Adjusting for bias due to variability of estimated recruitments in fishery assessment models

Richard D. Methot, Jr. and Ian G. Taylor

Abstract: Integrated analysis models provide a tool to estimate fish abundance, recruitment, and fishing mortality from a wide variety of data. The flexibility of integrated analysis models allows them to be applied over extended time periods spanning historical decades with little information from which to estimate the annual signal of recruitment variability to modern periods in which more information about recruitment variability exists. Across this range of data availability, the estimation process must assure that the estimated log-normally distributed recruitments are mean unbiased to assure mean unbiased biomass estimates. Here we show how the estimation method implemented in the integrated analysis model, Stock Synthesis, achieves this unbiased characteristic in a penalized likelihood approach that is comparable to the results from Markov chain Monte Carlo. The total variability in recruitment is decomposed into variability among annual recruitment estimates based on information in the data and a residual variability. Because data are never perfectly informative, we show that estimated recruitment variability will always be less than the true variability among recruitments and that the method implemented here can be used to iteratively estimate the true variability among recruitments.

Résumé : Les modèles d'analyse intégrée représentent un outil pour estimer l'abondance, le recrutement et la mortalité due à la pêche chez les poissons à partir d'une gamme étendue de données. La flexibilité des modèles d'analyse intégrée leur permet d'être utilisés sur de grandes périodes de temps, couvrant des décennies passées pour lesquelles il existe peu de données pour estimer le signal annuel de variabilité du recrutement, mais aussi sur des périodes contemporaines pour lesquelles il y a plus d'information sur la variabilité du recrutement. Sur cette gamme de données disponibles, le processus d'estimation doit assurer que les recrutements estimés selon une distribution log-normale n'ont pas de biais de moyenne afin de produire des estimations moyennes de biomasse non biaisées. Nous montrons ici comment la méthode d'estimation utilisée dans le modèle d'analyse intégrée Stock Synthesis réussit à obtenir cette caractéristique non biaisée dans une approche de vraisemblance pénalisée qui se compare aux résultats de la méthode de Monte Carlo par chaînes de Markov. La variabilité totale du recrutement est décomposée en estimations annuelles du recrutement d'après l'information contenue dans les données et en variabilité résiduelle. Parce que les données n'apportent jamais de l'information parfaite, nous montrons que la variabilité estimée du recrutement sera toujours inférieure à la véritable variabilité entre les recrutements et que la méthode que nous utilisons peut servir à estimer de façon itérative la vraie variabilité entre les recrutements.

[Traduit par la Rédaction]

Introduction

Fishery assessment models estimate time series of fish abundance and fishing mortality from data that ideally includes catch-at-age and fishery-independent surveys of fish abundance. Where there is complete, precise information on the age composition of the catch, age-structured (e.g., statistical catch-at-age) models (Ricker 1975; Deriso et al. 1985; Quinn and Deriso 1999) can accurately estimate the abundance of each annual recruitment in the modeled time series (Maunder and Deriso 2003). Where catch-at-age data are missing or incomplete, a more flexible class of fishery assessment models, termed integrated analysis, can be applied. Integrated analysis models provide a tool to estimate fish abundance, recruitment, and fishing mortality from a wide variety of data (Fournier and Archibald 1982; Methot 1990, 2000), including length composition, imprecise age composition with ageing error, and other data types. The flexibility of integrated analysis models allows them to be applied over long time series spanning historical decades, with little information from which to estimate the annual signal of recruitment variability, to modern periods in which more information about recruitment variability exists (e.g., Haltuch and Hicks 2009; Hamel 2009; Stewart 2009). However, when applied to such heterogeneous data situations, the variability among the estimated recruitments is also expected to be heterogeneous. Even in the data-rich years, the data will never be perfectly precise with regard to the true recruitment devia-

Received 5 November 2010. Accepted 4 June 2011. Published at www.nrcresearchpress.com/cjfas on 12 October 2011.
J2011-0194

Paper handled by Associate Editor Yong Chen.

R.D. Methot, Jr. NOAA Fisheries, Office of Science and Technology, 2725 Montlake Blvd. East, Seattle, WA 98112, USA.
I.G. Taylor. Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Box 955020, Seattle, WA 98195-5020, USA.

Corresponding author: Richard D. Methot, Jr. (e-mail: Richard.Methot@noaa.gov).

tions. Thus the estimate of each recruitment deviation will always be a compromise between the information in the data and the central tendency that pulls the log(recruitment) deviations towards zero. Here we describe a maximum likelihood method to address this situation in a way that is comparable to the unbiased estimation obtained from Bayesian integration (Maunder and Deriso 2003).

Maunder and Deriso (2003) explored a range of procedures to estimate recruitments in statistical catch-at-age models. These recruitments typically follow a lognormal distribution with standard deviation σ_R (Power 1996; Quinn and Deriso 1999; Haddon 2001). The expected arithmetic mean recruitment thus depends upon both the geometric mean and the variability of the lognormal distribution. This means that more of the biomass in the population, and in the long-term average potential yield, comes from the relatively infrequent large recruitments. Consequently, calculations based on the median or geometric mean recruitment will underestimate the mean stock abundance and mean potential yield. Maunder and Deriso (2003) found that a wide range of procedures, including a penalized likelihood method, worked adequately if the data were reasonably informative about recruitment deviations throughout the time series being analyzed. However, when the data were less informative during all or part of the time series, then a marginal likelihood or Bayesian integration method using Markov chain Monte Carlo (MCMC) performed better. This difference is due to the nature of the MCMC integration, which explores the full lognormal domain of each recruitment deviation, in contrast with the maximum likelihood procedure, which causes poorly informed recruitments to collapse to the central tendency, which is a geometric mean recruitment level. Unfortunately, the complexity of data and length of time series often included in the integrated analysis models is often not sufficient to estimate all recruitments well, but is sufficiently large to preclude routine use of MCMC because run times to achieve convergence can be several days.

In an assessment model in which log(recruitment) is penalized for deviating from 0.0, time periods during which the quantity of data regarding recruitment fluctuations is weak will have small fluctuations among the estimated log(recruitments), and the resultant arithmetic mean recruitment during those periods will be underestimated. This may then cause a bias in the estimate of other model parameters, such as spawner–recruitment steepness, that also affect the long-term trend in mean recruitment needed to provide the biomass from which the catch has been taken. The true annual variability among recruitments is also a key factor in forecasting the range of future fluctuations in stock abundance. These fluctuations will depend on the actual variability of recruitment, not the degree of variability estimated from our imperfect historical data (Maunder and Deriso 2003).

Here we extend the results of Maunder and Deriso (2003) to show how a penalized likelihood method can be modified to incorporate a time-varying bias correction to produce results that are equivalent to those produced using MCMC. The method is based on a modification of the bias adjustment, $R_y^* = R_y\, e^{r_y - \sigma_R^2/2}$, commonly used to adjust recruitment estimates to be mean unbiased. We start by showing how any imprecision in data will cause any assessment model to underestimate the variability among recruitments. Second, we establish a procedure for calculating the correct degree of bias adjustment for each recruitment deviation based upon the variance in the estimate of that recruitment deviation. This is basically a partitioning of the total variability in recruitment into a component that is the variability among the annual recruitments and a component that is the variance of the estimate of each recruitment. We demonstrate the performance of this procedure first with a hypothetical example in which the only data is a young-of-the-year survey that directly measures, with error, the fluctuations in actual recruitment. Then we show the performance using simulated data that are comparable to the types of data used in typical assessment situations. We compare the results with estimates of recruitment variability obtained using MCMC on the same data sets, then we show how the bias adjustment procedure can be used to improve estimates of the underlying σ_R associated with the true recruitment variability.

Materials and methods

A better understanding of how to model variability in recruitment was sought using both analytical derivations (using notation described in Table 1) and simulation experiments.

Lognormal recruitment variability

Variability in recruitment for fisheries stock assessment models is typically modeled using a lognormal distribution (Maunder and Deriso 2003). Values for the number of individuals recruited to the population in each year are often calculated as

$$(1) \qquad R_y^* = R_y\, e^{r_y - \sigma_R^2/2}$$

where R_y is the mean value of recruitment, often calculated as a function of spawning biomass, r_y is the recruitment deviation in year y, which is assumed to have a normal distribution, so that e^{r_y} is lognormally distributed, and, σ_R is the standard deviation for recruitment in log space.

The subtraction of the term $\sigma_R^2/2$ in the exponent is a bias adjustment that is made so that the mean of the resulting lognormally distributed recruitments R_y^* is equal to R_y. This results in a median recruitment value that is less than R_y. However, mean recruitment is a better representation of the long-term contribution to the population than median recruitment, because most of the population biomass comes from the numerous recruits in the upper tail of the lognormal distribution. As an illustration of this property, the average biomass under variable recruitment may be compared with the biomass at equilibrium. Assuming recruitment occurs at age 0, the total biomass B_0 associated with an equilibrium recruitment R_0 can be calculated as

$$(2) \qquad B_0 = \sum_{a=0}^{\infty} w_a\, e^{-aM} R_0$$

where w_a is the mean mass at age a, and M is the natural mortality, here set equal for all ages. A is a maximum age chosen to be high enough so that little growth occurs beyond this age. Numbers at age A are calculated so as to include older ages implicitly, so represented by ∞ here.

The average biomass across n years associated with variable recruitment may be calculated as

Table 1. Definitions of key symbols used in the analysis.

Symbol	Definition
R_y	Expected value of recruitment experienced by the population in year y
R_y^*	Realized recruitment after the application of a bias adjustment factor
r_y	True deviation in recruitment on a log scale experienced by the population in year y
σ_R	Standard deviation of the distribution from which true recruitment deviations are drawn
$\widehat{\boldsymbol{r}}_y$	Estimated recruitment deviation for year y
$\mathrm{SE}(\widehat{\boldsymbol{r}}_y)$	Standard error of the parameter estimate $\widehat{\boldsymbol{r}}_y$ as estimated from the Hessian matrix
$\overline{\mathrm{SE}(\widehat{\boldsymbol{r}})}$	Mean of the $\mathrm{SE}(\widehat{\boldsymbol{r}}_y)$ across some specified range of years
$\mathrm{SD}(\widehat{\boldsymbol{r}})$	Standard deviation among the set of $\widehat{\boldsymbol{r}}_y$ values across some specified range of years
$E[\mathrm{SD}(\widehat{\boldsymbol{r}}_y)]$	Expected standard deviation of distribution for the estimated recruitment deviation $\widehat{\boldsymbol{r}}_y$ in year y
μ_{dy}	Value for the recruitment deviation in year y that best fits the data
σ_{dy}	Standard deviation around μ_{dy} implied by the data
b_y	Fraction of bias adjustment applied in year y

$$(3)\qquad \overline{B_y} = \frac{1}{n}\sum_{y=1}^{n} B_y = \frac{1}{n}\sum_{y=1}^{n}\sum_{a=0}^{\infty} w_a\, \mathrm{e}^{-aM} R_{y-a} = \sum_{a=0}^{\infty} w_a\, \mathrm{e}^{-aM}\frac{1}{n}\sum_{y=1}^{n} R_{y-a}$$

where R_{y-a} is the recruitment in year $y - a$.

If the mean recruitment over n years, given by $\frac{1}{n}\sum_{y=1}^{n} R_{y-a}$, is equal to the equilibrium recruitment R_0, then eqs. 2 and 3 will match, and the mean biomass B_y will equal the equilibrium biomass B_0. It is straightforward to confirm via simulation that the bias adjustment is also needed under more complex scenarios, such as including a spawner–recruitment relationship and harvest.

Influence of data on recruitment estimation

Estimated logarithmic recruitment deviations, $\widehat{\boldsymbol{r}}_y$, are informed by two general sources of information. First, an objective function component for r_y (noted later in this text) pulls all deviations toward 0.0 with a strength determined by σ_R, representing the assumption that the distribution of r_y is approximately normal; thus large deviations are less common than small ones. Second, the data, especially age and length compositions (but other sources as well), provide information about individual year class strengths and thus pull the $\widehat{\boldsymbol{r}}_y$ toward whatever value maximizes the likelihood components for these other data sources. The influence of the data on the $\widehat{\boldsymbol{r}}_y$ will depend on the quality and quantity of the data, while the influence of σ_R is independent of data richness. Thus, for years with little information in the data on recruitment, the $\widehat{\boldsymbol{r}}_y$ will be close to 0. The data would have to be extremely precise for the influence of the data to completely overwhelm the influence of the objective function component for recruitment. Thus, in all practical applications, the estimated variability among $\widehat{\boldsymbol{r}}_y$ will be lower than σ_R.

A simple illustration of this point comes from the case where there is a time series of recruitment deviations and a survey (e.g., a recruitment survey, or young-of-the-year survey) that provides data, with measurement error, about these deviations. The likelihood for a random variable, informed by observations with a mean and standard deviation under the assumption of normal error, has a maximum that occurs where the estimate of the random variable is a weighted average of the estimates that would be calculated from each source of data independently and a standard error that is a combination of the uncertainty around these means. In the context of estimating deviations in recruitment, with two sources of information, μ_{dy} and σ_{dy} representing the information in the data about a given r_y, and $\mu_{Ry} = 0$ and σ_R representing the contribution of the assumptions about recruitment, the resulting estimated deviation is

$$(4)\qquad \widehat{\boldsymbol{r}}_y = \frac{\frac{\mu_{dy}}{\sigma_{dy}^2} + \frac{0}{\sigma_R^2}}{\frac{1}{\sigma_{dy}^2} + \frac{1}{\sigma_R^2}} = \frac{\sigma_R^2}{\sigma_{dy}^2 + \sigma_R^2}\mu_{dy}$$

with standard error

$$(5)\qquad \mathrm{SE}(\widehat{\boldsymbol{r}}_y) = \left(\frac{1}{\sigma_{dy}^2} + \frac{1}{\sigma_R^2}\right)^{-1/2}$$

If the data provide rich information about r_y, then σ_{dy} will be small relative to σ_R and the estimate $\widehat{\boldsymbol{r}}_y$ will approach μ_{dy}. If the data are uninformative about r_y, then σ_{dy} will be large relative to σ_R, and the estimated value will approach 0. Under no circumstances will the estimate $\widehat{\boldsymbol{r}}_y$ be of greater magnitude than μ_{dy}.

Because of sampling error, the values for μ_{dy} will vary around the true r_y. If the distribution of the μ_{dy} about the r_y has standard deviation equal to σ_{dy}, and, for simplicity, the r_y are assumed equally informed by the data (and thus all $\sigma_{dy} = \sigma_d$), then the standard deviation of the vector of μ_{dy} values, denoted $\mathrm{SD}\boldsymbol{\mu}_d$, is

$$(6)\qquad \mathrm{SD}(\boldsymbol{\mu}_d) = \left(\sigma_R^2 + \sigma_d^2\right)^{1/2}$$

and the standard deviation of the vector of $\widehat{\boldsymbol{r}}_y$ values, denoted $\mathrm{SD}(\widehat{\boldsymbol{r}})$, will be scaled toward zero following eq. 4 as

$$(7)\qquad \mathrm{SD}(\widehat{\boldsymbol{r}}) = \mathrm{SD}\left(\frac{\sigma_R^2}{\sigma_d^2 + \sigma_R^2}\boldsymbol{\mu}_d\right) = \frac{\sigma_R^2}{\sigma_d^2 + \sigma_R^2}\mathrm{SD}(\boldsymbol{\mu}_d) = \frac{\sigma_R^2}{\left(\sigma_R^2 + \sigma_d^2\right)^{1/2}} = \sigma_R\frac{\sigma_R}{\left(\sigma_R^2 + \sigma_d^2\right)^{1/2}} < \sigma_R$$

From this equation, it is clear that in the extreme cases, as σ_d approaches infinity (no information in the data) then $\mathrm{SD}(\widehat{\boldsymbol{r}})$ approaches 0 (with all $\widehat{\boldsymbol{r}}_y$ approaching 0), and as σ_d

approaches 0 (perfect information in the data) then $SD(\widehat{\boldsymbol{r}})$ goes to σ_R (with all $\widehat{r}_y$ approaching the true r_y).

Derivation of annual bias adjustment

In eq. 1, the term $\sigma_R^2/2$ is intended to adjust for bias in the lognormal distribution so that the distribution of exponentiated recruitment deviations, $e^{r_y-\sigma_R^2/2}$, has mean equal to 1.0. However, the true r_y values are never known, so models are projected using a time series of estimated $\widehat{r}_y$ values, which will be less variable than the true values, as shown in eq. 7. Thus, the mean of the lognormal distribution of recruitments, and the appropriate degree of bias adjustment associated with this mean, is dependent on the distribution of the estimated recruitments, not the distribution of true recruitments. If we shift from consideration of the variability among the estimated $\widehat{r}_y$ values over the full range of years to the expected standard deviation associated with a single year y, denoted $E\left[\mathrm{SD}(\widehat{r}_y)\right]$, then to have mean unbiased recruitment, the correction applied to the annual recruitment equation should be

$$R_y^* = R_y e^{\widehat{r}_y} - E\left[\mathrm{SD}(\widehat{r}_y)\right]^2/2 = R_y e^{\widehat{r}_y} - b_y\sigma_R^2/2 \quad (8)$$

where $b_y = \frac{E\left[\mathrm{SD}(\widehat{r}_y)\right]^2}{\sigma_R^2}$ is the bias adjustment fraction applied in year y.

In practice, data available for stock assessments is not uniformly informative for all years. Indeed, for most commercially exploited fish stocks, data on removals are available much further back in time than age and length-composition data, and the quality and quantity of these data typically increase over time. Consequently, the information about recruitment variability in an age-structured population model may be expected to be initially low, increase steadily for years where the cohorts are more and more represented in the data, remain higher for a period with similar data levels, and fall off at the end. Even during the most data-rich period, the estimated recruitment deviations will have uncertainty associated with the number of years for which each cohort can be observed, as well as factors such as ageing error, selectivity, and the identification of individual cohorts in length-composition data.

Integrated analysis models do not have a strong requirement for age-composition data, so can function across a range of years that span data-poor and data-rich periods. While it is possible to configure these models to not estimate recruitment deviations during the early years, this practice results in an underestimate in the uncertainty about population abundance during the early years because actual recruitment was fluctuating during that data-poor period just as much as it fluctuated during the subsequent data-rich era. However, allowing estimation of recruitments during the data-poor period results in estimated recruitment deviations that collapse towards 0.0 and will result in an underestimate of mean recruitment during that era. Thus, the bias adjustment fraction, b_y, that is appropriate for years without information should be close to 0.0, while the b_y for the years with informative data should be some value that is a function of their variability. A first approximation to time-varying bias adjustment is a pattern that linearly ramps up during an era of increasing information about recruitment variability, reaches a plateau, then decreases during the most recent years.

The magnitude of the b_y depends upon the degree to which the overall variability in recruitment is partitioned into variability among the recruitment estimates in the time series and residual variability of each recruitment estimate. Information about the expected distribution for the individual $\widehat{r}_y$ may be obtained using the following relation between the standard error of the individual $\widehat{r}_y$ estimates from eq. 5 and the standard deviation of the distribution of the vector of $\widehat{r}_y$ values from eq. 7 as

$$\mathrm{SE}(\widehat{r}_y)^2 + \mathrm{SD}(\widehat{\boldsymbol{r}})^2 = \left[\left(\frac{1}{\sigma_d^2}+\frac{1}{\sigma_R^2}\right)^{-1/2}\right]^2 + \left[\frac{\sigma_R^2}{(\sigma_R^2+\sigma_d^2)^{1/2}}\right]^2 = \sigma_R^2 \quad (9)$$

The equation above is based on the standard deviation of the vector of estimated $\widehat{r}_y$, but we can infer that as the number of years is reduced, the $SD(\widehat{\boldsymbol{r}})$ term may be replaced by the expected standard deviation associated with any individual year, $E\left[\mathrm{SD}(\widehat{r}_y)\right]$. Thus, an estimated or specified σ_R, combined with standard errors of the $\widehat{r}_y$, can be used to estimate the standard deviation of the distribution from which a single parameter estimate in need of bias correction was drawn, as

$$E\left[\mathrm{SD}(\widehat{r}_y)\right] = \sigma_R^2 - \mathrm{SE}(\widehat{r}_y)^2 \quad (10)$$

Thus, an alternative way to calculate the bias adjustment fraction values, b_y in eq. 8, is

$$b_y = \frac{E\left[\mathrm{SD}(\widehat{r}_y)\right]^2}{\sigma_R^2} = 1 - \frac{\mathrm{SE}(\widehat{r}_y)^2}{\sigma_R^2} \quad (11)$$

Refinement to the likelihood for recruitment deviations

The recruitment deviation penalty that contributes to the objective function along with the negative log likelihood of other objective function components is typically calculated as

$$\mathcal{L}_{\mathrm{recruit}} = 0.5\sum_y \left[\frac{\widehat{r}_y^2}{\sigma_R^2} + \ln(\sigma_R)\right] \quad (12)$$

However, this formulation is based on the assertion that the $\widehat{r}_y$ are distributed according to σ_R. Here we have shown that the $\widehat{r}_y$ are distributed according to the combination of σ_R and the amount of information about the recruitment deviations in the data and that their variability will change over the extent of a time series. Where there are weakly informed r_y, their values will collapse towards 0.0, and the value of σ_R that maximizes $\mathcal{L}_{\mathrm{recruit}}$ will underestimate the true σ_R. In the extreme case with no data to inform recruitments, the best estimate of σ_R would approach 0.0.

An alternative formulation can adjust for this problem:

$$\mathcal{L}_{\mathrm{recruit}} = 0.5\sum_y \left[\frac{\widehat{r}_y^2}{\sigma_R^2} + b_y\ln(\sigma_R)\right] \quad (13)$$

where $\widehat{r}_y$ is the estimated recruitment deviation in year y, σ_R is the standard deviation for the true recruitment deviations, and b_y is the bias adjustment fraction applied in year y.

With this approach, the contribution of the second term, ln

(σ_R), smoothly scales according to the value of b_y, ranging from b_y near 1.0 for data-rich years, thus equivalent to eq. 12, and near 0.0 for the data-poor years. But maintaining σ_R in the first term means that the estimated standard errors of the poorly informed r_y will approach σ_R. This refinement means that when the vector of estimated recruitment parameters is expanded to include years with no data on recruitment (such as early in the time series or future years in a forecast), the objective functions does not change — a desirable property that allows better estimation of uncertainty in population dynamics throughout the time series.

Experiments with simulated data

The experiments were conducted using the Stock Synthesis (SS) integrated analysis model (Methot 1989, 1990, 2009). The relevant features of this model are provided in an appendix to this paper. The model and its user manual, which describes the full suite of model options, are available from the NOAA Fisheries Toolbox (http://nft.nefsc.noaa.gov). Two simulation experiments were created to test the theory described above. The simulation procedures followed common practices for such experiments (e.g., Garrison et al. 2011), including the following four steps (all of which are described in greater detail below and in Appendix A): (1) Create 100 simulated populations using a set of assumed "true" parameter values, including a set of randomly generated recruitment deviations, r_y. (2) For each simulated population, create multiple data sets by simulating the process of sampling from the populations created in Step 1 under multiple assumptions of uncertainty and sample sizes. This involves generating stochastic data under an observation model using distributions appropriate to each data type. (3) For each simulated data set created in Step 2, apply multiple estimation models under different assumptions about how recruitment deviations are modeled. (4) Compare the estimated parameters and key derived quantities resulting from the estimation models with the true values and use these comparisons to evaluate the performance of the alternative estimation models and how the performance changes as a function of the uncertainty and sample sizes used to generate the simulated data.

Data simulation

The first experiment provides a straightforward test of the influence of a single data source on estimation of recruitment deviations, the theory for which is described using eqs. 4 through 9, above. The simulated population is simply a time series of 50 years of recruitment values r_y from a normal distribution with mean 0 and standard deviation $\sigma_R = 0.7$. The data are a recruitment survey that provides an estimate, μ_{dy}, of each year's recruitment with measurement error σ_{dy}. The survey values, μ_{dy}, are sampled from normal distributions with mean r_y and standard deviation σ_d. Within each simulation, one of four levels of sampling error were used, σ_d = 0.1, 0.5, 0.9, 5.0, ranging from highly informative to very uninformative. This process was repeated 100 times for each level of sampling error.

The second experiment is a fully age-structured configuration, with age-composition data informing recruitment. The simulated recruitments are as in the first experiment, but here the data about the recruitments is more complex. SS was used to generate 100 simulated populations, each with a different time series of normally distributed stochastic r_y. For each time series, two sets of data were generated, one data-rich and one data-poor, which differed only in the sample size of the age-composition data.

The population was modeled for 50 years, designated 1961 to 2010. The r_y were simulated for the years 1945 to 2010. These values were generated from a standard normal distribution and then standardized so that that over the period 1945–2004 they had mean 0 and standard deviation $\sigma_R = 0.7$. Prior to 1945, recruitment deviations were not used and recruitment was set to the mean level. Ages 20 and over were accumulated as a plus group and natural mortality was fixed at 0.2 year^{-1} (parameter values used in the analysis are collected in Appendix Table A1). Steepness of the spawner–recruit relationship (eq. A.7) was set to 0.99 to focus the analysis on estimating the central tendency of recruitment, rather than the more complex problem of estimating steepness as well.

The catch for all simulations increased linearly from 500 mt in 1961 to 10 000 mt for the years 1980–1983 and then decreased linearly to a constant level of 4000 mt for 1995–2010. Selectivity was a logistic function of age, with 50% selected at age 5 and 95% selection at age 7. The equilibrium recruitment was constant across simulations at a value that resulted in a median value for the minimum biomass level of 33% of B_0, with 90% of simulations having a minimum between 13% and 46% of B_0. The catch for the last 15 years of the model was sufficiently low to allow rebuilding of the stock, with a median 2010 spawning biomass value of 72% of B_0. These changes in catch and thus fishing mortality caused changes in the number of years that cohorts from different time periods would be represented in the age-composition data.

The simulated data were age compositions and a time series of catch per unit effort (CPUE). The age compositions covered the last 30 years of the model, from 1981 to 2010, with a sample size of either 200 (data-rich) or 20 (data-poor). The age compositions were simulated as samples from the multinomial distributions with probabilities equal to the projected population age composition for each corresponding year in the operating model. Bins were integer ages 1 to 20 for males and females, with the first bin including age 0 and the last bin including all members of the plus group. The CPUE data covered the last 40 years of the model, from 1971 to 2010. To increase the information about biomass trends and thus amplify any signal due to model mis-specification, the CPUE standard error was set to a low value of 0.1 in log scale (roughly equivalent to a coefficient of variation (CV) of 0.1).

Ageing imprecision was applied in both simulation and estimation models. The observed ages were assumed to have an unbiased distribution around the true age with variability increasing with age (eq. A.15). The same ageing imprecision was applied in both simulation and estimation models, although these matching assumptions do not eliminate the uncertainty in recruitment created by adding ageing imprecision to the composition data.

Estimation models

In the both the simple and age-structured experiments, σ_R was assumed known without error, except in the case of an additional analysis comparing methods for estimating σ_R.

The estimated r_y were calculated as a zero-centered vector of deviations. The zero-centering causes a small interaction between estimates. However, the influence of centering the values to have mean 0 is expected to be small. Standard errors, $\mathrm{SE}(\widehat{r}_y)$, for each $\widehat{r}_y$ were calculated from the Hessian matrix as implemented in the AD Model Builder software (ADMB Project 2009).

In the simple experiment, the objective function as a simple least-squares minimization of the differences between estimated $\widehat{r}_y$ and the observations of these values combined with a term related to the variability in estimated recruitments:

$$\mathcal{L} = 0.5\sum_y \frac{(\mu_{dy} - r_y)^2}{\sigma_{dy}^2} + 0.5\sum_y \frac{r_y^2}{\sigma_R^2} \quad (14)$$

where the components of this equation are as described for eq. 4 above.

For the age-structured experiment, six estimation models were applied (Table 2), differing in the pattern of the $b_{\max}$ values and in the weight applied to the objective function component for the recruitment deviations (ω in eq. A.19). The pattern with less weight applied to recruitment deviations was chosen to mimic the approach taken by the influential statistical catch-at-age (SCAA) model CAGEAN (Deriso et al. 1985), whereby the weighting of the contribution from the $\widehat{r}_y$ (ω in eq. A.19) was reduced from 1.0 to 0.1. This strongly reduces the degree to which the $\widehat{r}_y$ are pulled toward 0.

The annual bias adjustment fraction (b_y) was set to either a constant or to a series of values starting with 0 for the initial years, ramping up to a maximum bias adjustment fraction, $b_{\max}$, for an intermediate period and then back down to 0 for the final years of the model, according to the formula

$$b_y = \begin{cases} 0 & \text{for } y \le y_1^b \\ b_{\max}\left(1 - \dfrac{y - y_1^b}{y_2^b - y_1^b}\right) & \text{for } y_1^b < y < y_2^b \\ b_{\max} & \text{for } y_2^b \le y \le y_3^b \\ b_{\max}\left(1 - \dfrac{y_3^b - y}{y_4^b - y_3^b}\right) & \text{for } y_3^b < y < y_4^b \\ 0 & \text{for } y_4^b \le y \end{cases} \quad (15)$$

where the $y_1^b \ldots y_4^b$ are 4 years that form the break points for the piecewise linear relationship. The values for these break points were set based on a visual examination of the time trend of $\mathrm{SE}(\widehat{r}_y)$ in initial model runs. The estimated parameters were equilibrium recruitment (R_0), the two logistic selectivity parameters, and the $\widehat{r}_y$ for the years 1945–2010, a total of 69 estimated parameters. The annual fishing mortality, F, was calculated by an iterative routine that finds the F_y that matches the catch_y exactly. All other parameters were fixed at the true values. Initial values of the $\widehat{r}_y$ were 0, but the estimates of R_0 and the selectivity parameters were started at the true values. This is an artificially high level of information for an assessment model, but was intended to focus the results on the relationship between the amount of information about recruitment deviations and the method of modeling these deviations.

The estimation models were evaluated by comparing the estimated with the simulated values for initial biomass B_0 and final depletion B_{2010}/B_0. The relative error, $100\% \times \left(\theta_i - \theta_i^{\mathrm{true}}\right)/\theta_i^{\mathrm{true}}$, of these quantities was used as a performance measure for each of the methods described in Table 2. For plotting, the ratio $\theta_i/\theta_i^{\mathrm{true}}$ was used as a more visually informative quantity.

Table 2. Estimation models differing in the pattern of annual bias adjustment fraction, named by the maximum bias adjustment fraction value in each case.

Name	Description
0 (SCAA)	Bias adjustment fraction $b_y = 0$ for all years and $\omega = 0.01$ in eq. A.19
0	Bias adjustment fraction $b_y = 0$ for all years
0.52	Ramp for data-poor: following eq. 15 with $b_{\max} = 0.52$
0.72	Ramp for data-rich: following eq. 15 with $b_{\max} = 0.72$
1	Ramp to 1: following eq. 15 with $b_{\max} = 1.0$
1 (no ramp)	Bias adjustment fraction $b_y = 1.0$ for all years

Note: SCAA, statistical catch-at-age.

Estimating σ_R

The standard deviation σ_R of the r_y was not estimated while investigating alternative bias adjustment methods to focus the analysis on the best-case scenario in which the true value was correctly specified. However, in practical applications, this quantity is not known, so further analyses were conducted to explore its estimation. First, a profile approach was applied to the data-poor and data-rich data sets. Estimation models were applied with σ_R fixed at a range from 0.5 to 0.9 as well as the true value used in the simulations, $\sigma_R = 0.7$. Then, three methods were used to estimate σ_R using only the data-poor simulations. (1) In the first method σ_R was estimated as a parameter in the estimation model. (2) In the second approach, σ_R was determined using the bisection method (Conte and de Boor 1980), where repeated model runs were used to numerically estimate a value of σ_R within 0.02 of a value that would satisfy the equality $\sigma_R^2 = \mathrm{SD}(\widehat{r})^2$, where $\mathrm{SD}(\widehat{r})$ is the standard deviation of the vector of $\widehat{r}_y$ values for the years 1975–2004. This iterative approach has been used in cases where method 1 was not able to find a local minimum. (3) The last method also used the bisection method to tune σ_R, but this time it was brought to within 0.02 of the value that would satisfy the equality $\sigma_R^2 = \mathrm{SD}(\widehat{r}) + \overline{\mathrm{SE}(\widehat{r}_y)}^2$, where the additional term, $\overline{\mathrm{SE}(\widehat{r}_y)}^2$, is the square of the mean standard error estimates for the $\widehat{r}_y$ from the years 1975 to 2004. This method should perform better than method 2 based on the theory presented above.

In each of these investigations into σ_R, the bias adjustment values b_y were ramped up to and down from a maximum of 0.52 for the data-poor scenario and 0.72 for the data-rich scenario, respectively.

MCMC

In addition to maximum likelihood estimation, MCMC chains of parameter estimates were calculated for the 100 simulated data files from the data-poor scenario. For the MCMC, the bias adjustment fraction was fixed at $b_y = 1.0$ for all years according to the findings of Maunder and Deriso

(2003). The MCMC samples were calculated using the default algorithm included in AD Model Builder (ADMB Project 2009). Each chain was 1 000 000 MCMC samples, with the first 200 000 discarded as a burn-in period and thinned to every 1000th samples. Bayesian analyses implementation in AD Model Builder have often used longer chains with greater thinning (e.g., 5 000 000 thinned to every 5000th or 2500th sample; Punt et al. 2006; Stewart et al. 2011), but the need to perform these calculations for each of 100 simulated data sets necessitated shorter chains. Convergence diagnostics for the MCMC chains were calculated using the "coda" package (Plummer et al. 2010) in the software R (R Development Core Team 2011).

Results

Simple simulation model

The mean values for $\mathrm{SD}(\hat{\boldsymbol{r}})$ and $\overline{\mathrm{SE}(\hat{\boldsymbol{r}}_y)}$, calculated from 100 simulations for each level of sampling error, held very tightly to the relationship predicted by eq. 9 (Fig. 1). The observed parameter standard errors were lower than the predicted values by about 1% (Table 3), which is likely due to the zero-centering of the deviation vector in the estimation model. Variability in the sampling for each simulation leads to variability in the standard deviation of the 50 recruitment estimates, $\mathrm{SD}(\hat{\boldsymbol{r}})$, while the standard error of each recruitment estimate were constant across all years as a result of the constant σ_{dy} values in the time series, so in this case $\mathrm{SE}(\hat{\boldsymbol{r}}_y) = \overline{\mathrm{SE}(\hat{\boldsymbol{r}}_y)}$ for all y.

Age-structured simulation model

The recruitments that are most precisely estimated from the simulated data are for the cohorts born between 1975 and 2004 (Fig. 2). Age at 50% selectivity is age 5, so the 1975 cohort will be mostly selected when the age-composition data begins in 1981 and will be represented in the composition data through age 20 in 1995. Cohorts born before 1975 will be represented in fewer years, and those born prior to about 1965 will have very little potential to create a signal in the data. This is the basis for the selection of the ascending limb of the ramp function in eq. 15. Similarly, cohorts born after 2004 will have little representation in the data because of not being selected prior to the end of the data series in 2010. Although the representation in the data is not uniform for cohorts born over the range 1975–2004, the data include the year in which the majority of these cohorts are selected, which is the year where their representation in the composition data may be highest. Indeed, the $\mathrm{SE}(\hat{\boldsymbol{r}}_y)$ values are similar for these years, indicating similar precision of $\hat{\boldsymbol{r}}_y$ (Fig. 2).

The relationship between the mean values of $\mathrm{SD}(\hat{\boldsymbol{r}})$ and $\overline{\mathrm{SE}(\hat{\boldsymbol{r}}_y)}$ for the years 1975–2004 across the 100 simulations for each level of age-composition sample size was again very similar to that predicted by eq. 9 (Fig. 3), although to a lesser degree than in the idealized simple model (Fig. 1). The difference may be due to a variety of factors, including the zero-centering of the $\hat{\boldsymbol{r}}_y$ vector and the fact that the normal approximation assumed in deriving the $\mathrm{SE}(\hat{\boldsymbol{r}}_y)$ values from the Hessian matrix is less exact when the information in the data is not a point estimate as in the idealized scenario.

Fig. 1. Mean estimated uncertainty in estimated recruitment deviations (rec. devs.) vs. variability in the estimates for a simple model in which the data is a survey of recruitment deviations for 50 years. Four levels of sampling error in the survey are used (σ_d = 0.1, 0.5, 0.9, 5.0). Variability in simulated recruitment deviations is σ_R = 0.7 throughout. Grey points indicate results of 100 realizations of the model for each level of sampling error σ_d (indicated by text label associated with each cluster of points). Large Xs indicate observed average value for each quantity across the realizations. Large open circles indicate predicted values based on eq. 5 for mean parameter standard error and on eq. 7 for standard deviation of vector of estimates. Dotted arc indicates the predicted relationship between the two measurements, as described in eq. 9.

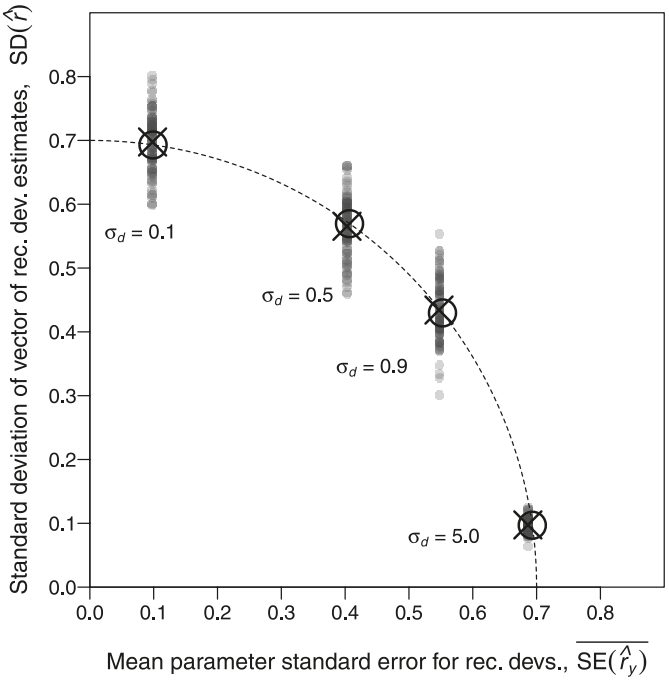


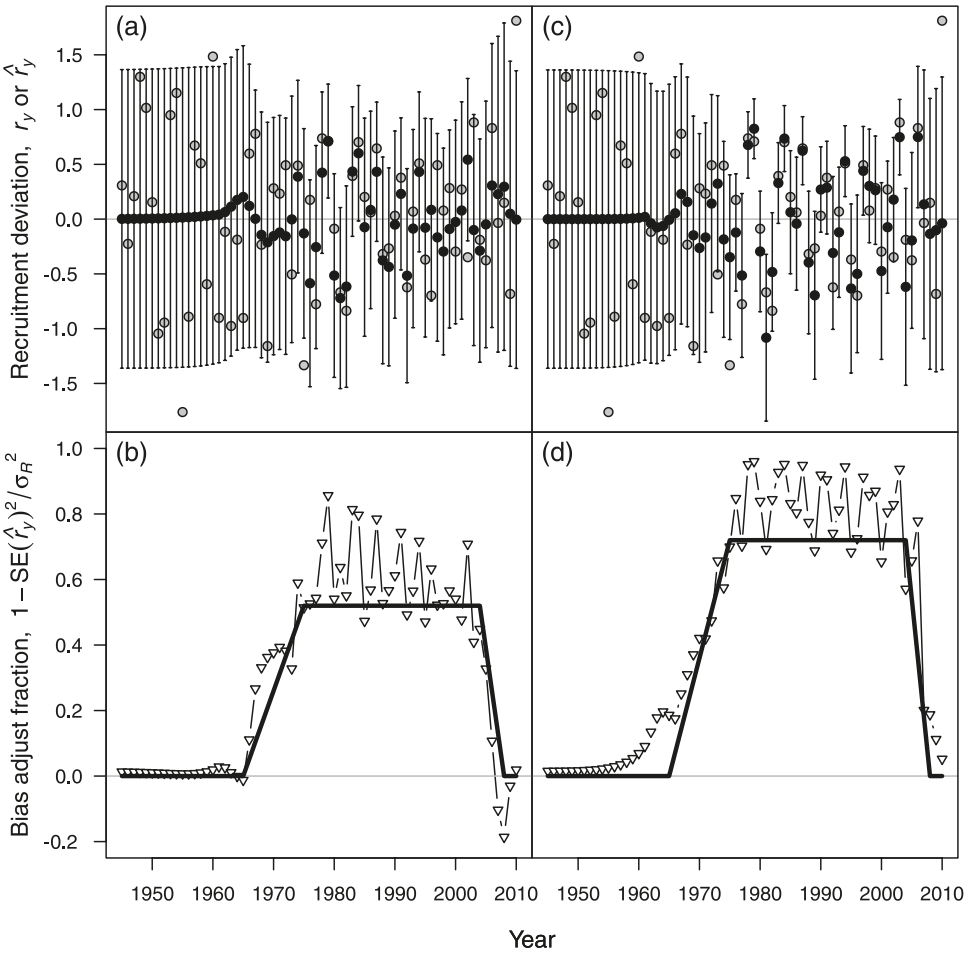
Table 3. Results of simple model corresponding to Xs and circles in Fig. 1.

Sampling error σ_d	Mean observed $\sigma(\hat{\boldsymbol{r}})$	Predicted $\sigma(\hat{\boldsymbol{r}})$	Mean observed $\overline{\mathrm{SE}(\hat{\boldsymbol{r}}_y)}$	Predicted $\overline{\mathrm{SE}(\hat{\boldsymbol{r}}_y)}$
0.1	0.691	0.693	0.098	0.099
0.5	0.560	0.570	0.403	0.407
0.9	0.431	0.430	0.547	0.553
5.0	0.097	0.097	0.686	0.693

The maximum bias adjustment values $b_{\max}$ 0.52 and 0.72 shown on panels (b) and (d), respectively, of Fig. 2 were the mean values of $\mathrm{Var}(\hat{\boldsymbol{r}})/\sigma_R^2$ for the years 1975–2004 calculated from an initial set of model runs with no bias adjustment applied (Fig. 3). Calculating the bias adjustment fraction from the values of $1 - \mathrm{SE}(\hat{\boldsymbol{r}}_y)^2/\sigma_R^2$ resulted in slightly higher $b_{\max}$ in each case but produced similar results.

Estimation models with ramped bias adjustment performed well in estimating B_0 and final depletion (Figs. 4 and 5), while those with constant bias adjustment performed poorly. For the data-poor simulation model with age-composition sample size of 20, the estimation model that was best suited for this data based on the theory above had $b_{\max}$ = 0.52. In

Fig. 2. Example realizations of full model with 40 years of age-composition data from 1971 to 2010. Model shown in panels (*a*) and (*b*) has 20 age samples in each year, while panels (*c*) and (*d*) have 200 samples per year. Plots (*a*) and (*c*) show recruitment deviations used in simulation model (grey circles), estimated recruitment deviations in estimation model (black circles), and 95% intervals based on standard error of estimated values (vertical bars). Plots (*b*) and (*d*) show transformed standard errors (triangles) based on eq. 10 and the bias adjustment fraction chosen for each year of the estimation model (thick line with peak at 0.52 for data-poor model and 0.72 for data-rich model).



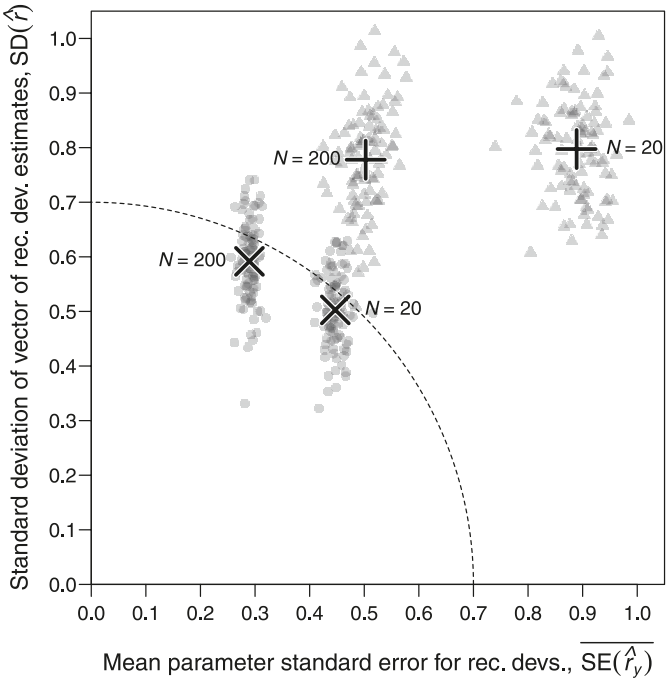
practice, this model had the least bias in B_0 of all estimation models (relative error of –1.3%), but performed second best at estimating final depletion (–2.2%), with the $b_{max} = 0.72$ model having slightly less relative error (1.2%). For the data-rich simulation model with sample size of 200, the estimation model with bias adjustment ramp based on the observed variability ($b_{max} = 0.72$) performed best in estimating B_0 (relative error of –0.1%), but was second best in estimating final depletion (3.1%), with the model with ramp to $b_{max} = 1$ having less relative error (–1.1%). The worst performing models were those with constant bias adjustment fractions at 0 or 1, which differed from the simulated values by at least 9% for both quantities in both data-poor and data-rich estimation models. The SCAA-like model overestimated final depletion by almost 30% and underestimated B_0 by over 20%.

Over 100 runs of the SCAA-like estimation model, the mean of the SD($\widehat{\boldsymbol{r}}$) values, which measure variability of the estimated deviations in recruitment, calculated for the years 1975–2004, was 0.78 for the data-rich model and 0.80 for the data-poor model (large pluses in Fig. 3). In both cases, these values are higher than the simulated variability, which had $\sigma_R = 0.7$. Furthermore, the model with less information in the data had more variability in the estimated parameters. This is in contrast with the non-SCAA estimation models (large Xs in Fig. 3), where the objective function component for recruitment deviations is sufficient to decrease variability in the estimates by pulling these values toward zero as the information in the data goes down.

Estimating σ_R

When σ_R was changed from the true value of 0.7 in the age-structured simulation model, the sum of the variance in estimated deviations and the average variance around each estimate provided a total variance that was closest to σ_R^2 when the value of σ_R was specified at the true value (Fig. 6). Specifying a value of σ_R that is too high is akin to the SCAA approach in that it reduces the penalty associated with more variability in $\widehat{\boldsymbol{r}}_y$. Conversely, a value of σ_R that is too low will create a pull on the $\widehat{\boldsymbol{r}}_y$ toward zero that will be stronger than necessary in relation to the signal in the data. This suggests that tuning σ_R to satisfy the relationship $\sigma_R^2 = \mathrm{Var}(\widehat{\boldsymbol{r}}) + \overline{\mathrm{SE}(\widehat{\boldsymbol{r}}_y)}^2$ could be a good way to estimate its value. Indeed, this new approach to tuning σ_R provides greater precision and less bias in the resulting value than either estimating σ_R as an additional parameter or tuning it to

Fig. 3. Mean estimated uncertainty in estimated recruitment deviations (rec. devs.) vs. variability in the estimates for the age-structured model. Variability in simulated recruitment deviations is $\sigma_R = 0.7$ throughout. Grey points indicate results of 100 realizations of the model for each combination of age-composition sample size (indicated by text label associated with each cluster of points) and estimation model (circles = non-SCAA, triangles = SCAA). Large Xs and pluses indicate observed mean values for each cluster. Dotted arc indicates the predicted relationship between the two measurements, as described in eq. 9. No bias adjustment was applied in any case ($b_y = 0$ for all years). Application of a ramped bias adjustment for the non-SCAA approach increased the mean values (position of the Xs) by less than 2% in each dimension.



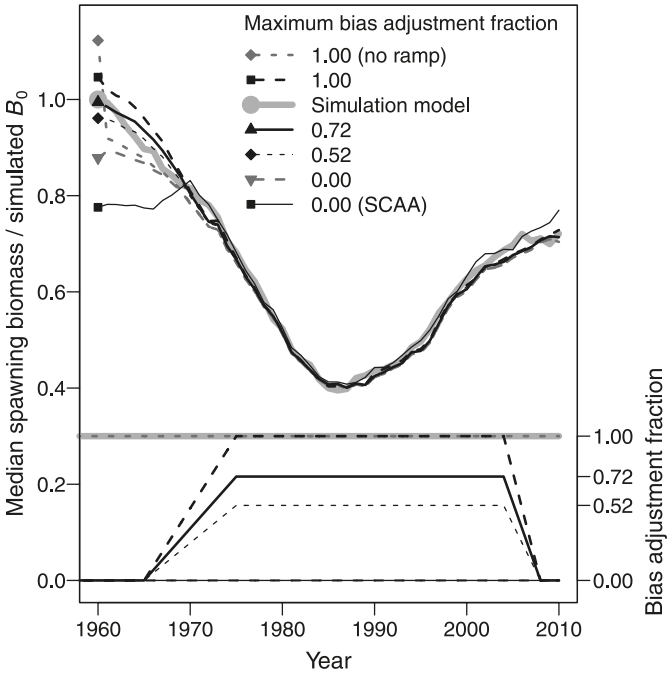
match the standard deviation of the $\hat{r}_y$ without considering the uncertainty around these estimates (Fig. 7).

MCMC

The MCMC results indicated that the relationship between the variability and uncertainty in $\hat{r}_y$ holds even more closely to the prediction from analytical derivations than in the maximum likelihood estimate (MLE) calculations (Figs. 8 and 9). The MCMC samples represent the combination of variability between and uncertainty within the $\hat{r}_y$ values (Fig. 8). This aggregate distribution has variability represented by σ_R throughout the time series of recruitments, indicating that a bias adjustment fraction of 1.0 is appropriate when integrating over the uncertainty in each $\hat{r}_y$. The distributions of MCMC samples of $\hat{r}_y$ from any given year were generally found to be left-skewed, especially for years where the simulated recruitment value was large (Fig. 10).

Convergence diagnostics indicated that chains longer than the 1 000 000 samples used in this analysis would be necessary to pass some tests of convergence. For the 6900 parameter chains (69 parameters estimated for each of 100 simulations), 99% of them passed the Heidelberger and Welch test of stationarity, but only 74% passed the associated

Fig. 4. Median biomass from the 100 simulations and for each of the six estimation models applied to the simulated data, represented as a fraction of the true B_0 value (left axis) and the associated bias adjustment fraction for each year (right axis). Only results for the data-rich scenario are shown, as the patterns are similar for the two cases. The points at 1960 indicate the median B_0 values in each case, while 1961 is the beginning of the time series that includes the nonequilibrium initial age composition derived from estimated $\hat{r}_y$ for the years 1945–1960. The B_0 points at 1960 correspond to the thick median lines within the light gray boxplots in Fig. 5.



half-width test of convergence (Plummer et al. 2010). However, the length of the MCMC chains appeared to be sufficient for purposes of illustrating the relationships between uncertainty and variability in recruitment estimates and the differences in treatment of recruitment bias adjustment required for MCMCs compared with MLE calculations.

Discussion

The results of the analytical derivations and the simulation experiments have some clear implications for the estimation of recruitment in age-structured population models. First, within the context of the simulation model used in this analysis, it is possible to allow recruitment deviations to be estimated parameters even in years with almost no information about recruitment and still achieve relatively unbiased estimates of initial biomass and final depletion. Second, without an objective function component penalizing recruitment deviations away from 0, the variability of $\hat{r}_y$ will increase as the information about the r_y goes down. That is, in a data-poor scenario, the $\hat{r}_y$ in a SCAA-like model will follow the noise in the data rather than accurately approximate the true variability in recruitment. Third, when an objective function component for recruitment deviations is used under the common assumption that recruitments are lognormally distributed, a bias adjustment is necessary to adjust the mean of the distribution, and this adjustment should be a function of the esti-

Fig. 5. Boxplots showing the performance of estimation models using six different bias adjustment estimation models. The dark and light boxes correspond to the data-poor and data-rich simulation models, respectively. 100 simulations are used for each combination of simulation and estimation models. Maximum bias adjustment fractions of 0.52 and 0.72 correspond to the curves shown in Fig. 2, while the value 1 indicates a similar ramp but with a peak at 1. Cases 0 (SCAA) and 0 have no bias adjustment, and 1 (no ramp) has 100% bias adjustment for all years. The SCAA case has low weight applied to the likelihood contribution for recruitment. Boxplots show median and 50% and 95% intervals. Boxes with thick outlines indicate the cases predicted to perform best based on the theory above.

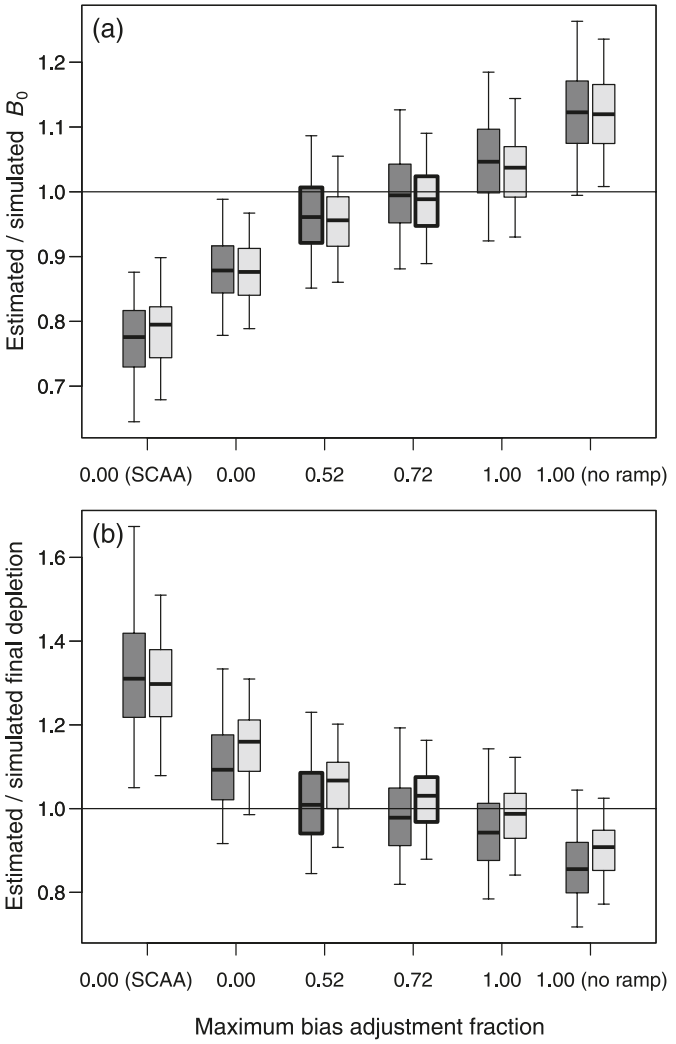
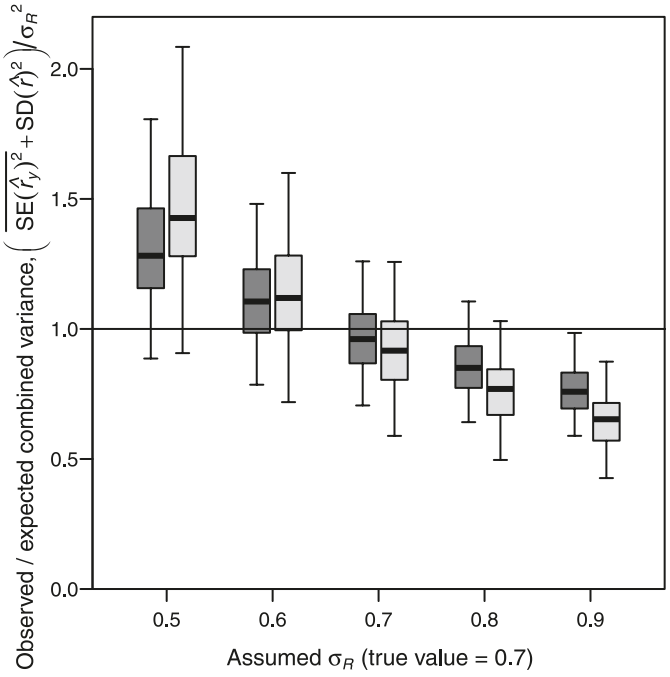


Fig. 6. Ratio of observed combined variance (mean squared standard error of parameter estimates + variance across recruitment deviations) to expected combined variance ($\sigma_R^2 = 0.7^2$) as described in eq. 9 over a range of assumed values for σ_R in the estimation model. The dark and light boxes correspond to the data-poor and data-rich simulation models, respectively. Boxplots show median and 50% and 95% intervals. Data-poor models use the bias adjustment pattern with $b_{max} = 0.52$, and data-rich models use $b_{max} = 0.72$ as described in Table 2.

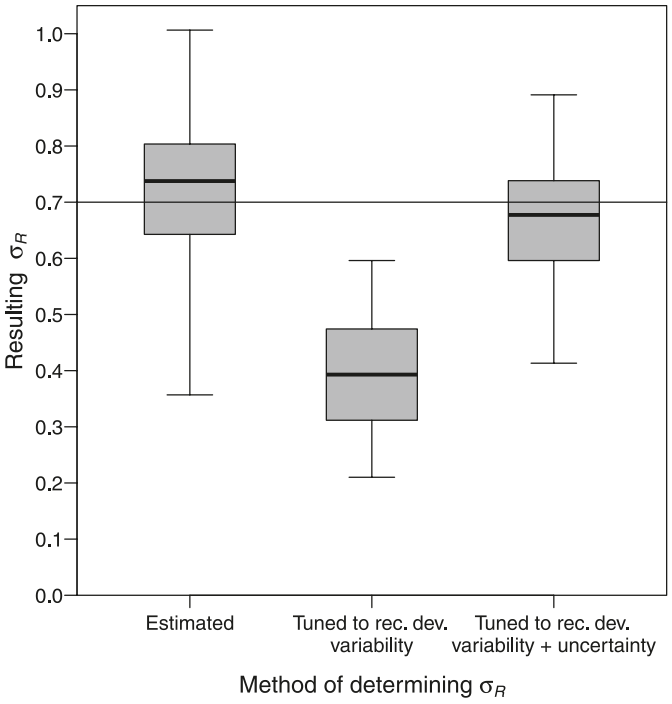


mated variability of the recruitment deviations rather than the true underlying variability in the population. This last conclusion leads to a suggested refinement to previous approaches to this problem; the bias adjustment applied to a lognormal distribution needs to change across years within a model because of heterogeneity in the information about recruitment in the data. A typical pattern will likely have zero bias adjustment for the early years in the model, a ramp up to a plateau, the height of which will be a function of the best level of information about recruitment in the data, and then drop down to zero bias adjustment in the final years of the model, where individuals typically have not yet been selected by any fishing gear, and thus the data provide no information about the strength of these cohorts. For models that include forecast years in the estimation (as in Maunder et al. 2006), the bias adjustment fraction for $\hat{r}_y$ in the forecast is $b_y = 0.0$ and these forecast parameters get an estimated uncertainty based solely on σ_R.

The results of both the analytical and simulation analysis indicate that the variability across estimated recruitments and the uncertainty in these estimates are tightly coupled according to a simple relationship. This relationship allows the uncertainty in recruitment estimates to inform the choice of bias adjustment pattern applied in the model. This requires running such models at least twice: once to get the point estimates of recruitment deviations and the variability around these values, which are then used to estimate a pattern of bias adjustment values for each year, and then a second time to apply the chosen pattern.

Among the patterns of bias adjustment applied in the age-structured simulation models, no one pattern performed best with regard to estimating B_0 and the final depletion. However, some results were unequivocal. For the data-rich simulations, a bias adjustment that ramped from 0 in the early years up to either 0.72 or 1.0 performed better than all other estimation models, including those with adjustment values of either 0 or 1.0 for all years. For data-poor simulations, ramped bias adjustment was again required to minimize bias in B_0 and depletion, but in this case, the maximum bias ad-

Fig. 7. Performance of three methods for determining σ_R in the data-poor model. Tuning methods use bisection method to match σ_R to either recruitment variability, $\sigma_R^2=\mathrm{Var}(\hat{r})$, or the combination of variability and uncertainty, $\sigma_R^2=\mathrm{Var}(\hat{r}) + \overline{\mathrm{SE}(\hat{r}_y)}^2$, as described in eq. 9, calculated over the years 1975–2004 in both cases. All cases had the bias adjustment pattern with $b_{max} = 0.52$ as described in Table 2. Estimated values are bounded between 0.1 and 1.5. Boxplots show median and 50% and 95% intervals. The horizontal at 0.7 indicates the true value of σ_R used in the simulations.



justment fraction needed to be less than 1, with the two cases of 0.52 and 0.72 outperforming all others.

The result of no single model outperforming the others may be due to various factors, including the relatively simple approximation of a linear ramp in bias adjustment up to and down from a constant value. However, application of this approach was able to achieve a median bias of at most 3.2% in key quantities for both data-rich and data-poor simulation models and may thus be recommended as an important refinement to the constant bias adjustment methods typically employed in such models.

Estimation of σ_R is known to be problematic in fisheries assessment models (Maunder and Deriso 2003) and is often attempted by iteratively tuning σ_R to match the standard deviation of the estimated recruitment deviations. For the data-poor simulation scenario examined here, tuning in this way performed the worst of three methods considered. Estimating σ_R directly had relatively unbiased results for this parameter, but high variability. The unbiased performance of this estimation approach was made possible by the adjustment introduced in eq. 13 to downweight the contribution of $-\ln(\sigma_R)$ according to the degree of bias adjustment. The best results came from iteratively tuning σ_R to match the combination of both variability among recruitment deviations and uncertainty about those estimates. This approach should be more robust than the estimation method in cases where lack of informa-

Fig. 8. Boxplots showing results of 100 MCMC chains for the data-poor case of the age-structured model. (*a*) Median values from each chain (between-simulation variability), (*b*) difference between samples from each chain and their respective median values (within-simulation variability), and (*c*) samples from all chains combined. Boxplots show median and 50% and 95% intervals.

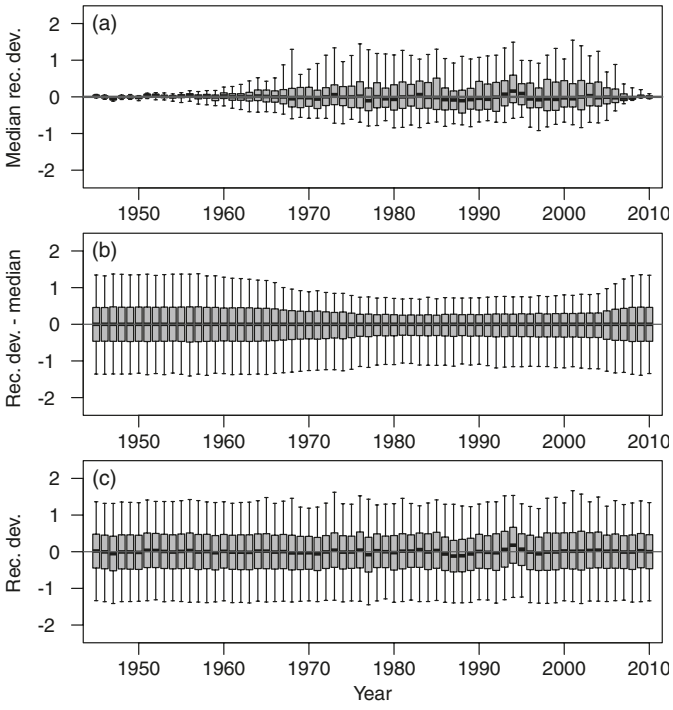


Fig. 9. Relationship between variability around median with chains to variability of medians across chains for MCMC samples. Vertical axis is standard deviation of points shown as boxplots in Fig. 8*a*. Horizontal axis is standard deviation of points shown in Fig. 8*b*. Shading of points indicates year from darkest (1945) to lightest (2010). Large X shows mean of values from 1975 to 2004, which is equivalent to the X for the data-poor stock in Fig. 3.

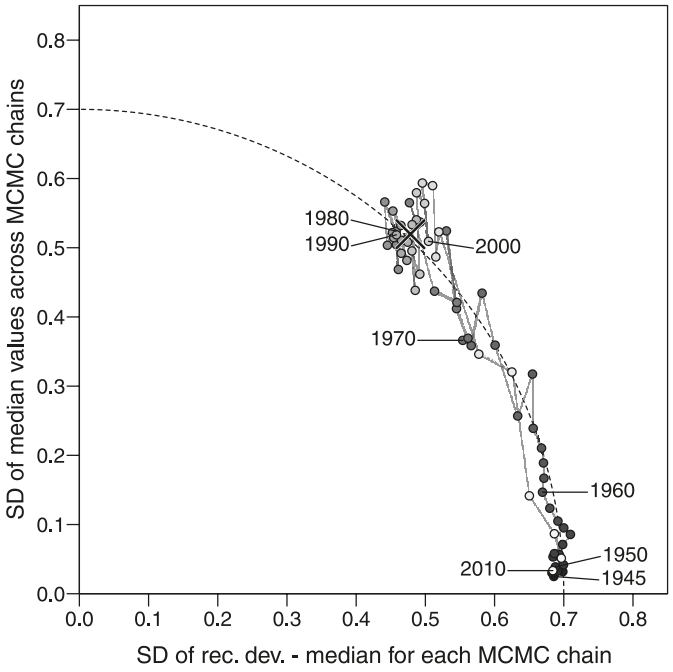
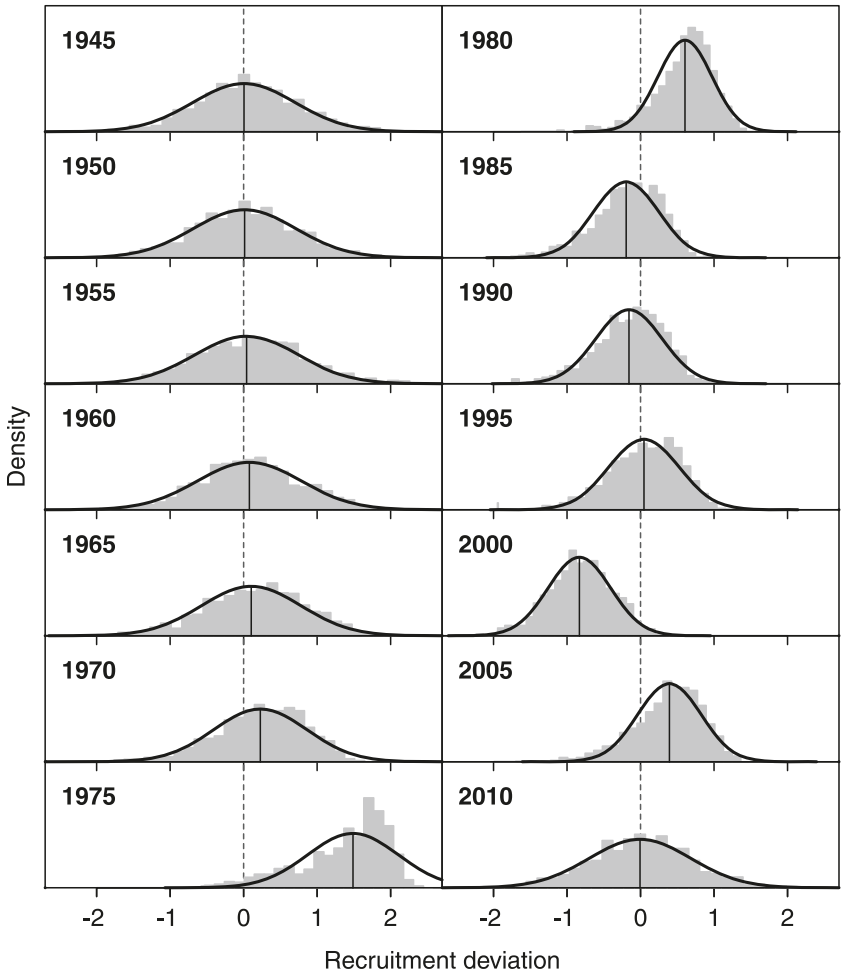


Fig. 10. Distribution of MCMC values and normal approximation from the maximum likelihood estimate (MLE) for $\widehat{r}_y$ from every fifth year in an example case within the data-poor scenario. Grey histograms show MCMC samples and black lines show normal distribution around point estimate based on Hessian matrix. Distribution of MCMC samples is left-skewed for all years after 1960, especially for years with larger recruitment levels.



tion about recruitment causes shrinkage of the estimated deviations toward 0.

This combination of variability and uncertainty is essentially what is found in Bayesian models, in which MCMC samples of the $\widehat{r}_y$ are the combination of variability among recruitment deviations and variability around the best fit to the data. Thus, when applying MCMC, the bias adjustment fraction should be equal to 1 for all years, regardless of the information in the data about recruitment. This result is consistent with the recommendations of Maunder and Deriso (2003) and Maunder et al. (2006), who suggest that Bayesian integration be used in the context of a constant bias adjustment based on σ_R applied in all years. In the SS model, bias adjustment levels and ramps used during the MLE phases are automatically turned to a full bias adjustment of 1.0 for the MCMC phase. In situations where time does not permit the use of MCMC for a full Bayesian integration, we suggest that for both data-poor and data-rich models, a ramped bias adjustment pattern can be used to allow the estimation of recruitment deviations in all years of an age-structured model to achieve both reasonable MLE estimates of key parameters and also accurately reflect the uncertainty in these parameters associated with variability in recruitment.

The small amount of skewness in the MCMC samples of $\widehat{r}_y$ suggests that the estimates of uncertainty from the normal approximation to the likelihood surface may be slightly inaccurate and may contribute to the minor differences between theory and practice in the MLE results presented above. The skewness is logical given the nature of the data and the penalized log-likelihood model formulation. Strong recruitments are well represented in the data, and this signal is strong relative to the penalty for deviating from median recruitment. This is consistent with the finding that large proportions are more precisely sampled than small proportions in fishery catch-at-age sampling programs (Crone and Sampson 1998). On the other hand, to estimate a weak recruitment of strength $-2\sigma_R$, the data would have to fit better with that estimate to overcome the penalty for such a large deviation. However, weak recruitments are not distinct in the data, especially if ageing error or use of length-composition data blur the information among adjacent recruitments, so weak recruitment estimates tend to accumulate at moderately weak levels rather than be distributed down to more extreme values. This skewness in the estimates deserves further investigation and may lead to refinements in the error distribution used in the estimation models.

Although this analysis was conducted with a fixed spawner–recruitment relationship that was essentially flat (steepness = 0.99), the results have strong implications for estimation of the degree of estimated depletion in stock abundance and for the estimated curvature in a spawner–recruitment relationship. The steepness parameter basically controls the expectation for mean recruitment as spawning biomass declines below the equilibrium unfished level. The findings presented here, as displayed in Fig. 4, clearly show that the degree of bias adjustment imposes an offset between the mean unfished biomass level and the mean level that occurs during the data-rich portion of the time series. If the assessment model has been set to estimate steepness, the estimate of steepness will be influenced by the degree of bias adjustment. Inaccurate specification of the bias adjustment pattern will result in biased estimates of steepness or biased estimates in the degree of stock depletion.

Not only would one expect an effect of the degree of bias adjustment on estimates of steepness, it seems likely that when steepness is much below 1.0, then an effect of steepness on recruitment variability could occur (Minto et al. 2008; Methot 2009). This effect is due to the observation that in natural systems some of the variability in young fish survival happens before the density-dependent life stage, thus introducing randomness in the effective spawner output. This before-density-dependent variability is dampened at high spawner abundance, but magnified by steepness at low spawner abundance.

Our results provide an explanation for the finding that MLE estimates of stock abundance sometimes differ from the mean of the posterior distribution from MCMC applied to the same data set. Typically, assessment practitioners using MLE approaches have set the bias adjustment to 0 or 1 throughout the modeled time series under the expectation that the bias was constant and would be simply offset by the estimated mean level of recruitment. Our results show that when there is heterogeneous information about recruitment variability, the use of bias adjustment of 0 or 1 will result in a biased trend in estimated stock biomass. Using our recommended procedure for a ramp in the bias adjustment, we were able to obtain MLE results that are consistent with the MCMC results.

Stock assessment estimation models that are being fit to long time series with changing degrees of information about recruitment variability bridge the gap between age-aggregated, biomass dynamics models (Prager 2002) during data-poor periods and highly age-specific SCAA-like models during data-rich periods. In the real world, recruitment variability occurs throughout the time series, but in the biomass dynamics estimation model it is a hidden process. In Bayesian implementations of data-poor models (Walters and Martell 2004; Maunder and Deriso 2003), the estimaton process can integrate across the variability in recruitment and achieved unbiased results. Here we have shown that proper attention to the expected distribution of estimated recruitments allows for achievement of unbiased results in a maximum likelihood context also.

The protocols for recruitment bias adjustment described here have been implemented in the SS assessment program, which is available from the NOAA Fisheries Toolbox (http://nft.nefsc.noaa.gov). The implementation includes ability to specify a ramp and plateau for the level of the bias adjustment and model outputs from eq. 11 to enable user adjustment of the level of the bias adjustment. As more experience with use of this bias adjustment approach is obtained, it is likely that more flexible alternatives to the linear ramp and plateau will be needed. The principles are not unique to SS, and the protocols for bias adjustment should be implemented in any model that uses a penalized likelihood approach to estimate annual recruitment deviations.

Acknowledgements

We thank our many colleagues, especially Jim Ianelli, Mark Maunder, André Punt, and Ian Stewart, who have contributed to the development of the concepts we have presented in this paper. We also thank Allan Hicks, Owen Hamel, and André Punt for reading an earlier version and providing suggestions for clarification of the presentation. Ian Taylor's substantial contributions were made possible through NOAA Award NA17RJ1232 to the University of Washington's Joint Institute for the Study of the Atmosphere and Ocean.

References

ADMB Project. 2009. AD Model Builder: automatic differentiation model builder. Developed by David Fournier and freely available from http://admb-project.org [updated 22 June 2011].

Conte, S.D., and de Boor, C. 1980. Elementary numerical analysis: an algorithmic approach (third edition). McGraw-Hill, New York.

Crone, P.R., and Sampson, D.B. 1998. Evaluation of assumed error structure in stock assessment models that use sample estimates of age composition. *In* Fishery stock assessment models. *Edited by* F. Funk, T.J. Quinn II, J. Heifetz, J.N. Ianelli, J.E. Powers, J.F. Schweigert, P.J. Sullivan, and C.-I. Zhang. Alaska Sea Grant College Program Rep. AK-SG-98–01, University of Alaska Fairbanks, Fairbanks, Alaska. pp. 355–370.

Deriso, R.B., Quinn, T.J., II, and Neal, P.R. 1985. Catch-age analysis with auxiliary information. Can. J. Fish. Aquat. Sci. **42**(4): 815–824. doi:10.1139/f85-104.

Fournier, D.A., and Archibald, C.P. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. **39**(8): 1195–1207. doi:10.1139/f82-157.

Garrison, T.M., Hamel, O.S., and Punt, A.E. 2011. Can data collected from marine protected areas improve estimates of life-history parameters? Can. J. Fish. Aquat. Sci. **68**(10): 1761–1777. doi:10.1139/f2011-073.

Haddon, M. 2001. Modeling and quantitative methods in fisheries. Chapman & Hall/CRC Press, Boca Raton, Fla.

Haltuch, M.A., and Hicks, A. 2009. Status of the U.S. petrale sole resource in 2008. Pacific Fishery Management Council, Portland, Ore.

Hamel, O.S. 2009. Status and future prospects for the Pacific ocean perch resource in waters off Washington and Oregon as assessed in 2009. Pacific Fishery Management Council, Portland, Ore.

Maunder, M.N., and Deriso, R.B. 2003. Estimation of recruitment in catch-at-age models. Can. J. Fish. Aquat. Sci. **60**(10): 1204–1216. doi:10.1139/f03-104.

Maunder, M.N., Harley, S.J., and Hampton, J. 2006. Including parameter uncertainty in forward projections of computationally intensive statistical population dynamic models. ICES J. Mar. Sci. **63**: 969–979.

Methot, R.D. 1989. Synthetic estimates of historical abundance and mortality for northern anchovy. *In* Mathematical analyses

of fish stock dynamics: review and current applications. *Edited by* E.F. Edwards and B. Megrey. Am. Fish. Soc. Symp. 6. pp. 68–82.
Methot, R.D. 1990. Synthesis model: an adaptable framework for analysis of diverse stock assessment data. Int. North Pac. Fish. Comm. Bull. **50**: 259–277.
Methot, R.D. 2000. Technical description of the stock synthesis assessment program. National Marine Fisheries Service, Seattle, Wash. NOAA Tech. Memo. NMFS-NWFSC-43.
Methot, R.D. 2009. Stock assessment: operational models in support of fisheries management. *In* The future of fishery science in North America. *Edited by* R.J. Beamish and B.J. Rothschild. Fish and Fisheries Series 31. pp. 137–165.
Minto, C., Myers, R.A., and Blanchard, W. 2008. Survival variability and population density in fish populations. Nature, **452**(7185): 344–347. doi:10.1038/nature06605. PMID:18354480.
Plummer, M., Best, N., Cowles, K., and Vines, K. 2010. coda: Output analysis and diagnostics for MCMC. R package version 0.14-2 [online]. Available from http://cran.r-project.org/web/packages/coda/index.html.
Power, M. 1996. The testing and selection of recruitment distributions for North Atlantic fish stocks. Fish. Res. **25**(1): 77–95. doi:10.1016/0165-7836(95)00395-9.
Prager, M.H. 2002. Comparison on logistic and generalized surplus production models applied to swordfish, *Xiphias gladius*, in the north Atlantic Ocean. Fish. Res. **58**(1): 41–57. doi:10.1016/S0165-7836(01)00358-7.
Punt, A.E., Hobday, D.K., and Flint, R. 2006. Bayesian hierarchical modelling of maturity-at-length for rock lobsters, *Jasus edwardsii*, off Victoria, Australia. Mar. Freshw. Res. **57**(5): 503–511. doi:10.1071/MF05261.
Quinn, T.P., II, and Deriso, R.B. 1999. Quantitative fish dynamics. Oxford University Press, New York.
R Development Core Team. 2011. R: a language and environment for statistical computing [online]. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3–900051–07–0. Available from http://www.R-project.org/.
Ricker, W.E. 1975. Computation and interpretation of biological statistics of fish populations. Bull. Fish. Res. Board Can. 191.
Stewart, I.J. 2009. Status of the U.S. canary rockfish resource in 2009 (update of 2007 assessment model). Pacific Fishery Management Council, Portland, Ore.
Stewart, I.J., Forrest, R.E., Grandin, C., Hamel, O.S., Hicks, A.C., Martell, S.J.D., and Taylor, I.G. 2011. Status of the Pacific hake (whiting) stock in U.S. and Canadian Waters in 2011. Pacific Fishery Management Council, Portland, Ore.
Walters, C., and Martell, S.J.D. 2004. Fisheries ecology and management. Princeton University Press, Princeton, N.J.

Appendix A. Population dynamics, observation, and statistical models as implemented in Stock Synthesis

The model description below, and with parameters shown in Table A1, illustrates a simple subset of the wide range of options available in Stock Synthesis. For a user manual describing the full range of model options, see the NOAA Fisheries Toolbox (http://nft.nefsc.noaa.gov).

Population model

Numbers at age, mortality, and catch

The number at age a in gender g in the beginning of each year is incremented as

$$N_{yga} = \begin{cases} 0.5R_y\, e^{r_y} - b_y\sigma_R^2/2 & \text{for } a = 0 \\ N_{yga-1}\, e^{-Z_{yga-1}} & \text{for } 0 < a < A \\ N_{yga-1}\, e^{-Z_{yga-1}} + N_{yga}\, e^{-Z_{yga}} & \text{for } a = A \end{cases} \quad \text{(A.1)}$$

where

R_y is the expected recruitment derived from the spawner–recruit curve (described below),
r_y is the deviation in recruitment in year y,
b_y is bias adjustment fraction applied in year y, as described in eq. 15 and associated text,
σ_R is the standard deviation for recruitment deviations,
N_{yga-1} is the number at age $a - 1$ in gender g at the beginning of year y,
Z_{yga-1} is the total mortality rate for age $a - 1$ in gender g, in year y, and
A is the accumulator age.

The total mortality in year y for age a in gender g is calculated as

$$Z_{yga} = M + F_y\beta_a \quad \text{(A.2)}$$

where

M is the natural mortality (assumed constant across ages and genders),
F_y is the fishing mortality in year y, and
β_a is the selectivity at age a.

The mortality values are used to calculate the catch in numbers at age as

$$C_{yga} = N_{yga}\frac{F_y\beta_a}{Z_{ya}}(1 - e^{-Z_{ya}}) \quad \text{(A.3)}$$

where the values of F_y in the two equations above are determined by iteratively solving for the values that make the total expected catch match the observed catch in each year (with starting value for the F search based upon a Pope's (1972) approximation so that the search itself is differentiable). This total expected catch (in biomass) for year y is

$$\widehat{C}_y = \sum_{a=0}^{A}\beta_a\sum_{g=1}^{2}C_{yga}\sum_{l=1}^{A_l}\widetilde{\phi}_{al}w_l \quad \text{(A.4)}$$

where

A_l is the number of length bins,
$\widetilde{\phi}_{al}$ is the proportion of numbers at age a within length bin l in the middle of the year, as described below in eq. A.11, and
w_l is the mean mass of individuals in length bin l, calculated from L_l, the middle length of length bin l, as $w_l = 2L_l^3 \times 10^{-6}$ for both males and females. The length bins used for the calculations in this example are the 2 cm intervals between 2 and 90 cm.

Virgin age structure

The age structure of the virgin population (divided equally among males and females) is given by

Table A1. Parameter values used in the age-structured simulation analysis.

Symbol	Description	Value in simulation
A	Accumulator age (years)	20
A_l	Number of length bins	45
M	Natural mortality (year^{-1})	0.2
a_1	Reference age for growth parameterization	1
a_2	Reference age for growth parameterization	12
L_1	Length at age a_1 (cm)	30
L_2	Length at age a_2 (cm)	70
K	Growth coefficient (year^{-1})	0.25
CV_1	Length CV at age a_1	0.1
CV_2	Length CV at age a_2	0.128
x	Constant added to the standard deviation of all ages	0.1
Ω_1	Mass coefficient (kg·cm^{-3} × 10^{-6})	2
Ω_2	Mass exponent	3
Ω_3	Length at 50% maturity (cm)	50
Ω_4	Maturity slope (cm^{-1})	–0.25
β_1	Age at 50% selectivity (years)	5
β_2	Age at 95% selectivity – age at 50% selectivity	2
R_0	Initial recruitment (log)	10.5
h	Stock–recruit steepness	0.99
σ_R	Recruitment variability	0.7
$r_{y_1} \ldots r_{y_2}$	Recruitment deviations for years y_1 to y_2	Stochastic
$y_1^b \ldots y_4^b$	Break points for piecewise linear function for annual bias adjustment fraction, b_y	1944, 1945, 2010, and 2011; or 1965, 1975, 2004, and 2008

$$N'_{0a} = \begin{cases} R_0\, e^{-Ma + r_{y_1-a} - b_{y_1-a}\sigma_R^2} & \text{for } 0 < a < A \\ R_0 \dfrac{e^{-aM}}{1 - e^{-M}} & \text{for } a = A \end{cases} \quad \text{(A.5)}$$

where

R_0 is the initial recruitment,
y_1 is the initial year of the model,
r_{y1-a} is the recruitment deviation in the year y_{1-a}, and
b_{y1-a} is the bias adjustment fraction for the year y_{1-a}, as described in eq. 15 and associated text.

Spawning biomass

Spawning biomass is calculated at the beginning of the year as

$$S_y = \sum_{a=0}^{A} N_{y,g=1,a} w'_a \quad \text{(A.6)}$$

where

$N_{y,g=1,a}$ is the number of females of age a in year y, and
w'_a is the average spawning output for females of age a defined in eq. A.12 below.

Spawner–recruit relationship

Expected recruitment is based on the Beverton–Holt spawner–recruit curve, as modified by Mace and Doonan (1988), but with constant spawning biomass above the equilibrium value

$$R_y = \begin{cases} \dfrac{4hR_0S_y}{S_0(1-h) + S_y(5h-1)} & \text{for } S_y < S_0 \\ R_0 & \text{for } S_y \geq S_0 \end{cases} \quad \text{(A.7)}$$

where

h is the parameter for steepness of the stock–recruitment function,
S_0 is the unfished equilibrium spawning biomass corresponding to R_0, and
S_y is the spawning biomass at the beginning of the spawning season in year y.

Growth

Growth was assumed to follow the von Bertalanffy growth curve, parameterized in terms of reference ages a_1 and a_2 (Schnute and Fournier 1980). The mean length at age a, designated either L_a when a is the integer age at the beginning of the year or $\widetilde{L}_a$ when a is the real age in the middle of the year, is calculated as

$$L_a = L_\infty + (L_1 - L_\infty)\, e^{-K(a-a_1)} \quad \text{(A.8)}$$

where

a_1 is the first reference age,
L_1 is the mean length at age a_3,
K is the growth coefficient, and
L_∞ is the mean asymptotic length, calculated from

$$L_\infty = L_1 + \frac{L_2 - L_1}{1 - e^{-K(a_2-a_1)}} \quad \text{(A.9)}$$

where

a_2 is the second reference age, and
L_2 is the mean size at age a_2.

The standard deviation of length at age a increases with mean length at age as

$$\text{(A.10)}\qquad \sigma_a = \begin{cases} \widetilde{L}_a \mathrm{CV}_1 + x & \text{for } a \le a_1 \\ \widetilde{L}_a \left[\mathrm{CV}_1 + \dfrac{(\widetilde{L}_a - L_1)}{(L_2 - L_1)} (\mathrm{CV}_2 - \mathrm{CV}_1) \right] + x & \text{for } a_1 < a < a_2 \\ \widetilde{L}_a \mathrm{CV}_2 + x & \text{for } a \ge a_2 \end{cases}$$

where

$\widetilde{L}_a$ is the mean length in the middle of the year for age a,
CV_1 is the coefficient of variation for length at age a_1,
CV_2 is the coefficient of variation for length at age a_2, and
x is a constant added to the standard deviation of all ages.

The resulting standard deviation in length at age is smoothly increasing from $\sigma_{a=0}$ at 1.7 to $\sigma_{a=20} = 9.40$.

Age–length population structure

The numbers at age for each gender are distributed across the defined length bins. The proportion in length bin l for age a, designated either ϕ_{al} when a is the integer age at the beginning of the year or $\widetilde{\phi}_{al}$ when a is the real age in the middle of the year, is calculated as

$$\text{(A.11)}\qquad \phi_{al} = \begin{cases} \Phi\left(\dfrac{L'_{\min} - L^*_a}{\sigma_a} \right) & \text{for } l = 1 \\ \Phi\left(\dfrac{L'_{l+1} - L^*_a}{\sigma_a} \right) - \Phi\left(\dfrac{L'_l - L^*_a}{\sigma_a} \right) & \text{for } 1 < l < A_l \\ 1 - \Phi\left(\dfrac{L'_{\max} - L^*_a}{\sigma_a} \right) & \text{for } l = A_l \end{cases}$$

where

Φ is the standard normal cumulative density function,
L'_l is the lower limit of length bin l,
$L'_{\min}$ is the lower limit of the smallest length bin,
$L'_{\max}$ is the lower limit of the largest length bin,
L^*_a is the mean length of age a, either L_a at the start of the year or $\widetilde{L}_a$ in the middle of the year, and
σ_a is the standard deviation of length at age a.

Maturity and fecundity

The fecundity at age is given by

$$\text{(A.12)}\qquad w'_a = \sum_{l=1}^{A_l} \phi_{al} \varphi_l w_l$$

where

ϕ_{al} is the proportion in length bin l for females of age a at the start of the year,
φ_l is the fraction mature, given by the logistic function $\varphi_l = \left[1 + \mathrm{e}^{-0.25(L_l - 50)}\right]^{-1}$, and
w_l is the mass of females in length bin l.

Selectivity

Selectivity was assumed to be a logistic function of age

$$\text{(A.13)}\qquad \beta_{fgal} = \left[1 + \mathrm{e}^{-\log(19)(a - \beta_1)/\beta_2}\right]^{-1}$$

where

β_1 is the age at 50% selectivity and
β_2 is the difference between the age at 95% selectivity and the age at 50% selectivity.

Observation model

Ageing imprecision

The proportion of age a assigned to age a' is

$$\text{(A.14)}\qquad \Omega_{a'a} = \begin{cases} \Phi\left(\dfrac{a'}{\sigma_a} \right) & \text{for } a' = 1 \\ \Phi\left(\dfrac{a' + 1}{\sigma_a} \right) - \Phi\left(\dfrac{a'}{\sigma_a} \right) & \text{for } 1 < a' < A \\ 1 - \Phi\left(\dfrac{a'}{\sigma_a} \right) & \text{for } a' = A \end{cases}$$

where σ_a is the standard deviation of ageing imprecision at age a given by the formula

$$\text{(A.15)}\qquad \sigma_a = 0.525 + 0.05a$$

Age compositions

The expected proportion at age a' for gender g in year y is calculated as

$$\text{(A.16)}\qquad \widehat{p}_{yga'} = \frac{\sum_{a=0}^{A} \Omega_{a'a} \sum_{l=1}^{A_l} C_{ygal} + \varepsilon}{\sum_{a'=1}^{A} \left(\sum_{a=0}^{A} \Omega_{a'a} \sum_{l=1}^{A_l} C_{ygal} + \varepsilon \right)}$$

where $\varepsilon = 10^{-4}$ is added to make the calculations more robust to zero values.

Index of abundance

The CPUE for the model is calculated as catchability and vulnerable biomass

$$G_y = QB_y \tag{A.17}$$

where Q is the catchability of the fishery.

For simplicity, catchability was fixed at 1.0 in the operating model. In the estimation model, the catchability was calculated as the median unbiased value:

$$\widehat{Q} = \mathrm{e}\left[\frac{\sum_t \ln\left(\widehat{G}_y/\widehat{B}_y\right)/\sigma_1^2}{\sum_y 1/\sigma_1^2}\right] \tag{A.18}$$

where σ_1 is the standard deviation in log space of the index observations for all years, $\ln(G_y)$.

Statistical model

Objective function

The objective function L is the weighted sum of the individual components:

$$\mathcal{L} = \mathcal{L}_{\text{comp}} + \mathcal{L}_{\text{index}} + \omega\mathcal{L}_{\text{recruit}} \tag{A.19}$$

where

$\mathcal{L}_*$ are the negative log-likelihood components for the components for the age compositions, index of abundance, and recruitment deviations, as described below, and

ω is a weighting factor for the recruitment deviations.

Age compositions

The observed age compositions are assumed to have a multinomial distribution. The contribution to the objective function for the age compositions is

$$\mathcal{L}_{\text{comp}} = \sum_y \sum_{g=1}^{2} n_{yg} \sum_{a'=1}^{A} p_{yga'} \ln\left(p_{yga'}/\widehat{p}_{yga'}\right) \tag{A.20}$$

where n_{yg} is the number of observed ages in the catch in year y.

Index of abundance

The objective function component for the index of abundance is

$$\mathcal{L}_{\text{index}} = 0.5\sum_y \left[\frac{\ln(G_y) - \ln(\widehat{G}_y)}{\sigma_1}\right]^2 \tag{A.21}$$

Recruitment deviations

The objective function component for deviations in recruitment is

$$\mathcal{L}_{\text{recruit}} = 0.5\sum_y \left[\frac{r_y^2}{\sigma_R^2} + b_y \ln(\sigma_R)\right] \tag{A.22}$$

References

Mace, P.M., and Doonan, I.J. 1988. A generalised bioeconomic simulation model for fish population dynamics. New Zealand Fish. Assess. Res. Doc. 88/4.

Pope, J.G. 1972. An investigation of the accuracy of Virtual Population Analysis using cohort analysis. ICNAF Res. Bull. 9. pp. 65–74. [Also available in Cushing, D.H. (*Editor*). 1983. Key papers on fish populations. IRL Press, Oxford, UK. pp. 291–301.]

Schnute, J.T., and Fournier, D. 1980. A new approach to length–frequency analysis: growth structure. Can. J. Fish. Aquat. Sci. **37**(9): 1337–1351. doi:10.1139/f80-172.

This article has been cited by:

1. Catherine M. Dichmont, Roy A. Deng, Andre E. Punt, Jon Brodziak, Yi-Jay Chang, Jason M. Cope, James N. Ianelli, Christopher M. Legault, Richard D. Methot, Clay E. Porch, Michael H. Prager, Kyle W. Shertzer. 2016. A review of stock assessment packages in the United States. *Fisheries Research* **183**, 447-460. [CrossRef]
2. Kelli F. Johnson, Elizabeth Councill, James T. Thorson, Elizabeth Brooks, Richard D. Methot, André E. Punt. 2016. Can autocorrelated recruitment be estimated using integrated assessment models and how does it affect population forecasts?. *Fisheries Research* **183**, 222-232. [CrossRef]
3. Cole C. Monnahan, Kotaro Ono, Sean C. Anderson, Merrill B. Rudd, Allan C. Hicks, Felipe Hurtado-Ferro, Kelli F. Johnson, Peter T. Kuriyama, Roberto R. Licandeo, Christine C. Stawitz, Ian G. Taylor, Juan L. Valero. 2016. The effect of length bin width on growth estimation in integrated age-structured stock assessments. *Fisheries Research* **180**, 103-112. [CrossRef]
4. Timothy J. Miller, Jonathan A. Hare, Larry A. Alade. 2016. A state-space approach to incorporating environmental effects on recruitment in an age-structured assessment model with an application to southern New England yellowtail flounder. *Canadian Journal of Fisheries and Aquatic Sciences* **73**:8, 1261-1270. [Abstract] [Full Text] [PDF] [PDF Plus] [Supplemental Material]
5. Ian J. Stewart, Cole C. Monnahan. 2016. Implications of process error in selectivity for approaches to weighting compositional data in fisheries stock assessments. *Fisheries Research* . [CrossRef]
6. André E. Punt, Malcolm Haddon, L. Richard Little, Geoffrey N. Tuck. The effect of marine closures on a feedback control management strategy used in a spatially aggregated stock assessment: a case study based on pink ling in Australia. *Canadian Journal of Fisheries and Aquatic Sciences*, ahead of print1-14. [Abstract] [Full Text] [PDF] [PDF Plus] [Supplemental Material]
7. Jonathan J. Deroba, Timothy J. Miller. 2016. Correct in theory but wrong in practice: Bias caused by using a lognormal distribution to penalize annual recruitments in fish stock assessment models. *Fisheries Research* **176**, 86-93. [CrossRef]
8. André E. Punt, Malcolm Haddon, L. Richard Little, Geoffrey N. Tuck. 2016. Can a spatially-structured stock assessment address uncertainty due to closed areas? A case study based on pink ling in Australia. *Fisheries Research* **175**, 10-23. [CrossRef]
9. James T. Thorson, Kasper Kristensen. 2016. Implementing a generic method for bias correction in statistical models using random effects, with spatial and population dynamics examples. *Fisheries Research* **175**, 66-74. [CrossRef]
10. Claudio Castillo-Jordán, Neil L. Klaer, Geoffrey N. Tuck, Stewart D. Frusher, Luis A. Cubillos, Sean R. Tracey, Michael J. Salinger. 2016. Coincident recruitment patterns of Southern Hemisphere fishes. *Canadian Journal of Fisheries and Aquatic Sciences* **73**:2, 270-278. [Abstract] [Full Text] [PDF] [PDF Plus]
11. André E. Punt. 2016. Some insights into data weighting in integrated stock assessments. *Fisheries Research* . [CrossRef]
12. James T. Thorson, Jason M. Cope. 2015. Catch curve stock-reduction analysis: An alternative solution to the catch equations. *Fisheries Research* **171**, 33-41. [CrossRef]
13. André E. Punt, Malcolm Haddon, Geoffrey N. Tuck. 2015. Which assessment configurations perform best in the face of spatial heterogeneity in fishing mortality, growth and recruitment? A case study based on pink ling in Australia. *Fisheries Research* **168**, 85-99. [CrossRef]
14. J. T. Thorson, A. O. Shelton, E. J. Ward, H. J. Skaug. 2015. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. *ICES Journal of Marine Science* **72**:5, 1297-1310. [CrossRef]
15. McGilliard Carey R., Punt André E., Methot Richard D. Jr., Hilborn Ray. 2015. Accounting for marine reserves using spatial stock assessments. *Canadian Journal of Fisheries and Aquatic Sciences* **72**:2, 262-280. [Abstract] [Full Text] [PDF] [PDF Plus] [Supplemental Material]
16. J. T. Thorson, A. C. Hicks, R. D. Methot. 2015. Random effect estimation of time-varying factors in Stock Synthesis. *ICES Journal of Marine Science* **72**:1, 178-185. [CrossRef]
17. M. Dickey-Collas, N. T. Hintzen, R. D. M. Nash, P.-J. Schon, M. R. Payne. 2015. Quirky patterns in time-series of estimates of recruitment could be artefacts. *ICES Journal of Marine Science* **72**:1, 111-116. [CrossRef]
18. S. Subbey, J. A. Devine, U. Schaarschmidt, R. D. M. Nash. 2014. Modelling and forecasting stock-recruitment: current and future perspectives. *ICES Journal of Marine Science* **71**:8, 2307-2322. [CrossRef]
19. Momoko Ichinokawa, Hiroshi Okamura, Yukio Takeuchi. 2014. Data conflict caused by model mis-specification of selectivity in an integrated stock assessment model and its potential effects on stock status estimation. *Fisheries Research* . [CrossRef]
20. James T. Thorson, Ian G. Taylor, Ian J. Stewart, André E. Punt. 2014. Rigorous meta-analysis of life history correlations by simultaneously analyzing multiple population dynamics models. *Ecological Applications* **24**:2, 315-326. [CrossRef]
21. André E. Punt, Martin Dorn. 2014. Comparisons of meta-analytic methods for deriving a probability distribution for the steepness of the stock–recruitment relationship. *Fisheries Research* **149**, 43-54. [CrossRef]

22. Hui-Hua Lee, Kevin R. Piner, Richard D. Methot, Mark N. Maunder. 2014. Use of likelihood profiling over a global scaling parameter to structure the population dynamics model: An example using blue marlin in the Pacific Ocean. *Fisheries Research* . [CrossRef]
23. James T. Thorson, Ian G. Taylor. 2013. A comparison of parametric, semi-parametric, and non-parametric approaches to selectivity in age-structured assessment models. *Fisheries Research* . [CrossRef]
24. Robyn E. Forrest, Murdoch K. McAllister, Steven J.D. Martell, Carl J. Walters. 2013. Modelling the effects of density-dependent mortality in juvenile red snapper caught as bycatch in Gulf of Mexico shrimp fisheries: Implications for management. *Fisheries Research* **146**, 102-120. [CrossRef]
25. JT Thorson, IJ Stewart, IG Taylor, AE Punt. 2013. Using a recruitment-linked multispecies stock assessment model to estimate common trends in recruitment for US West Coast groundfishes. *Marine Ecology Progress Series* **483**, 245-256. [CrossRef]
26. Athol R. Whitten, Neil L. Klaer, Geoffrey N. Tuck, Robert W. Day. 2013. Accounting for cohort-specific variable growth in fisheries stock assessments: A case study from south-eastern Australia. *Fisheries Research* **142**, 27-36. [CrossRef]
27. Jason M. Cope. 2013. Implementing a statistical catch-at-age model (Stock Synthesis) as a tool for deriving overfishing limits in data-limited situations. *Fisheries Research* **142**, 3-14. [CrossRef]
28. Ian G. Taylor, Richard D. Methot. 2013. Hiding or dead? A computationally efficient model of selective fisheries mortality. *Fisheries Research* **142**, 75-85. [CrossRef]
29. Ian J. Stewart, Allan C. Hicks, Ian G. Taylor, James T. Thorson, Chantell Wetzel, Sven Kupschus. 2013. A comparison of stock assessment uncertainty estimates using maximum likelihood and Bayesian methods implemented with the same model framework. *Fisheries Research* **142**, 37-46. [CrossRef]
30. Ian G. Taylor, Vladlena Gertseva, Richard D. Methot, Mark N. Maunder. 2013. A stock–recruitment relationship based on pre-recruit survival, illustrated with application to spiny dogfish shark. *Fisheries Research* **142**, 15-21. [CrossRef]
31. Sally E. Wayte. 2013. Management implications of including a climate-induced recruitment shift in the stock assessment for jackass morwong (Nemadactylus macropterus) in south-eastern Australia. *Fisheries Research* **142**, 47-55. [CrossRef]
32. Richard D. Methot, Chantell R. Wetzel. 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research* **142**, 86-99. [CrossRef]
33. Jiangfeng Zhu, Yong Chen, Xiaojie Dai, Shelton J. Harley, Simon D. Hoyle, Mark N. Maunder, Alexandre M. Aires-da-Silva. 2012. Implications of uncertainty in the spawner–recruitment relationship for fisheries management: An illustration using bigeye tuna (Thunnus obesus) in the eastern Pacific Ocean. *Fisheries Research* **119-120**, 89-93. [CrossRef]
34. GarrisonThomas M., HamelOwen S., PuntAndré E.. 2011. Can data collected from marine protected areas improve estimates of life-history parameters?. *Canadian Journal of Fisheries and Aquatic Sciences* **68**:10, 1761-1777. [Abstract] [Full Text] [PDF] [PDF Plus]