q , and the initial biomass depletion $\Psi = B_1/K$. In the Bayesian approach, we need to impose prior distributions on these four parameters, as well as the noise variances σ_B^2 , σ_C^2 and σ_I^2 . It is worth noting that, while Equation (4) is needed to provide a complete description of the data generating process, it is not needed in stock assessment.

3 MULTI-PASS BAYESIAN ESTIMATION

As in the standard Bayesian approach, we consider a parametric probabilistic model $p(D | \theta)$ with parameter vector $\theta \in \mathbb{R}^d$, and we assume a prior distribution $p_1(\theta)$ on the parameter θ . In the standard Bayesian approach, we combine the prior p_1 and a dataset D by using the Bayes rule to compute the posterior

$$p_2(\theta) = P(\theta | D) \propto p_1(\theta) \ell(\theta), \quad (6)$$

where $\ell(\theta) = p(D | \theta)$ is the likelihood function given data D .

To motivate our approach, we describe how our method can be derived as a Bayes filter with identical observations at different time steps. Consider a dynamic model with parameter vector θ_k at time step k . An observation D_k is sampled from $p(D_k | \theta_k)$, and then the parameter vector transitions to θ_{k+1} according to a transition model $p(\theta_{k+1} | \theta_k)$. The initial parameter vector θ_1 follows a prior distribution $p_1(\theta_1)$. This can be seen as a hidden Markov model as illustrated in Figure 2a.

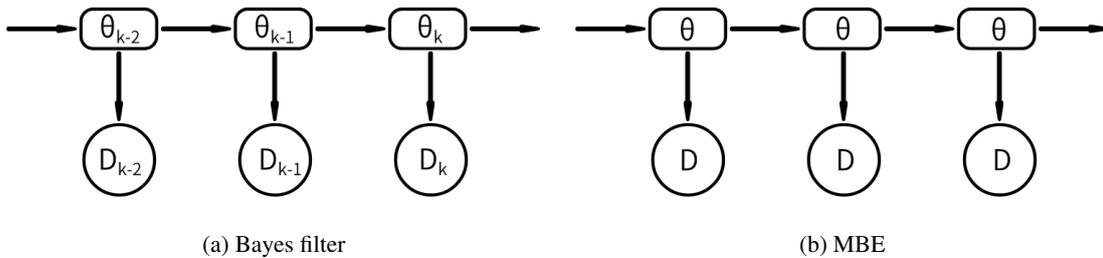


Figure 2. Bayes filter and MBE

The Bayes filter provides a recursive approach to compute the posterior distribution at time k , $p(\theta_k | D_{1:k})$, using the previous posterior $p(\theta_{k-1} | D_{1:k-1})$ (Chen 2003):

$$p(\theta_k | D_{1:k}) \propto \int p(\theta_k | \theta_{k-1})p(D_k | \theta_k)p(\theta_{k-1} | D_{1:k-1})d\theta_{k-1}. \quad (7)$$

We obtain the MBE when (i) the model is static with a fixed parameter vector θ , that is, $p(\theta_k | \theta_{k-1}) = \mathbb{I}(\theta_k = \theta_{k-1})$, where \mathbb{I} is the indicator function, and (ii) the same observation data D are received at each time step, that is, $D_k = D$ for all k as shown in Figure 2b.

For completeness, the pseudo-code of our *multi-pass Bayesian estimation* (MBE) method is shown in Algorithm 1. Essentially, the MBE adjusts the relative importance of the prior and the data by repeatedly performing Bayesian update using the same given dataset. We explain each step below.

Algorithm 1 Multi-pass Bayesian Estimation

- 1: Start with some prior distribution $p_1(\theta)$.
- 2: **for** $j = 1$ to J **do**
- 3: Compute the posterior distribution

$$q_j(\theta) \propto p_j(\theta)\ell(\theta). \quad (8)$$

- 4: Calculate an approximation $p_{j+1}(\theta)$ from $q_j(\theta)$ using a suitable criterion.
-

We first initialize the priors as in a standard Bayesian model (line 1), then we perform J approximate Bayesian updates using the same given dataset D (line 2-4). In each approximate Bayesian update, we first perform a standard Bayesian update to compute $q_j(\theta) \propto p_j(\theta)\ell(\theta)$ (line 3), then computes $p_{j+1}(\theta)$ as an approximation of $q_j(\theta)$ (line 4). Usually, $q_j(\theta)$ cannot be computed in a closed form, and Monte Carlo methods such as Gibbs sampling are used, and $q_j(\theta)$ is represented as sample of θ values drawn from $q_j(\theta)$. To make the approximation of $p_{j+1}(\theta)$ from $q_j(\theta)$ simple, we choose all $p_j(\theta)$'s to have the same parametric form, and use techniques such as moment matching to estimate the parameters of $p_{j+1}(\theta)$ using a sample drawn from $q_j(\theta)$.

The value of J controls the relative importance of the prior and the data: as J becomes larger, the prior becomes less important, and the data becomes more important. In practice, we do not know the true θ value, thus we cannot tune J so that we obtain parameters which are closest to the true values. Instead, a fishery expert can examine the inference results for different J values, and use her domain expertise to select the one that yields most sensible results.

We illustrate how we can use moment matching to estimate the parameters of $p_{j+1}(\theta)$ using a sample drawn from $q_j(\theta)$. If θ contains multiple parameters, we can simply perform moment matching separately for each parameter. Below, we assume that θ contains just one parameter, and $p_{j+1}(\theta)$ is a log-normal distribution, $\text{lognormal}(\mu, \sigma^2)$, and the posterior $q_j(\theta)$ is represented by sample $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(n)}$. In this case, $p_{j+1}(\theta)$ has two parameters, mean μ and variance σ^2 , and we choose to match the expectation and variance of $p_{j+1}(\theta)$ with the mean and variance of a sample drawn from q_j :

$$e^{\mu+\sigma^2/2} = \bar{\theta}, \quad (e^{\sigma^2} - 1)e^{2\mu+\sigma^2} = S^2, \quad (9)$$

where $e^{\mu+\sigma^2/2}$ and $(e^{\sigma^2} - 1)e^{2\mu+\sigma^2}$ are the expectation and variance of $p_{j+1}(\theta)$, and $\bar{\theta} = \frac{\sum \theta^{(i)}}{n}$ and $S^2 = \frac{\sum (\theta^{(i)} - \bar{\theta})^2}{n}$ are the sample mean and the variance of sample for $q_j(\theta)$. Solving the above two equations, we have

$$\mu = \log \left(\frac{\bar{\theta}^2}{\sqrt{S^2 + \bar{\theta}^2}} \right), \quad \sigma^2 = \log \left(\frac{S^2}{\bar{\theta}^2} + 1 \right). \quad (10)$$

4 SIMULATION STUDIES

We perform simulation studies to compare the performance of the MBE and the standard Bayesian method for stock assessment for different levels of error in the CPUE. We focus on the Schaefer model with unknown

K , but known r , q , and Ψ values. For the MBE, we perform $J = 25$ Bayesian updates, that is, we give a higher importance to the data as compared to the standard Bayesian model. The value 25 is somewhat arbitrarily chosen to make sure that the MBE with $J = 25$ exhibits a qualitatively different emphasis on the data across all settings considered in the simulation studies from the standard Bayesian method. A single Bayesian update is implemented using R2JAGS, an R interface to the JAGS (Plummer et al. 2003) library. We selected a relatively large iteration (i.e. 300000) per chain when implement to make sure each chain convergences properly. Parameters of $p_{j+1}(\theta)$ are determined by matching its mean and variance with that of $q_j(\theta)$, as illustrated in Section 3.

4.1 DATA GENERATION

We first deterministically generate yearly effort values (E_t) over a 30-year period: the effort values increases 10% each year in the first 10 years, then decreases by 10% for the next 10 years, and remains stable in the last 10 years. Such a trend represents a typical fishery where effort increases in the developing stage, leading to an overfished stock. With the intervention of management, fishing effort decreases and stabilizes at more sustainable levels. Each effort value is then multiplied with e^{ϵ_E} , where $\epsilon_E \sim N(0, 0.2^2)$. We then used the model in Equation (3)-Equation (5) to generate the catches and CPUE time series. The values of the seven parameters K , r , q , Ψ , σ_C , σ_B , and σ_I are listed in Table 1. Note that we considered 10 different values of σ_I , which correspond to different levels of noise in the CPUE values.

Table 1. The true values of the parameters used to generate the data, and the priors used in the simulation studies.

Parameter	Values	Prior
K	2000	$K \sim \text{lognormal}(\mu = \log(8C_{\max}), CV = 1)$ $K \sim \text{lognormal}(\mu = \log(8C_{\max}), CV = 0.1)$
r	0.4	known parameter
q	0.004	known parameter
Ψ	0.8	known parameter
σ_C	0.1	$\sigma_C^2 \sim \text{invgamma}(\alpha = 0.001, \beta = 0.001)$
σ_B	0.05	$\sigma_B^2 \sim \text{invgamma}(\alpha = 0.001, \beta = 0.001)$
σ_I	0.1, 0.2, ..., 1	$\sigma_I^2 \sim \text{invgamma}(\alpha = 0.001, \beta = 0.001)$

4.2 PRIORS

For a fair comparison, we used the same priors for the standard Bayesian method and our MBE method, as listed in Table 1. In this study, we only estimate the carrying capacity K among the four parameters K , r , q , Ψ , thus the actual values for the other three parameters are provided to the model. Following common practice (Winker et al. 2018), we use a log-normal prior for K , and inverse Gamma priors for the variances σ_C^2 , σ_B^2 and σ_I^2 . We consider two different priors for K : the first one uses a lognormal prior with $\mu = \log(8C_{\max})$ and

a coefficient of variation (CV) of 1, where C_{\max} is the largest catch value and has a $8C_{\max}$ value of around 3800 in our experiments; the second one is a lognormal prior with $\mu = \log(8C_{\max})$ and $CV = 0.1$. The first prior is less concentrated than the second one. The prior for the noise variances σ_C^2 , σ_B^2 and σ_I^2 are all $\text{invgamma}(\alpha = 0.001, \beta = 0.001)$, which are widely used non-informative priors (Winker et al. 2018).

4.3 RESULTS

Figure 3 shows how the standard Bayesian method and the MBE with $J = 25$ perform for different σ_I values when using the weak prior ($\mu = \log(8C_{\max})$ and $CV = 1$) for K . For each σ_I value, we run the two methods on 40 random datasets generated according to Section 4.1. For each dataset and each method, we compute the relative error $(K_{est} - K_{true})/K_{true}$, where K_{est} is the estimated posterior median and K_{true} is the true K value. A boxplot is shown for the 40 relative error values. As expected, when σ_I increases, it becomes harder to estimate the true K value, and thus the performances of both methods drop. Both methods generally provide fairly accurate estimates for the K values: the median errors are less than 0.35% when $\sigma_I = 0.1$, and are less than 9.5% even when $\sigma_I = 1$. On the other hand, the performance of the MBE with $J = 25$ is better than the standard Bayesian method at all noise levels. In particular, the improvement of the MBE gradually

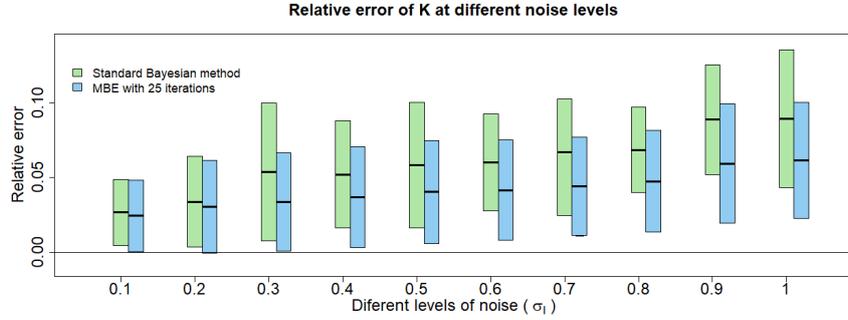


Figure 3. Relative error of the estimated carrying capacity K for the first experiment

increase from less than 0.5% to around 4% as the σ_I changes from 0.1 to 1. It is worth noting that the MBE 25 performance much better than the standard Bayesian method when the $\sigma_I > 0.2$.

Figure 4 shows how the standard Bayesian method and the MBE with $J = 25$ perform for different σ_I values when using the highly informative (but biased) prior ($\mu = \log(8C_{\max})$ and $CV = 0.1$) for K . We also run the two models on 40 different synthetic datasets for different σ_I values using this prior setting. The boxplots of relative errors are shown in Figure 4. The performance of the MBE is significantly better than the standard Bayesian method with this prior setting. As we apply the MBE with 25 iterations, the prior becomes less important. For the standard Bayesian method, when $\sigma_I = 0.1$, the data are less noisy, and with a single Bayesian update, we can differentiate models that are more consistent with the data from models which are less consistent with the data, even when the less consistent models have higher probability in the prior. On the other hand, when data become more noisy, it becomes harder to differentiate the more consistent models and the less consistent models with a single Bayesian update, while the MBE with 25 iterations allows the data to “speak for themselves” even when the prior is incorrect.

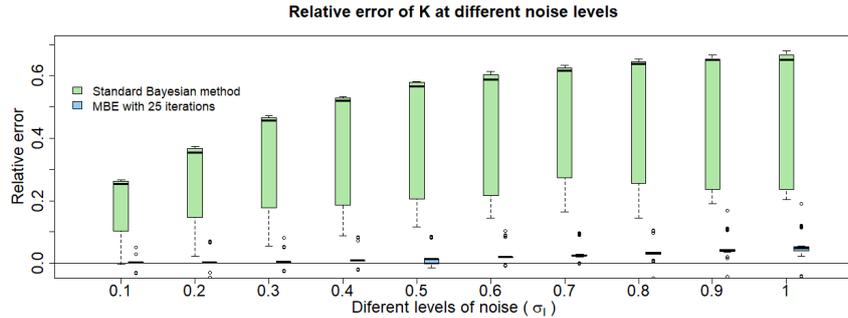


Figure 4. Relative error of the estimated carrying capacity K for the second experiment

To further understand how the MBE works in this setting, we examine how the posteriors evolve in the MBE using a $\sigma_I = 0.7$ dataset as example. Figure 5 shows the first 4 iterations of the MBE with each panel representing one single iteration. In each panel, the dark probability density plot is the input prior of that iteration, and the lighter is the posterior density plot. As we mentioned in Section 3 that the new prior is approximated from previous posterior, hence Figure 5 can be understood as the evolution of the posterior of the first 4 iterations of the MBE. For the first 4 iterations, the change in the posterior within one iteration is quite large. In addition, the posterior is moving away from the initial prior as we keep increasing the number of iterations. This is because the prior is initially concentrated around a value that is somewhat different from the true value, and the plots suggest that by giving a larger relative importance to the data (i.e., performing more Bayesian updates), we can gradually reduce the influence of the prior, and rely more on the data to find out a more accurate estimate on K in the case of inaccurate or less informative initial prior.

5 DISCUSSION AND CONCLUSION

We proposed the MBE method as a new Bayesian method that supports adjusting the relative importance of prior and data. This is particularly useful when we want to give a higher importance to the data. In our simulation studies, we demonstrated that the MBE can provide more accurate results than the standard Bayesian method.

The simulation studies consider various noise levels in the CPUE time series, but they are preliminary, as they focus only on the setting where only the carrying capacity K is unknown among the four key parameters K , r , q , Ψ for the Schaefer model. However, the results suggest that the MBE is more advantageous in cases with higher uncertainty, thus in the cases when we are uncertain about all the four parameters, the MBE may still

perform better than the standard Bayesian method. We plan to perform extensive study using simulated data and real data for estimating multiple parameters together.

Our choice of J in the simulation study is somewhat arbitrary. We mentioned that in practice, fishery experts can examine the stock assessment results produced using different J values and choose the best one based on her domain expertise. It is nevertheless a tedious process to look at the outcomes for many J values. We observed that the estimated parameters often change in a monotonic way as J increase, thus we plan to develop helpful graphic aids to allow fishery experts to focus on parameters which produce more expected estimates.

Finally, we conjecture that the MBE with $J = \infty$ is in fact equivalent to maximum likelihood estimation. This is intuitively reasonable, and we are currently investigating a rigorous formal proof — this will provide a strong theoretical underpinning for the MBE method, if successful.

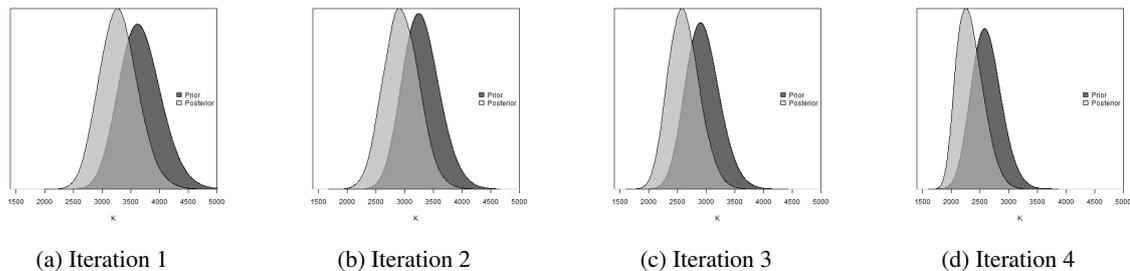


Figure 5. Posterior versus Prior plot for first 4 iterations for a $\sigma_I = 0.7$ dataset.

ACKNOWLEDGEMENT

We would like to thank Jerzy Filar for many helpful discussions and comments.

REFERENCES

- Breen, P. A., Kim, S. W. & Andrew, N. L. (2003), 'A length-based bayesian stock assessment model for the new zealand abalone *haliotis iris*', *Marine and Freshwater Research* **54**(5), 619–634.
- Chen, Z. (2003), 'Bayesian filtering: From kalman filters to particle filters, and beyond'.
- Cooper, A. B. (2006), *A guide to fisheries stock a sssessment: from data to r ecommendations*, University of New Hampshire, Sea Grant College Program.
- McAllister, M. K. & Ianelli, J. N. (1997), 'Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm', *Canadian Journal of Fisheries and Aquatic Sciences* **54**(2), 284–300.
- McAllister, M. K. & Kirkwood, G. P. (1998), 'Bayesian stock assessment: a review and example application using the logistic model', *ICES Journal of Marine Science* **55**(6), 1031–1060.
- McAllister, M. K., Pikitch, E. K. & Babcock, E. A. (2001), 'Using demographic methods to construct bayesian priors for the intrinsic rate of increase in the schaefer model and implications for stock rebuilding', *Canadian Journal of Fisheries and Aquatic Sciences* **58**(9), 1871–1890.
- Method, R. D., J. & Wetzel, C. R. (2013), 'Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management', *Fisheries Research* **142**, 86–99.
- Meyer, R. & Millar, R. B. (1999), 'Bugs in bayesian stock assessments', *Canadian Journal of Fisheries and Aquatic Sciences* **56**(6), 1078–1087.
- Pella, J. J. & Tomlinson, P. K. (1969), 'A generalized stock production model', *Inter-American Tropical Tuna Commission Bulletin* **13**(3), 416–497.
- Plummer, M. et al. (2003), Jags: A program for analysis of bayesian graphical models using gibbs sampling, in 'Proceedings of the 3rd international workshop on distributed statistical computing', Vol. 124, Vienna, Austria., pp. 1–10.
- Prager, M. (1994), 'A suite of extensions to a nonequilibrium surplus-production model', *Fishery Bulletin* **92**, 374–389.
- Punt, A. E. & Hilborn, R. (1997), 'Fisheries stock assessment and decision analysis: the bayesian approach', *Reviews in fish biology and fisheries* **7**(1), 35–63.
- Schaefer, M. B. (1954), 'Some aspects of the dynamics of populations important to the management of the commercial marine fisheries', *Inter-American Tropical Tuna Commission Bulletin* **1**(2), 23–56.
- Winker, H., Carvalho, F. & Kapur, M. (2018), 'Jabba: just another bayesian biomass assessment', *Fisheries Research* **204**, 275–288.