SCRS/2013/100 Collect. Vol. Sci. Pap. ICCAT, 70(4): 1725-1757 (2014)

A GENERALIZED BAYESIAN SURPLUS PRODUCTION STOCK ASSESSMENT SOFTWARE (BSP2)

Murdoch K. McAllister[1]

SUMMARY

A generalized Bayesian surplus production stock assessment software (BSP2) is presented as an update to ICCAT's current BSP software. BSP2 differs from BSP in a few different respects. Most importantly, BSP2 provides a state-space implementation of the deterministic Bayesian generalized surplus production model found in BSP. BSP2 models both process error in the dynamics equations to account for the effects of, e.g., interannual variation in recruitment, and error in predicted observations. BPS2 provides outputs to enable the computation of Bayes factors (BFs). BFs show Bayes posterior weights for different models and can be particularly important when structurally different models suggest different interpretations of stock status. A generalized production function implementation (i.e., the Fletcher model) is incorporated in which the ratio of the most productive stock size to carrying capacity can be set at hypothesized values other than the Schaefer model value of 0.5. The software can accommodate a variety of different priors for key parameters including carrying capacity (K), the maximum rate of population increase (r), and the ratio of stock biomass in the initial year to carrying capacity (B_{init}/K).

RÉSUMÉ

Le programme d'évaluation des stocks de production excédentaire bayésienne généralisée de l'ICCAT (BSP) est présenté en tant qu'actualisation du programme BSP actuel de l'ICCAT. Le BSP2 diffère du BSP à différents égards. Plus important encore, le BSP2 fournit une mise en œuvre état-espace du modèle bayésien déterministe de production excédentaire généralisée rencontrée dans le BSP. Le BSP2 modélise à la fois l'erreur de processus dans les équations dynamiques pour tenir compte des effets de, p.ex. la variation interannuelle dans le recrutement, et l'erreur dans les observations prédites. Le BSP2 fournit des résultats afin de permettre le calcul des facteurs bayésiens (BF). Les BF montrent les pondérations bayésiennes a posteriori pour différents modèles et peuvent s'avérer particulièrement importants quand des modèles structurellement différents suggèrent différentes interprétations de l'état du stock. Une mise en œuvre de la fonction de production généralisée (p.ex. le modèle de Fletcher) est incorporée en vertu de laquelle le ratio de la taille du stock le plus productif par rapport à la capacité de charge peut être établi à des valeurs postulées autres que la valeur de 0,5 du modèle de Schaefer. Le programme peut intégrer plusieurs priors pour les paramètres-clés, dont la capacité de charge (K), le taux maximum d'accroissement de la population (r) et le ratio de la biomasse du stock de l'année initiale par rapport à la capacité de charge (Binit/K).

RESUMEN

Se presenta un programa de evaluación de stock de producción excedente bayesiano generalizado (BSP2) como actualización para el actual programa BSP de ICCAT. El BSP2 difiere del BSP en unos cuantos aspectos. El más importante, el BSP2 proporciona una implementación estado-espacio del modelo de producción excedente bayesiano generalizado determinista de BSP. El BSP2 modela tanto el error de proceso en las ecuaciones dinámicas para tener en cuenta los efectos de, por ejemplo, la variación interanual en el reclutamiento, y el error en las observaciones predichas. El BSP2 proporciona resultados para permitir el cálculo de los factores Bayes (BF). Los BF muestran las ponderaciones posteriores bayesianas para diferentes modelos y pueden ser especialmente importantes cuando modelos estructuralmente diferentes sugieren distintas interpretaciones del estado del stock. Se incorpora la implementación de una función de producción generalizada (es decir, el modelo Fletcher) en la que la ratio entre el tamaño más productivo del stock y la capacidad de transporte puede establecerse en valores hipotetizados diferentes al valor de 0,5 del modelo Schaefer. El programa puede integrar una amplia variedad de distribuciones previas diferentes

[1] Fisheries Centre, AERL, 2202 Main Mall, University of British Columbia, Vancouver, B.C., Canada, V6T 1Z4. m.mcallister@fisheries.ubc.ca

para parámetros clave, lo que incluye la capacidad de transporte (K), la tasa máxima de incremento de la población (r) y la ratio de la biomasa del stock en el año inicial con respecto a la capacidad de carga (B_{init}/K).

KEYWORDS

Bayesian surplus production, State-space model, Generalized production function, Bayes factors

Introduction

This paper presents an updated version of ICCAT's deterministic Bayesian surplus production model (BSP) software documented in McAllister *et al.* (2001) and McAllister and Babcock (2006). The updated software offers a state-space model version that permits the user to specify a fixed value for the standard deviation in annual deviates from the state dynamics equations. Observation error variance is obtained via iterative re-weighting. The updated software enables the computation of Bayes factors for different model runs when structurally different models are fitted to the same data sets (Kass and Raftery 1999). Bayes factors provide an empirical probabilistic basis to evaluate the credibility of different model runs and can be applied to summarize uncertainty in results between different model runs and combine these results to form mixture distributions that represent the joint uncertainty between different models (McAllister and Kirchner 2002). The updated software includes as before a generalized production function in which the ratio of biomass at maximum production to carry capacity can be fixed at values other than the constant of 0.5 in the Schaefer production model (McAllister *et al.* 2000). An implementation in which fishing effort for a given fishing fleet is used as a covariate for fishing mortality rate and the catchability coefficient is estimated from the prediction of one or more catch observations is incorporated (Stanley *et al.* 2009).

The updated software has been developed through implementations in peer-reviewed assessments of numerous Canadian fish stocks (e.g., Stanley *et al.* 2009, 2012, Yamanaka *et al.* 2012a, McAllister and Duplisea 2011). The updated BSP software has offered a flexible framework with which to explore alternative hypotheses about how growing seal and sea lion populations have impacted rockfish populations (Yamanaka *et al.* 2012a), about technological creep in abundance indices derived from long time series of commercial trawl data (King *et al.* 2012), and about how bycatch in spatially extensive non-target fisheries have contributed to stock declines in low productivity rockfish stocks (Stanley *et al.* 2009, 2012). The updated BSP software (BSP2) is illustrated with an application to data from the 2013 stock assessment of North Atlantic swordfish (ICCAT 2010).

Surplus production model equations

The updated BSP software (BSP2) applies a Bayesian surplus production model (Prager 1994; McAllister and Babcock 2006) that utilizes Sampling Importance Resampling (Rubin 1987, 1988). The surplus production model is updated using a conventional state-space formulation (Millar and Myer 2000; Stanley *et al.* 2009). The version of the model applied in this updated BSP software package have been developed for and applied in several peer-reviewed stock assessments that have been conducted by Fisheries and Oceans Canada. These include B.C. Bocaccio rockfish (Stanley *et al.* 2009, 2012), inside waters yelloweye rockfish (Yamanaka *et al.* 2012a), inside and outside waters quillback rockfish (Yamanaka *et al.* 2012b), four offshore lingcod stocks (King *et al.* 2012) and four Atlantic redfish stocks (McAllister and Duplisea 2011, 2012). Required inputs for the program are a time series of catch biomass starting from near the beginning of the fishery, at least one catch rate (CPUE) index of abundance with coefficients of variation (CV) for model fit deviations and fixed values for the state-space process error variance ($\sigma_{process}$) and autocorrelation coefficient for future process error deviates (ρ). Estimated parameters include carrying capacity (K), the maximum intrinsic rate of population growth (r), the biomass in the first modeled year defined as a ratio of K (p_0), variance parameters for each CPUE series, and constant of proportionality (q) for each CPUE series. Prior probability distributions (priors) were specified for all of the estimated parameters.

Deterministic model components

The default surplus production model in the software is Prager's instantaneous *F* version of the Schaefer production model (Schaefer 1954; Prager 1994). State dynamics are modelled by assuming that biomass in a given year is a function of biomass in the previous year, the instantaneous fishing mortality rate, and two parameters that describe the impact of earlier biomass in growth, *r* and *K*:

$$B_{y+1} = B_y + rB_y\left(1 - \frac{B_y}{\mathrm{K}}\right) - F_y B_y \quad \text{(F1)}$$

where y is the year, B_y the stock biomass at the start of year y, r the intrinsic rate of increase, K the carrying capacity and F_y the instantaneous fishing mortality rate during year y. For the initial year, an additional parameter, p_0, is estimated which gives the ratio of initial stock biomass to carrying capacity (e.g., for north Atlantic swordfish, $p_0 = B_{1950}/K$).

Abundance indices are assumed to be directly proportional to stock biomass. The deterministic observation equation is:

$$\hat{I}_{j,y} = q_j B_y \quad \text{(F2)}$$

where q_j is the constant of proportionality for the abundance index j, $I_{j,y}$ the observed abundance index j in year y and is the model $\hat{I}_{j,y}$ predicted value for $I_{j,y}$.

Stochastic model components

The state-space approach allows for deviations from model predictions (i.e., random variability) in both (i) the data (e.g., relative biomass indices) and (ii) the unobserved state of the system of interest (e.g., annual stock biomass that has recruited to the fishery) (Millar and Meyer, 2000). These two components of the system are modelled within a single probabilistic framework that can be highly flexible (Rivot *et al.*, 2004). Fisheries modellers tend to choose multiplicative lognormal errors (Millar and Meyer, 2000), which is what is applied in the presented model. The abundance index data are assumed to be lognormally distributed:

$$I_{j,y} \sim \text{lognormal}\left(\ln\left(\hat{I}_{j,y}\right), \sigma^2_{\text{obs},\,j}\right) \quad \text{(F3)}$$

where $I_{j,y}$ is the observed index of abundance for series j in year y, is the $\hat{I}_{j,y}$ predicted index for series j and $\sigma_{\text{obs},\,j}$ is the standard deviation in the error deviation between the log predicted index and the log observed index j.

The stochastic form equation F1 (i.e., the process equation) is:

$$\log\left(B_{y+1}\right) = \log\left(B_y + rB_y\left(1 - \frac{B_y}{\mathrm{K}}\right) - F_y B_y\right) + \varepsilon_{process,y} - \frac{\sigma^2_{process}}{2}, \quad \text{(F4a)}$$

where, $\varepsilon_{process,\,y} \sim \text{Normal}\left(0, \sigma^2_{process}\right)$.

Given these equations, the expected value for B_{y+1} is:

$$E\left(B_{y+1}\right) = B_y + rB_y\left(1 - \frac{B_y}{\mathrm{K}}\right) - F_y B_y \quad \text{(F4b)}$$

Also, under unfished conditions the posterior mean of B_y is K and under the maximum sustainable harvest rate the posterior mean of B_y is $K/2$.

The stochastic form of equation F2 (i.e., the observation equation) is:

$$\text{(F5)} \qquad \log\left(I_{j,y}\right) = \log\left(q_j\right) + \log\left(B_y\right) + \varepsilon_{obs,j,y}$$

where $\varepsilon_{obs,\,j,y} \sim \text{Normal}\left(0, \sigma^2_{\text{obs},j}\right)$.

Both $\varepsilon_{process,y}$ and $\varepsilon_{obs,j,y}$ are i.i.d. random variables in all modelled years up to 2012. For each future year in the projections, I have modelled $\varepsilon_{process,y}$ to be positively autocorrelated with a correlation coefficient, ρ (see Stanley *et al.* (2009) for details on the autocorrelation equations). There were too few years in which it was possible to estimate the correlation in process error deviates because non-zero estimates of process error only became non-zero after 2000. I therefore applied the commonly applied default value for ρ of 0.5. The sensitivity of results to different values for ρ was evaluated in the BSP application to bocaccio (Stanley *et al.* 2009) and projection results were found to be relatively insensitive to values between 0.5 and 0.7 but more pessimistic than assuming that ρ = 0.

A summary of key parameters estimated in the surplus production model is provided in **Table 1**. A summary of derived management parameters is provided in **Table 2**.

A summary of prior distributions for estimated parameters is given in **Table 3**. A more detailed description of the methods used to determine each prior is provided below.

Computing a prior density function for the maximum intrinsic rate of increase (*r*)

The prior for *r* for North Atlantic swordfish that had been formulated by McAllister *et al.* (2000) and had been applied in assessments since then was updated using an updated methodology and inputs. McAllister *et al.* (2000) applied a stochastic life-table approach to compute a prior for the maximum intrinsic rate of increase (*r*) for North Atlantic swordfish. This is the base case prior for *r* that has been applied in ICCAT's BSP applications to Atlantic swordfish since then. McAllister *et al.*'s (2000) methodology includes input distributions for the survival rate of pre-recruits and the natural mortality rate of fish that have recruited to the exploitable stock biomass. It presumes fixed assumptions about fecundity at age and the fraction mature at age. This methodology has been extended in Canadian BSP2 stock assessments (e.g., Stanley *et al.* 2009; Yamanaka *et al.* 2012a) to a more easily parameterized protocol that applies the Euler-Lotka method to compute r (Lotka 1907) which offers a near exact approximation of the Leslie matrix approach (McAllister *et al.* 2001). Stanley *et al.* (2009) replaces the prior for egg to age 1 survival rate with a prior for the steepness stock-recruit parameter. Steepness is a unitless parameter that reflects the fraction of average unfished recruitment achieved when spawning potential is reduce to 20% of unfished conditions. While it was possible to use literature based estimates of batch fecundity and frequency of spawning bouts to formulate a prior for egg to age zero survival rate (McAllister *et al.* 2000), the quantification of uncertainty in survival rate was entirely arbitrary. The formulation of a prior for the steepness parameter is more accessible due to its common usage in stock assessments and the numerous meta-analyses of stock-recruit data sets that have formulated priors for steepness whereby the central tendency and variance can have a rigorous empirical basis (Dorn 2002; Michielsens and McAllister 2004; Forrest *et al.* 2010). The updated methodology includes uncertainty the stock-recruit steepness (*h*) parameter and the rate of natural mortality (*M*), means and variances for the female growth parameter estimates, the length to weight conversion factors, and parameters for the fraction maturity-at-age schedule (the prior covariances in parameter values can be assumed to be zero or can be empirically based). See appendix A for details on the updated methodology and inputs. The resulting prior was lognormal with a mean of 0.424 and a CV of 0.39.

Carrying capacity (*K*)

The prior for carrying capacity (K) in BSP has been commonly assumed to be uniform over a large range of values, e.g., between 5,000 tonnes and 1,000,000 tons in the current illustration, to enable equal credibility for small and large possible values for K. The upper bound could be set at the highest unfished stock size of any related fish stock worldwide or at some large value beyond the support of the data (i.e., where the posterior density is very close to zero relative to the mpd).

In some stock assessments, (e.g., McAllister and Duplisea 2011; King *et al.* 2012) a uniform prior on K has appeared to be unsuitable because posterior distributions for some assessed stock units had a thick tail at the upper bound. This problem has previously been noted by Millar and Meyer (2000). An alternative approach is to apply a uniform prior over the log of K with the same upper and lower bounds (see McAllister and Duplisea 2011 and King *et al.* 2011). This alternative thins out the flat upper tail in posteriors for K and initial stock size, but has relatively little influence on posterior median results for all quantities of interest.

Chris Francis (pers. commn) has suggested that uniform on log K or uniform on log(B_0) is more consistent with logical reasoning about uncertainty than a uniform on K prior. For example, it would appear to be more logically consistent to suggest that relative ranges of values are equally credible at different point values for K. Take for example a hypothesized value for K of e.g. 100,000 tons. A biologist could claim quite fairly that he or she is uncertain about K plus or minus 50% whether it was 100,000 tons, 500,000 or 5,000,000 tons. It would seem to be less credible to claim that he or she was equally uncertain whether it was 100,000 tons plus or minus 50,000 tons (plus or minus 50%) or 500,000 tons, plus or minus 50,000 tons (i.e., plus or minus 10%) or 5,000,000 tons plus or minus 50,000 tons (plus or minus 1%). Presuming a constant fixed range of values at increasingly higher values for K would thus appear to be increasingly more certain at higher values of K. In contrast, a uniform on log K prior would imply the same relative (or percentage) uncertainty at increasingly higher values of K. The uniform prior over the log of K is thus suggested as an alternative reference case prior for K.

Ratio of initial biomass to carrying capacity (P_0)

The first year of the total catch time series for North Atlantic swordfish is 1950. The prior distribution for p_o from previous Atlantic swordfish assessments suggests the stock biomass in 1950 (B_{1950}) was at lightly fished conditions at this time. The prior for p_0 was assumed to be log-normal with a prior mean of 0.85 and a SD in log(p_0) of 0.25 (ICCAT 2010).

Process error variance

To illustrate the implementation of a state-space BSP model, the standard deviation of $\varepsilon_{process,y}$, σ_{process}, was arbitrarily set at different values ranging from 0.005 to 0.15 in several different model runs to evaluate the effect on stock assessment results of applying different values for σ_{process}. Implementing non-zero values for σ_{process} allows the model to account for interannual variability in stock biomass due to variability in stock dynamics processes that were not explicitly modeled (e.g., interannual variability in recruitment, variation in growth and the rate of natural mortality). A value for σ_{process} of 0.05 would result in interannual changes in total recruited stock biomass of about 5% on average. A continued run of 5% positive errors would result in a net 28% increase in stock biomass in five years, all other things being equal. This would be 45% with σ_{process} set at 0.075, 65% at 0.1 and about 111% at 0.15. As the value for σ_{process} increases, numerical integration becomes less computationally efficient. This is due to more of the importance draws resulting in trajectories that are inconsistent with the data and that cause the population to crash prior to the current year. The number of draws required to achieve reasonably precise estimates of the target posterior density functions thus increased markedly as σ_{process} was increased to 0.15 and final posterior approximations, though sufficiently precise were slightly more bumpy with the largest values tried for σ_{process}. As σ_{process} was increased, it would be expected that the outputted value for $\sigma_{\text{obs},j}$ at the mpd would decrease. In this instance, the model fit value for $\sigma_{\text{obs},j}$ at the posterior mode was insensitive to the inputted value for σ_{process} and remained at about 0.25.

As in Stanley *et al.* (2009) and McAllister and Duplisea (2012), I will evaluate the sensitivity of results to this parameter. I have applied a range of values including 0.005, 0.01, 0.05, 0.075, 0.1 and 0.15. The lowest setting for σ_{process}, i.e., at 0.005, approaches an observation error only model and the model behaved nearly identically to the original BSP model version (McAllister and Babcock 2006).

Observation error variance

Values for $\sigma_{obs,j}$ (i.e., the standard deviation of $\varepsilon_{obs,j}$, from equation F-5) using this methodology are obtained by implementing a procedure of iterative reweighting for each model run. For each abundance index, a trial inputted value for $\sigma_{obs,j}$ is applied and a model fit value for $\sigma_{obs,j}$ is obtained at the posterior mode or maximum posterior density (mpd). This value is taken and inputted and the model is rerun subsequently to obtain the mpd of parameter values. This iterative process is repeated until the inputted and model fit values are practically identical (usually only one or two iterations is sufficient). The last outputted value for $\sigma_{obs,j}$ is increased by about 20% and applied as an input to importance sampling to accommodate model fits that deviate from the posterior mode and to thus avoid producing posterior distributions that are too narrow, i.e., based on values for $\sigma_{obs,j}$ obtained at the mpd. The outputted values obtained for $\sigma_{obs,j}$ tended to be similar across different model runs (at about 0.25), with 0.3 being the value that was applied in all runs.

It is presumed that values for $\sigma^2_{obs,j}$ are the sum of (i) the variance for each index j, determined analytically from the construction of the survey indices ($\sigma^2_{ind,j}$) and (ii) the variance presumably due to interannual processes affecting the annual availability of the fish stock to the gear or fleet associated with the abundance index ($\sigma^2_{int,j}$) (e.g., due to variation in the spatial distribution of the fish stock, $\sigma^2_{obs,j} = \sigma^2_{ind,j} + \sigma^2_{int,j}$). Thus in the iterative reweighting, when values for $\sigma_{ind,j}$ were available, the values for $\sigma_{int,j}$ were adjusted so that the inputted values for $\sigma_{obs,j}$ exceeded by about 20% the values for $\sigma_{obs,j}$ that were outputted from the stock assessment model.

Constant of proportionality (*q*)

The prior probability density function (pdf) for the constant of proportionality for abundance indices, q_j, is by default treated as uniform on the log of q_j over the interval [-20,2]. Thus, the prior for q is treated as non-informative over a wide range of potential values. This is the so-called "Jeffrey's prior" which is commonly applied to scale parameters such as q (Box and Tiao 1973). When there's been no scientific research devoted to formulating and informative prior for q, it is commonly accepted that the most defensible prior for q is a non-informative one that ranges from values less than one to values above one (McAllister *et al.* 1994; McAllister and Ianelli 1997).

The computational shortcut of Walters and Ludwig (1994) for integrating the joint posterior pdf with respect to scale parameters such as q is implemented in BSP (McAllister and Babcock 2006) and BSP2. BSP2 implements this shortcut for the lognormal likelihood function of abundance indices when $\sigma_{obs,j}$ is either constant for an index or varies by year. The software also implements this shortcut should a normal density function of the abundance indices be applied, e.g., when there are one or more abundance index values that are zero (Stanley *et al.* 2009).

Method of approximation of the posterior distribution

The sampling/importance resampling (SIR) algorithm was used to compute marginal posterior distributions for BSP model parameters and quantities of interest (Rubin 1987, 1988; McAllister *et al.* 1994). Importance sampling can be shown via the strong law of large numbers (Ross 2009) and a five line statistical proof (Appendix 2) that as the number of samples increases, the importance sampling distribution approaches the posterior density function for model parameters and derived quantities of interest (Berger 1985, p.263; McAllister *et al.* 1994). With the application of an importance function with good properties (Oh and Berger 1992), importance sampling can be reasonably swift and numerically efficient for problems with up to about 110 combined key and nuisance parameters (e.g., Stanley *et al.* 2012). An additional convenient feature of importance sampling is that it provides an efficient numerical approximation of the probability of the data given the model, which is required for the computation of Bayes factors in the evaluation of the credibility of alternative models given the data (Kass and Raftery 1995).

The key output statistics computed include marginal posterior distributions of current stock biomass (B_{2012}), current stock biomass to carrying capacity (B_{2012}/K), the ratio of current stock biomass to stock biomass at MSY (B_{2012}/B_{MSY}), the replacement yield in 2012 ($RepY_{2012}$), the ratio of the replacement yield in 2012 to the catch biomass in 2012 ($RepY_{2012}/C_{2012}$), and the ratio of fishing mortality rate in 2012 to fishing mortality rate at *MSY* (F_{2012}/F_{MSY}). Posterior samples of B_{2012}/B_{MSY}, F_{2012}/F_{MSY}, and posterior medians for B_y/B_{MSY}, F_y/F_{MSY}, are also outputted for the formulation of a Kobe plot to indicate stock status.

For the majority of runs, precise and consistent approximations of posterior distributions were obtained with runs that took only about 5 minutes on a notebook computer (i.e., about one million draws from the importance function). In a few instances, longer runs were required (i.e., up to 36 million importance samples) since the importance functions applied were for some parameters, e.g., *K*, considerably wider than their marginal posteriors, and the strong correlation patterns between some of the parameters, e.g., were only partially accounted for in the importance function applied. However, a run of 36 million however took only a few hours of computing. The marginal posteriors for the quantities of interest were reliably estimated with the maximum importance ratio for any one draw taking no more than about 0.5% in each of the runs conducted. Runs using alternative importance functions, (e.g., with different variances in the key parameters), yielded practically identical marginal posterior estimates. Post model, pre-data runs were carried out using a few different priors for *K*, to evaluate the effect on the model output distributions for key quantities of interest of running the model with the priors and the catch data. The marginal prior and posterior *pdf*s of *r* and *K* are plotted below to show the extent to which priors have been updated. SIR was also applied to compute Bayes factors when comparing the credibility of alternative model settings to the reference case runs (see below). Due to major differences in interpretation, Bayes factor comparisons should not be made between runs with different priors for *K*, e.g., between runs with uniform on *K* and uniform on log *K* priors.

Diagnostics for importance sampling

BSP and BSP2 software provides both numerical and graphical outputs to help diagnose how importance sampling is performing in approaching an approximation of the posterior distribution. These different diagnostic methods are detailed in McAllister and Kirchner (2002), McAllister *et al.* (2002) and McAllister and Babcock (2006). The two most important diagnostics are the ratio of the maximum importance weight to the sum of the importance weights from a given set of draws, n, from the importance function, %mwt. As n increases, %mwt should drop systematically, with occasional small increases when a draw that resets the maximum weight occurs. No one draw should take more than a few percentage of the sum of the weights compiled from importance sampling. The maximum weights should not consistently occur in the tails of the marginal posterior distribution for any one estimated parameter. Take for example, *K*. A good choice of an importance function would provide maximum weights occurring more in the area of highest posterior density. A poor choice of an importance function would tend to result in maximum weights coming from draws in the upper tail of the posterior distribution with the value for *K* at the maximum weight getting progressively larger as importance sampling increases. This would indicate that the marginal density for *K* in the importance function has tails that are sharper than those of the marginal posterior distribution for *K* and that importance sampling should be stopped and adjustments should be made to thicken the tails of the importance function for *K*.

The BSP2 software provides live updated values of %mwt, a graph of %mwt against the number of importance samples, a table printing values for all parameters at the posterior mode and a table next to it printing values for all parameters obtained at the draw with the maximum weight. Should the value for e.g. *K* at the maximum weight be consistently several times higher than the value at the mode and should this value at the maximum weight increase as %mwt are updated, then this would indicate that importance sampling should be stopped and the importance function readjusted to improve sampling efficiency.

A second useful diagnostic is the ratio of the coefficient of variation (CV) in the weights to the CV in the product of the prior and the likelihood function from the draws taken from the importance function (CV(w)/CV(LP)). A ratio of less than one suggests that the variation in the weights from the importance samples taken is less than the variation in the posterior surface and that the importance function applied is providing a stable approximation of the posterior distribution of interest. For example, if the importance function was actually the posterior density function itself, the value of the importance weights from importance draws would be constant and CV(W) would be zero. However, the variation in LP would reflect the variation in the relative posterior density of values of parameters taken from the importance function. The values of CV(w) and CV(LP) are also printed in an on-screen table, live with updates as importance sampling is carried out. With the most efficient importance functions for BSP2 and good datasets, CV(w) can remain as low as less than five. However, should a large value for $\sigma_{process}$ be applied, e.g., 0.15, CW(w) could be in the low to mid 200s.

Definition of reference case

I develop and present results using a reference case set of inputs and assumptions. For the reference case run, all inputs, assumptions and settings are to be based where possible on the best available information and scientific judgment. Prior distributions used in the reference case have been described above. The following list summarizes the key settings for the North Atlantic swordfish case study application:

- Prior mean *r* formulated for this stock (Appendix A).
- Stock trend index was the base case stock trend index in ICCAT (2013).
- Likelihood function of the abundance index data follows a lognormal distribution as in ICCAT (2009)
- Schaefer surplus production function (B_{MSY}/K=0.5) (as in ICCAT 2010)
- Prior mean p_0 (B_{1950}/ K) = 0.85, prior SD(log(p_0))=0.25 as in ICCAT (2009).
- $\sigma_{process}$ set at 0.05.
- Uninformative prior for *q*
- Lag 1 autocorrelation with the autocorrelation coefficient, ρ, set at 0.5 starts in 2012 (see Stanley *et al.* 2009 for the equations).
- CVs for stock trend indices obtained by iterative reweighting, with fixed observation error from survey imprecision and process error components determined by fitting the BSP model to the data to find the parameter values that give the maximum posterior density (mpd).

I allowed for the possibility of updating the reference case settings based on Bayes factor results obtained after fitting the model to the data in the different sensitivity analyses. I applied conservative criteria for updating the reference case settings to reduce the possibility of making excessively frequent and numerous changes or poorly justified changes that could result from random variation in the data when reference case settings are actually better approximations than the alternative settings. I would consider suggesting a revision of the reference case settings only if there was a very strong weight of evidence (e.g., a Bayes factor of less than 1/50 (see below)) against the reference case setting compared to the most credible alternative setting for some model component) in the posterior results.

Sensitivity analyses

Sensitivity tests were conducted to evaluate the effect of stock assessment model assumptions on stock status and projection results. A summary of the additional model runs carried out in this assessment is provided in **Table 4**, and a brief description of each analysis is provided below.

Prior distribution for *K* - To evaluate the sensitivity of model results to the prior distribution for *K*, two additional runs were conducted: one with a uniform on log K prior and a lognormal prior for K with a mean of 200,000 tons and a SD in the log of *K* of 0.8.

Prior distribution for *r* - To evaluate the sensitivity of model results to the informative prior distribution for *r*, two additional runs were conducted: one with a high prior mean for *r* and one with a low prior mean for r. The low r prior was obtained by applying a prior mean for *r* (0.28) that was two thirds of the reference case prior mean, while the high *r* prior was obtained by using a prior mean (0.56) that was one third higher than the reference case prior mean (0.42). The prior CVs were held constant at 0.49.

Prior distribution on B_{1950}/K (p_0 or B_{init}/K) - p_0 typically cannot be estimated from available data and when catch records are available from near to the beginning of the fishery or lightly exploited conditions it is commonly assumed that B_{init}/K falls at 90-100% of *K*. It has been found that if the catch series is more than a few decades, the final results are insensitive to the value assumed for p_0, provided it is over about 50%. In the BSP model, alternative prior means of 0.7 and 1.0 were considered.

Uncertainty in catch estimates - The influence of uncertainty in historic catch is evaluated by conducting runs where annual fixed catch values for all fisheries combined are set at 75% and then 150% (i.e., 0.5 and 2.0 times) of the time series of compiled fixed catch values. It is presumed that catch records earlier in the series were less well determined than those later in the series. Arbitrarily, 1985 was presumed as the cut-off point between less and more well determined catch records. Thus one alternative set of catch records was obtained by multiplying catch records 1950-1985 by 0.75 and leaving the catch records in subsequent years as is. A second alternative set of catch records was obtained by applying a catch multiplier of 1.5 also only to records 1950-1985.

Uncertainty in the standard deviation (SD) in process error (σ_p) deviates in annual stock biomass – Due to having one only time series of abundance, it is not possible to jointly estimate σ_p and the standard deviation in observation error deviates for the different abundance indices (σ_o). I thus evaluated the sensitivity of results to applying a lower and higher value for σ_p. The values applied in this sensitivity analysis were 0.005, 0.01, 0.05, 0.075, 0.10 and 0.15.

Uncertainty in the form of the surplus production function – it is typically not possible to estimate the third parameter in generalized surplus production functions such as the Fletcher or Pella Tomlinson models (Quinn and Deriso 1999). It is common thus to apply only the Schaefer surplus production model for which B_{msy}/K is fixed at 0.5. McAllister *et al.* (2000) provide a variant of the Fletcher production function that can incorporate an informative prior for r and avoids the infinite slope at the origin of the Fletcher and Pella Tomlinson functions when B_{msy}/K approaches and drops below 1/e (about 0.368). The original BSP and updated BSP2 software packages include this Fletcher model variant. I evaluate the sensitivity of results to setting B_{msy}/K at 0.3, 0.4 and 0.6.

Evaluation of credibility of alternative sensitivity analysis scenarios

To compare the credibility of competing model runs given the data in sensitivity analyses, I computed Bayes factors (Kass and Raftery 1995) for the reference case and for each of the related sensitivity runs. Bayes factors account for both the relative goodness of fit of the model to the data and the parsimony for each of the alternative models. They are calculated as the ratio of the marginal probability of the data for one model to that for another model. Bayes factors were computed by approximating the marginal posterior probability of the data given the model using the average value of the importance weights obtained from each model run (Kass and Raftery 1995; McAllister and Kirchner 2002). In all instances I referenced Bayes factors to our reference case model settings, i.e., the probability of the data for the reference case model was placed in the denominator and that for the model to which it was compared in the numerator. It is commonly held that nothing should be made of Bayes factor unless the value for it departs substantially from 1. Even fairly large or small Bayes factors can come from random chance in the data and possible misspecification of probability models for the data, e.g., treating annual errors for each observed index value as independent when they may not be independent. Thus, while a factor of 1/10 may appear to provide strong evidence against a model, the difference in fits of the model to the data could still have resulted from random chance in the data. Intermediate values for Bayes factor (e.g., between about 1/100 and 100) should be interpreted with restraint. Models with Bayes factors of about 1/100 could be interpreted as unlikely but not discredited. When Bayes factor is less than 1/1000, the model with lower credibility can be viewed as highly unlikely relative to the other.

Bayes factors can be used in decision tables to indicate the relative weight of results from different scenarios that are presented in a decision table (e.g., Yamanaka 2012a, 2012b, King *et al.* 2012). As has been noted previously it is recommended that the posterior probabilities that can be derived from Bayes factors for different model runs not be applied for model averaging of results especially for stock status indicators. Taking the average of results obtained from different models to provide a management recommendation could be nonsensical and possibly dangerous depending on how the combined results are presented. For example, the mean or median value of a bimodal mixture distribution could yield an intermediate result between the two alternative scenarios that had close to zero probability density, i.e., close to zero credibility compared to values closer to the modes of the mixture distribution. This could happen when the modes were separated and there was close to zero probability density between the modes in the mixture distribution. In such case, it would be inappropriate to report the model averaged, e.g., mean or median value from the combined outputs from two different models. This point has been made numerous times in the past and there should be some conventions adopted for instances in which there is multi-modality in the combined output distributions from different models, e.g., if there is multi-modality it would appear to report output values only from the different modes, and not from the medians, or means of the combined distribution.

The outputted value that is to be used in Bayes factor computations is the natural logarithm of the average importance weight taken from all of the samples from the importance function. Bayes factor for a given model run, i, is computed as:

$$BF(m_i) = \frac{exp\,(ln(\overline{w_i}) - const)}{exp\,(ln(\overline{w_r}) - const)}$$

where $\overline{w_i}$ is the average value for the importance weights from model i, $\overline{w_r}$ is the average value for the importance weights from the reference case model and *const* is a constant value that prevents the operation from producing excessively large or small values. The value for $ln(\overline{w_i})$ can be found in the histogram output file beside the label, log(average_wt).

Results

Post-Model Pre-Data Runs

A useful diagnostic from a Bayesian numerical integration is the so-called "post-model, pre-data" distribution. In this instance, the post model, pre-data distribution shows how the priors interact with the BSP model, and fixed inputs for catch before the model is fitted to the abundance index data. The post-model, pre-data distributions show the effect of the priors for model parameters, when applied in combination with the inputted values for catch, on the output distributions for the model parameters and quantities of interest such as current stock biomass and replacement yield. It is possible to find out whether the priors acting in combination and with the catch data within the simulated model structure lead to output distributions for, e.g., *r* and *K* that are markedly different from their priors, and thereby impact the shape of the output distributions for initial and current stock biomass. This allows analysts to evaluate the extent to which fitting the model to abundance index data updated the distributions determined by the interaction of the priors and inputted catch records within the stock assessment model formulation.

When a uniform on K prior was applied, the marginal output distributions for K and r, from the post-model pre-data run, resembled the prior pdfs with only minor updates in the shape of the priors. The biomass quantities showed mostly relatively flat or dome-shaped distributions as a result of the lower and upper bound cut-off points for the prior for *K*, and the minimum value for *K* that could allow population persistence to the present with the available catch records (**Figure 1**). In contrast, when a uniform on logK prior was applied, the marginal output distribution for *K* and other stock biomass quantities of interest reflected the prior down-weighting of larger values for *K* in the uniform on logK prior (**Figure 1**). The outputted distributions for stock status quantities, e.g., B_{cur}/B_{msy} and F_{cur}/F_{msy} however were influenced very little by the form of the prior for *K* (**Figure 1**). It is noted that the post-model pre-data output distributions were invariant to whether a data point from either one abundance index series or a data point from each of more than one abundance index series was included in the computation.

SIR diagnostics

For the reference case and all runs tried, importance sampling provided numerically stable results and sufficiently precise approximations of the marginal posterior distributions for parameters within a few hours of importance sampling. Importance sampling, however, became noticeably less efficient for the runs with the largest values for the standard deviation in process error (e.g., when $\sigma_{process}$ error was set at 0.15.). For all runs except those where when $\sigma_{process}$ error was set at 0.15, the maximum weight from any one draw from the importance function dropped rapidly to less than 0.5% within an hour of importance sampling. In runs with $\sigma_{process}$ error was set at 0.15 the maximum weight dropped below t 1%, after a few hours when importance sampling. In all runs, the CV in the weights was less than half of the value for CV in the likelihood times the prior, and the maximum weights were not consistently in the tails of the marginal posterior density functions for key parameters.

Reference Case Run

The reference case run results were very similar to results obtained with the application of ASPIC and BSP in the 2013 assessment. The model provided a fairly good fit to the abundance index data, except for one extreme outliers in the early part of the time series (**Figure 2a**). The estimated process error deviates differed relatively little from each other and without any significant difference from the prior distribution for process error (**Figure**

2b). Marginal posterior distributions for key parameters and quantities of interest showed a marked update from the priors and the post model-pre-data distributions (**Figure 3**). The prior for r was updated to some extent with the posterior mean for r shifting to about 0.37 (CV=0.28) down from the prior mean of 0.424 (CV=0.39) (Table 5). The update was made possible by the considerable drop in total catch biomass in the 1980s and the subsequent turn around of the stock trend index. The index increased less quickly than supported by the central tendency of the prior for *r*, and this lead to an update in the posterior to support lower values for r. **Figure 4** shows a Kobe scatter plot of 5000 draws from the posterior density function for the quantity F_{2012}/F_{msy}, B_{2012}/B_{msy}. The median value is centered just below the 1,1 point indicating that the current fishing mortality rate is just below the F_{msy} level and the stock biomass is very close to the B_{msy} level. The plotted trajectory of F/F_{msy} against B/B_{msy} shows that the stock decreased to below B_{msy} in the 1970s with *F* exceeding F_{msy} but then increased back up to the B_{msy} level as *F* levels decreased in the 1980s to the present.

Projections were carried out with future constant catch levels being set at 10,000, 11,000, 12,000, 13,000, 14,000 and 15,000 tons. Summary results are tabulated in **Table 6**. For 5, 10 and 20 year horizons, the model predicts that stock biomass can be expected to decrease for constant total catch policies of larger than 13,000 tons. A constant catch of 13,000 tons is expected to maintain the stock at close to it's *Bmsy* level.

Evaluation of sensitivity of results to different model settings

Stock status results were most sensitive to applying different assumptions about the value for B_{msy}/K (**Table 7**). See **Figure 5** for a plot of the alternative surplus production functions considered. The smallest values considered, i.e., 0.3 and 0.4 resulted in a considerably posterior higher median values for r (i.e., 0.43, and 0.41, respectively) compared to the reference case in which the posterior median was 0.36. Stock status was the most optimistic under these scenarios compared to other scenarios with e.g. the posterior median value for B_{2012}/B_{msy} at 1.46 and 1.21 compared to all other scenarios which ranged between about 0.89 and 1.15 (**Table 7**). The Kobe plot showing the time trajectory of stock status shows considerable sensitivity to the assumed value for B_{msy}/K with the smallest values showing the most optimistic endpoints (**Figure 6**). The smallest value for B_{msy}/K of 0.3 had considerably higher credibility given the data, with a Bayes factor of 7.9 relative to the reference case where B_{msy}/K was set at 0.5 (**Table 8**).

Runs with different prior means for *r*, different scenarios for possible systematic bias in early catch records, different priors for *K*, and different prior means for p_0, had relatively little impact on stock status results (**Table 7**, **Figure 7**). The scenario in which historic catches prior to 1986 were considered to be systematically low or high suggested somewhat different perceptions of stock status. For example, the version with a catch multiplier of 1.5 suggested a slightly more optimistic stock status results (**Table 7**) and had 1.8 times the credibility of the reference case run (**Table 8**).

Runs with different fixed values for the standard deviation in the process error ($\sigma_{\text{process error}}$), showed considerably different historic trajectories of stock biomass and increasingly variable and less precise time series of estimated process error deviates (**Figure 8**). The highest values for $\sigma_{\text{process error}}$ scaled up the values for *K* and stock biomass and resulted in slightly more optimistic assessments of stock status (B_{2012}/B_{msy}) and considerably lower values for F_{2012}/F_{msy} (**Table 7, Figure 9**). The estimated stock biomass trends fitted the stock trend data progressively better as $\sigma_{\text{process error}}$ was increased but lead to much wider posterior intervals for annual stock biomass (**Figure 8**). Bayes factors were not very different between runs with different values for $\sigma_{\text{process error}}$. Under a uniform on *K* prior, ranged from values as low as 0.3 for the lowest values for $\sigma_{\text{process error}}$ and as high as 3.2 for $\sigma_{\text{process error}}$ of 0.15. Under a uniform on log(*K*) prior, Bayes factors favoured slightly lower values for $\sigma_{\text{process error}}$. However, none of the Bayes factors would suggest that any of the hypothesized values for $\sigma_{\text{process error}}$ had notably lower credibility than other values.

Using Bayes factors for model averaging

It is becoming increasingly common for outputs from different stock assessment model runs to be combined using the so-called approach of "model averaging". This involves formulating a mixture distribution from the outputs of different assessment models which could be considered to be alternative interpretations of stock status and stock and fishery dynamics. This would account for uncertainty in model outputs better than considering outputs from a single run of a single stock assessment model (McAllister and Kirchner 2002). If I had a probability density function (pdf), e.g., for B_{cur}/B_{msy} from model 1 (pdf(1)), a pdf for the same quantity from model 2 (pdf(2)), and so on up to n different models, then the resulting mixture pdf of e.g. B_{cur}/B_{msy} would be

pdf_mixture(1,2,…,n) = w(1) x pdf(1) + w(2) x pdf(2) + …+ w(n) x pdf(n) where the weights, w(i) sum to 1. With BSP2, the weights are obtained by renormalizing Bayes factor for the different model runs of interest. This could produce a multi-modal density function if the modes of models 1 to n were sufficiently far apart, e.g., if there were two models and the mode of the pdf from model 1 was at 0.5 (and the CV was e.g. 0.3) and the mode of the pdf of model 2 was at 1.5 (and the CV was also small at e.g. 0.3), these two modes would be retained in the resulting mixture distribution and the mixture distribution would be much wider than either of the density functions from the two different models. Graphically, this would represent uncertainty much more effectively than showing the output distribution from only a single run of a single model.

Figure 9 shows the marginal posterior pdfs for r, B_{cur}/B_{msy} , and F_{cur}/F_{msy} obtained when the value for B_{msy}/K was fixed at 0.3, 0.4, 0.5 and 0.6 in different model runs. The marginal posterior pdfs show markedly different central tendencies for each of these quantities. The Bayes factors computed for the different runs show that the Bayes factors are markedly different across the four hypothesized values for B_{msy}/K and suggest nearly two orders of magnitude difference between the lowest and highest values for B_{msy}/K (**Table 8**). Rather than treating the results from different model runs as equally likely in considering different model runs, as is normally the case, it is argued that it would be more appropriate to take into account the probability of the data given the model, from the different model runs to formulate model weightings (McAllister and Kirchner 2002; McAllister 2013). This is especially important where Bayes factors are noticeably different between different model runs. Considerable differences in the interpretation of different model runs can be obtained in such circumstances as shown in the resulting mixture distributions for these quantities when either equal weights or Bayes factors are applied as the weights to combine the different outputs from the different model runs (**Figure10**).

It is also argued that Bayes factors should also be presented in decision tables to show how projection results can be sensitive to how the stock assessment model is formed and most importantly the credibility of the different stock assessment models considered given the stock trend data. An example decision table that shows projection results from production models with different assumed values for B_{msy}/K is shown in **Table 9**. Here it is important to show that the results from the different models are not equally likely and the least likely of these that would normally be considered to be equally likely can be down-weighted given its low Bayes factor.

Discussion

An updated version of the BSP2 software has been presented and illustrated with an application to the 2013 stock assessment data for North Atlantic swordfish. The main updates in the software include the introduction of a Bayesian state-space version of the generalized surplus production model, the provision of summary statistics to enable computation of Bayes factors, and mixture distributions of outputs from different model runs using Bayes factors. The software offers a statistically and mathematically consistent framework with which to account for both parameter and structural uncertainties in the application of time dynamic surplus production models.

The updated software also enables the computation of an additional set of diagnostic output distributions to evaluate the effect of priors for model parameters and the inputted catch records on output distributions for model parameters, stock biomass and other quantities of interest before the model is fitted to data. This allows analysts to evaluate the impact of fitting the model to the stock trend data on the output distributions of interest.

Under the reference case model settings that included a relatively conservative value for the standard deviation in process error ($\sigma_{process}$), results obtained were not very different from those obtained from applications of the Prager model in the 2013 stock assessment (ICCAT 2010). These indicate that the stock is close to B_{msy} and fishing mortality rates are a little less than F_{msy}. Very substantial updates in marginal posterior distributions resulted from fitting the model to the stock trend index with the prior for r being updated considerably to a lower posterior mean than the prior mean (0.31 compared to the prior mean of 0.42). This update was produced by the lowering of catches in the 1970s and the subsequent reversal of the declining trend in the stock trend index.

The illustration of the updated software with an application to data from the 2012 assessment of North Atlantic swordfish shows that considerably different estimates of historic stock biomass can result when different values for the standard deviation in process error are applied. As standard deviation in process error was increased from very low values, i.e., 0.005, approximating a deterministic values to moderately high values, i.e., up to 0.15, the estimates of historic stock biomass nearly doubled. Stock status estimates however were less sensitive to the value assumed for $\sigma_{process\ error}$, though also became slightly more optimistic as the value for $\sigma_{process}$ was increased. Bayes factors were not very different between models with different values for $\sigma_{process}$, though values of 0.075 and 0.10 had slightly higher Bayes factors than other values considered. This suggests that careful a priori judgment needs to be applied in determining a reference case setting for this parameter.

Results were most sensitive to the value assumed for the ratio of stock biomass at the most productive level to unfished stock size (B_{msy}/K). Considerably higher values for r and $B_{current}/B_{msy}$ and lower values for $F_{current}/F_{msy}$ were favoured by the lowest setting for B_{msy}/K. The projection results were also highly sensitive to this assumed parameter value with the most optimistic projections coming from the runs with lowest assumed value for B_{msy}/K. Bayes factors in this instance most strongly favoured the lowest value applied for B_{msy}/K (i.e., 0.3) which indicated that this version was about 8 times more credible given the stock trend data than the reference case value of 0.5. The illustration showed that the common practice of giving different model runs the same weight could have resulted in considerably different conclusions about stock status than the more statistically rigorous approach of applying Bayes factors to obtain weights for different model runs.

Acknowledgements

Elizabeth Babcock is thanked for her meticulous work in helping to develop the first version of the BSP software package and developing a user-friendly interface and user manual for the BSP software that were not available in versions of the software that I had originally developed. Members of the ISC North Pacific blue shark working group are thanked for their discussions and feedback on an earlier beta version of the BSP2 software that I distributed to them in 2012. Groundfish scientists with Fisheries and Oceans Canada, particularly Rick Stanley, Lynne Yamanaka, Jackie King, Kendra Holt and Daniel Duplisea are all thanked for their helpful feedback on earlier versions of the BSP2 software that has since 2012 been applied to several Canadian groundfish stocks. Paul Starr is thanked for helpful comments and discussions that have stimulated numerous improvements to the software. Charles Edwards, Alistair Dunn, and Chris Francis are also thanked for comments and discussion on the methodological approach. Laurie Kell is thanked for helping to make available data for the case study application reported in this paper.

References

Berger, J.O., 1985. Statistical decision theory and Bayesian analysis. 2nd ed. Springer-Verlag, New York, N.Y. 617p.

Box, G.E.P., and Tiao, G.C. 1973. Bayesian inference in statistical analysis. Addison-Wesley, Reading, Mass.

Courtney, D., and Piner, K., 2009. Age Structured Stock Assessment of North Pacific Swordfish (*Xiphias gladius*) with Stock Synthesis under a Single Stock Scenario. ISC/09/BILLWG-3/08.

Dorn, M.W. 2002. Advice on West Coast rockfish harvest rates from Bayesian meta-analysis of stock–recruit relationships. N. Am. J. Fish. Manage. 22(1): 280–300.

Harley, S., Kleiber, P., and Hoyle, S., 2012. Progress Towards a Stock Assessment for Swordfish in the southern WCPO including Standardized CPUE for Spanish Swordfish Fleet. WCPFC-SC8-2012/ SA-WP-08.

Hinton, M.G., and Maunder, M.N., 2009. Status of Swordfish in the Eastern Pacific Ocean In 2010 And Outlook For The Future. SAC-02-09.

ICCAT. 2010. Report of the 2009 Atlantic Swordfish Stock Assessment Session. Collect. Vol. Sci. Pap. ICCAT, 65(1): 1-123.

ICCAT. 2013. Report of the 2013 Atlantic Swordfish Stock Assessment Session. Collect. Vol. Sci. Pap. ICCAT, *in press*.

Kass, R. E. and Raftery, A. E. 1995. Bayes factors. J. Am. Stat. Assoc. 90: 773–795.

King, J.R., McAllister, M.K., Holt, K.R., and Starr, P.J. 2012. Lingcod (Ophiodon elongates) stock assessment and yield advice for outside stocks in British Columbia. DFO Can. Sci. Advis. Sec. Res. Doc. 2011/124. 177 p.

McAllister, M.K. 2013. Using Bayes factors to evaluate the credibility of stock-recruitment relationships for Atlantic bluefin tuna. ICCAT Standing Committee on Research and Science Document SCRS/ 2012/54, 26 pp.

McAllister, M.K., and Ianelli, J.N. 1997. Bayesian stock assessment using catch-age data and the sampling/ importance resampling algorithm Can. J. Fish. Aquat. Sci. 54, 284-300.

McAllister, M.K. and Babcock, E.A. 2002. "Importance sampling issues in the 1998 large coastal shark assessment ". National Marine Fisheries Service 2002 Shark Evaluation Workshop, Panama City, Florida. SB-02-25.

McAllister, M.K. and Kirchner, C.H. 2002. Accounting for structural uncertainty to facilitate precautionary fishery management: illustration with Namibian orange roughy. In "Targets, Thresholds, and the Burden of Proof in Fisheries Management" Mangel, M. ed. Bull. of Mar. Sci., 70(2). 499-540.

McAllister, M.K. and Babcock, E.A. 2006. Bayesian surplus production model with the Sampling Importance Resampling algorithm (BSP): A user's guide. Available from: http://www.sefsc.noaa.gov/library.jsp (January 27, 2010).

McAllister, M. and Duplisea, D.E. 2011. Production model fitting and projection for Atlantic redfish (*Sebastes fasciatus* and *Sebastes mentella*) to assess recovery potential and allowable harm. DFO Can. Sci. Advis. Sec. Res. Doc. 2011/057 vi + 75 p.

McAllister, M. and Duplisea, D.E. 2012. Production model fitting and projection for Atlantic redfish (*Sebastes fasciatus*) in Units 1 and 2. DFO Can. Sci. Advis. Sec. Res. Doc. 2012/103. iii + 34 p.

McAllister, M.K., Babcock, E.A., Pikitch, E.K. and Prager, M.H. 2000. Application of a non-equilibrium generalized production model to South and North Atlantic Swordfish: combining Bayesian and demographic methods for parameter estimation. International Commission for the Conservation of Atlantic Tunas, Madrid. ICCAT Col. Vol. Sci. Papers SCRS/99/85.

McAllister, M.K. Pikitch, E.K. and Babcock, E.A. 2001. Using demographic methods to construct Bayesian priors for the intrinsic rate of increase in the Schaefer model and implications for stock rebuilding. Can. J. Fish. Aquat. Sci. 58(9): 1871-1890.

McAllister, M.K., Babcock, E.A., Pikitch, E.K. and Prager, M.H., 2000, Application of a non-equilibrium generalized production model to South and North Atlantic swordfish: Combining Bayesian and demographic methods for parameter estimation). Collect. Vol. Sci. Pap. ICCAT, 51(5): 1523-1550.

McAllister, M.K., Pikitch, E.K., Punt, A.E. and Hilborn, R. 1994. A Bayesian approach to stock assessment and harvest decisions using the Sampling/ Importance Resampling Algorithm. Can. J. Fish. Aquat. Sci. 51, 2673-2687.

Millar, R.B. and Meyer, R. 2000. Non-linear state space modelling of fisheries biomass dynamics by using Hastings–Metropolis within-Gibbs sampling. Appl. Stat., 49: 327–342.

Prager, M. H. 1994. A suite of extensions to a nonequilibrium surplus-production model. Fish. Bull. (U.S.) 92: 374–389.

Nishida, T., and Wang, S.P. 2009. Preliminary stock assessment of swordfish (*Xiphias gladius*) in the Indian Ocean by the Age Structured Production Model (ASPM) (1952-2007). IOTC-2009--WPB7.

Oh, M.S., and Berger, J.O. 1992. Adaptive importance sampling in Monte Carlo integration. J. Stat. Comput. Simul. 41: 143-168.

Quinn, T.R. and Deriso, R.B. 1999. Quantitative Fish Dynamics. Oxford University Press.

Rivot, E., Prévost, E., Parent, E. and Baglinière, J.L. 2004. A Bayesian state-space modelling framework for fitting a salmon stage-structured population dynamic model to multiple time series of field data. Ecological Modelling. 179: 463–485.

Ross, Sheldon (2009). A first course in probability (8th ed.). Prentice Hall press.

Rubin, D. B. 1987. Comment on "The calculation of posterior distributions by data augmentation." J. Am. Stat. Assoc. 82: 543–546.

Rubin, D. B. 1988. Using the SIR algorithm to simulate posterior distributions. Pages 385–402 in J. M. Bernardo, M. H. Degroot, D. V. Lindley and A. M. Smith, eds. Bayesian statistics 3: Proceedings of the Third Valencia International Meeting, June 1–5, 1987. Clarendon Press, Oxford. 805 p.

Schaefer, M.B. 1954. Some aspects of the dynamics of populations important to the management of commercial marine fisheries. Bulletin of the Inter-American Tropical Tuna Commission, 1: 27-56.

Stanley, R. D., McAllister, M., Starr, P. and Olsen, N. 2009. Stock assessment for bocaccio (*Sebastes paucispinis*) in British Columbia waters. DFO Can. Sci. Advis. Sec. Res. Doc. 2009/055: xiv + 200 p.

Stanley, R. D., M. McAllister, P. Starr. 2012. Updated stock assessment for bocaccio (*Sebastes paucispinis*) in British Columbia waters for 2012. DFO Can. Sci. Advis. Sec. Res. Doc. 2012/ 109: ix+73p.

Walters, C.J., Ludwig, D., 1994. Calculation of Bayes posterior probability distributions for key population parameters. Can. J. Fish. Aquat. Sci. 51, 713–722.

Yamanaka, K. L., McAllister, M. K., Olesiuk, P. F. Etienne, M.-P., Obradovich, S. and Haigh, R. 2012a. Stock Assessment for the inside population of yelloweye rockfish (*Sebastes ruberrimus*) in British Columbia, Canada for 2010. DFO. Can. Sci. Advis. Sec. Res. Doc. 2011/129. 131 p.

Yamanaka, M.K. McAllister, M.-P. Etienne and R. Flemming. 2012b. Stock Assessment and Recovery Potential Assessment for Quillback Rockfish (*Sebastes maliger*) on the Pacific Coast of Canada. DFO Can. Sci. Advis. Sec. Res. Doc. 2011/135. 151p.

Table 1. Summary of estimated parameters.

• Parameter	• Description
• r	• Intrinsic rate of increase
• K	• Carrying capacity
• p_0	• Ratio of initial stock biomass in first year to carrying capacity
• $e_{process,y}$	• Process error deviate in year y

Table 2. Summary of derived management parameters of interest for the Schaefer model.

• Maximum Sustainable Yield (MSY)	• $rK/4$
• Stock size for MSY (B_{msy})	• $K/2$
• Rate of exploitation at MSY	• $r/2$
• Replacement yield	• $rB_y\left(1-\frac{B_y}{K}\right) \quad for \quad B_y < K$ $0 \quad for \quad B_y \geq K$
• Threshold extirpation rate of exploitation	• r

Table 3. Prior distributions for surplus production model parameters for North Atlantic Swordfish. Biomass values are shown in t.

Parameter	Prior density function
K	Uniform(5000, 1,000,000) or
ln(K)	Uniform(log(5000),log(1,000,000))[1]
ln(q_j)	Uniform(-200,10)
p_0	Lognormal(log(0.875),0.25^2)[2]
r	logNormal(log(0.424),0.405^2)[3]
$\square_{process,y}$	Normal(0, 0.05^2)

[1] Uniform on logK is suggested as an alternative prior for K, as explained above.
[2] Prior taken from ICCAT (2010).
[3] When the prior coefficient of variation (CV) in X is specified for a lognormal density function, the parameter sigma for this density function is given by sigma = sqrt(ln(1+CV(X)^2).

Table 4. Summary of sensitivity test runs.

Code	Category Description	Code	Run Description
Ref	Reference run	Ref	Reference run
A	B_{msy}/K	A.1	$B_{msy}/K = 0.3$
		A.2	$B_{msy}/K = 0.4$
		A.3	$B_{msy}/K = 0.6$
B	r prior mean	B.1	low r (mean = 0.28, CV = 0.39) (equivalent to steepness of 0.71)
		B.2	High r (mean = 0.56, CV = 0.39) (equivalent to steepness of 0.94)
C	Catch assumptions	C.1	Catch records 1950-1985 x 0.75
		C.2	Catch records 1950-1985 x 1.5
D	Prior for K	D.1	Uniform on log K prior.
		D.2	Lognormal prior for K with mean of 200,000t, and SD in log(K) of 0.8.
E	p_0 prior	E.1	Prior mean set at 0.7.
		E.2	Prior mean set at 1.0.
F	Standard deviation in process error	F.1a,b	$\sigma_{process} = 0.005$ (a=uniform on K prior, b=uniform on logK prior)
		F.2a,b	$\sigma_{process} = 0.01$
		F.3a,b	$\sigma_{process} = 0.05$ (reference case)
		F.4a,b	$\sigma_{process} = 0.075$
		F.5a,b	$\sigma_{process} = 0.10$
		F.6a,b	$\sigma_{process} = 0.15$

Table 5. Posterior mean, standard deviation (SD), coefficient of variation (CV) and 5th, Median and 95th percentiles for parameters and management quantities of interest for North Atlantic swordfish using reference case model specifications. Biomass values are in tons. The referenced current year is 2012.

Variable	Mean	SD	CV	5th Percentile	Median	95th Percentile
r	0.37	0.10	0.28	0.21	0.36	0.54
K	162215	48606	0.3	104466	152388	254959
MSY	13721	958	0.07	12342	13750	14958
B_{msy}	81107	24303	0.3	52233	76194	127479
B_{1950}	147304	65611	0.445	80843	132667	255720
B_{2012}	81554	26302	0.323	50464	77446	131488
B_{2012}/B_{msy}	1.01	0.12	0.12	0.808	1.01	1.21
B_{2012}/B_{1950}	0.59	0.16	0.27	0.37	0.58	0.88
B_{2012}/K	0.50	0.06	0.12	0.40	0.50	0.61
F_{MSY}	0.18	0.05	0.28	0.10	0.18	0.27
F_{2012}	0.16	0.04	0.28	0.09	0.16	0.24
F_{2012}/F_{MSY}	0.90	0.14	0.16	0.70	0.88	1.14
$REPY_{2012}$	13511	838	0.06	12062	13546	14784
$Cat_{2012}/REPY_{2012}$	0.90	0.06	0.07	0.82	0.90	1.01
$P(B_{2012} > 0.4B_{msy})$	1					
$P(B_{2012} > 0.8B_{msy})$	0.96					
$P(B_{2012} > B_{msy})$	0.52					
$P(F_{2012} < F_{msy})$	0.79					

Table 6. Projected outcomes from a set of constant total catch policy options at 5, 10 and 20 horizons computed using the BSP reference case model with standard deviation in process error set at 0.05.

Horizon	Policy option Constant catch	$Median(B_{fin}/B_{msy})$	$P(B_{fin}>0.8\ B_{msy})$	$P(B_{fin}>B_{2012})$	$P(B_{fin}>B_{msy})$
5 -year	10000	1.23	0.98	0.89	0.85
	11000	1.17	0.96	0.82	0.78
	12000	1.12	0.93	0.72	0.70
	13000	1.06	0.88	0.60	0.61
	14000	1.00	0.81	0.46	0.50
	15000	0.94	0.72	0.34	0.40
10 -year	10000	1.36	0.98	0.93	0.92
	11000	1.28	0.96	0.86	0.85
	12000	1.19	0.91	0.75	0.75
	13000	1.08	0.82	0.60	0.61
	14000	0.96	0.68	0.40	0.45
	15000	0.82	0.52	0.25	0.29
20 -year	10000	1.5	0.99	0.95	0.96
	11000	1.4	0.96	0.89	0.90
	12000	1.2	0.88	0.76	0.78
	13000	1.09	0.74	0.55	0.59
	14000	0.84	0.53	0.33	0.37
	15000	0.33	0.31	0.15	0.18

Table 7. Medians and 80% credibility intervals drawn from the posterior distributions for seven parameters taken from the Bocaccio assessment for the reference run and 17 sensitivity runs. Codes used for each run along with a run description can be found in **Table 4**. Biomass values are in tons. The referenced current year is 2012.

Run	***r***			B_{msy}			$B_{current}$			$RepY_{current}$			$B_{current}/B_{msy}$			$F_{current}/F_{msy}$			$Catch_{curr}/RepY$		
	10%	**Median**	**90%**	**10%**	**Median**	**90%**	**10%**	**Median**	**90%**	**10%**	**Median**	**90%**	**10%**	**Median**	**90%**	**10%**	**Median**	**90%**	**10%**	**Median**	**90%**
Ref.	0.21	0.36	0.54	52233	76194	127479	50464	77446	131488	12062	13546	14784	0.81	1.01	1.21	0.70	0.88	1.15	0.82	0.90	1.01
	B_{msy}/K																				
A.1	0.25	0.43	0.64	42893	62364	103682	57800	90161	169552	10965	12540	13966	1.12	1.46	1.91	0.44	0.62	0.88	0.87	0.97	1.11
A.2	0.25	0.41	0.61	45576	65915	105665	52428	79507	135253	11703	13126	14274	0.94	1.21	1.53	0.56	0.74	1.00	0.85	0.93	1.04
A.3	0.18	0.31	0.46	47134	68703	108889	47711	70547	108898	12769	14502	15739	0.79	1.03	1.30	0.61	0.79	1.08	0.77	0.84	0.95
	***r* prior mean**																				
B.1	0.15	0.29	0.46	60940	94635	168800	58835	94571	175597	11419	13299	14866	0.79	0.99	1.21	0.71	0.91	1.22	0.82	0.91	1.06
B.2	0.25	0.40	0.59	47763	69337	107803	48348	71009	114143	12377	13665	14806	0.82	1.02	1.23	0.69	0.86	1.11	0.82	0.89	0.98
	Catch assumptions																				
C.1	0.21	0.36	0.55	50097	74499	123756	52585	80288	136823	11656	13155	14393	0.87	1.07	1.28	0.67	0.85	1.11	0.85	0.92	1.04
C.2	0.20	0.32	0.47	64375	91266	144718	56023	81925	135481	12693	14524	15948	0.70	0.89	1.10	0.73	0.92	1.20	0.76	0.84	0.96
	K priors																				
D.1	0.23	0.38	0.57	49796	71948	114220	49711	73403	118503	12200	13576	14759	0.82	1.01	1.23	0.70	0.87	1.13	0.82	0.90	1.00
D.2	0.24	0.38	0.56	50680	71572	110972	50034	73569	115611	12246	13574	14777	0.82	1.01	1.23	0.70	0.88	1.12	0.82	0.90	0.99
	p_0 priors																				
E.1	0.22	0.37	0.55	52085	75153	122516	50920	77314	126901	12094	13548	14796	0.81	1.01	1.23	0.69	0.88	1.14	0.82	0.90	1.00
E.2	0.20	0.36	0.54	52703	77179	129225	51671	79086	133682	11996	13530	14803	0.81	1.01	1.23	0.70	0.88	1.15	0.82	0.90	1.01
	Standard deviation in process error																				
F.1a	0.29	0.41	0.56	50527	65575	87982	50569	63339	83136	12297	13394	14054	0.80	0.97	1.12	0.78	0.93	1.17	0.87	0.91	0.99
F.2a	0.28	0.41	0.57	50268	65359	89965	50088	63385	85571	12325	13429	14143	0.81	0.97	1.12	0.78	0.93	1.17	0.86	0.91	0.99
F.3a	0.21	0.36	0.54	52233	76194	127479	50464	77446	131488	12062	13546	14784	0.81	1.01	1.21	0.70	0.88	1.15	0.82	0.90	1.01
F.4a	0.18	0.32	0.52	54745	87646	162438	54446	92523	179276	11705	13965	16262	0.80	1.04	1.32	0.60	0.82	1.13	0.75	0.87	1.04
F.5a	0.16	0.30	0.51	56638	99857	228949	56371	107214	276840	11587	14496	19568	0.79	1.06	1.43	0.45	0.76	1.10	0.62	0.84	1.05
F.6a	0.15	0.26	0.47	68168	153922	447175	68853	166578	547197	11331	17575	35522	0.76	1.15	1.59	0.23	0.56	0.99	0.32	0.68	1.01

Table 8. Bayes factors for alternative mode runs. These reflect the ratio of the probability of the stock assessment data based on a sensitivity run to the probability of the data obtained from the reference case. For runs with alternative process error Bayes factors are shown for runs with uniform on K and uniform on log(K) priors. NA indicates no results produced.

				Bayes factor	
Category Code	**Category Description**	**Code**	**Run Description**	**U(K)**	**U(log(K))**
A	B_{msy}/K	A.1	B_{msy}/K = 0.3	7.9	NA
		A.2	B_{msy}/K = 0.4	6.7	NA
		Ref	B_{msy}/K = 0.5	1.0	NA
		A.3	B_{msy}/K = 0.6	0.4	NA
B	*r* prior mean	B.1	low *r* (mean = 0.28, CV = 0.49)	1.0	NA
		Ref	ref. prior (mean = 0.42, CV = 0.49)	1.0	NA
		B.2	high *r* (mean = 0.56, SD = 0.49)	0.7	NA
C	Catch	C.1	Total catch for years <1985 x 0.75	0.4	NA
		Ref.		1.0	NA
		C.4	Total catch for years <1985 x 1.5	1.8	NA
E	p_0 prior	E.1	Prior mean p_0 = 0.7	0.9	NA
		Ref.	Prior mean p_0=0.875	1.0	NA
		E.2	Prior mean p_0=1.0	1.0	NA
F	Process error SD	F.1a,b	$\square_{\text{process error}}$=0.005	0.3	0.4
		F.2a,b	$\square_{\text{process error}}$=0.01	0.4	0.4
		Ref., D.1	$\square_{\text{process error}}$=0.05	1.0	1.0
		F.3a,b	$\square_{\text{process error}}$=0.075	2.0	1.8
		F.4a,b	$\square_{\text{process error}}$=0.10	3.2	2.6
		F.6a,b	$\square_{\text{process error}}$=0.15	3.2	1.7

Table 9. Summary decision table for the probability that stock biomass exceeds B_{msy} in 10 years under each alternative constant TAC policy (t) and under each alternative hypothesized value for B_{msy}/K.

	Hypothesized B_{msy} to K ratio			
			Reference	
B_{msy}/K	**0.3 (A.1)**	**0.4 (A.2)**	***0.5***	**0.6 (A.3)**
Bayes factor	7.9	6.7	1.0	0.4
TAC				
10000	0.99	0.97	0.92	0.95
11000	0.97	0.93	0.85	0.92
12000	0.93	0.86	0.74	0.86
13000	0.87	0.76	0.61	0.75
14000	0.76	0.61	0.45	0.59
15000	0.62	0.45	0.29	0.42

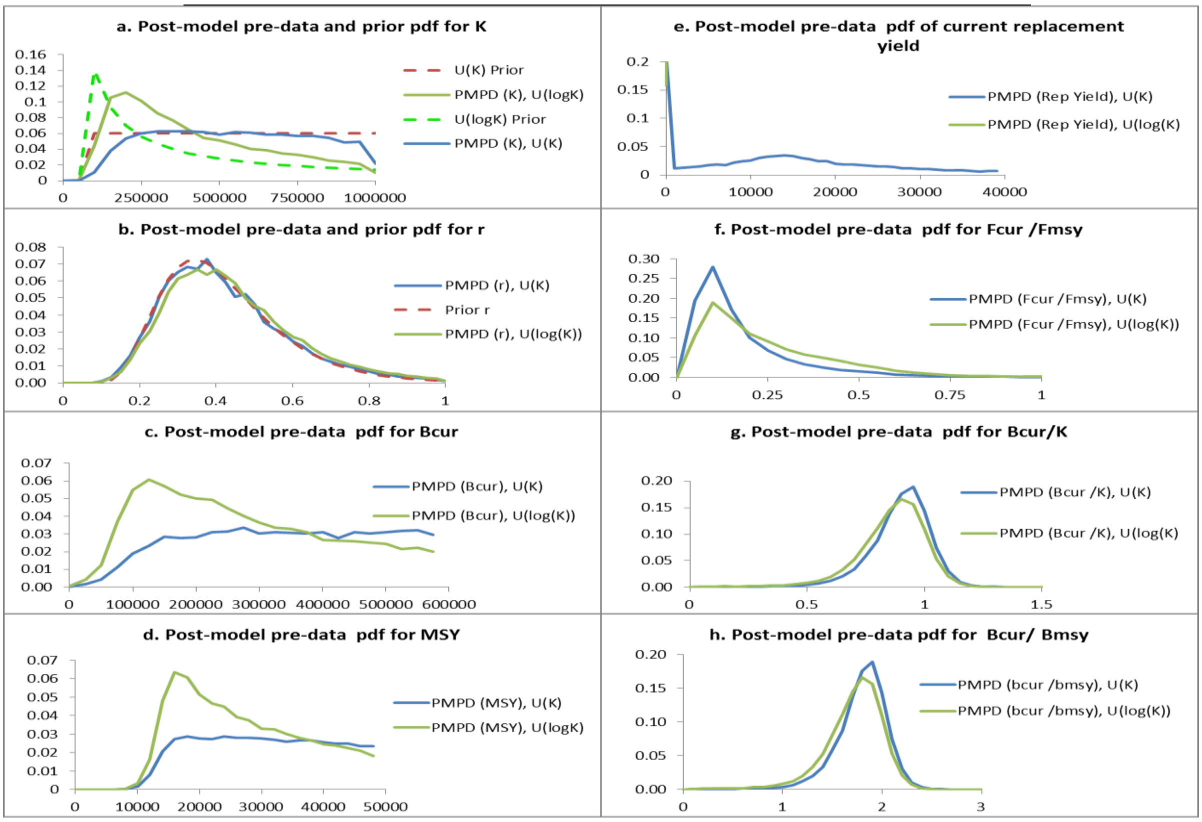

Figure 1. Post-model, pre-data distributions for quantities of interest for North Atlantic swordfish when either a uniform on K and uniform on log K prior is used.

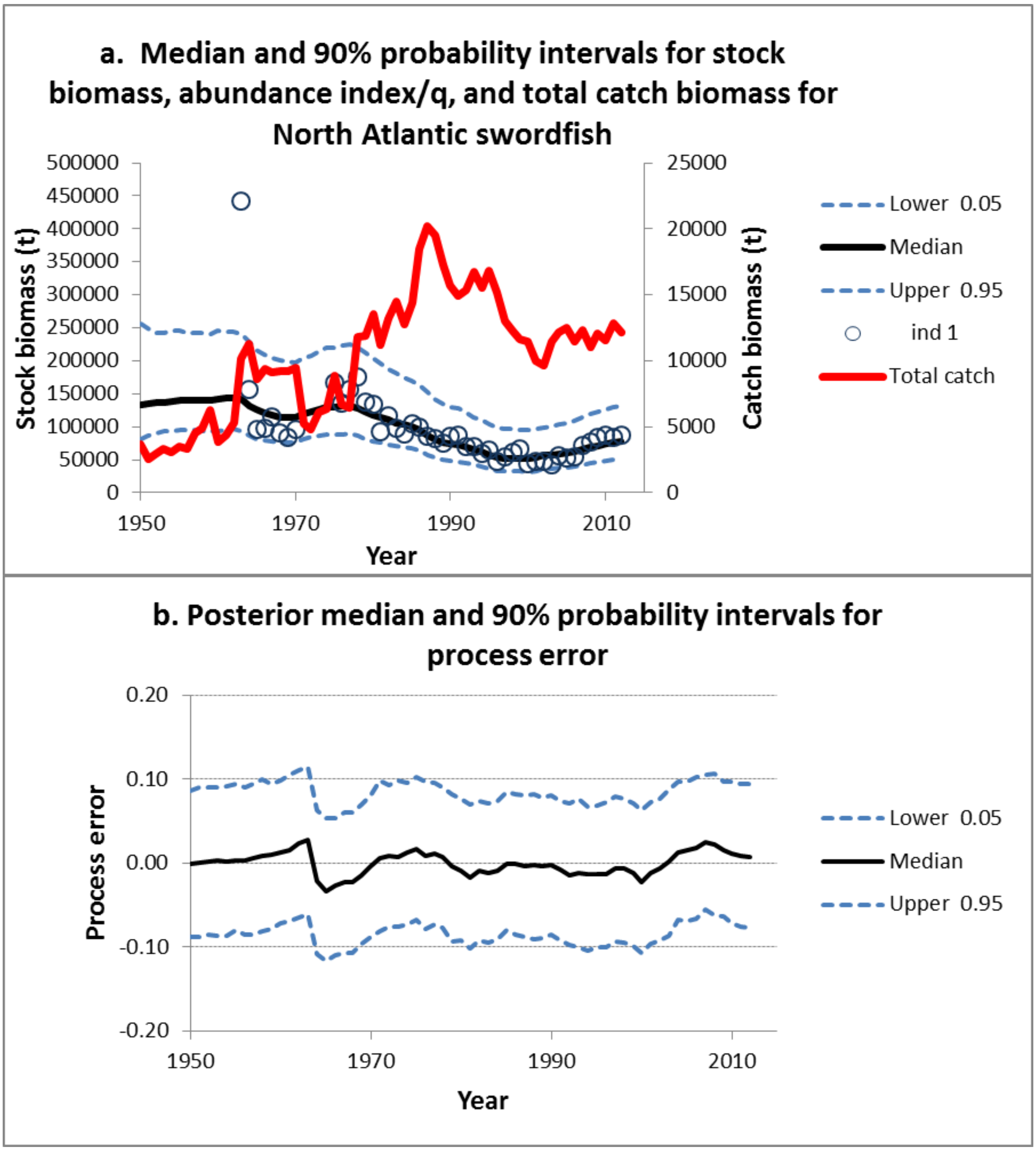

Figure 2. Plots of a. posterior median and 90% probability intervals for stock biomass, the stock trend index divided by the posterior median q, and total catch biomass, and b. the posterior median and 90% probability intervals for annual process error deviates. This is for the reference case BSP2 run for North Atlantic swordfish fitted to data to 2012 from the 2013 stock assessment.

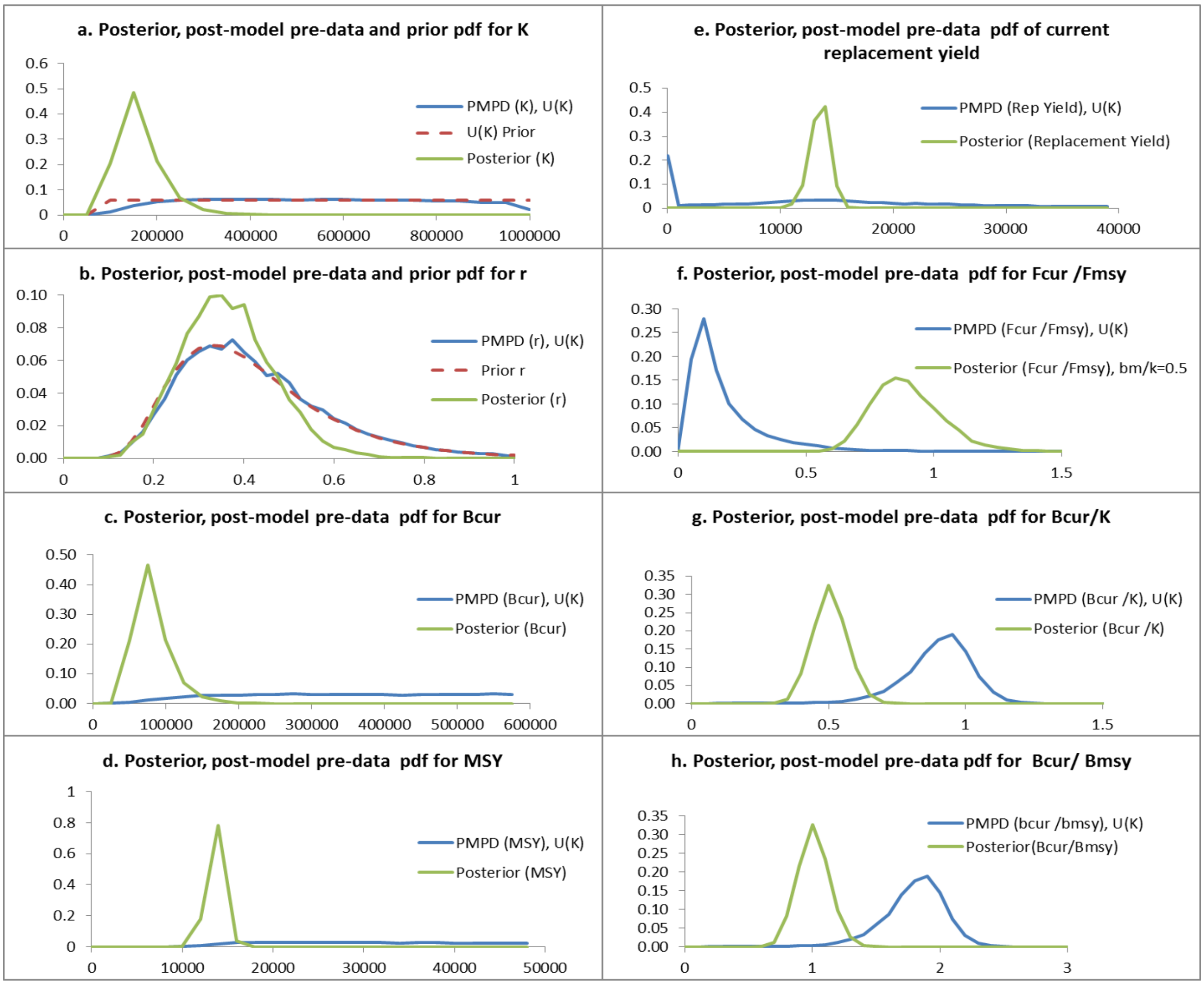

Figure 3. Plots of marginal posterior, and post-model, pre-data distributions for quantities of interest for North Atlantic swordfish when either a uniform on K prior is used.

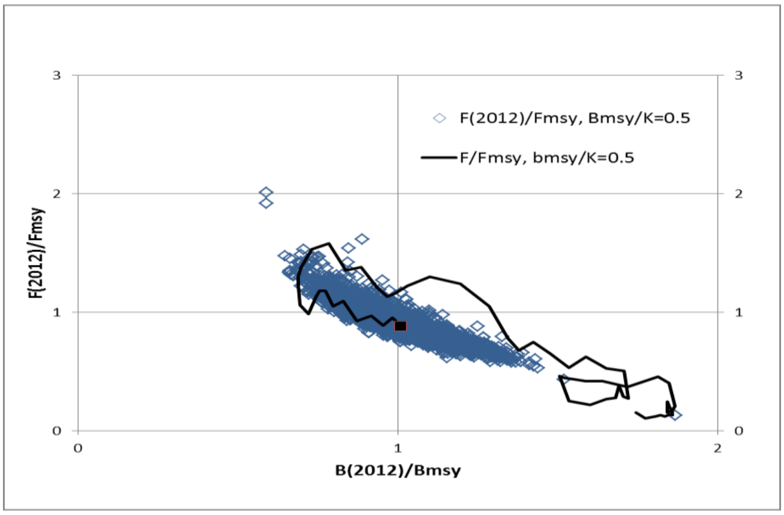

Figure 4. Kobe plot of stock status under reference case run with uniform on K prior, and the standard deviation in the process error set at 0.05. The black trajectory shows the progression of the status of the fishery from 1950 to 2012 from right to left, using the posterior median values for B_y/B_{msy}, F_y/F_{msy}..

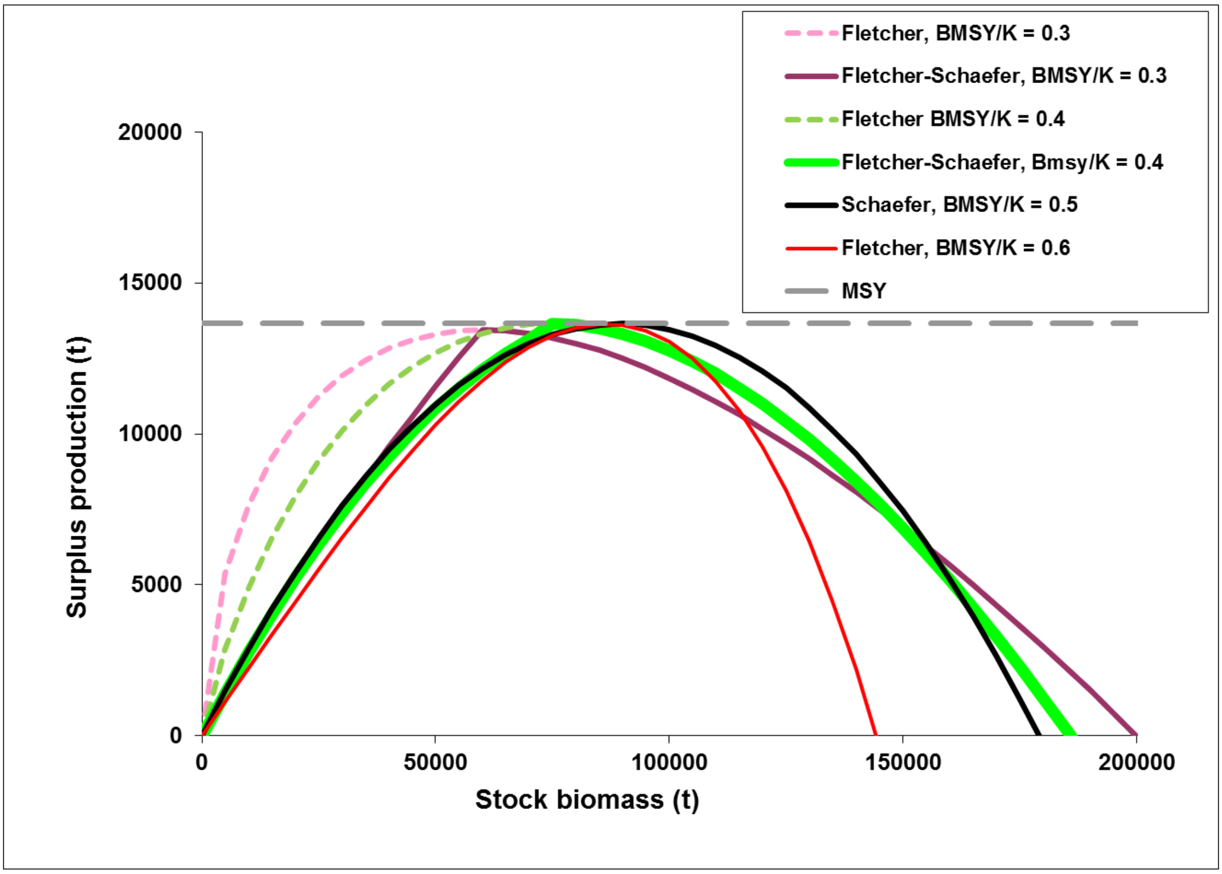

Figure 5. A plot of the reference case Schaefer and rescaled versions of the three alternative production functions applied in evaluations of the sensitivity of results to different model settings. All plotted production functions are referenced to approximately the same MSY value to highlight differences in the shape of the production functions.

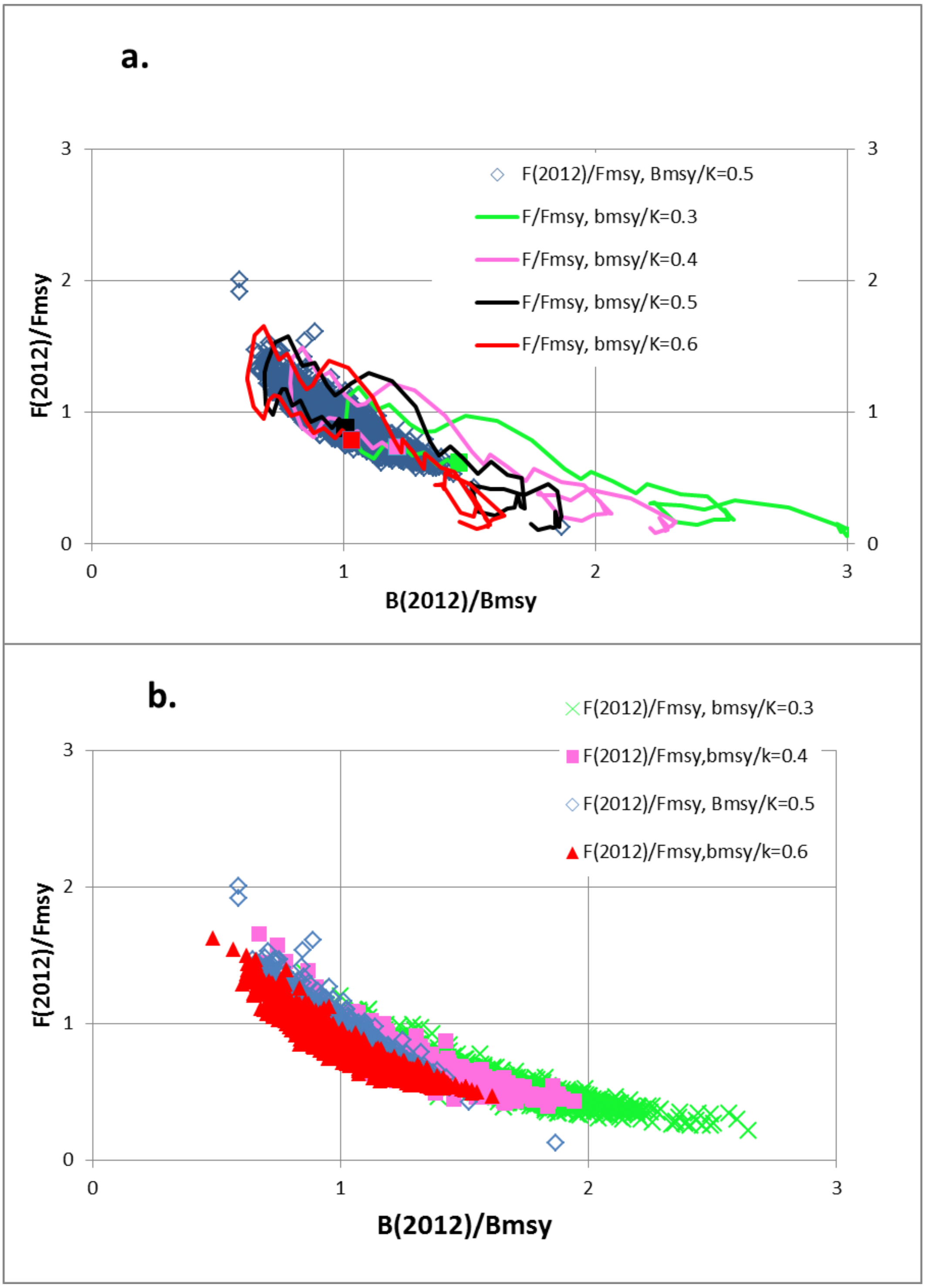

Figure 6. Kobe plots of stock status under reference case run with uniform on K prior, and the standard deviation in the process error set at 0.05. a. The trajectories show the progression of the status of the fishery from 1950 to 2012 from right to left, using the posterior median values for B_y/B_{msy}, F_y/F_{msy} from the runs with the B_{msy}/K point set at 0.3, 0.4, 0.5 (base case), 0.6. Also plotted are 5000 draws from the posterior for F_{2012}/F_{msy}, B_{2012}/B_{msy}. b. Scatter plots of F_{2012}/F_{msy}, B_{2012}/B_{msy} from runs with B_{msy}/K point set at 0.3, 0.4, 0.5 (base case), 0.6.

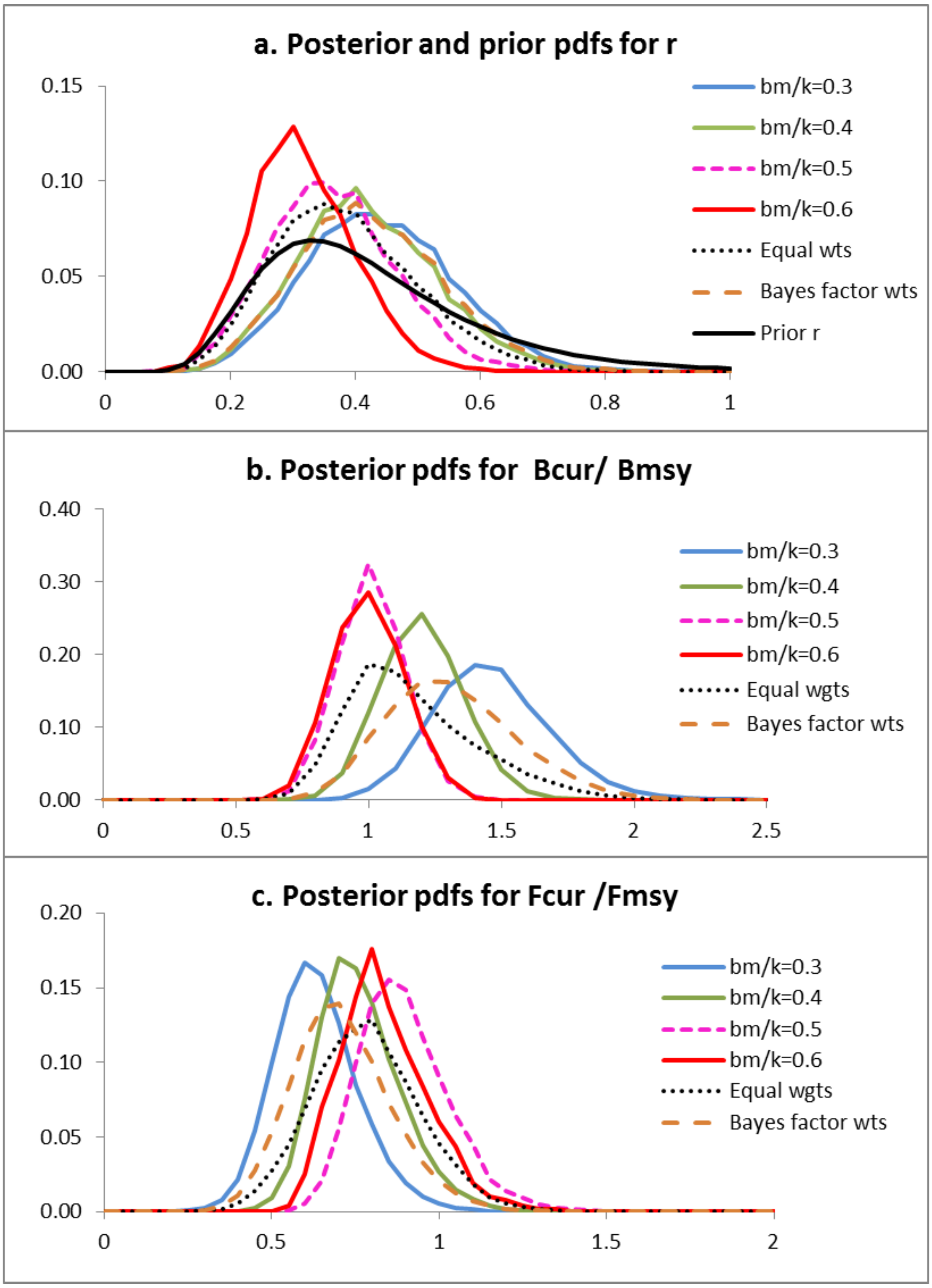

Figure 7. Marginal posterior distributions for (a) the maximum rate of increase, r, (b) stock status (B_{cur}/B_{msy}) and (c) F_{cur}/F_{msy} for North Atlantic swordfish obtained from model runs where B_{msy}/K is fixed at 0.3, 0.4, 0.5 and 0.6. Results are shown for the marginal posterior density functions from the separate runs, using equal weights, and using Bayes factors to weight the results from different runs.

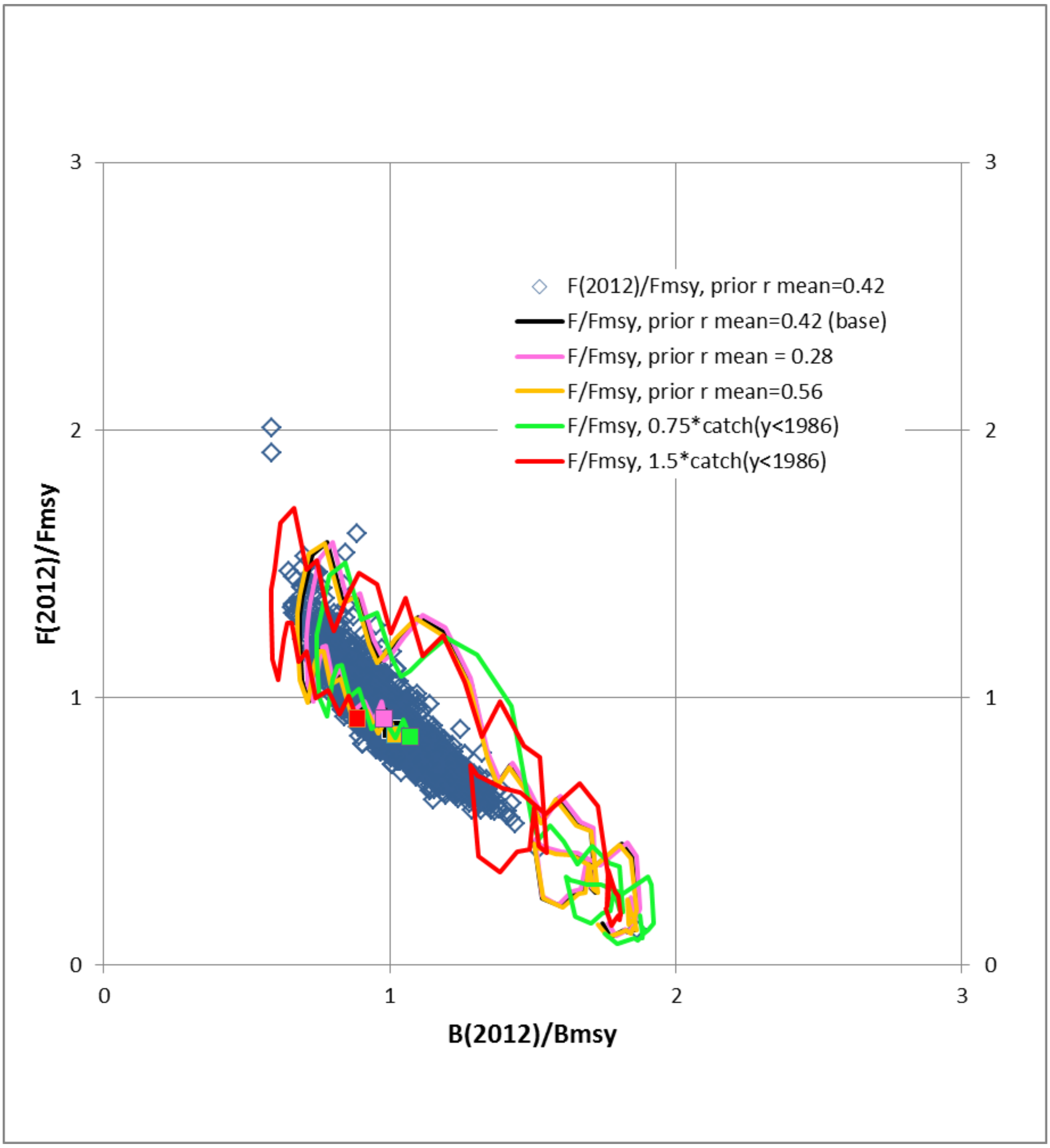

Figure 8. Kobe plot of stock status under reference case run with uniform on K prior, and the standard deviation in the process error set at 0.05. The trajectories shows the progression of the status of the fishery from 1950 to 2012 from right to left, using the posterior median values for B_y/B_{msy}, F_y/F_{msy} from the runs with the prior mean for r set at 0.28, and 0.56, and the catches prior to 1986 multiplied by either 0.75 or 1.5 to account for uncertain in historic catch records.

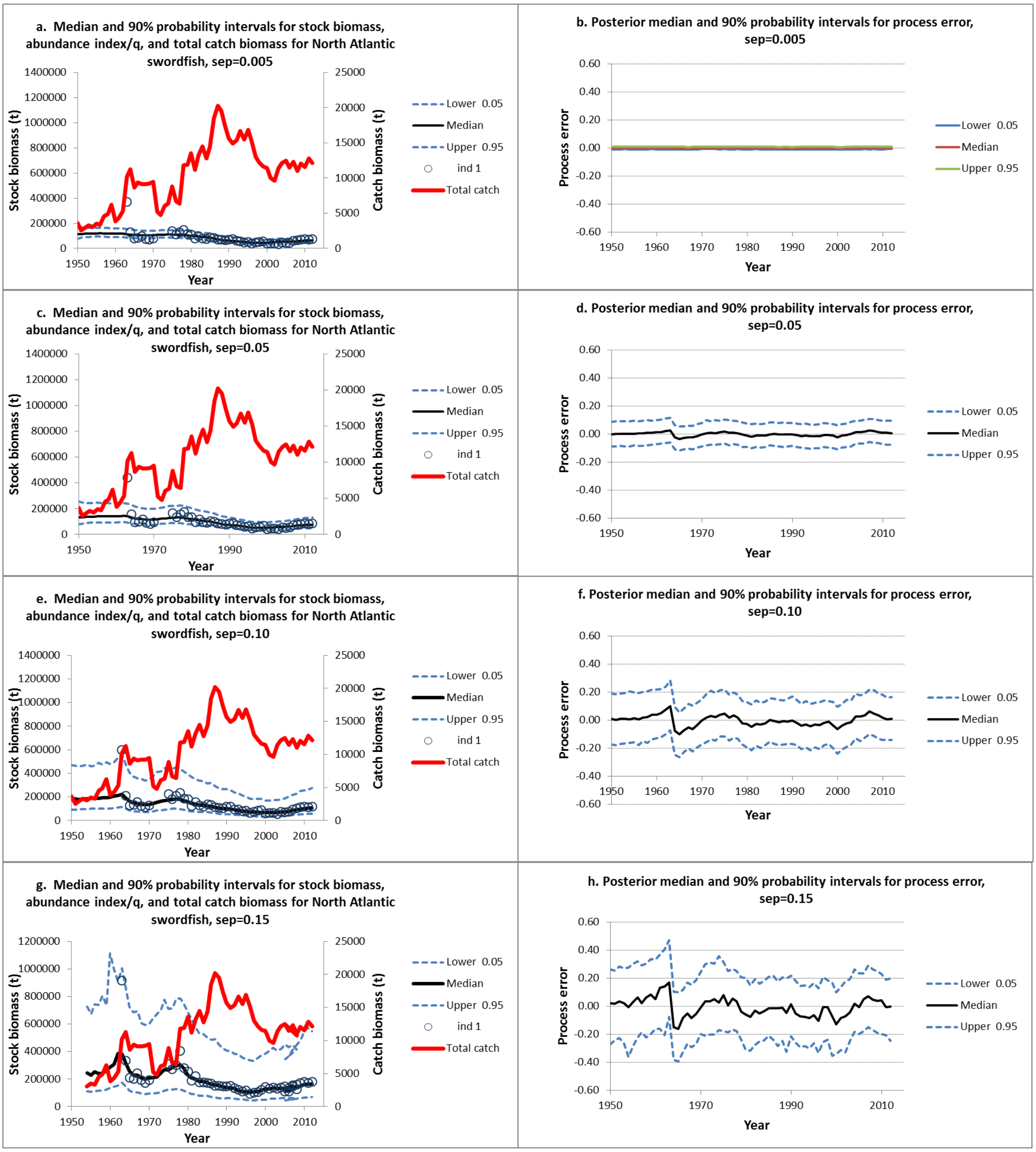

Figure 9. Fits of BSP2 to the data and medians, and 90% probability intervals under values for process error SD of 0.005, 0.05, 0.10, and 0.15 for North Atlantic swordfish. sep refers to the prior standard deviation in annual process error deviates ($\sigma_{\text{process error}}$).

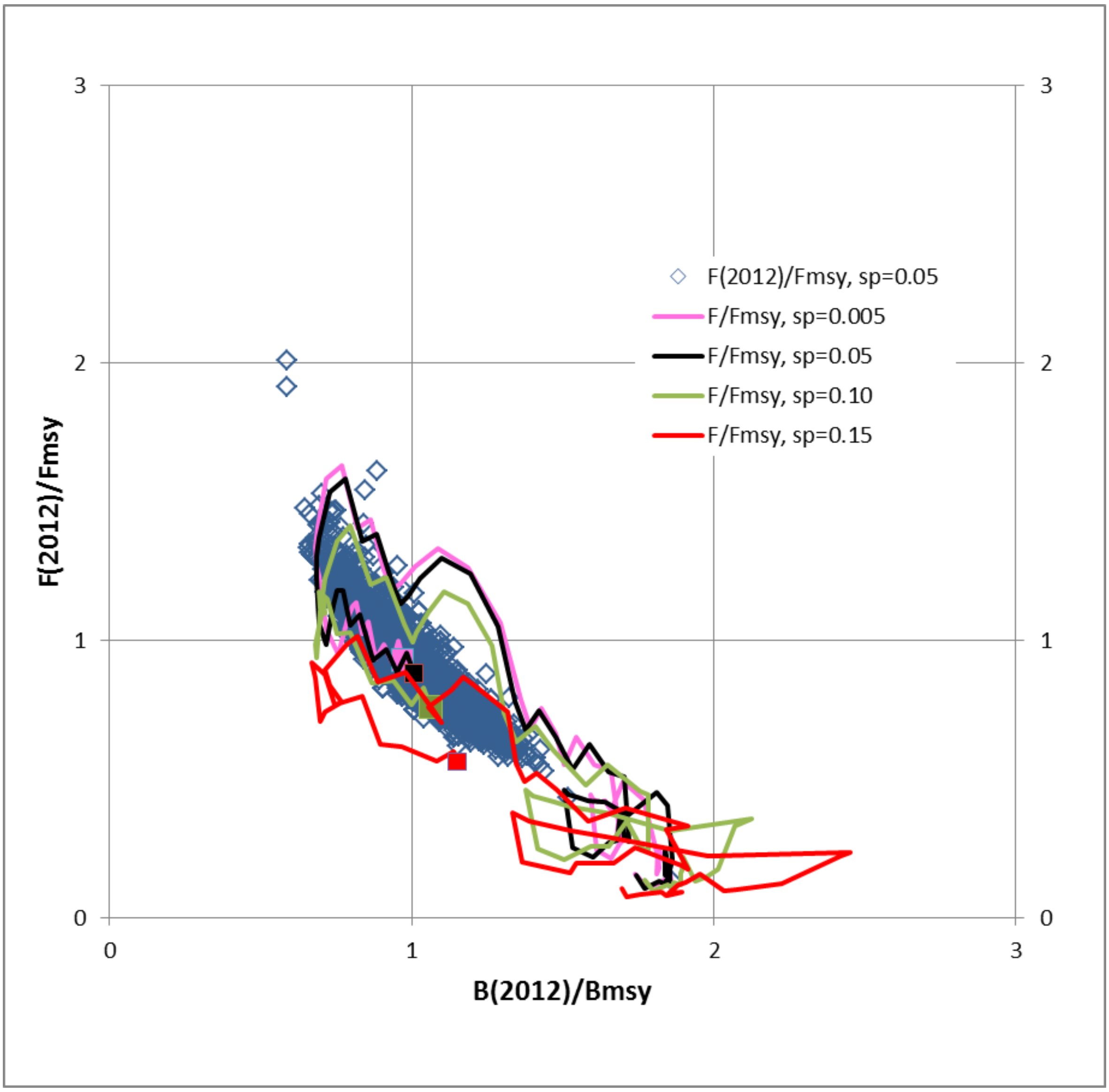

Figure 10. Kobe plot of stock status under reference case run with uniform on K prior, and the standard deviation in the process error set at 0.05. The trajectories shows the progression of the status of the fishery from 1950 to 2012 from right to left, using the posterior median values for B_y/B_{msy}, $F_y/\ F_{msy}$ from the runs with the standard deviation in process error set at 0.005, 0.05, 0.1 and 0.15. sp refers to the standard deviation in log process error deviates $\sigma_{process}$.

Appendix A: Reformulating the Euler-Lotka method to compute a prior for r for North and South Atlantic swordfish

The prior for the maximum rate of increase (r) for North Atlantic swordfish as reformulated using the Euler Lotka method (McAllister *et al.* 2001) that had been adapted for teleost species (Stanley *et al.* 2009; Yamanaka *et al.* 2012). The Euler Lotka method to compute r, is based on the following identify:

$$\sum_{x=0}^{A} e^{-r_m x} l_x m_x = 1$$

where r_m is the maximum rate of increase, l_x is the fraction surviving to age x, m_x is the number of age 1 individuals produced at age x, A is the oldest age class. A Monte Carlo procedure is applied in which the life history parameters that determine l_x and m_x are repeatedly drawn from prior distributions for these parameters. Given l_x and m_x, a value for r_m is solved for numerically. The Monte Carlo procedure is repeated numerous times to formulation a frequency distribution for the parameter r_m. The frequency distribution is then approximated using a parametric distribution. The terms l_x and m_x are determined by the following life history parameters for females and population dynamics parameters together with estimates of coefficients of variation (CV) for them to represent uncertainty in them:

1. the means and CVs for the von Bertalanffy growth parameters, K, L_{inf}, t_0,
2. the means and CVs for the rate of natural mortality (M) at age for recruited animals (e.g., either a constant value M for recruited animals or a Lorenzen schedule for M at age),
3. the mean and CV for the length-weight conversion parameters (a,b),
4. the means and CVs for parameters for the fraction mature at age (e.g., for the logistic function),
5. a prior median and uncertainty specifications for the Beverton-Holt steepness parameter (*h*) that could be applicable for North and South Atlantic swordfish.

Where possible these parameter values were obtained for the North and South Atlantic stocks. Since only point estimates are available for most of these parameters, the prior CVs were subjectively determined. The CVs considered include 10-20% for the growth parameters, 20% for the age at maturity parameters, 10% for the length-weight parameters, 25% for the natural mortality rates, and about 15% for steepness. In all instances, it was assumed that the prior covariance was zero. These parameter estimates were entered into existing computer software (Stanley *et al.* 2009; Yamanaka *et al.* 2012) to compute an updated prior for *r* for the North Atlantic stock.

Values for the rate of natural mortality were obtained from ICCAT's summary information on Atlantic swordfish. The point value provided was 0.2 yr^{-1} for all ages older than age 1 (ICCAT 2006) (**Table A.1**). Growth parameter values were also obtained from ICCAT (2006). Only a single set of growth parameter estimates were provided for the north and south Atlantic swordfish stocks (**Table A.1**). Due to the unfamiliar parameterization, the equation provided was applied to predict length at age. Following this, the von Bertalannfy parameters that most closely predicted these length at age values were obtained using least squares (**Table A.1**).

Length to weight conversion parameters for the north and southern Atlantic stocks were obtained from values synthesized by the ICCAT secretariat. Three different estimates of these parameters were available for the northeast, north central and northwestern Atlantic Ocean and were synthesized to produce a single best estimate of the a-b parameters for both stocks (**Table A.1**).

The median body length at maturity for female swordfish for the north Atlantic stock was reported to be 179cm (ICCAT 2006). The minimum body size at maturity was reported as 146 cm. From the growth curve these correspond to ages 5 years, and 3 years. Given the minimal information available on fraction mature at age, the median age at maturity (A_{50}) was taken as 5 years, the minimum at 3 years, and the 95th percentile at 7 years. A logistic function was presumed for fraction mature at age. The prior CV assumed for the fraction mature at age parameters was 20%.

The Beverton-Holt steepness parameter was the most difficult parameter for which to derive a prior for swordfish. It appears that there are no reliable stock-recruit datasets for swordfish, though there are some VPA constructions for north Atlantic swordfish in previous ICCAT assessments. McAllister *et al.* (2000) computed values for egg to age 1 survivorship using available literature based estimates of daily mortality rates, batch fecundity and the average number of spawning bouts per year per fish. The Leslie Matrix approach to computing a prior was applied to give a median value of r of 0.405. It is possible to solve numerically for steepness given the life history parameter estimates and a fixed value for r. When the life history parameter estimates compiled in this document were applied together with the point value for r of 0.405, the point value for Beverton-Holt steepness was 0.85. Since the egg to age 1 survivorship estimate was based on a VPA and VPA has been considered to be unreliable, a web search was conducted to identify point values for steepness that have been applied in age-structured stock assessments of swordfish elsewhere and to see where this point value fell with respect to these estimates.

In seeking out values for Beverton-Holt steepness that had been applied in other swordfish assessments a wide range of values was found to have been applied. Base case values ranged from about 0.8 to 1. Values applied in sensitivity analyses varied considerably also. For example, in the ISC stock assessments of North Pacific swordfish, only a single value for steepness of 0.9 appears to have been applied (Courtney and Piner 2009). The IOTC has also applied a fixed value of 0.9, though the choice of this value may have been influenced by the value chosen for steepness in ISC assessments of North Pacific swordfish (Nishida and Wang 2009). Stock assessments of swordfish in the southwestern Pacific Ocean have used a value of 0.8 as the reference case value and have run sensitivity analyses using values of 0.65 and 0.95 (Harley *et al.* 2012). The IATTC has in contrast applied a value of 1 as the reference case value for steepness, asserting that there appears to be no apparent relationship between recruitment and spawning stock size (Hinton and Maunder 2009). A sensitivity analysis was run using a value of 0.75. To cover the range of steepness values presumed to be plausible for swordfish stocks around the world, I formulated a wide prior distribution for steepness that derived from a Beta density function (**Table A.1**). The median value was 0.85 and the SD was 0.11. This gave a 2.5^{th} percentile of 0.56 and a 97.5^{th} percentile of 0.98.

Table A.1. Values for prior distributions for input parameters to the algorithm to compute a prior for r. N means north, S means south, SD(X) is standard deviation, CV is the coefficient of variation, SD(ln(X)) is the standard deviation in the natural logarithm in instances where a lognormal prior is applied. Note that the Table in this early draft is as yet incomplete.

Parameter	Stock	Mean	SD(X)	CV(X)	SD(ln(X))	Minimum	2.5%	Median	97.5%	Probability distribution
Rate of natural mortality (*M*) (yr^{-1})	N,S	0.206	0.0515	0.25	0.25	NA	0.12	0.20	0.33	Log normal
Growth parameters										
L_∞ (cm)	N,S	263.92	26.39	0.1	0.1	NA	218.03	265.24	322.67	Log normal
K (yr^{-1})		0.137	0.027	0.2	NA	NA	0.08	0.137	0.19	Normal
t_0 (yr)		-2.82	-1.41	0.5	NA	NA	-5.59	-2.82	-0.06	Normal
Length-weight (cm to kg)						NA				
a	N	4.48E-06	4.48E-07	0.1	0.1	NA	3.66E-06	4.45E-06	5.42E-06	Log Normal
b	N	3.2038	0.32	0.1	NA	NA	2.58	3.2038	3.83	Normal
	S	4.96E-06	4.96E-07	0.1	0.1	NA	3.66E-06	4.96E-06	5.42E-06	Log Normal
	S	3.188	0.32	0.1	NA	NA	2.56	3.188	3.81	Normal
Fraction mature						NA				
a_{50} (yr)	N	5.1	1.02	0.2	0.2	NA	3.38	5.0	7.40	Log Normal
d_{95} (yr)	N	2.0	0.41	0.2	0.2	3	1.35	2.0	2.96	Log Normal
Steepness (h)	N,S	0.83	0.11	0.14	NA	0.2	0.56	0.85	0.98	h~0.2 + 0.8 Beta(5.86, 1.59)

Results

The prior for r that was derived and subsequently applied based on Leslie Matrix approach in McAllister *et al.* (2000) was lognormal with a mean of 0.42, and a CV of 0.50. This was a mildly informative prior distribution whose large spread was driven by the wide input prior distribution for the survivorship from egg to age 1 fish (i.e., age zero survivorship). The algorithm applied in this update analysis, applied instead the Euler-Lokta approach (McAllister *et al.* 2002; Stanley *et al.* 2009; Yamanaka *et al.* 2012). This required in place of the prior for egg to age 1 survivorship, a prior for steepness, i.e., either Beverton-Holt or Ricker steepness, depending on which stock-recruit function is deemed to be most appropriate for the stock. In this case, the Beverton-Holt steepness was adopted since there's no known instance of cannibalism for swordfish and where stock-recruit estimates have been plotted in swordfish assessments, they've resembled a Beverton-Holt relationship much more so than a dome-shaped Ricker relationship (reference to be added). Using the values for north and south Atlantic swordfish listed in **Table 1**, the prior mean for r was the same to three decimal places for both stocks at 0.424, and the prior CV for both stocks was 0.39 (**Figure A.1**). These values are very close because the only difference in input values is the length-weight conversion factors and these are very similar (**Table A.1**). The prior values for r were fairly close to those obtained from McAllister *et al.* (2000). This is largely because the input distributions to the compu tation of the prior for *r* changed relatively little between 1999 and 2012. The main difference was that the prior for steepness applied in this analysis tended to bound the value for r more so than did the prior for age zero survivorship in McAllister *et al.* (2000). The empirical frequency distribution for r this time also corresponded very closely to a log normal distribution, as it usually does when a Beverton-Holt function is applied (Stanley *et al.* 2009). The median value for r was 0.392 and the standard deviation in the natural logarithm of r was 0.405.

Table A.2. Summary statistics for Monte Carlo outputs in generating a prior for r for north and south Atlantic swordfish.

Stock	Average	Median	Variance	SD	CV	SD(log(X))	Var(log(X))
North	0.4240	0.3922	0.02769	0.1664	0.3925	0.4050	0.1640
South	0.4244	0.3926	0.02772	0.1665	0.3924	0.4048	0.1639

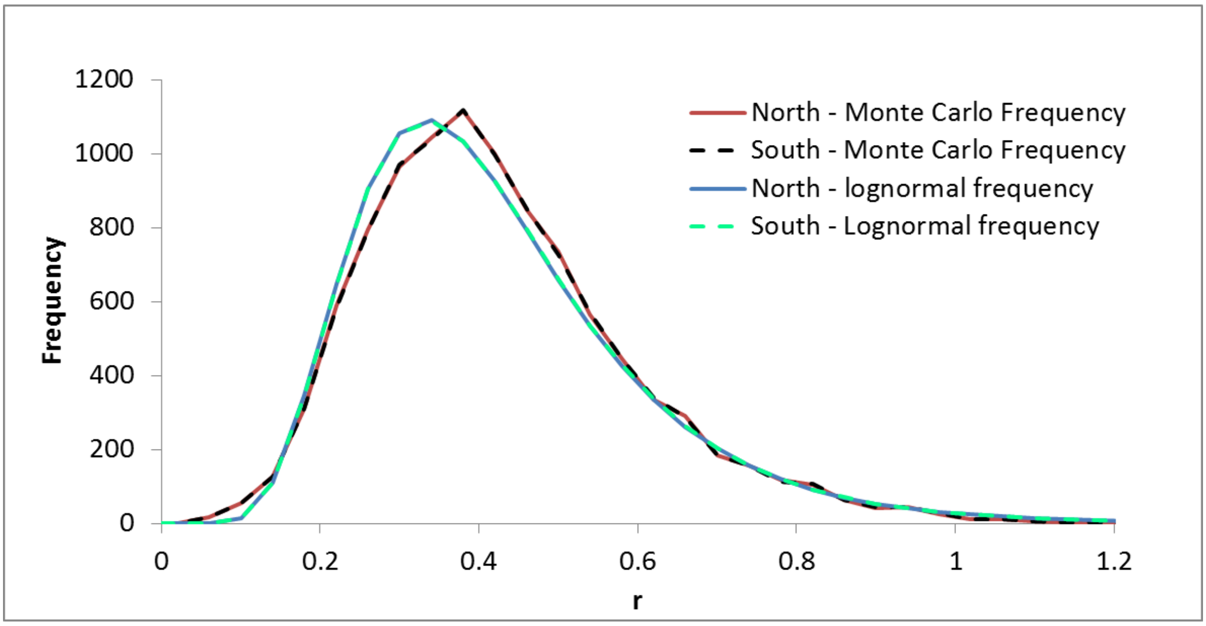

Figure A. 1. Plots of the Monte Carlo frequency distributions and lognormal approximations for the prior distributions for r for North and South Atlantic swordfish.

Appendix B: Proof for Importance Sampling for Approximation of Posterior Distributions

I provide here the statistical proof for why importance sampling can be expected via the strong law of large numbers to provide an unbiased estimate of the posterior distribution. To make the proof accessible to a wider audience, the proof is elaborated from the ones shown in Berger (1985) and McAllister *et al.* (1994). The posterior probability density for a parameter or parameter vector θ given the data obtained is given via Bayes theorem:

$$(B.1) \quad P(\theta \mid data) = \frac{P(data \mid \theta) p(\theta)\, d\theta}{\int P(data \mid \theta) p(\theta)\, d\theta}$$

where $P(data \mid \theta)$ is the probability of obtaining the data given a particular realization for θ, $p(\theta)$ is the prior density evaluate at θ. The posterior expectation or the posterior mean for some variable of interest, $g(\theta)$, which is a function of the parameter or parameters θ, is obtained by the following integral:

$$(B.2) \quad E^{P(\theta|data)}[g(\theta)] = \frac{\int g(\theta) P(data \mid \theta) p(\theta)\, d\theta}{\int P(data \mid \theta) p(\theta)\, d\theta}$$

$g(\theta)$ could represent for example the maximum sustainable yield, or stock biomass in the current year.

We can take the numerator on the right of equation B.2 and multiply it and divide it by a third density function, $h(\theta)$, which we will call the importance function of θ. $h(\theta)$ has the same set of parameters θ as the prior and posterior. For computational efficiency $h(\theta)$ is formulated to approximate the central tendency, covariance and spread of θ as the posterior density function that we are trying to approximate.

$$(B.3) \quad \int g(\theta) P(data \mid \theta) p(\theta)\, d\theta = \int \frac{g(\theta) P(data \mid \theta) p(\theta)\, h(\theta)}{h(\theta)}\, d\theta$$

with this rearrangement, the integral on the left of equation B.3 can be interpreted to be the posterior expectation of the product,

$$\frac{g(\theta) P(data \mid \theta) p(\theta)}{h(\theta)},$$

over the density function of h(θ).

$$(B.4) \quad \int g(\theta) P(data \mid \theta) p(\theta)\, d\theta = E^{h(\theta)}\left[\frac{g(\theta) P(data \mid \theta) p(\theta)}{h(\theta)}\right]$$

Using some independent and identically distributed (i.i.d.) sequence of *m* random variables θ_1, $\theta_2, \ldots, \theta_m$ drawn from the common density, $h(\theta)$, and the strong law of large numbers it follows that

$$(B.5) \quad \int g(\theta)\, P(data \mid \theta)\, p(\theta)\, d\theta = \lim_{m \to \infty} \sum_{k=1}^{m} \left[\frac{g(\theta_k)\, P(data \mid \theta_k)\, p(\theta_k)}{h(\theta_k)}\right]$$

Following from equation B.2, and equation B.5, the posterior expectation of g(θ) can be approximated via importance sampling:

$$(B.6) \quad E^{P(\theta|data)}[g(\theta)] = \frac{\sum_{k=1}^{m} g(\theta_k)\, w(\theta_k)}{\sum_{k=1}^{m} w(\theta_k)}$$

where

$$w(\theta_k) = \frac{P(data \mid \theta_k) p(\theta_k)}{h(\theta_k)}$$

and numerous i.i.d. draws of θ have been obtained from the importance function h(θ).