Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Including parameter uncertainty in forward projections of computationally intensive statistical population dynamic models

Mark N. Maunder, Shelton J. Harley, and John Hampton

The increased computational demands of modern statistical stock assessment models have made the standard methods to provide uncertainty estimates for forward projections impractical for timely results in many applications. However, forward projections and their associated estimates of uncertainty are an important and popular piece of management advice. We describe a less computationally intense method to estimate uncertainty in forward projections that includes both parameter uncertainty and future demographic stochastic uncertainty. This frequentist method uses penalized likelihood as an approximation to mixed effects and can be viewed as treating the future projection period as part of the estimation model rather than performing stochastic projections. This allows confidence intervals to be calculated using normal approximation based on the delta method. The method is tested using simulation analysis and compared with Bayesian analysis and with projections based on point estimates of the parameters. The method is applied to yellowfin tuna in the eastern Pacific Ocean.

M. N. Maunder and S. J. Harley: Inter-American Tropical Tuna Commission, 8604 La Jolla Shores Drive, La Jolla, CA 92037-1508, USA. J. Hampton: Oceanic Fisheries Programme, Secretariat of the Pacific Community, B.P. D5, Noumea, New Caledonia. S. J. Harley also at Ministry of Fisheries, PO Box 1020, Wellington, New Zealand. Correspondence to M. N. Maunder: tel: +1 858 5467027; fax: +1 858 5467133; e-mail: mmaunder@iattc.org.

Introduction

Modern quantitative methods used to assess the status of commercial fisheries have become very complex and computationally intensive (Quinn, 2003). Such methods are usually statistically non-linear, and can have hundreds or thousands of parameters (e.g. Fournier *et al.*, 1998; Hampton and Fournier, 2001; Butterworth *et al.*, 2003; Maunder and Watters, 2003). Sophisticated iterative non-linear function optimizers are needed to estimate the model parameters and to provide management advice. These analyses often take days to estimate parameters, which limit the number and type of analyses that can be performed. For example, only a limited number of sensitivity analyses can be performed to investigate a model's robustness to structural assumptions.

A current trend in fisheries stock assessment is to estimate and present uncertainty (Francis and Shotton, 1997; Punt and Hilborn, 1997). Consideration of uncertainty is essential when making management decisions and is particularly important for applying the precautionary approach to fisheries

management. There are several methods that can be used to estimate uncertainty (Hilborn and Mangel, 1997). Three of the most common, profile likelihood, bootstrap, and Bayesian analysis, are computationally intense. The profile likelihood and bootstrap methods require the objective function to be optimized numerous times, while Bayesian analysis often requires model equations to be calculated millions of times. Because of limited availability of computational resources, these methods are not practical for complex stock assessment models, and less computationally intense methods are required. For example, the normal approximation method based on the delta method is often used (Fournier *et al.*, 1998). These methods are traditionally used only to estimate uncertainty in model parameters and current and historical stock status.

A common piece of scientific advice considered in fisheries management is the predicted outcome of future management actions and the uncertainty in these predictions. This requires projecting the population into the future. Unlike estimates of model parameters and historical biomass, estimates of uncertainty for forward projections should also

include stochastic variation of demographic processes in the future (e.g. annual variation in recruitment caused by environmental stochasticity). However, the main methods used to represent the full uncertainty, both parameter and demographic, the bootstrap and Bayesian analysis, are too computationally intense for complex stock assessment models. Less computationally intense methods that are used are often inadequate. For example, methods that project from parameter point estimates ignore parameter uncertainty, which can, in some cases, represent the majority of the uncertainty. Other methods ignore future demographic uncertainty (Gavaris, 1993).

Here we describe a method for calculating uncertainty in forward projections that includes both parameter and demographic uncertainty that is much less computationally intensive than the bootstrap and Bayesian analysis. This frequentist method uses penalized likelihood as an approximation to mixed effects, and can be viewed as treating the future projection period as part of the estimation model rather than performing stochastic projections. This allows confidence intervals to be calculated using normal approximation based on the delta method, as would be done for historical biomass. This method is in the flavour of the likelihoodist framework of statistics (Tanner, 1993; Royall, 1997), an alternative to Bayesian and frequentist (e.g. bootstrap) frameworks. However, it can also be used in a frequentist or objective Bayesian context (Schweder and Hjort, 2002). It extends Gavaris' (1993) "Function of Parameters Method" to include future uncertainty in recruitment and catchability. We use simulation analysis to investigate and test the method and to compare the results with the results of Bayesian analysis and with forward projections from point estimates of the model's parameters. The method is applied to yellowfin tuna (*Thunnus albacares*) in the eastern Pacific Ocean.

Methods

Random effects

To understand the methods introduced in this paper, it is important first to understand the concept of random effects in statistics. Traditional parameters of a stock assessment model are assumed to be fixed effects (e.g. the carrying capacity of a surplus production model), i.e. to have an unknown but fixed true value (Pawitan, 2001). Uncertainty in these parameters comes about by incomplete information in the data. For prediction, the true value of these parameters does not change. In contrast, parameters that are considered random effects may change for predictions (i.e. the value for the prediction may differ from the estimate and among predictions). An obvious example of a random effect in fisheries stock assessment models is annual recruitment, which differs from year to year. Another example is process error in the relationship between fishing mortality and effort (i.e. catchability).

Inclusion of random effects in an estimation model might be somewhat confusing. Take, for example, the annual deviations in recruitment. For each year, there will be a realized value for the recruitment deviation. This will be a number, and it is not a random variable when the year has been completed. However, because recruitment is not directly observed, there is no direct measurement of the deviation. Information about the deviation comes from other data (e.g. catch-at-age data). When random effects are included in an estimation model, they have additional contextual information attributed to the distributional assumption made for the random effect (Pawitan, 2001). Even in the absence of data about these parameters, there is information available from the distributional assumption. One way to view random effects in estimation models is as a method to deal with large numbers of parameters, i.e. the realized values of all random effects (Pawitan, 2001), or a way to share information among parameters.

When viewed conditional on the values of unobserved random effects, the likelihood of observed data is a function of all parameters and the hypothetical values of the random effects. To obtain the likelihood function in the fixed effects, including possibly parameters for variances of random effects, this conditional likelihood can be averaged over the joint distribution of the random effects. The same averaging is not possible for fixed effects because they do not have a distributional assumption.

If both the estimated parameters and the predictions are treated as random effects, there is no difference between the two. Therefore, by way of an example, if recruitment in a stock assessment model is treated as a random effect, future recruitments (predictions) are no different from historical recruitments. We exploit this to include demographic uncertainty in projections using a less computationally intensive method than those often used. Unfortunately, the appropriate method to deal with random effects, integration across the random effects, is often not practical for the type of non-linear dynamic models used for stock assessment (see Maunder and Deriso, 2003). Therefore, using an approach traditional in fisheries stock assessment, we approximate a mixed effect model using penalized likelihood, the penalty being based on the random effect distributional assumption. This approach is equivalent to maximizing the posterior density with a normal prior on the random effects parameters and uniform priors on the fixed effect parameters. In some cases it is possible to obtain reasonable estimates of the standard deviation of the random effects distribution using penalized likelihood, but this estimator is negatively biased and inconsistent (Berger *et al.*, 1999; Maunder and Deriso, 2003). Therefore, if practical, integrating out the random effects parameters is the appropriate method to estimate the standard deviation (Maunder and Deriso, 2003).

Stock assessment model

For illustrative purposes, a simple catch-at-age stock assessment model is developed. The model is age-structured with two fisheries. Recruitment is related to stock size using a Beverton−Holt stock-recruitment relationship. The model is fitted to total catch and catch-at-age data for each year

and fishery. Catch is predicted on the basis of known effort, following Fournier *et al.* (1998).

The population is assumed to be in an unexploited equilibrium at the start of the modelling time period:

$$N_{1,a} = R_0 \exp[-M(a-1)], \tag{1}$$

with the last age (A) acting as an accumulating plus group

$$N_{1,A} = \frac{R_0 \exp[-M(A-1)]}{1 - \exp(-M)}, \tag{2}$$

where $N_{t,a}$ is the number of individuals at time t in age class a, M the natural mortality rate, and R_0 is the average recruitment in an unexploited population.

Recruitment at age one is assumed to follow the Beverton–Holt stock-recruitment relationship:

$$N_{t+1,1} = \frac{S_t}{\alpha + \beta S_t} \exp\left(\varepsilon_{R,t} - 0.5\sigma_R^2\right) \tag{3}$$

$$\alpha = \frac{S_0(1-h)}{4hR_0} \tag{4}$$

$$\beta = \frac{5h-1}{4hR_0} \tag{5}$$

$$S_t = \sum_a N_a m_a w_a, \tag{6}$$

where $\varepsilon_{R,t}$ is the annual iid $N(0,\sigma_R^2)$ recruitment deviate for year t, $-0.5\sigma_R^2$ the lognormal bias correction factor that makes the expected recruitment equal to the stock-recruitment relationship (otherwise the median recruitment would be equal to the stock-recruitment relationship), h the recruitment as a proportion of R_0 achieved when the spawning biomass (S_t) is 20% of the average spawning biomass in an unexploited population (Francis, 1992), S_0 the spawning biomass for an unexploited population, w_a the weight at age a, and m_a is the proportion mature at age a.

The abundance is modelled using a simple exponential decay model, which is a function of fishing mortality ($F_{t,a}$) at time t and age a and natural mortality (M), which is constant over time and age:

$$N_{t+1,a+1} = N_{t,a} \exp[-(F_{t,a} + M)], \tag{7}$$

with the last age (A) acting as an accumulating plus group:

$$\begin{aligned} N_{t+1,A} &= N_{t,A-1} \exp[-(F_{t,A-1} + M)] \\ &\quad + N_{t,A} \exp[-(F_{t,A} + M)]. \end{aligned} \tag{8}$$

Fishing mortality is assumed to be separable into age- and time-specific components and proportional to effort.

Fishing mortality is allowed to vary from this relationship using an annual estimated deviate that is penalized using a distributional assumption (Fournier *et al.*, 1998):

$$F_{g,t,a} = q_g E_{g,t} s_{g,a} \exp\left[\varepsilon_{E,g,t} - 0.5\sigma_{E,g}^2\right] \tag{9}$$

$$F_{t,a} = \sum_g F_{g,t,a}, \tag{10}$$

where $s_{g,a}$ is the selectivity of individuals in age class a to gear g, $E_{g,t}$ the effort in time t for gear g, q_g the catchability coefficient for gear g, $\varepsilon_{E,g,t}$ the effort deviate for gear g in year t, $-0.5\sigma_{E,g}^2$ the lognormal bias correction factor for gear g, and $\sigma_{E,g}$ is the standard deviation of the logarithm of the annual effort deviates for gear g.

The following describes the objective function used to fit the model to the data. L is used to represent the likelihood and P is used to represent a prior distribution or penalty. The model is fitted to catch data in weight assuming that the observed catch (C^{obs}) is lognormally distributed around the expected catch (C) with negative log-likelihood (ignoring constants):

$$-\ln L\left(\theta \big| C^{\text{obs}}\right) \propto \sum_{g,t} \frac{\left[\ln(C_{g,t}^{\text{obs}}) - \ln(C_{g,t})\right]^2}{2\sigma_{C,g}^2} \tag{11}$$

$$C_{g,t} = \sum_a C'_{g,t,a} w_a \tag{12}$$

$$C'_{g,t,a} = \frac{F_{g,t,a}}{F_{t,a} + M} N_{t,a}[1 - \exp(-F_{t,a} - M)], \tag{13}$$

where $F_{t,a} = \sum_g F_{g,t,a}$, $C'_{g,t,a}$ is the catch in number of individuals from age class a in time t for gear g, and θ represents the model parameters.

The model is fitted to catch-at-age data using a multinomial-based negative log-likelihood (ignoring constants):

$$-\ln L\left(\theta \big| C^{\text{obs}}\right) \propto -\sum_{g,t,a} C_{g,t,a}^{\text{obs}} \ln(p_{g,t,a}) \tag{14}$$

$$p_{g,t,a} = \frac{C'_{g,t,a}}{\sum_a C'_{g,t,a}}. \tag{15}$$

Abundance information for catch-and-effort data is modelled using a penalty on the annual effort deviates (ignoring constants and assuming that $\sigma_{E,g}$ is known; see Fournier *et al.*, 1998):

$$-\ln P(\varepsilon_E) \propto \sum_{g,t} \frac{\varepsilon_{E,g,t}^2}{2\sigma_{E,g}^2}. \tag{16}$$

The size of $\sigma_{E,g}$ determines how much influence the catch-and-effort data for gear g has on the biomass estimates. In general, if $\sigma_{E,g}$ is small, the estimated biomass changes to predict the observed catch; if $\sigma_{E,g}$ is large, the estimated fishing mortality, through more flexibility in $\varepsilon_{E,g,t}$, changes to predict the observed catch.

A penalty is included based on the assumption that recruitment is lognormally distributed around the stock-recruitment relationship (ignoring constants):

$$-\ln P(\varepsilon_R) \propto n \ln(\sigma_R) + \sum_t \frac{\varepsilon_{R,t}^2}{2\sigma_R^2}, \tag{17}$$

and is equivalent to a prior on the annual recruitment deviate $\varepsilon_{R,t}$. Here, σ_R is the standard deviation of the logarithm of the annual recruitment deviates and n is the number of time periods. Unlike the penalty for the effort deviates, we have to include the term $\ln(\sigma_R)$, because σ_R is estimated.

The model estimates the values of R_0, $\varepsilon_{R,t}$, $\varepsilon_{E,g,t}$, σ_R, h, and q_g. The values for $s_{g,a}$, M, w_a, $\sigma_{E,g}$, and m_a are assumed to be known (Table 1). To aid the estimation procedure, σ_R was restricted between 0.2 and 2.0. This is because the global maximum likelihood (without integration over ε_R) is at $\sigma_R = 0$, but for informative data, there is a local maximum around the true value (Maunder and Deriso, 2003).

Normal approximation

We here present a method where the demographic uncertainty in forward projections that is traditionally represented by randomly selecting these parameters from the appropriate distribution is represented by parametric uncertainty. This approach extends Gavaris' (1993) "Function of Parameters Method" to include future uncertainty in recruitment and catchability. This frequentist method uses penalized

Table 1. Parameter values used in the simulator and the estimator. If a value for a parameter is assumed for the estimator it is set equal to the value for the simulator.

Parameter	Simulator value	Estimated
R_0	1 000	y
ε_R	$N(0,\sigma_R)$	y
$\varepsilon_{E,1}$	$N(0,\sigma_{E,1})$	y
$\varepsilon_{E,2}$	$N(0,\sigma_{E,2})$	y
σ_R	1.0	y
h	0.6	y
q_1	0.1	y
q_2	0.05	y
$\sigma_{E,1}$	0.4	n
$\sigma_{E,2}$	0.6	n
$\sigma_{C,1}$	0.07	n
$\sigma_{C,2}$	0.07	n
s_1, s_2	Table 2	n
M	0.3	n
w	Table 2	n
m	Table 2	n

likelihood as an approximation to mixed effects, and can be viewed as treating the future projection period as part of the estimation model, rather than performing stochastic projections. This allows confidence intervals to be calculated using normal approximation based on the delta method, as would be done for historical biomass. The normal approximation method is implemented by extending the timeframe of the model for estimation to include the future. As for the past, recruitment for the future, which can be viewed as a random effect, is estimated as an annual deviate with a prior distribution (penalty). The prior distribution is the same as used in the historical period, and it describes the uncertainty in future recruitment. However, there are no data in the future to provide information about recruitment. The only additional data included in the model are those related to the effort assumed in the future. Effort deviates for the future, which can be viewed as a random effect, are also estimated with the same prior used in the historical period. The parameters of the stock assessment model are estimated using penalized maximum likelihood estimation (MLE), with the addition of the recruitment deviate and effort deviate penalties to approximate the mixed effect model. This is equivalent to finding the mode of a joint posterior distribution with locally uniform priors on all other parameters.

The confidence intervals for the future abundance are calculated using normal approximation and the delta method (Ratkowsky, 1983; see Dupont, 1983, and Fournier *et al.*, 1998, for examples of the use of this method in fisheries applications). Assuming that the MLE is asymptotically efficient, the Cramér-Rao lower bound can be used as an approximation to the true variance of the MLE (Casella and Berger, 1990):

$$\widehat{\text{Var}}(f(\hat{\boldsymbol{\theta}})|\boldsymbol{\theta}) \approx \frac{\partial f(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\Big|_{\theta=\hat{\theta}} \mathbf{I}^{-1} \frac{\partial f(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}^T \Big|_{\theta=\hat{\theta}}, \tag{18}$$

where $\mathbf{I} = -(\partial/\partial\boldsymbol{\theta}_i \partial\boldsymbol{\theta}_j) \log L(\boldsymbol{\theta}|\mathbf{X})|_{\theta=\hat{\theta}}$, $f(\boldsymbol{\theta})$ is the quantity of interest, here future abundance, defined as a function of the parameters of the model $\boldsymbol{\theta}$, and $\mathbf{X}$ is the data. The estimate of the variance can then be used based on asymptotic and regularity conditions to form the confidence interval (Casella and Berger, 1990) based on the normal distribution:

$$f(\hat{\boldsymbol{\theta}}) - z_{\alpha/2}\sqrt{\text{Var}(f(\hat{\boldsymbol{\theta}})|\boldsymbol{\theta})} \leq f(\boldsymbol{\theta})$$
$$\leq f(\hat{\boldsymbol{\theta}}) + z_{\alpha/2}\sqrt{\text{Var}(f(\hat{\boldsymbol{\theta}})|\boldsymbol{\theta})}. \tag{19}$$

These confidence intervals with $\alpha = 0.05$ and $z = 1.96$ are used to represent the 95% projection interval of future abundance. We define the projection interval as the interval where there is a 95% probability that the true future value lies within this interval. The frequentist-based confidence

interval is used as an approximation of the projection interval.

Bayesian analysis

We use Markov Chain Monte Carlo (MCMC) to estimate the joint posterior distribution of model parameters (see Punt and Hilborn, 1997). Uniform priors are assumed for all model parameters except for the annual recruitment and effort deviates, which have normally distributed priors as described above. For each sample of the model parameters from the posterior distribution, we project the population forward in time sampling a recruitment deviate each year from the distribution $N(0, \sigma_R)$, and iid effort deviates, $\varepsilon_{E,g,t}$, from the distribution $N(0, \sigma_{E,g})$. The stock-recruitment relationship and bias correction factors are used when projecting recruitment and effort. The 5th and 95th percentiles of these projections are used to represent the 95% Bayesian credibility interval of the abundance. More information on Bayesian analysis relevant to this application can be found in Punt and Hilborn (1997) and Maunder and Deriso (2003).

Projections from parameter point estimates

First, the parameters of the stock assessment model are estimated using penalized maximum likelihood with the addition of the recruitment deviate and effort deviate penalties. As mentioned above, this is equivalent to finding the mode of the joint posterior distribution. Then, the population is projected forward in time using the MLE parameter values, and sampling a recruitment deviate each year from the distribution $N(0, \sigma_R)$ and effort deviates from $N(0, \sigma_{E,g})$. The stock-recruitment relationship and bias correction factors are used when projecting recruitment and effort. The forward projection procedure is repeated for a sample $(n = 500)$ of recruitment and effort deviates, each time keeping model parameters at the MLE values, and the 5th and 95th percentiles of these projections are used to represent the 95% projection interval of the abundance.

Simulation analysis

A simulation model is developed using the same population dynamics model (Equations (3)−(11)) used in the estimation (see Maunder and Deriso, 2003, for the general approach). The simulation model is run for 35 years (20 for estimation, 15 for projections), based on known effort (Figure 1), to produce total catch and catch-at-age data using the parameter values in Tables 1 and 2. Only the first 20 years of catch and catch-at-age data are fitted in the estimation procedures. Catchability is modelled as coming from a lognormal distribution with a standard deviation on the log of the catchability of 0.4 and 0.6 for gears 1 and 2, respectively. The total observed catch is normally distributed around the true catch with a standard deviation of 0.07, and the catch-at-age data are generated using a multinomial distribution with sample size of 50 each year.
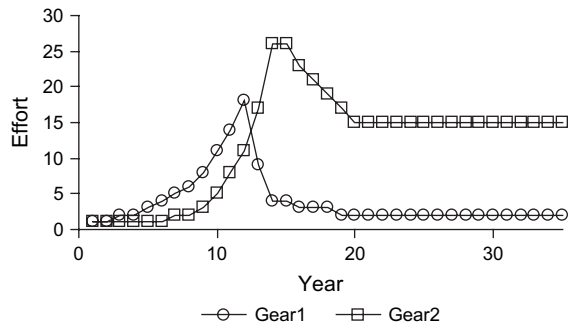

Figure 1. Effort time-series used for the two fisheries in the simulation analysis.

In all, 500 (200 for the Bayesian analysis) artificial data sets are generated, and the estimation methods are applied to these data sets. We investigate three methods: (i) normal approximation; (ii) Bayesian analysis; and (iii) forward projections from point estimates. The point-estimate method only includes uncertainty in the future recruitment and the future effort−fishing mortality relationship. The Bayesian analysis and normal approximation methods also include uncertainty in parameter estimates.

The normal approximation method is further investigated by testing two different versions of the recruitment and effort deviate penalties: (i) including the full penalty for all years including the future (normal approximation method); (ii) eliminating $\ln(\sigma_R)$ from the penalty of the future

Table 2. Age-specific model parameters used in the simulation study. The values are based on a generic species.

| Age | Selectivity | | Maturity | Weight |
	Gear 1	Gear 2		
1	0.33	0	0	0.000597
2	0.66	0	0	0.00315
3	1	0.25	1	0.007168
4	0.66	0.5	1	0.011705
5	0.33	0.75	1	0.016082
6	0	1	1	0.019947
7	0	1	1	0.023179
8	0	1	1	0.025786
9	0	1	1	0.027839
10	0	1	1	0.029428
11	0	1	1	0.030643
12	0	1	1	0.031565
13	0	1	1	0.032259
14	0	1	1	0.03278
15	0	1	1	0.03317
16	0	1	1	0.03346
17	0	1	1	0.033676
18	0	1	1	0.033837
19	0	1	1	0.033957
20	0	1	1	0.034046

recruitments, and eliminating the bias correction factors for the future recruitments and future effort deviates (adjusted normal approximation method). Dropping $\ln(\sigma_R)$ from the penalty of the future recruitments is motivated by the fact that there is no information in the future to estimate the recruitments, so minimizing the log-likelihood that includes $\ln(\sigma_R)$ for these years will cause σ_R to be estimated at a lower value. Not using the bias correction factor for the future recruitments (and effort deviates) is motivated by the fact that when there is no information to estimate the recruitments, the recruitment deviate is estimated to be zero and the bias correction factor will cause the recruitments to be lower than expected by the stock–recruitment relationship.

We present the percentage of times that the true value lies within, above, and below the 95% projection interval. We also present the % bias when future abundance is estimated by the MLE estimates from the normal approximation method, the average of the projections for the point estimation method, and the average of the projections for the Bayesian method.

Application

The stock assessments of yellowfin, bigeye (*Thunnus obesus*), and skipjack (*Katsuwonus pelamis*) tuna in the eastern Pacific Ocean are some of the most computationally intensive and highly parameterized stock assessment models (Maunder and Watters, 2003). The stock assessment model has more than 1000 parameters for these applications. Historically, forward projections using this model have been based on parameter point estimates, and only include demographic uncertainty for recruitment in the future projections. Therefore, the results ignore parameter uncertainty. Initial analyses using the normal approximation method have been presented for bigeye tuna (Maunder and Harley, 2002), but σ_R was fixed, both $\ln(\sigma_R)$ and the bias correction factor were used for the future recruitments, and the normal approximation method had not been tested. We apply the adjusted normal approximation method to the stock assessment of yellowfin tuna in the eastern Pacific Ocean (see Maunder, 2002, for a description of the assessment, and Maunder and

Watters, 2003, for further technical details) and compare it with the point estimation method. In both models we ignore future variation in catchability represented by effort deviates. This is because fishing mortality is not proportional to effort for some of the fisheries, so the standard deviation for the effort deviates in these fisheries is fixed at a relatively large value to remove any biomass information from the catch-and-effort data. Not including effort deviates in the future will underestimate uncertainty. This model has 1919 parameters and takes 7 h to converge on a 2.8-GHz Pentium 4 desktop computer. The majority of the model parameters are quarterly recruitment and effort deviates.

Results

Simulations

The coverage of the biomass projection interval for the point estimation method is very poor in the first few years, a large number of true biomass points being outside the projection interval (Table 3). The coverage becomes close to the desired coverage (95%) as the projection timeframe is increased, but it is still slightly lower than the desired level. The normal approximation method has better coverage for the first few years, but has poorer coverage as the projection timeframe is increased. The projection intervals are asymmetrical, with a greater chance of the true value being above the projection interval. The coverage is greatly improved and close to the desired level if $\ln(\sigma_R)$ is removed from the penalty on the recruitment deviations and the bias correction factors are removed in the projection time period. However, the projection intervals remain asymmetrical in their coverage. The coverage of the Bayesian method is higher than the desired level, and higher than that of the other methods. The coverage performance is symmetrical, with equal numbers of true biomasses above and below the projection intervals.

The size of the projection intervals is smaller for the normal approximation method than for the point-based method, except for the first few years (Table 3). However, the projection intervals are similar in size for the adjusted normal approximation method compared with the point-based

Table 3. Estimates of projection interval coverage for the spawning biomass for projections of 2, 5, and 10 years. "Below" means the proportion of times the true value is less than the lower bound of the projection interval, "inside" the proportion of times the true value is inside the projection interval, "above" the proportion of times the true value is greater than the upper bound of the projection interval, and "size" means the size of the projection interval. The required coverage is 95%.

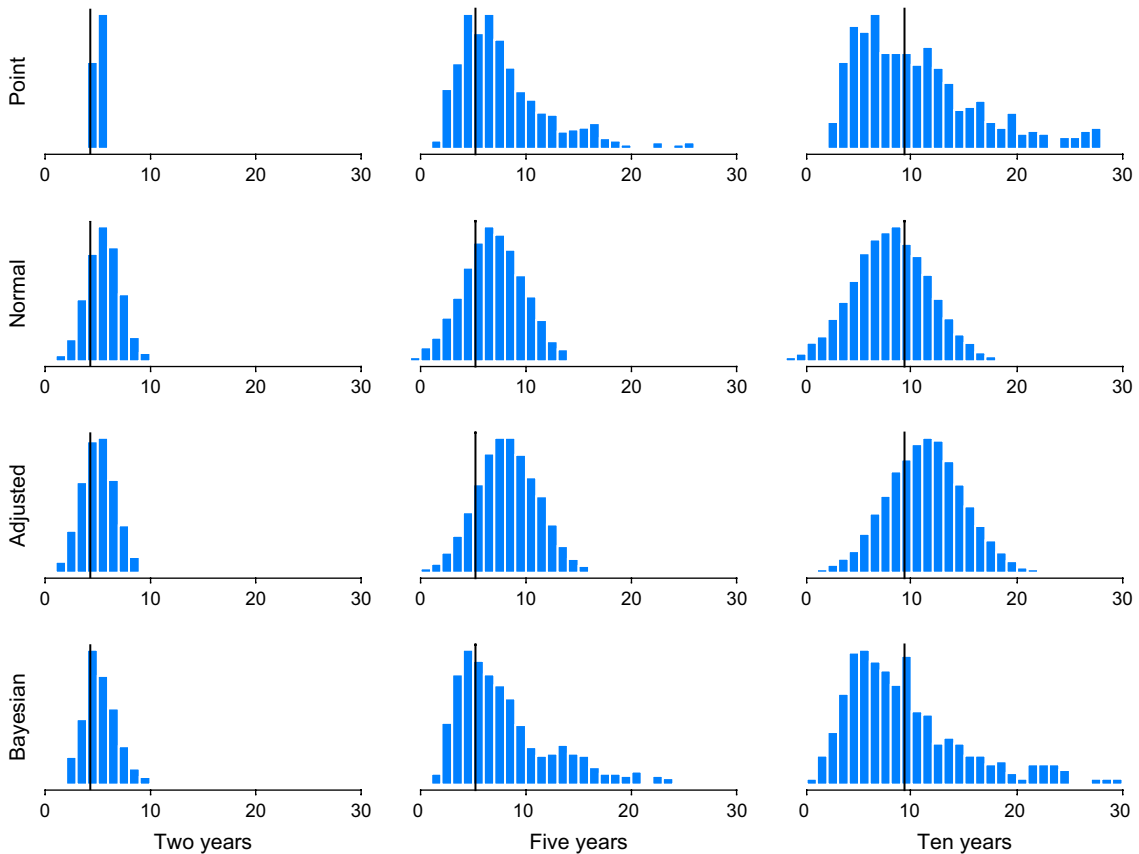|	Point			Normal			Adjusted			Bayesian		
	2 Years	5 Years	10 Years	2 Years	5 Years	10 Years	2 Years	5 Years	10 Years	2 Years	5 Years	10 Years
Below	0.16	0.05	0.04	0.02	0.02	0.02	0	0	0	0.02	0.02	0.02
Inside	0.53	0.94	0.94	0.92	0.84	0.8	0.92	0.95	0.94	0.96	0.97	0.97
Above	0.31	0.02	0.03	0.06	0.14	0.18	0.08	0.05	0.06	0.02	0.02	0.02
Size	48	287	298	131	177	173	129	263	278	135	343	348

Figure 2. Comparisons of the projection distribution of spawning-stock biomass from the different estimation methods for a single simulated data set. The dark vertical lines represent the true values.

method, except for the first few years. The confidence intervals are largest for the Bayesian method. Figure 2 shows the projection distribution for the different methods for a single simulated data set, and highlights the difference in projection uncertainty and the asymmetrical nature of the Bayesian and point estimation methods.

There is significant negative bias in the estimates by the normal approximation methods of future spawning biomass (Figure 3). When the bias correction factors and $\ln(\sigma_R)$ are removed, there is a smaller but positive bias in the spawning biomass (Figure 3). This bias can be attributed to the stock-recruitment relationship (Figure 3). Bias in estimated recruitment follows a similar pattern to bias in spawning biomass (Figure 3), and recruitment does not appear to be biased when the steepness, h, is set to 1 (recruitment is independent of stock size) for the adjusted normal approximation method (Figure 3).

Estimates of σ_R are negatively biased for all methods except the Bayesian (Figure 4). The negative bias is much greater, and R_0 is also negatively biased (Figure 4) if $\ln(\sigma_R)$ is included in the recruitment deviate penalty in the projection period.

Application

The spawning biomass confidence intervals of the adjusted normal approximation method were much larger in the future projections than for the historical period (Figure 5). The mean of the point-estimate-method projections was essentially identical to the MLE of the fully adjusted normal approximation method. The confidence interval from the point-estimate method took about two years to become as wide as the adjusted normal approximation method. The lower bound of the confidence interval estimated by the point-estimate method was not as low as the adjusted normal approximation method.

Discussion

We introduce a method to perform forward projections that accounts for both parameter uncertainty and demographic stochastic uncertainty that is less computationally intensive than Bayesian and bootstrap methods. The method is a simple extension of current methods based on penalized likelihood, and it performs reasonably well as long as it is
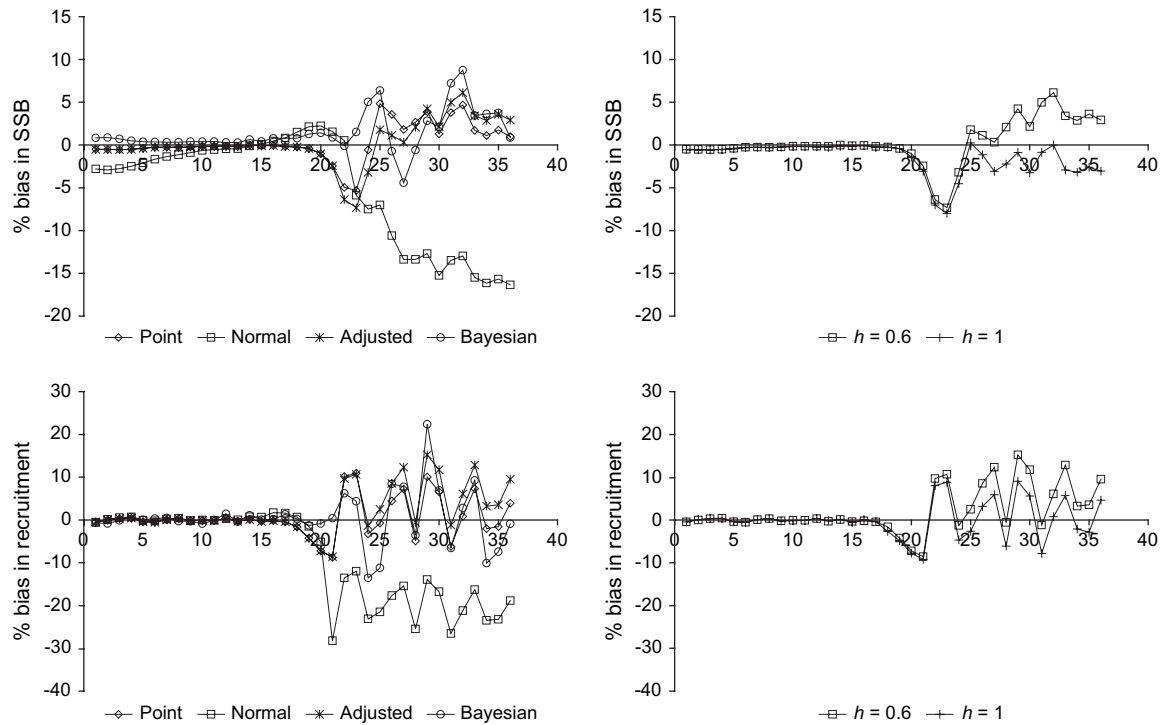
Figure 3. Bias in the estimates of spawning-stock biomass (SSB) and recruitment from the four estimation methods (left panels) and using the adjusted normal approximation method when the model has a Beverton−Holt stock-recruitment relationship ($h = 0.6$), and when recruitment is independent of stock size ($h = 1$) (right panels).

adjusted to remove potential bias. The estimation routine minimizes the negative log-likelihood which includes the term $\ln(\sigma_R)$. However, for future recruitment, there is no additional information to provide estimates of the recruitment deviation. Therefore, the model parameter estimates should be the same as those achieved when the model is fitted without the projection period. This is not the case, because minimizing the negative log-likelihood with $\ln(\sigma_R)$ for these years causes the σ_R to be estimated lower. This is because there are additional years where $\varepsilon_{R,t}^2/2\sigma_R^2 = 0$ because there is no information in the data about $\varepsilon_{R,t}$, so that a smaller value for σ_R reduces the penalty by reducing $\ln(\sigma_R)$. This result highlights a consequence of the approach commonly used in fisheries to estimate the recruitment

residuals as free parameters. If the model includes years where there is no or little information, the estimate of σ_R would be biased low. This is consistent with the results of Maunder and Deriso (2003), and methods that integrate out the recruitment deviations (e.g. Bayesian analysis or mixed effects) should be better estimators of σ_R (see Maunder and Deriso, 2003).

A bias correction factor for the lognormal distribution is added to the recruitment to keep the expected recruitment equal to that of the stock-recruitment relationship. This assumes that all the recruitment deviations from the underlying stock-recruitment relationship are iid lognormal. However, if there is little information in the data about recruitment for certain years, these years will be estimated at
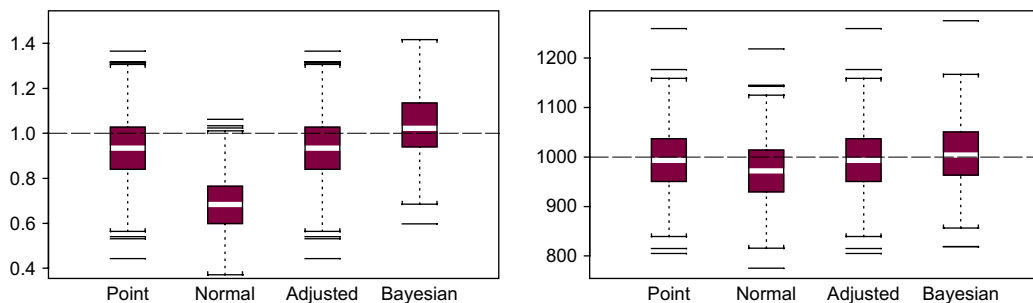

Figure 4. Estimates of σ_R (left panel) and R_0 (right panel) for each simulated data set using the different estimation methods.
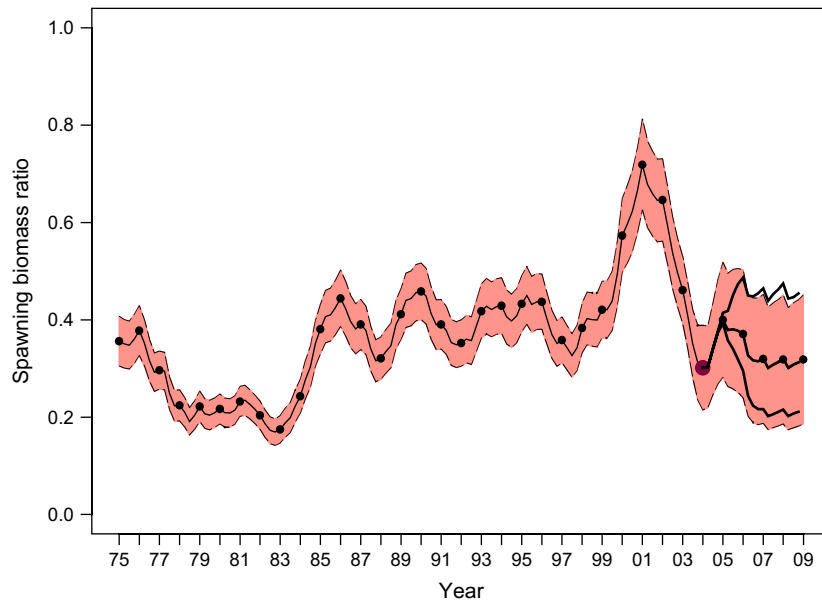
Figure 5. Comparisons of estimates and projection intervals for spawning biomass ratio (spawning biomass divided by the unexploited spawning biomass) from the point-estimate and adjusted normal approximation estimation methods for yellowfin tuna in the eastern Pacific Ocean. The shaded area represents the 95% projection intervals for the fully adjusted estimation method, the thin line with solid points represents the MLE estimates from the fully adjusted estimation method, the large solid circle represents the end of the historical estimation period, and the solid lines represent the 2.5% and 97.5% percentiles and the mean from the point estimation method.

values below the expected value owing to the effect of the bias correction factor and that the penalty on the recruitment deviates is centred at zero. Therefore, because there is no information in the future, the maximum likelihood estimate of recruitment in the future will be biased low if the bias correction factor is used. This result highlights an additional consequence of the approach commonly used in fisheries to estimate the recruitment deviates as free parameters. If the model includes years where there is no or little information, the estimate of recruitment will be biased low for these years. This is particularly important for the most recent years, for which the catch-at-age data provide little information and these recruitments are influential for biomass projections in the near future. Again, methods that integrate out the recruitment deviations (e.g. Bayesian analysis or mixed effects) should be better estimators.

Deterministic projections of a population dynamics model will underestimate the effect of a stock-recruitment relationship. In a stochastic projection when random recruitment is lower than that expected from the stock-recruitment relationship (i.e. a negative annual recruitment deviate), the spawning biomass will also drop more than expected, causing the stock-recruitment relationship to make the expected recruitment lower in following years. The opposite also occurs when the recruitment deviate is positive. However, owing to the shape of the Beverton−Holt stock-recruitment curve, this has more impact for lower-than-expected recruitment. Without stochastic recruitment this does not occur, so

the normal approximation method will, if corrected for other biases, on average overestimate recruitment in future if there is a Beverton−Holt stock-recruitment relationship.

Maunder and Deriso (2003) found that estimation of σ_R was possible when maximizing the penalized likelihood, but estimability was reduced when the catch-at-age data were not available for all years. They showed that a local optimum often occurs close to the true value of σ_R, but that a global optimum occurs at zero. Therefore, putting reasonable bounds on σ_R and initiating the estimation routine at a reasonable value for σ_R, as in this analysis, may provide reasonable results. If the estimate of σ_R is at the lower bound, estimation methods that integrate over the annual recruitment deviates (e.g. Bayesian analysis or mixed effects) may be needed (see Maunder and Deriso, 2003). In collaboration with Dave Fournier (Otter Research) we have shown that the Laplace approximation (Barndorff-Nielsen and Cox, 1989; as implemented in ADMB) performs identically to the numerical integration method of Maunder and Deriso (2003), but is an order of magnitude faster and may be a promising method to estimate σ_R in these cases. The Laplace approximation method may also improve the normal approximation method for projections (note that the adjustments presented here may not be appropriate if the Laplace approximation is used).

Application to yellowfin tuna highlighted a problem with the normal approximation method linked to the method used to include abundance information from catch-and-effort

data. As explained previously, the amount of information about abundance contained in catch-and-effort data is controlled by the standard deviation of the penalty on the effort deviates. However, this standard deviation is also used to determine variability in catchability in the future. Therefore, downweighting the information on abundance contained in the catch-and-effort data by increasing the standard deviation will increase variability in catchability in the future. The main reasons to downweight the information about abundance contained in the catch-and-effort data are because catchability changes over time. To overcome this, catchability could be modelled as a random walk process (Hampton and Fournier, 2001; Maunder and Watters, 2003), or the assumption that fishing mortality is proportional to effort could be modified.

Normal approximation requires the confidence intervals to be symmetrical. In many cases symmetry may not be appropriate (see Figure 2). For example, biomass cannot be less than zero, and this may contribute to the asymmetry coverage of the normal approximation methods in this analysis. However, calculating the confidence intervals on the natural logarithm of biomass, then using the appropriate transformation, as done by Hampton and Fournier (2001), may improve coverage. The approach of including the projection period in the estimation model can be used in a profile likelihood context to produce asymmetrical confidence intervals. However, this requires optimizing the objective function numerous times, which may be impractical because of computational or time limitations.

The normal approximation and profile likelihood methods to carry out projections are in the flavour of the likelihoodist framework (Tanner, 1993; Royall, 1997), which is an alternative to the Bayesian and frequentist (e.g. bootstrap) frameworks. However, the method can also be used in a frequentist framework. The distribution represented by the normal approximation can be used to approximate a confidence distribution (Schweder and Hjort, 2002), which in turn can be used as an objective Bayesian posterior distribution with inferred reference priors and good coverage properties.

We used the normal approximation method to reduce the computational demands of performing forward projections while still allowing for both parameter uncertainty and future demographic stochasticity. Alternative methods to reduce computational demands are possible. For example, the parameter estimates could be sampled from a multivariate normal distribution based on the delta method, and these can be used to carry out stochastic forward projections (Patterson *et al.*, 2001). This method is essentially a Bayesian analysis that uses a multivariate normal approximation to the posterior distribution, with uniform priors on all parameters.

Methods to reduce computational demand that can be used with Bayesian analysis or bootstrap procedures are also available. The stock assessment method used for our analyses follows the method of Fournier *et al.* (1998), allowing for continuous fishing and natural mortality

throughout the year using the Baranov catch equation and estimating annual effort deviates. An alternative method to that of implementing the Baranov catch equation is iteratively to solve the catch equation within each function evaluation of the parameter estimation procedure. However, this method may also be computationally demanding. The computational demands can be reduced by approximating continuous fishing and natural mortality, by removing catch in the middle of the year (Pope's approximation; Pope, 1972). This removes the need to estimate the annual effort deviates, but it also changes the method used to include the information from catch-and-effort data. Maunder and Starr (2001) used this method to implement a Bayesian analysis for a catch-at-age model with several gears. The reduced computational demands come at the price of possible bias attributable to the approximation. However, it should be noted that fishing and natural mortality do not take place at a constant rate throughout the year, so the method of Fournier *et al.* (1998) is also only an approximation. Also, the normal approximation method for forward projections using constant catch strategies may not be appropriate, and could produce different parameter estimates for the historical period. This is due to the population size getting too small in the projection period for the catch to be removed, and, depending on the implementation, estimated model parameters may have values to compensate for the small population size.

Acknowledgements

References

Barndorff-Nielsen, O. E., and Cox, D. R. 1989. Asymptotic Techniques for Use in Statistics. Chapman & Hall, London. 252 pp.

Berger, J., Liseo, B., and Wolpert, R. 1999. Integrated likelihood methods for eliminating nuisance parameters. Statistical Science, 14: 1–28.

Butterworth, D. S., Ianelli, J. N., and Hilborn, R. 2003. A statistical model for stock assessment of southern bluefin tuna with temporal changes in selectivity. South African Journal of Marine Science, 25: 331–362.

Casella, G., and Berger, R. L. 1990. Statistical Inference. Duxbury Press, Belmont. 650 pp.

Dupont, W. D. 1983. A stochastic catch-effort method for estimating animal abundance. Biometrics, 39: 1021–1033.

Fournier, D. A., Hampton, J., and Sibert, J. R. 1998. MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore, *Thunnus alalunga*. Canadian Journal of Fisheries and Aquatic Sciences, 55: 2105–2116.

Francis, R. I. C. C. 1992. Use of risk analysis to assess fishery management strategies: a case study using orange roughy (*Hoplostethus atlanticus*) on the Chatham Rise, New Zealand. New Zealand Journal of Marine and Freshwater Research, 49: 922−930.

Francis, R. I. C. C., and Shotton, R. 1997. 'Risk' in fisheries management: a review. Canadian Journal of Fisheries and Aquatic Sciences, 54: 1699−1715.

Gavaris, S. 1993. Analytical estimates of reliability for the projected yield from commercial fisheries. *In* Risk Evaluation and Biological Reference Points for Fisheries Management, pp. 185−191. Ed. by S. J. Smith, J. J. Hunt, and D. Rivard. Canadian Special Publication in Fisheries and Aquatic Sciences, 120.

Hampton, J., and Fournier, D. A. 2001. A spatially disaggregated, length-based, age-structured population model of yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. Marine and Freshwater Research, 52: 937−963.

Hilborn, R., and Mangel, M. 1997. The Ecological Detective: Confronting Models with Data. Princeton University Press, Princeton. 315 pp.

Maunder, M. N. 2002. Status of yellowfin tuna in the eastern Pacific Ocean. Inter-American Tropical Tuna Commission Stock Assessment Report, 3: 47−134.

Maunder, M. N., and Deriso, R. B. 2003. Estimation of recruitment in catch-at-age models. Canadian Journal of Fisheries and Aquatic Sciences, 60: 1204−1216.

Maunder, M. N., and Harley, S. J. 2002. Status of bigeye tuna in the eastern Pacific Ocean. Inter-American Tropical Tuna Commission Stock Assessment Report, 3: 201−311.

Maunder, M. N., and Starr, P. J. 2001. Bayesian assessment of the SNA1 snapper (*Pagrus auratus*) stock on the northeast coast of New Zealand. New Zealand Journal of Marine and Freshwater Research, 35: 87−110.

Maunder, M. N., and Watters, G. M. 2003. A-SCALA: an age-structured statistical catch-at-length analysis for assessing tuna stocks in the eastern Pacific Ocean. Inter-American Tropical Tuna Commission Bulletin, 22: 433−582.

Patterson, K., Cook, R., Darby, C., Gavaris, S., Kell, L., Lewy, P., Mesnil, B., Punt, A., Restrepo, V., Skagen, D. W., and Stefansson, G. 2001. Estimating uncertainty in fish stock assessment and forecasting. Fish and Fisheries, 2: 125−157.

Pawitan, Y. 2001. In All Likelihood: Statistical Modelling and Inference Using Likelihood. Oxford University Press, Oxford, UK. 528 pp.

Pope, J. G. 1972. An investigation of the accuracy of virtual population analysis using cohort analysis. Research Bulletin International Commission for the Northwest Atlantic Fisheries, 9: 65−74.

Punt, A. E., and Hilborn, R. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Reviews in Fish Biology and Fisheries, 7: 35−63.

Quinn, T. J. 2003. Ruminations on the development and future of population dynamics models in fisheries. Natural Resource Modeling, 16: 341−392.

Ratkowsky, D. A. 1983. Nonlinear Regression Modeling: A Unified Practical Approach. Marcel Dekker, New York. 276 pp.

Royall, R. M. 1997. Statistical Evidence: A Likelihood Paradigm. Chapman & Hall, New York. 192 pp.

Schweder, T., and Hjort, N. L. 2002. Confidence and likelihood. Scandinavian Journal of Statistics, 29: 309−332.

Tanner, M. A. 1993. Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions. Springer, New York. 207 pp.