BAYESIAN STATE-SPACE SURPLUS PRODUCTION MODEL *JABBA* ASSESSMENT OF ATLANTIC BIGEYE TUNA (*THUNNUS OBESUS*) STOCK

H. Winker[1,2*], S. Kerwath[1,3], G. Merino[4] and M. Ortiz[5]

SUMMARY

*Previous stock assessment advice for Atlantic bigeye tuna (*Thunnus obesus*) originated from a combination of 'A Stock Production Model Incorporating Covariates' (ASPIC) and Stock Synthesis (ss3). This contribution aims to extend the assessment toolbox for this stock by the Bayesian State-Space Surplus Production Model software 'JABBA'. While Bayesian priors for the key parameters are kept uninformative, we specifically focus on developing an informative prior for approximating the expected process error from a stochastic age-structured simulation model. The model diagnostics provided ample support for use of the newly developed split, Joint-Research CPUE index in the reference case. Based on multi-model inference from JABBA runs over a small uncertainty grid, we predict with 85.5% probability that the stock remains overfished. Whereas JABBA appears sufficiently robust for inference about the stock status, we caution against the use of JABBA projections for specific quota recommendations in the case of bigeye tuna, because the relative impact of the different fleets can currently not be explicitly accounted for with (aggregated) biomass dynamic models.*

RÉSUMÉ

*Un précédent avis découlant de l'évaluation du stock de thon obèse de l'Atlantique (*Thunnus obesus*) était fondé sur la combinaison d'un modèle de production de stocks incorporant des covariables (ASPIC) et de Stock Synthesis (SS3). Cette contribution vise à étendre la boîte à outils d'évaluation de ce stock au moyen du logiciel « JABBA », le modèle de production excédentaire état-espace de type bayésien. Bien que les priors bayésiens pour les paramètres clés restent non informatifs, nous nous sommes concentrés spécifiquement sur la mise au point d'un prior informatif permettant d'estimer l'erreur de processus escomptée à partir d'un modèle de simulation stochastique structuré par âge. Les diagnostics du modèle viennent appuyer l'utilisation de la nouvelle division de l'indice de CPUE, fruit de la recherche conjointe, dans le cas de référence. Sur la base de l'inférence multimodèle des scénarios JABBA sur une petite grille d'incertitude, nous prédisons avec une probabilité de 85,5% que le stock reste surexploité. Alors que JABBA semble suffisamment robuste pour pouvoir déduire l'état du stock, nous recommandons de ne pas utiliser les projections de JABBA pour des recommandations de quota spécifiques dans le cas du thon obèse, car l'impact relatif des différentes flottilles ne peut actuellement pas être explicitement reflété (valeur agrégée) avec des modèles de dynamique de la biomasse.*

RESUMEN

*El asesoramiento de la evaluación de stock anterior para el patudo del Atlántico (*Thunnus obesus*) se generó a partir de una combinación de "un modelo de producción de stock que incorpora covariables" (ASPIC) y de Stock Synthesis (ss3). Esta contribución tiene como objetivo ampliar la herramienta de evaluación para este stock mediante un software de modelo de producción excedente bayesiano estado espacio "JABBA". Mientras que las distribuciones previas bayesianas para los parámetros clave se mantuvieron como no informativas, la atención se centró específicamente en el desarrollo de una distribución previa informativa para*

[1] DAFF, Department of Agriculture, Forestry and Fisheries, Private Bag X2, Rogge Bay 8012, South Africa.
[2] Centre for Statistics in Ecology, Environment and Conservation (SEEC), Department of Statistical Sciences, University of Cape Town, Rondebosch 7700, South Africa
[3] Department of Biological Sciences, University of Cape Town, Rondebosch 7700, South Africa
[4] AZTI, Herrera Kaia Portualdea, 20110, Pasaia, Spain
[5] ICCAT Secretariat. Calle Corazón de Maria 8, Madrid Spain 28002.
*Corresponding Author: henningW@daff.gov.za

realizar una aproximación del error de proceso previsto a partir de un modelo de simulación estocástico estructurado por edad. Los diagnósticos del modelo facilitaron un amplio respaldo para el uso del desglose recientemente desarrollado del índice de CPUE de investigación conjunta en el caso de referencia. Basándose en la inferencia de modelos múltiples a partir de ensayos de JABBA sobre una pequeña matriz de incertidumbre, se predijo con una probabilidad del 85,5% que el stock sigue estando sobrepescado. Mientras que JABBA parece suficientemente robusto para la inferencia sobre el estado del stock, desaconsejamos el uso de las proyecciones de JABBA para recomendaciones de cuota específicas en el caso del patudo, debido a que el impacto relativo de las diferentes flotas no pude tenerse explícitamente en cuenta actualmente en los modelos de dinámica de biomasa (agregada).

KEYWORDS

Tropical tunas, stock status, diagnostics,
process error, stochastic biomass dynamics, steepness

1. Introduction

In 2015, the International Commission for the Conservation of Atlantic Tunas (ICCAT) carried out the last stock assessment for Atlantic bigeye tuna (*Thunnus obesus*) based on data for the period 1950-2014. During this assessment several analysis approaches were explored to estimate the stock status, including catch curve analysis, (2) fitting surplus production models (SPMs), (3) Virtual Population Analysis (VPA) and (4) Integrated Assessment models. The scientific advice on the stock status was based of three alternative surplus production model runs using the open source software 'A Stock Production Model Incorporating Covariates' (ASPIC; Prager, 1994) and twelve alternative scenarios based on Integrated Assessments using Stock Synthesis (ss3; Methot and Wetzel, 2013).

ASPIC implements a generalized Pella and Tomlinson (1969) production function, which can be fitted to individual or multiple indices of abundance and conditions the stock population dynamics on either catch or effort. Fitting to data is done using negative maximum log-likelihood minimization and uncertainty in ASPIC can be evaluated with bootstrapping and sensitivity tests. The more recent release of the open-source *mpb* package (Kell 2016; https://github.com/laurieKell/mpb) for fitting surplus production models has been shown to be an efficient tool for reproducing ASPIC results (Kell et al., 2017; Merino et al., 2018a) and provides the additional option to estimate parameter posteriors using Monte Carlo Markov Chain (MCMC) simulations. Together with FLR libraries (Kell et al., 2007a; Merino et al., 2018b), *mpb* can be used to produce a variety of diagnostics and quality checks for the assessment (Kell et al., 2017). Perhaps the greatest shortcoming of both ASPIC and *mpb* is their inability to estimate process error, thus providing reference point estimates wholly reliant on deterministic population dynamics to fit abundance trends based on observation errors alone (Ono et al., 2012; Punt, 2003; Thorson and Minto, 2015). During the 2015 bigeye tuna assessment, three ASPIC runs based on separate fits to standardized abundance indices for Japanese, Taiwanese and US longline were selected to provide advice on stock status, biomass levels, and harvest rate.

In contrast to surplus production models, statistical age-structured models, such as ss3, allow separating between spawning-biomass (*SB*) and fleet specific exploitable biomass (*EB*), where *SB* is the biomass fraction of mature fish (or females) in the population, and *EB* is the exploitable (vulnerable) biomass fraction of the total biomass that is selected by each fleet. Age-structured models therefore explicitly account for the lag-effect of the biomass response of *SB* and *EB*, where the latter is related to the observed abundance index and can be informed by size or age composition data. In the case of the 2015 bigeye tuna assessment, catch, standardized catch-per-unit-effort (CPUE) indices and size composition data were assigned to 15 'fleets' related to Purse Seine (PS), Longline (LL) and Baitboat (BB) fleets operating in different areas. Each fleet is characterized by a selectivity function. This approach comes with the trade-off of a large number of stock parameters to model the population dynamics. Density-dependent processes are typically limited to a spawner-recruitment relationship (SRR) while natural mortality (M) is mostly treated as time invariant (Thorson et al., 2012). The form and steepness (h) of the SRR and estimates of M are highly uncertain, yet these parameters have a large effect on the outcome of statistical age-structured modelassessments. Because it is often not possible to estimate h and M from the data, scientists commonly explore alternative fixed values for one or both (Lee et al., 2012; Mangel et al., 2013). The substantial complexity inherent to multi-fleet ss3 models poses challenges during time constrained assessment meetings when it comes to adequately capturing the uncertainty around model structure related to confounding effects among growth, selectivity, natural mortality and recruitment.

The objective of this contribution is to extend the assessment toolbox for Atlantic bigeye tuna by the Bayesian State-Space Surplus Production Model software 'JABBA' (Winker et al. 2018; Just Another Bayesian Biomass Assessment). JABBA is implemented as a flexible, user-friendly open-source tool that is hosted on GitHub (https://github.com/jabbamodel). The fairly fast convergence properties for a Bayesian model allows for extensive use of model run permutations as required for diagnostics such as retrospective analysis, simulation testing or 'grid' runs for evaluation of uncertainty of model structure. As for SPMs in general, somatic growth, reproduction, natural mortality and associated density-dependent processes are inseparably captured within the surplus production function, which is governed by the intrinsic growth rate *r*, unfished biomass *K* and the shape parameter *m*. Due to these minimal parameter requirements, JABBA offers a parsimonious 'control' for the more parameter demanding ss3 assessment model. In contrast to ASPIC, the Bayesian state-space formulation for JABBA provides the potential for a more accurate representation of uncertainty by accounting for both process and observation error (Ono et al., 2012). An additional advantage appears to be the superior convergence properties of Bayesian state-space implementations in the presence of conflicting CPUE indices or stochasticity in the biomass dynamics. This has been evidenced by recent assessments of, for example, South Atlantic swordfish (ICCAT, 2017) or Atlantic blue marlin (ICCAT, 2018a).

Here, we apply JABBA to four initial scenarios based on alterative sets of CPUE indices, which we evaluate using a variety of model diagnostics. While priors for the key parameters *r* and *K* are kept purposefully uninformative, we specifically focus on developing an informative prior for approximating the expected range of process error for year-to-year biomass variation using a stochastic age-structured simulation model. We explore the structural model uncertainty of the reference case by implementing a small, one-dimensional grid of B_{MSY}/K values as function of the shape *m* being the only fixed input parameter. To facilitate comparability between JABBA and ss3 results, the range of B_{MSY}/K was chosen based on the equivalent ss3 output ratios of SB_{MSY} to unfished spawning biomass (SB_0), which can be directly related to alternative steepness values considered in 2015 and 2018 assessment. The results are discussed in the context of model robustness and multi-model inference for Atlantic bigeye stock assessment advice.

2. Material and Methods

2.1 Fishery input data

In addition to the ICCAT's Task I total catch time series (**Figure 1**), the following CPUE time series were considered for this analysis: two options for a standardized joint fleet CPUE index from the collaborative study (JR2), a new CPUE index from the Dakar baitboat (DAK_BB) fisheries and three longline CPUE indices provided by Japan (JP_LL), Taiwan (TW_LL) and the US (US_LL) (**Table 1**). Index CVs for the JR2 CPUEs and DAK_BB were scaled so that they averaged 0.2, while preserving the inter-annual variability. The index CPUEs for JP_LL, TW_LL and US_LL were fixed at 0.2 for all years.

We considered four initial scenarios based on alterative sets of CPUE indices (**Table 1)**. Following the recommendations from the 2018 ICCAT bigeye tuna data preparatory meeting (ICCAT, 2018b), the JR2_early and JR2_late CPUE indices were selected for the reference case S1 (**Figure 2**). Scenario S2 was fitted to the long (combined) JR2_long CPUE and S3 represent a variation of S1 by including the DAK_BB CPUE index (**Figure 2**). Considering that JP_LL, TW_LL and US_LL were probably the most influential CPUE indices in the 2015 assessments, we formulated a fourth, additional scenario S4 by fitting all three indices simultaneously (**Figure 2**).

2.2 JABBA stock assessment model

This Atlantic bigeye tuna stock assessment is implemented using the Bayesian state-space surplus production model framework JABBA (Winker et al., 2018a). JABBA implements the generalized three-parameter Pella-Tomlinson surplus production function of the form:

$$SP = \frac{r}{m-1} B \left(1 - \left(\frac{B}{K} \right)^{m-1} \right) \quad (1)$$

where *r* is the intrinsic rate of population increase at time *t*, *K* is the unfished biomass at equilibrium and *m* is a shape parameter that determines at which the *B*/*K* ratio maximum surplus production is attained. If $0 < m < 2$, *SP* attains MSY at biomass levels smaller than *K*/2 (Schaefer); the converse applies for values of *m* greater than 2. In line with previous surplus production model based bigeye tuna assessments, we initially assumed a Fox production function by setting $m = 1.0001$, so that maximum surplus production is predicted at ~ 0.37*K*.

In principle, JABBA follows the Bayesian state-space estimation framework proposed by Meyer and Millar (1999). The biomass B_y in year y is expressed as a proportion of K (i.e. $P_y = B_y /K$) to improve the efficiency of the estimation algorithm. The model is formulated to accommodate multiple CPUE series i. The initial biomass in the first year of the time series is scaled by introducing model parameter to estimate the ratio of the spawning biomass in the first year to K (Carvalho et al., 2014). The stochastic form of the process equation is given by:

$$P_y = \begin{cases} \varphi e^{\eta_y} & \text{for } y = 1 \\ \left(P_{y-1} + \frac{r}{(m-1)} P_{y-1}\left(1 - P_{y-1}^{\,m-1}\right) - \frac{\sum_f C_{f,y-1}}{K} \right) e^{\eta_y} & \text{for } y = 2,3,...,n \end{cases} \quad (2)$$

where $C_{f,y}$ is the catch in year y by fishery f and η_y is the process error, with $\eta_y \sim N(0, \sigma^2_{proc})$, where the process variance σ^2_{proc} can be either fixed or estimated using inverse-gamma distributions. .

The corresponding biomass for year y is:

$$B_y = P_y K \quad (3)$$

The observation equation is given by:

$$I_{i,y} = q_i B_y e^{\varepsilon_{y,i}} \quad (4)$$

where q_i is the estimable catchability coefficient associated with the abundance index i, and is $\varepsilon_{y,i}$ the observation error, with $\eta_y \sim N(0, \sigma^2_{\varepsilon_{i,y}})$, where $\sigma^2_{\varepsilon_{i,y}}$ is the observation variance in year y for index i. A full JABBA model description, including variance and prior specification options is available in Winker et al (2018).

2.3 Priors

Priors for r and K were kept uninformative to convey minimal prior information on the parameter estimates. For K, a lognormal was implemented using the JABBA 'range' option to set lower and upper 95% CI's located at 500,000 t and 5,000,000 t, respectively, which resulted in a mean value of 1,581,138.8 t and a CV of 172.7% (**Figure 3)**. Similarly, we set the lower and upper 95% CI's of an assumed lognormal distribution for r to 0.05 and 5, respectively, which resulted in a mean of r = 0.5 and an associated CV of 166% (**Figure 3)**. As for previous ASPIC runs, we assumed an initial B_{1950}/K ratio of 0.95 but admitted some uncertainty by assigning a lognormal prior with a CV of 5% (**Figure 3)**. Uninformative, uniform priors were specified for all CPUE specific catchability parameters q_i. Process error variance was assumed to follow inverse-gamma distribution (Brodziak and Ishimura, 2012; Meyer and Millar, 1999; Winker et al., 2018a). The approach for the deriving the shape and rate parameter for specifying the inverse-gamma prior for the process error variance is described in the next section.

2.4 Developing a process variance prior from age-structured simulations

To derive an informative prior for approximating expected process variance of the sampled biomass for Atlantic bigeye tuna, we developed a simple age-structured population model to simulate stochastic biomass dynamics in the absence of fishing. For this purpose, we assume that the natural stochasticity in biomass dynamics is mostly driven by variations in (1) recruitment (Thorson et al., in press) and (2) time-varying age-dependent natural mortality (Millar and Meyer, 2000a). To account for recruitment variation, we generated random deviates of lognormal recruitment standard deviations σ_R from a lognormal with a mean of log(0.55) and a CV of 20% to cover the spectrum of plausible values (σ_R= 0.4 – 0.8; **Figure 4a**) for bony fishes (Mertz and Myers, 1996; Rose et al., 2001). For natural mortality, we allowed for additional year-to-year deviations from the deterministic age-

dependent natural mortality M_a by assuming a lognormal error with a CV of 15% (**Figure 4b**). We note, however, that despite including time-varying natural mortality as additional source of variation, the resulting process noise on the biomass may still represent a conservative approximation for low biomass levels as our approach based on unfished biomass simulations does not account for the juvenescent effect on heavily fished, age-truncated populations.

The stock parameters used for this simulation experiment were sourced from the 2015 ICCAT bigeye assessment results and the report of the 2018 data preparatory session (**Table 2**). Assuming unfished, stochastic population dynamics, the numbers at age a for year y, $N_{a,y}$, are given by:

$$N_{a,y} = \begin{cases} R_y & \text{if } a = 0 \\ N_{a-1,y-1} \exp(-M_{y-1,a-1}) & \text{if } a > 0 \text{ and } a < a_{\max} \\ N_{a-1,y-1} \exp(-M_{y-1,a-1}) / (1 - \exp(-M_{y-1,a})) & \text{if } a = a_{\max} \end{cases} \quad (5)$$

where R_y the number of recruits (age = 0) in year y, $M_{y,a}$ is the age-specific instantaneous rate of natural mortality for year y, and $a_{\max}$ denotes the maximum age that is treated here as a plus group.

In the absence of fishing, we can ignore the existence of the spawning recruitment relationship, by assuming that annual recruitment is a lognormally distributed random variable with an expected value given by the mean unfished recruitment R_0, such that:

$$\log(R_t) = Normal\left(\log(R_0) - \frac{\sigma_R^2}{2}, \sigma_R^2\right) \quad (6)$$

where R_0 can be set to an arbitrary value (here R_0 = 1000) without loss of generality. For the initial year, we assume that abundance at age (excluding the plus group) is function of variation in recruitment from previous years, such that:

$$\log(N_{a,y=1}) \sim Normal\left(\log(\tilde{N}_a) - \frac{\sigma_R^2}{2}, \sigma_R^2\right) \quad \text{if } a > 0 \text{ and } a < a_{\max} \quad (7)$$

where $\tilde{N}_a$ denotes the deterministic numbers at age at for an unfished equilibrium recruitment R_0.

The sampled (exploitable) biomass EB_y of interest is given as a function of $N_{a,y}$, weight-at-age (w_a), selectivity-at-age (s_a), so that:

$$EB_y = \sum_a w_a s_a N_{a,y} \quad (8)$$

Weight-at-age is described as function of the weight to length conversion parameters ω and δ and length-at-age L_a (**Figure 4c**), such that

$$w_a = \omega L_a^{\delta} \quad (9)$$

The corresponding L_a was calculated based on the Bertalanffy growth function parameters as:

$$L_a = L_\infty (1 - e^{-\kappa (a - a_0)}) \quad (10)$$

where L_∞ is the asymptotic length, κ is the growth coefficient and a_0 is the theoretical age at zero length for sex *s*, respectively (**Figure 4d**). Selectivity-at-age was calculated as a function of length-at-age, L_a, using a two parameter logistic model of the form (**Figured 3e**):

$$s_a = \frac{1}{1 + e^{-(L_a - L_{50})/\delta_s}} \tag{11}$$

where s_a is the proportion of fish selected in the age *a*, L_{s50} is the length at which 50% of the fish are retained and δ_S is the inverse slope of the logistic ogive. The parameters L_{s50} and δ_S were chosen to approximate the estimated logistic selectivity functions for the Japanese long-line fishery based on the 2015 ss3 reference model run (**Figure 4f**).

The expected lognormal process error deviation for the natural variation of the (sampled) biomass can then be calculated as the standard deviation of the year-to-year change in the stochastic log(*EB*):

$$\sigma_{proc} = \sqrt{\frac{\left(\log(EB_y) - \log(EB_{y-1})\right)^2}{n_y - 1}} \tag{12}$$

where n_y is the number of years.

For this simulation experiment, we simulated the stochastic bigeye tuna biomass dynamics over a period of 68 years (**Figure 5)**, which corresponds to the currently available catch time series (1950-2018). We conducted a total of 10,000 Monte-Carlo simulation runs and noted σ_{proc} after each iteration. To obtain the shape and rate for the desired inverse-gamma prior for the process variance σ^2_{proc}, we fitted a gamma distribution to the inverse of the simulated process variance vector, such that $1/\sigma^2_{proc} \sim gamma(shape, rate)$, which produced the maximum likelihood estimates of *shape* = 0.9381 and *rate* = 0.03 (**Figure 5)**. The corresponding mean and CV of σ_{proc} are 0.06 and 0.174, respectively.

2.5 Model fitting and Diagnostics

JABBA is implemented in R (R Development Core Team, https://www.r-project.org/) with JAGS interface (Plummer, 2003) to estimate the Bayesian posterior distributions of all quantities of interest by means of a Markov Chains Monte Carlo (MCMC) simulation. The JAGS model is executed from R using the wrapper function jags() from the library r2jags (Su and Yajima, 2012), which depends on rjags. In this study, two MCMC chains were used. Each model was run for 30,000 iterations, sampled with a burn-in period of 5,000 for each chain and thinning rate of five iterations. Basic diagnostics of model convergence included visualization of the MCMC chains using MCMC trace-plots as well as Heidelberger and Welch (Heidelberger and Welch, 1992) and (Geweke, 1992) and Gelman and Rubin (1992) diagnostics as implemented in the coda package.

To evaluate CPUE fits, JABBA provides three types of plots. The first is referred to as "JABBA-residual plot" which includes: (1) colour-coded lognormal residuals of observed versus predicted CPUE indices by fleet, (2) boxplots indicating the median and quantiles of all residuals available for any given year; the area of each box indicates the strength of the discrepancy between CPUE series (larger box means higher degree of conflicting information), (3) a loess smoother through all residuals which highlights systematically auto-correlated residual patterns and (4) the Root-Mean-Squared-Error (RMSE) to quantitatively evaluate the randomness of model residuals. The second type of plot is similar to those obtained from the Stock Synthesis output (Methot and Wetzel, 2013) using the program r4ss (Taylor *et al.*, 2013) and shows the observed and predicted CPUE values in log scale, as well as the 95% credibility interval (CI). A third plot type shows observed CPUE and the model predicted mean CPUE and associated 95% confidence intervals over the entire catch time series.

Parameter posteriors relative to their priors were inspected visually. Here we focused specifically on prior and posterior distributions for the process variance. For this purpose, we assumed that an inflated process variance posteriors relative to the expected prior (Section 2.4) would point towards model misspecification for the given set of fitted CPUE indices. In addition, we provide process error plots for η_y (Eq. 2) to explore systematic patterns in process deviations of the predicted biomass over the assessment period.

To evaluate the robustness of important stock status quantities (biomass, surplus production, B/B_{MSY} and F/F_{MSY}) for use in projections, we conducted a retrospective analysis (Cadigan and Farrell, 2005; Hurtado-Ferro *et al*., 2014; Mohn, 1999) for the reference case S1 by sequentially removing the most the recent year and refitting the model over a period of ten years (i.e. 2017 back to 2007). To further address the model sensitivity to the trend in the JR2_early CPUE, we proceeded by a sequentially removing the earliest CPUE observations of the JR2_early starting in 1959 and ending in 1978, which we subsequently refer to as 'progressive pattern analysis'.

2.6 Evaluating structural model uncertainty with reference to age structured models

First recall that for SPMs structural and biological uncertainty is represented in the form of *K, r* and the shape *m* of the production function, where Schaefer and Fox formulations with fixed values of *m* being probably the most common choices. The shape parameter *m* is directly related to the inflection point of the surplus curve at B_{MSY}/K by:

$$\frac{B_{MSY}}{K} = m^{\left(-\frac{1}{m-1}\right)}$$

Given that *r* and *K* are estimated with uninformative priors, we assumed that the main source of structural uncertainty of the model can be attributed to the choice of *m*. To capture this uncertainty, we fitted the JABBA reference case model over a range of fixed B_{MSY}/K input values.

We noted that, by assuming a Fox model, the choice m = 1.001 (B_{MSY}/K~0.37) may not necessarily be comparable with the ss3 reference case steepness value of h = 0.8 agreed for 2015 and 2018 bigeye tuna assessments. Similar to *m*, *h* is a strong determinant for SB_{MSY}/SB_0 (Mangel et al., 2013). The linear relationship between *h* and SB_{MSY}/SB_0 is illustrated by the regression of the three runs from the 2018 ss3 assessment (**Figure 6)**. To facilitate comparability between JABBA and ss3 results, we therefore chose the range of B_{MSY}/K values from the equivalent ss3 averages of SB_{MSY}/SB_0 ratios for each the respective steepness values (h = 0.9, 0.8, 0.7). However, because $B_{MSY}/K < 0.37$ (Fox) is not defined for either mpb or ASPIC, we also present the stock status results for the initial Fox model for comparison with mpb (Merino *et al.*, 2018b).

3. Results

3.1 Model diagnostics

The visual inspection of trace plots for the key model parameters showed adequate mixing of the two chains (i.e., moving around the parameter space), which is indicative for convergence of the MCMC chains (**Appendix: Figure A1**). All key parameter MCMC chains passed both Heidelberger and Welch (Heidelberger and Welch, 1992) and Gelman and Rubin (1992) test diagnostics further providing evidence for adequate model convergence.

Model residuals for S1 and S2 showed only minor systematic residual pattern and indicated overall good fits to the JR2 indices (**Figure 7)**. This changed for S3, where major discrepancies were evident between the JR2_late and DAK_BB CPUE index. By comparison, there was again fairly little conflict evident among the JP_LL, TW_LL and US_LL CPUEs (**Figure 7)**. Judging by the RMSE, as a goodness-of-fit criterion, the reference case S1 produced the best fit (RMSE = 0.123), followed by S2 (RMSE = 0.127). The inclusion of the DAK_BB CPUE severely deteriorated the fit (RSME = 0.385), whereas S4 fitted the three independent longline indices comparably well (RSME = 0.170). In general, all individual fits to individual CPUE indices, shown on log-scale, appear adequate (**Figures 8-11)**, except for DAK_CPUE that showed a poor fit in the case of S3 (**Figure 10).** Additional figures illustrating the fits of the untransformed observed and predicted CPUE are provided in Appendix A (**Figures A2-A5).**

The estimates of the absolute quantities *K, MSY, B_{MSY}* and *F_{MSY}* alternated notably across the four scenarios (**Figure 12)**. The reference case S1 produced the lowest *K* and *B_{MSY}* and the highest *F*, while the MSY estimate closely approximated mean MSY across all four scenarios. The posterior medians for the stock status estimates were all located below *B_{2017}/B_{MSY}* and above *F_{2017}/F_{MSY}*, with S2 representing the most optimistic scenario, and S1 and S4 producing the most similar estimates. Perhaps noteworthy is that the reference case S1 estimates were associated with substantially higher precision than the three other scenarios, which can be inferred from the narrower boxplots (**Figure 12).** In general, absolute biomass estimates, such as *K* and *B_{MSY}* were estimated with poor precision, whereas the MSY estimate was reasonably precise, especially in S1.

Comparisons between prior and posterior distribution indicate that the priors for *K* and *r* conveyed minimal information to the parameters estimates (**Figure 13)**. Of particular interest were the locations of process error variance (σ^2_{proc}) posteriors relative to the informative prior, which was derived from the age-structured simulation experiment. Scenarios S1 and S2 resulted in σ^2_{proc} posteriors that closely approximated the prior, which indicated that the admitted σ^2_{proc} was sufficient to adequately describe the predicted population dynamics fitted to the JR2 CPUE indices. This was in sharp contrast to the strongly inflated σ^2_{proc} relative to the prior for S3 as a result of the inclusion of the DAK_BB CPUE index. By comparison, S4 resulted in only moderate inflation of the σ^2_{proc} posterior relative to the prior (**Figure 13)**.

The corresponding process error deviance (η_y) plots are shown in **Figure 14**. Here, η_y indicated a systematic, positive deviation from the zero mean over the period 1970-1989 for S2, which was less pronounced for the reference case S1. For S3, the sharp negative spike of η_y is a clear indication of forcing the predicted biomass to fit the DAK_BB index (**Figure 2)**, which also appears to retrospectively cause η_y inflation over the early time period 1960-1930. Similarly, the notable increase in positive η_y between 1965 and 1978 for S4 can be attributed to the increasing trend in the JP_LL CPUE (**Figure 2)**, which appears to be in conflict with the simultaneous increase in total catch (**Figure 1).**

The retrospective analysis conducted for the reference case revealed some retrospective patterns (**Figure 15)**. However, the observed departures from the reference case point estimates (S1) fell mostly within the 95% credibility intervals of the reference case posteriors and therefore appeared to be adequately represented by the estimated uncertainty (**Figure 15)**. Exceptions were the strong positive departures of *F/F_{MSY}* for the years 2011-2013 from the reference trajectory, which were associated with a steeper negative in *B/B_{MSY}*. On closer inspections of residuals (**Figures 7-8),** this period is preceded by a sequence of negative CPUE residuals that show increasing deviations from the zero mean for the years 2008-2012.

The progressive pattern analysis indicated that both stock status and MSY estimates were robust to sequential removal of the JR_early CPUE (**Figure 16)**. Even complete removal of the JR_early CPUE (1959-1978) would not result in any meaningful changes for the current stock status estimates.

3.4. Structural uncertainty and multi-model inference

First, the intermediate grid run, with *B_{MSY}/K* = 0.306 (*h* = 0.8), was used to conduct further sensitivity tests. Simultaneously decreasing the precision for the priors of *r* and *K* (CV 200% - 500%) did not have any notable influence on the parameter estimates and predictions, conforming that the priors were uninformative as intended (**Figure A6**). Similarly, sensitivity analysis showed that, in the case of the split CPUE JR2, it was possible to 'freely' estimate the process variance, using an uninformative inverse-gamma prior (Jeffrey's prior) with the scale and shape set to 0.001. The uninformative process variance prior resulted in even lower posteriors densities than predicted by the initial informative prior (**Figure**. **A7**). This had, however, no discernable effect on the stock status predictions, other than slightly improved precision estimates for model parameters and stock status quantities. As a result, the uninformative inverse-gamma prior was adopted for all final runs. A summary of the final JABBA-uncertainty grid model specifications is provided in **Table 3**.

The JABBA runs over the range of fixed *B_{MSY}/K* input values (0.26, 0.324 and 0.386) produced similar trajectories for fishing mortality (*F*) and biomass relative to unfished biomass (*B/K*) (**Figure 17)**. Over the initial period 1950-1990, total biomass estimates were higher for *B_{MSY}/K* = 0.278 compared to input values of 0.306 and 0.332, but similar thereafter. Harvest rates started to exceed sustainable levels in 1994 (**Figure 17)** when catches reached their historical maximum of about 135,000 t (**Figure 1)** and remained above levels that can produce MSY (*F/F_{MSY}* > 1) thereafter. Both MSY and *B_{MSY}* estimates fell within close proximity to each other as illustrated by the location of the maxima of the surplus production curve (**Figure 17; Table 4)**. The MSY point among for three grid models ranged from 77.493 – 76,768 t, whereas the Fox model run was slightly higher at

80,288 t (**Table 4)**. Point estimates of B_{2017}/K for the year 2017 ranged from 0.244-0.252, where B_{MSY}/K = 0.278 (high h = 0.9) resulted in the most pessimistic B_{2017}/K. The opposite is the case for the trajectories of B/B_{MSY} and F/F_{MSY}. Here the B_{MSY}/K = 0.278 (h = 0.9) produced the most optimistic stock status trajectories for B/B_{MSY} and F/F_{MSY} that showed notably more contrast to runs with B_{MSY}/K input values of 0.306 (h = 0.8) and 0.332 (h = 0.7). This can attributed to predetermining the maximum of the surplus production curve (MSY) along the B_{MSY}/K (x-)axis by the choice of the shape parameter m (and steepness h) which appears to be compensate by increased estimates of K as the reference point B_{MSY}/K is decreased (**Table 4)**. Unsurprisingly, the model results based on the low h = 0.7 came closest to the initial Fox model (**Table 4)**.

Degrees of stock depletion and overfishing are illustrated the three alternative B_{MSY}/K input values and the Fox reference case model in the form of individual Kobe plots (**Figure 18)**. The cumulative probability of the red and yellow regions suggests that current biomass levels are below B_{MSY} with probabilities ranging between 77.4% for B_{MSY}/K = 0.278 to 92.5% for B_{MSY}/K = 0.332. The combined uncertainty about the stock status reference points B_{2017}/B_{MSY} and F_{2017}/F_{MSY}, for the three alternative B_{MSY}/K (and thus m) reference case runs is presented as a combined Kobe posterior plot (**Figure 19)**. The combined posteriors from the three grid runs predict with 85.5% probability that the stock remains overfished.

4. Discussion

The presented Atlantic bigeye tuna assessment represents the first application of the Bayesian state-space surplus production model JABBA to a major tuna stock. We found that all scenarios convergences with two fairly short MCMC chains, which allowed completing a JABBA run in less than 2 minutes. This makes it possible to reproduce the entire presented analysis in less 40 min, including all scenarios, retrospective- and progress analysis, sensitivity analysis on B_{MSY}/K, projections and figure outputs. The four considered sets CPUE indices fitted reasonably well with the exception of the DAK_BB CPUE. Overall, our analysis provided evidence that supports the choice of the split JR2 CPUE for the reference case model based on the goodness-of-fit, parameter precision and more desirable residual and process error patterns compared to the alternative CPUE scenarios. Both retrospective and progress patterns performed overall adequate. The exceptions were the fairly strong retrospective patterns for the years 2011-2013, which warrants closer examination. In the context of the past 2015 and 2010 bigeye tuna assessments, the retrospective analysis showed similar MSY estimates for data fitted through 2014, but would have suggested a more conservative MSY estimate and more severe overfishing for data fitted through 2009, respectively.

The approach to develop a process variance prior through stochastic age-structured simulations was originally inspired by a request of process error diagnostic plot during the 2017 International Fisheries Stock Assessment Review Workshop in Cape Town (http://www.maram.uct.ac.za/ maram/workshops/2017). The panel suggested that, ideally, the posterior for the process error variance should be comparable to the variation in biomass obtained by projecting the age-structured model forward without catches, but with process error (Cox et al., 2017). Here, we build on this concept by directly using stochastic age-structured projection to formulate an informative for the process variance. Interestingly, the simulated process variance range for bigeye tuna closely matches the process variance prior developed for South Atlantic albacore tuna (*Thunnus alalunga*) by Millar and Meyer (2000b). Based on information on recruitment variation in age-3 fish (Myers et al., 1999) and the relative contribution of age-3 fish to the biomass as a function of somatic growth and natural mortality Millar and Meyer (2000b), predicted a process error range between 0.04 and 0.08. In general, this agrees with the typically considered process errors (< 0.15), where Bayesian state-space surplus models appear to perform generally well (Froese et al., 2017; Ono et al., 2012; Thorson et al., 2014; Winker et al., 2018a).

The grid-approach covering alternative B_{MSY}/K inputs highlights the importance of capturing the model uncertainty associated with the shape parameter m, which determines the position of B_{MSY} relative to K. We found that MSY and biomass depletion estimates relative to K were fairly robust to the choice of B_{MSY}/K, but that the reference points B/B_{MSY} and F/F_{MSY} essential become a 'moving target' if m is varied. However, the situation is no different for age-structured models (**Figure 6)**, where the choice of steepness h typically predetermines the position of the target reference point SB/SB_{MSY} (Mangel et al., 2013). Further, we illustrated that the Fox model parameterization for Atlantic bigeye tuna would be best approximated by a steepness h = 0.56 as this would achieve comparable values of SB_{MSY}/SB_0 and B_{MSY}/K for ss3 and JABBA, respectively. This also suggest that the Fox bigeye tuna model is more directly comparable to the low steepness sensitivity run (h = 0.7) than the ss3 reference case (h = 0.8).

The implicit relationships among *h*, *m* and the functional form of the stock's productivity curve are often ignored when comparing outputs from age-structured and surplus production models. Maunder (2003) highlighted issue by pointing out that the Schaefer model, in predicting MSY at 50% unfished biomass, rarely matches the typical range of steepness values of $h = 0.6 – 0.95$ considered in age-structured assessments for most tuna and billfishes, which would imply MSY at biomass depletion levels that are notably below 50%. To facilitate comparability between JABBA and ss3 results, we therefore chose the range of B_{MSY}/K (0.278, 0.306 and 0.332) that corresponded to the equivalent ss3 ratios of SB_{MSY}/SB_0 based on the steepness values ($h = 0.9, 0.8, 0.7$) considered in 2015 assessment. We suggest capturing the model uncertainty associated with *m* by combining the posterior from multiple JABBA runs as a means to draw multi-model inference about the stock status.

The number of age-structured assessments of tuna and tuna-like species has been continuous increasing over the last three decades (Thorson et al., 2018), with stock synthesis having been on the forefront of this development in recent years. Yet, surplus production models persist as a routine assessment tool within their traditional realm of large pelagic tuna, billfish and shark assessments (Carvalho et al., 2014; Punt et al., 2015; Winker et al., 2018a). Judging between the two model types is not always straight forward and will dependent on data available as well as fisheries dynamics. For example, in the cases of swordfish and blue marlin most of the catch is typically taken by long-line fleets with overall similar selectivity (ICCAT, 2017; Winker et al., 2018b) so that at it may sufficient describe the stock in the form of aggregated biomass dynamics. This situation is different for bigeye tuna where purse-seine and longline fisheries target entirely different age classes of the population. For Pacific bigeye tuna it has been shown the shift towards an increasing purse seine catch directly influences the attainable MSY and surplus production (Wang et al., 2009). Nevertheless, this initial JABBA assessment appears sufficiently robust for reliable inference about the stock status. However, we caution against the use of JABBA-based projections for specific quota recommendations in the case of bigeye tuna, because the likely strong difference in the relative impact of the different fleets can currently not be explicitly accounted for with (aggregated-) biomass dynamic models.

References

Brodziak, J., Ishimura, G., 2012. Development of Bayesian production models for assessing the North Pacific swordfish population. Fish. Sci. 77, 23–34. doi:10.1007/s12562-010-0300-0

Cadigan, N.G., Farrell, P.J., 2005. Local influence diagnostics for the retrospective problem in sequential population analysis. ICES J. Mar. Sci. 62, 256–265. doi:10.1016/j.icesjms.2004.11.015

Carvalho, F., Ahrens, R., Murie, D., Ponciano, J.M., Aires-da-silva, A., Maunder, M.N., Hazin, F., 2014. Incorporating specific change points in catchability in fisheries stock assessment models : An alternative approach applied to the blue shark (*Prionace glauca*) stock in the south Atlantic Ocean. Fish. Res. 154, 135–146. doi:10.1016/j.fishres.2014.01.022

Cox, S.P., Howell, D., Punt, A.E., 2017. International Review Panel Report for the 2017 International Fisheries Stock Assessment Workshop. Cape Town.

Froese, R., Demirel, N., Coro, G., Kleisner, K.M., Winker, H., 2017. Estimating fisheries reference points from catch and resilience. Fish Fish. 18. doi:10.1111/faf.12190

Gelman, A., Rubin, D.B., 1992. lnference from Iterative Simulation Using Multiple Sequences. Stat. Sci. 7, 457–472. doi:10.2307/2246093

Geweke, J., 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments., in: Berger, J.O., Bernardo, J.M., Dawid, A.P., Smith, A.F.M. (Eds.), Bayesian Statistics 4: Proceedings of the Fourth Valencia International Meeting. Clarendon Press, Oxford, pp. 169–193.

Heidelberger, P., Welch, P.D., 1992. Simulation run length control in the presence of an initial transient. Oper. Res. 31, 1109–1144. doi:10.1287/opre.31.6.1109

Hurtado-Ferro, F., Szuwalski, C.S., Valero, J.L., Anderson, S.C., Cunningham, C.J., Johnson, K.F., Licandeo, R., McGilliard, C.R., Monnahan, C.C., Muradian, M.L., Ono, K., Vert-Pre, K.A., Whitten, A.R., Punt, A.E., 2014. Looking in the rear-view mirror: Bias and retrospective patterns in integrated, age-structured stock assessment models, in: ICES Journal of Marine Science. pp. 99–110. doi:10.1093/icesjms/fsu198

ICCAT, 2018a. Report of the 2018 ICCAT blue marlin stock assessment meeting. ICCAT-SCRS 18–22.

ICCAT, 2018b. Report of the 2018 ICCAT bigeye tuna data preparatory meeting. ICCAT-SCRS 1–44.

ICCAT, 2017. Report of the 2017 ICCAT Atlantic swordfish stock assessment session. Collect. Vol. Sci. Pap. ICCAT 74, 841–967.

Kell, L., Arrizabalaga, H., De Bruyn, P., Merino, G., Mosqueira, I., Sharma, R., de Urbina, J.-M.O., 2017. Validation of the biomass dynamic stock assessment model for use in a management procedure. Col. Vol. Sci. Pap. ICCAT 73, 1354–1376.

Kell, L.T., Mosqueira, I., Grosjean, P., Fromentin, J., Garcia, D., Hillary, R., Jardim, E., Mardle, S., Pastoors, M.A., Poos, J.J., Scott, F., Scott, R.D., 2007a. FLR : an open-source framework for the evaluation and development of management strategies 640–646.

Kell, L.T., Mosqueira, I., Grosjean, P., Fromentin, J., Garcia, D., Hillary, R., Jardim, E., Mardle, S., Pastoors, M.A., Poos, J.J., Scott, F., Scott, R.D., 2007b. FLR : an open-source framework for the evaluation and development of management strategies. ICES J. Mar. Sci. 64, 640–646.

Lee, H.H., Maunder, M.N., Piner, K.R., Methot, R.D., 2012. Can steepness of the stock-recruitment relationship be estimated in fishery stock assessment models? Fish. Res. 125–126, 254–261. doi:10.1016/j.fishres.2012.03.001

Mangel, M., MacCall, A.D., Brodziak, J., Dick, E., Forrest, R.E., Pourzard, R., Ralston, S., Chang, Y., Lee, H., Mangel, M., MacCall, A.D., Brodziak, J., Dick, E., Forrest, R.E., Pourzard, R., Ralston, S., 2013. A Perspective on Steepness, Reference Points, and Stock Assessment. Can. J. Fish. Aquat. Sci. 940, 930–940. doi:10.1139/cjfas-2012-0372

Maunder, M.N., 2003. Is it time to discard the Schaefer model from the stock assessment scientist's toolbox? Fish. Res. 61, 145–149. doi:10.1016/S0165-7836(02)00273-4

Merino, G., Murua, H., Urtizberea, A., Santiago, J., Winker, H., Walter, J., 2018a. Continuity stock assessment for Atlantic bigeye using a biomass production model. SCRSSCRS/2018/099.

Merino, G., Murua, H., Urtizberea, A., Santiago, J., Winker, H., Walter, J., 2018b. Alternatives for the stock assessment for Atlantic bigeye using a biomass production model. SCRS/2018/100.

Mertz, G., Myers, R. a, 1996. Influence of fecundity on recruitment variability of marine fish. Can. J. Fish. Aquat. Sci. 53, 1618–1625. doi:10.1139/f96-089

Methot, R.D., Wetzel, C.R., 2013. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fish. Res. 142, 86–99. doi:http://dx.doi.org/10.1016/j.fishres.2012.10.012

Meyer, R., Millar, R.B., 1999. BUGS in Bayesian stock assessments. Can. J. Fish. Aquat. Sci. 56, 1078–1086. doi:10.1139/cjfas-56-6-1078

Millar, R.B., Meyer, R., 2000a. Bayesian state-space modeling of age-structured data: fitting a model is just the beginning. Can. J. Fish. Aquat. Sci. 57, 43–50. doi:10.1139/cjfas-57-1-43

Millar, R.B., Meyer, R., 2000b. Non-linear state space modelling of fisheries biomass dynamics by using Metropolis-Hastings within-Gibbs sampling. J. R. Stat. Soc. Ser. C (Applied Stat. 49, 327–342. doi:10.1111/1467-9876.00195

Mohn, R., 1999. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. ICES J. Mar. Sci. 56, 473–488. doi:10.1006/jmsc.1999.0481

Myers, R. a, Bowen, K.G., Barrowman, N.J., 1999. Maximum reproductive rate of fish at low population sizes. Can. J. Fish. Aquat. Sci. 56, 2404–2419. doi:10.1139/f99-201

Ono, K., Punt, A.E., Rivot, E., 2012. Model performance analysis for Bayesian biomass dynamics models using bias, precision and reliability metrics. Fish. Res. 125, 173–183. doi:10.1016/j.fishres.2012.02.022

Plummer, M., 2003. JAGS: A Program for Analysis of Bayesian Graphical Models using Gibbs Sampling, 3rd International Workshop on Distributed Statistical Computing (DSC 2003); Vienna, Austria.

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

Punt, A.E., 2003. Extending production models to include process error in the population dynamics. Can. J. Fish. Aquat. Sci. 60, 1217–1228. doi:10.1139/f03-105

Punt, A.E., Su, N.-J., Sun, C.-L., 2015. Assessing billfish stocks: A review of current methods and some future directions. Fish. Res. 166, 103–118. doi:10.1016/j.fishres.2014.07.016

Rose, K.A., Cowan Jr, J.H., Winemiller, K.O., Myers, R.A., Hilborn, R., Jr, J.H.C., Winemiller, K.O., Myers, R.A., Hilborn, R., 2001. Compensatory density dependence in fish populations: importance, controversy, understanding and prognosis. Fish Fish. 2, 293–327.

Su, Y.-S., Yajima, M., 2012. R2jags: A Package for Running jags from R.

Taylor, I.G., Stewart, I.J., Hicks, A.C., Garrison, T.M., Punt, A.E., Wallace, J.R., Wetzel, C.R., Thorson, J.T., Takeuchi, Y., Ono, K., Monnahan, C.C., Stawitz, C.C., Teresa, Z.A., Whitten, A.R., Johnson, K.F., Emmet, R.L., Anderson, S.C., Iantaylornoaagov, M.I.T., 2013. Package 'r4ss': R Code for Stock Synthesis.

Thorson, J.T., Cope, J.M., Branch, T.A., Jensen, O.P., Walters, C.J., 2012. Spawning biomass reference points for exploited marine fishes, incorporating taxonomic and body size information. Can. J. Fish. Aquat. Sci. 69, 1556–1568. doi:10.1139/f2012-077

Thorson, J.T., Minto, C.C., 2015. Mixed effects: a unifying framework for statistical modelling in fisheries biology. ICES J. Mar. Sci. 72, doi:10.1093/icesjms/fsu213. doi:10.1093/icesjms/fst048

Thorson, J.T., Ono, K., Munch, S.B., 2014. A Bayesian approach to identifying and compensating for model misspecification in population models. Ecology 95, 329–341.

Thorson, J.T., Rudd, M.B., Winker, H., 2018. (in press) The case for estimating recruitment variation in data-moderate and data-poor age-structured models. Fish. Res. https://doi.org/10.1016/j.fishres.2018.07.007.

Wang, S.-P., Maunder, M.N., Aires-da-Silva, A., Bayliff, W.H., 2009. Evaluating fishery impacts: Application to bigeye tuna (Thunnus obesus) in the eastern Pacific Ocean. Fish. Res. 99, 106–111. doi:10.1016/j.fishres.2009.05.010

Winker, H., Carvalho, F., Kapur, M., 2018a. JABBA: Just Another Bayesian Biomass Assessment. Fish. Res. 204, 275–288.

Winker, H., Carvalho, F., Sow, F.N., Ortiz, M., 2018b. Unifying parameterizations between age-structured and surplus production models: An application to Atlantic blue marlin (*Makaira nigricans*). ICCAT-SCRS/2018/092 1–16.

Table 1. Summary of four scenarios S1-S4 that were formulated based on alternative sets of standardized CPUE indices for Atlantic bigeye tuna.

Scenario	CPUE indices and period	Abbreviation
S1 (reference case)	Joint R2 Early no vessel id (1959-1978)	JR2_early
	Joint R2 Late vessel id (1979-2017)	JR2_late
S2	Joint R2 Long term no vessel id (1959-2017)	JR2_long
S3	Joint R2 Early no vessel id (1959-1978)	JR2_early
	Joint R2 Late vessel id (1979-2017)	JR2_late
	Dakar Baitboat (2005-2017)	DAK_BB
S4	Japan Longline (1961-2017)	JP_LL
	US longline (1986-2017)	US_LL
	Taiwan longline (1995-2017)	TW_LL

Table 2. Stock parameters for bigeye tuna used as input for the age-structured simulation model to develop an informative prior for the process error variance (see text for parameter definition).

Parameter	Value	Unit	Age	M_a
L_∞	217.3	cm	0	0.72
κ	0.18	year^{-1}	1	0.486
t_0	-0.709	years	2	0.383
a	0.00002396	kg	3	0.326
b	2.9774	kg cm^{-1}	4	0.29
$a_{\min}$	0	years	5	0.265
$a_{\max}$	10	years	6	0.248
S_L	107	cm	7	0.235
$\square_S$	9	cm^{-1}	8	0.225
M_a (CV = 0.15)	Lorenzen	year^{-1}	9	0.218
			10+	0.212

Table 3. Summary JABBA uncertainty grid model specifications for Atlantic bigeye tuna.

Quantity	Specification	Abbreviation
CPUE	Joint R2 Early no vessel id (1959-1978)	JR2_early
	Joint R2 Late vessel id (1979-2017)	JR2_late
Unfished biomass	K ~ lnorm(log(1,581139),1.726) with 0.025th = 500,000 and 0.975th = 5,000,000	K
Intrinsic rate of population increase	r ~ lnorm(log(0.5),1.66) with 0.025th = 0.05 and 0.975th = 5	r
Initial biomass depletion (B_{1950}/K)	φ ~ lnorm(1,0.05)	φ
Biomass at MSY relative to the unfished biomass	B_{MSY}/K = 0.278 B_{MSY}/K = 0.306 B_{MSY}/K = 0.332	B_{MSY}/K
Process variance	σ^2_{proc}~ inverse-gamma (0.001,0.001)	σ^2_{proc}
catchability coefficient	q ~ uniform (10^{-30},1000)	q

Table 4. Summary of posterior quantiles denoting the median and the 95% confidence intervals of parameter estimates for the JABBA reference grid runs and the Fox model run for Atlantic bigeye tuna.

	B/B_{MSY} = 0.306 (Ref: h = 0.8)			B/B_{MSY} = 0.332(low: h = 0.7)		
Estimates	Median	2.50%	97.50%	Median	2.50%	97.50%
K (t)	1349994	958351	2270464	1262803	899757	2018226
r	0.132	0.07	0.199	0.154	0.088	0.229
ψ (*psi*)	0.934	0.853	1.019	0.937	0.853	1.021
σ_{proc}	0.032	0	0.089	0.045	0	0.084
m	0.706	0.706	0.706	0.82	0.82	0.82
F_{MSY}	0.188	0.099	0.281	0.188	0.107	0.279
B_{MSY} (t)	413106	293261	694776	419299	298754	670128
MSY (t)	77493	65695	86427	78608	68454	87446
B_{1959}/K	0.926	0.814	1.015	0.927	0.813	1.021
B_{2017}/K	0.252	0.185	0.338	0.257	0.191	0.339
$B_{2017}/B_{\mathrm{MSY}}$	0.822	0.606	1.106	0.775	0.575	1.02
$F_{2017}/F_{\mathrm{MSY}}$	1.214	0.848	1.738	1.272	0.923	1.792
	B/B_{MSY} = 0.278 (high: h = 0.9)			Fox (B_{MSY}/K = 0.37)		
Estimates	Median	2.50%	97.50%	Median	2.50%	97.50%
K (t)	1408989	1017500	2224121	1176284	815980	3583657
r	0.117	0.067	0.171	0.185	0.056	0.283
ψ (*psi*)	0.936	0.852	1.019	0.936	0.854	1.021
σ_{proc}	0.045	0	0.084	0.045	0	0.089
m	0.597	0.597	0.597	1.001	1.001	1.001
F_{MSY}	0.196	0.112	0.287	0.185	0.055	0.283
B_{MSY} (t)	391753	282904	618391	432947	300332	1319013
MSY (t)	76768	66141	85521	80288	66674	89365
B_{1959}/K	0.927	0.811	1.017	0.925	0.808	1.018
B_{2017}/K	0.244	0.183	0.323	0.267	0.191	0.356
$B_{2017}/B_{\mathrm{MSY}}$	0.879	0.659	1.163	0.724	0.519	0.968
$F_{2017}/F_{\mathrm{MSY}}$	1.148	0.816	1.615	1.327	0.949	2.077

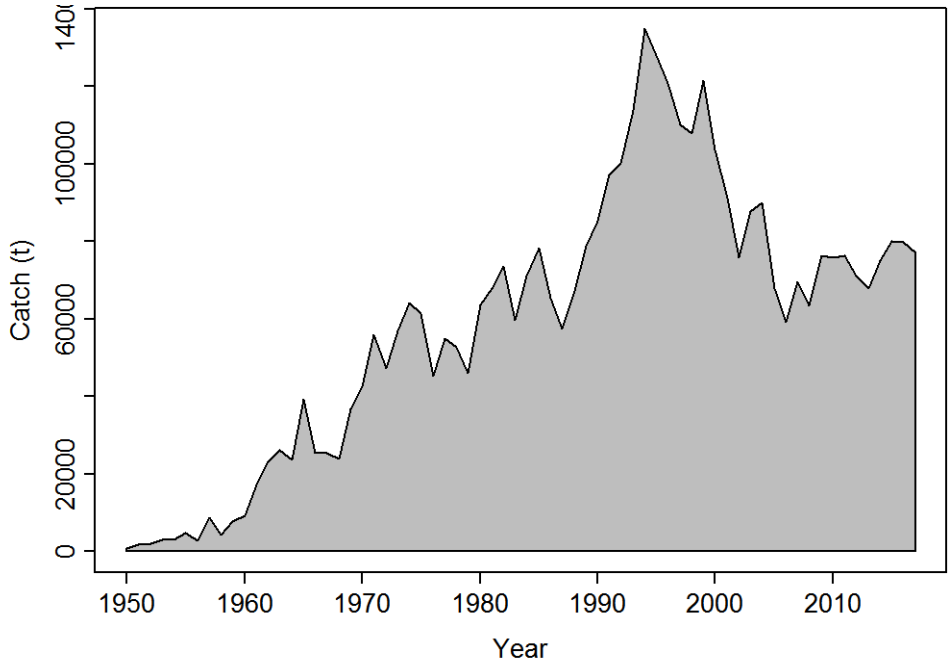

Figure 1. Time-series of catch in metric tons (t) for Atlantic bigeye tuna (1950-2017).

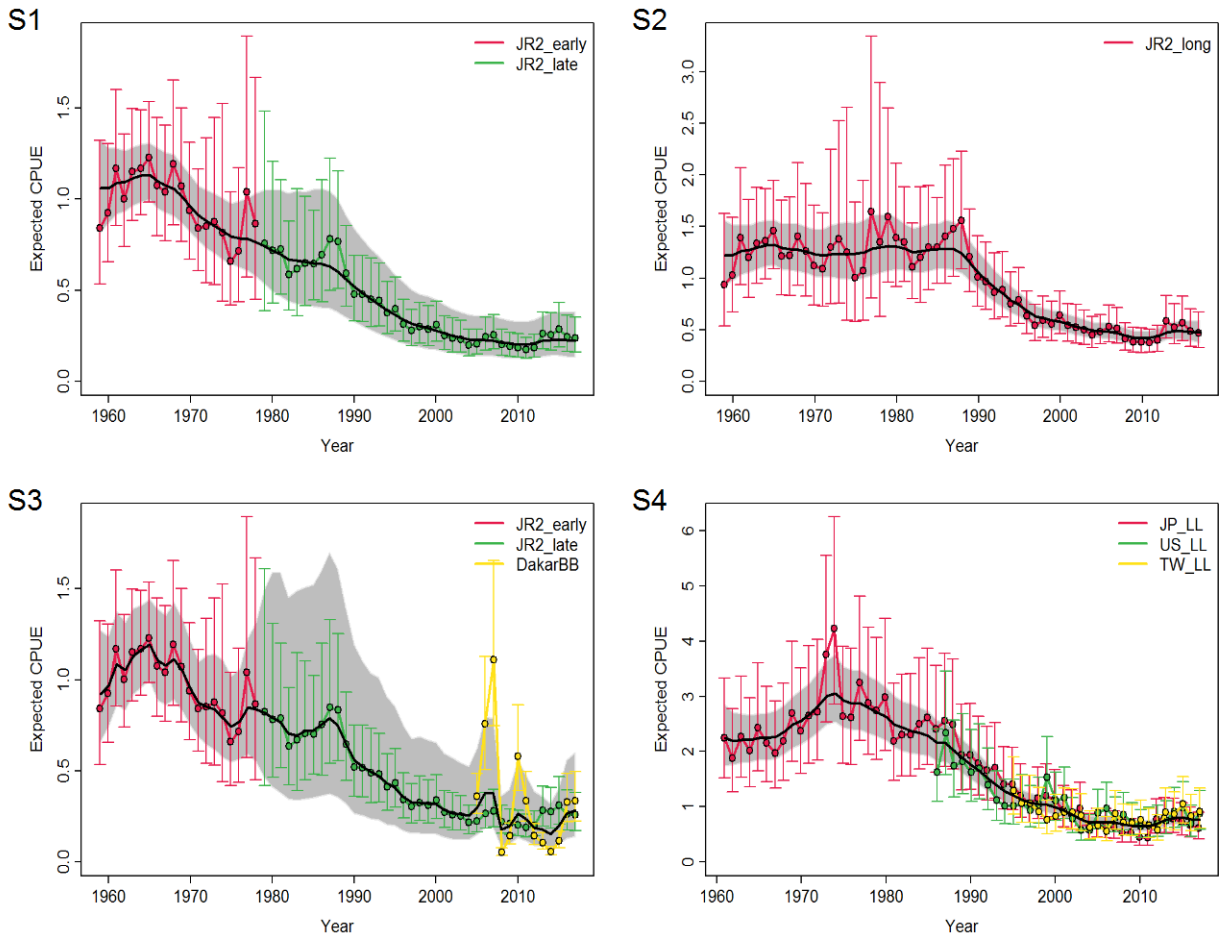

Figure 2. Illustrations of trends in CPUE indices according to Scenarios S1-S4 for Atlantic bigeye tuna, which were produced using the state-space CPUE averaging tool implemented in JABBA. The underlying abundance trend is treated as an unobservable state variable that follows a log-linear Markovian process, so that the current mean relative abundance was assumed to be a function of the mean relative abundance in the previous year, an underlying mean population trend and lognormal process error term. The CPUE indices are aligned with the base index via estimable catchability scaling parameters.

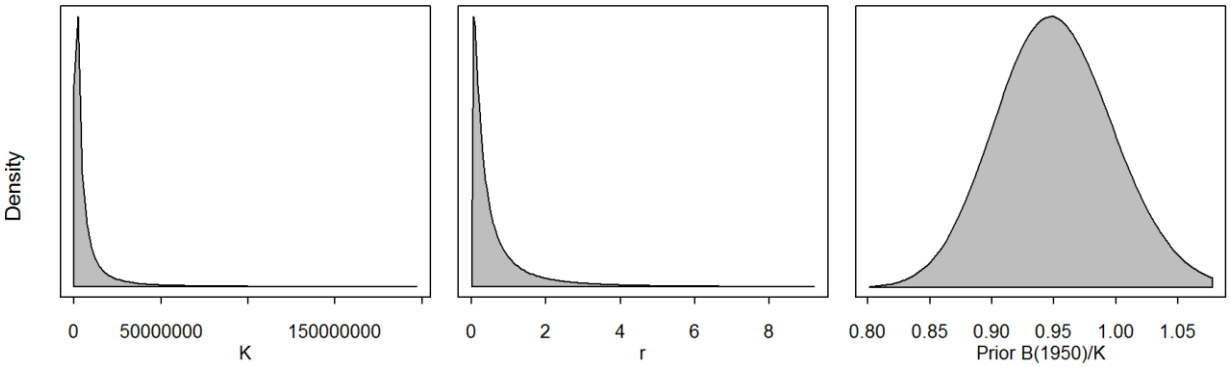

Figure 3. Illustration of the 'flat' lognormal priors for *K* (CV=173%) and *r* (CV = 166%) and the informative lognormal prior for the initial biomass depletion $P_{1950} = B_{1950}/K$.

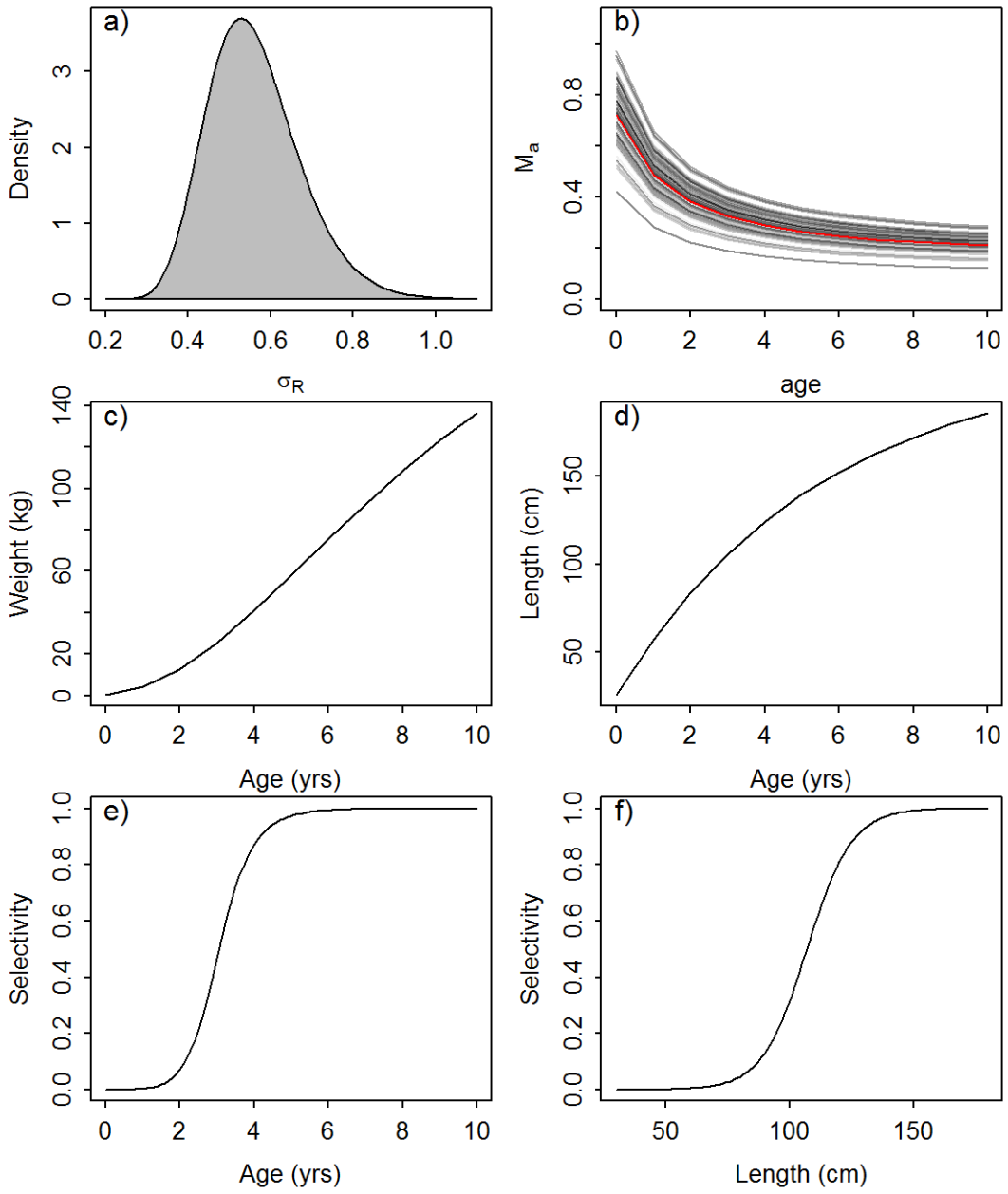

Figure 4. Illustrating the assumed (a) lognormal standard deviations determining the variation in recruitment (σ_R) variation, (b) inter-annual variation in the natural mortality-at-age (M_a) and functional relationships of (c) weight-at-age (d) length-at-age, (e) selectivity-at-age and (f) selectivity-at-length used to simulated the natural variation in an unfished population of Atlantic bigeye tuna.

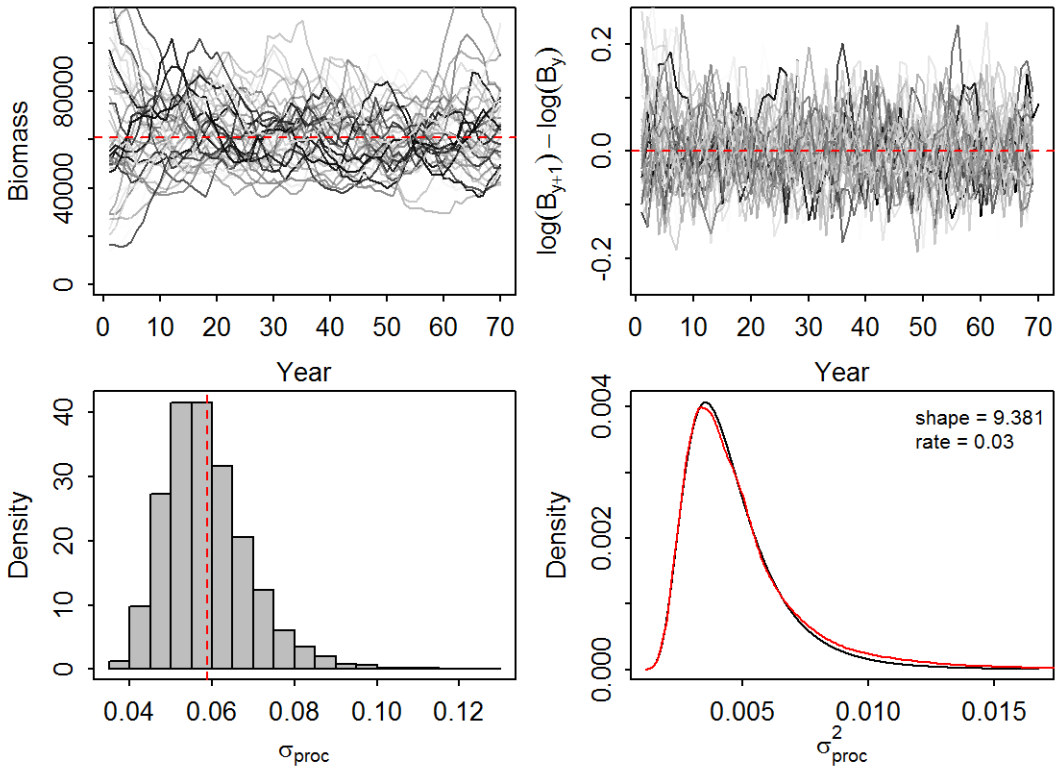

Figure 5. Showing simulated (exploitable) variation in biomass trajectories (top left) of an unfished population of Atlantic bigeye tuna over a period of 68 years; log-biomass deviations for sequential years (top right); the resulting distribution of 10,000 random process error deviates (bottom left); and the observed distribution for the process error variance (black line), with shape = 0.9381 and rate = 0.03 representing the maximum likelihood estimates of the fitted (red line) inverse-gamma distribution (bottom right).

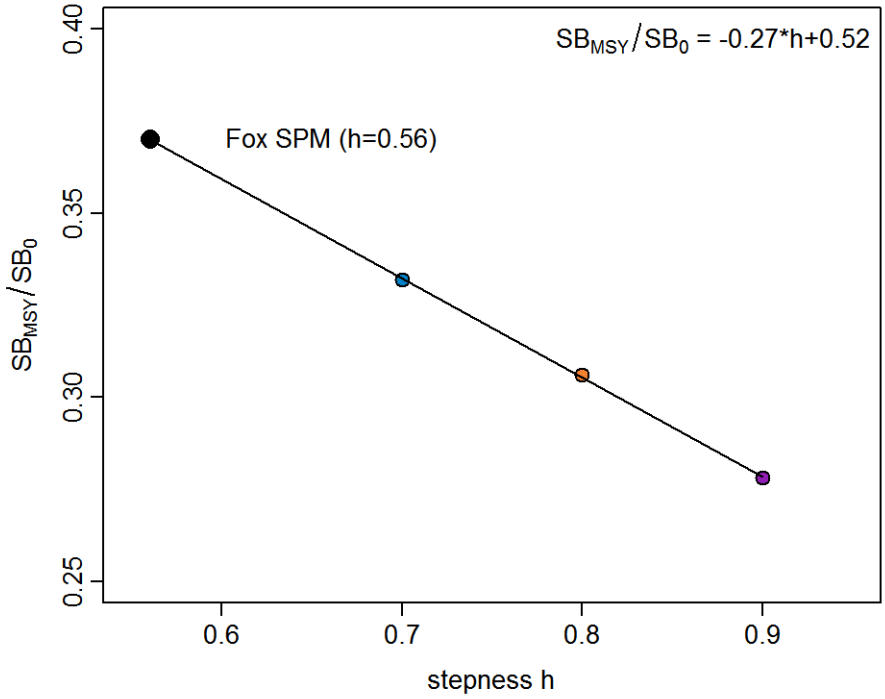

Figure 6. Linear relationship between steepness *h* in values and predicted SB_{MSY}/SB_0 ratio from the stock synthesis (ss3) model runs for the range of steepness values (*h* = 0.7-0.8) conducted during 2018 ICCAT bigeye tuna stock assessment. The solid black circle denotes the predicted position of *h* = 0.56 that would correspond to B/B_{MSY} ~ 0.37 for the Fox surplus production model.

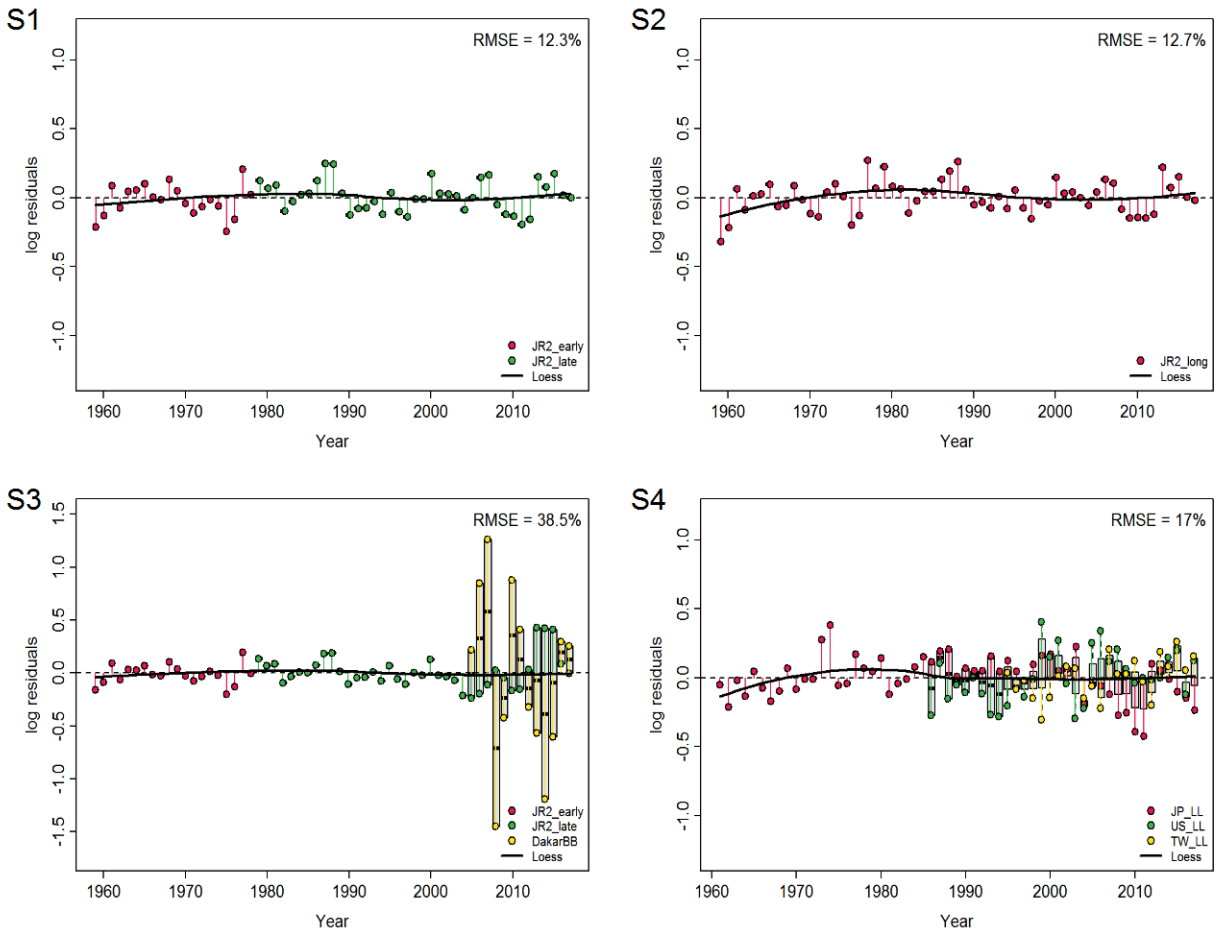

Figure 7. JABBA residual diagnostic plots for alternative sets of CPUE indices examined for Scenarios S1-S4. Boxplots indicate the median and quantiles of all residuals available for any given year, and solid black lines indicate a loess smoother through all residuals.

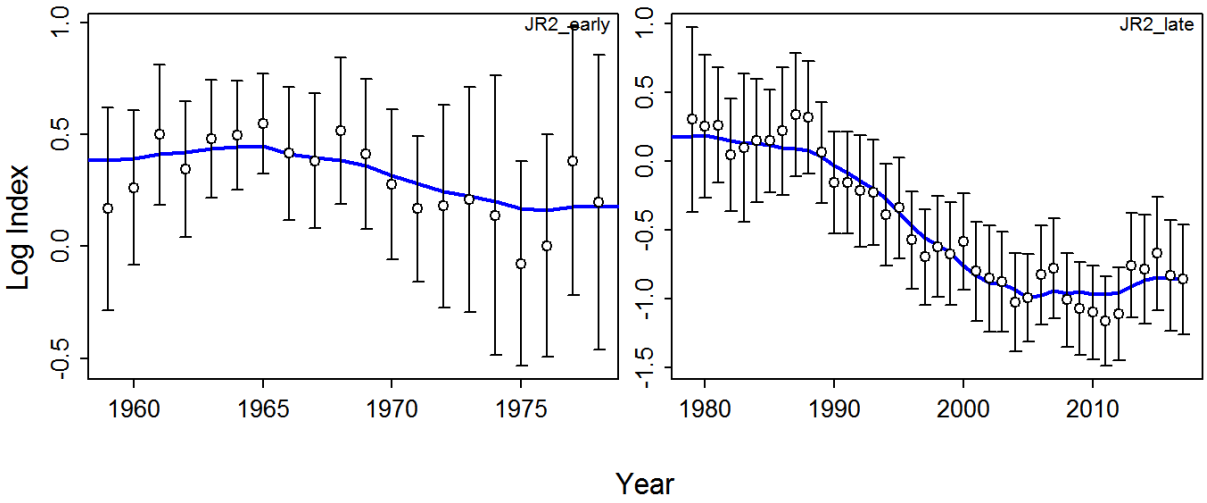

Figure 8. JABBA fits to the standardized JR2_early and JR2_late CPUE (in log scale) for the reference case scenario S1. The solid blue line is the model predicted value and the circles are observed data values. Vertical black lines represent the estimated 95% confidence intervals around the CPUE values.

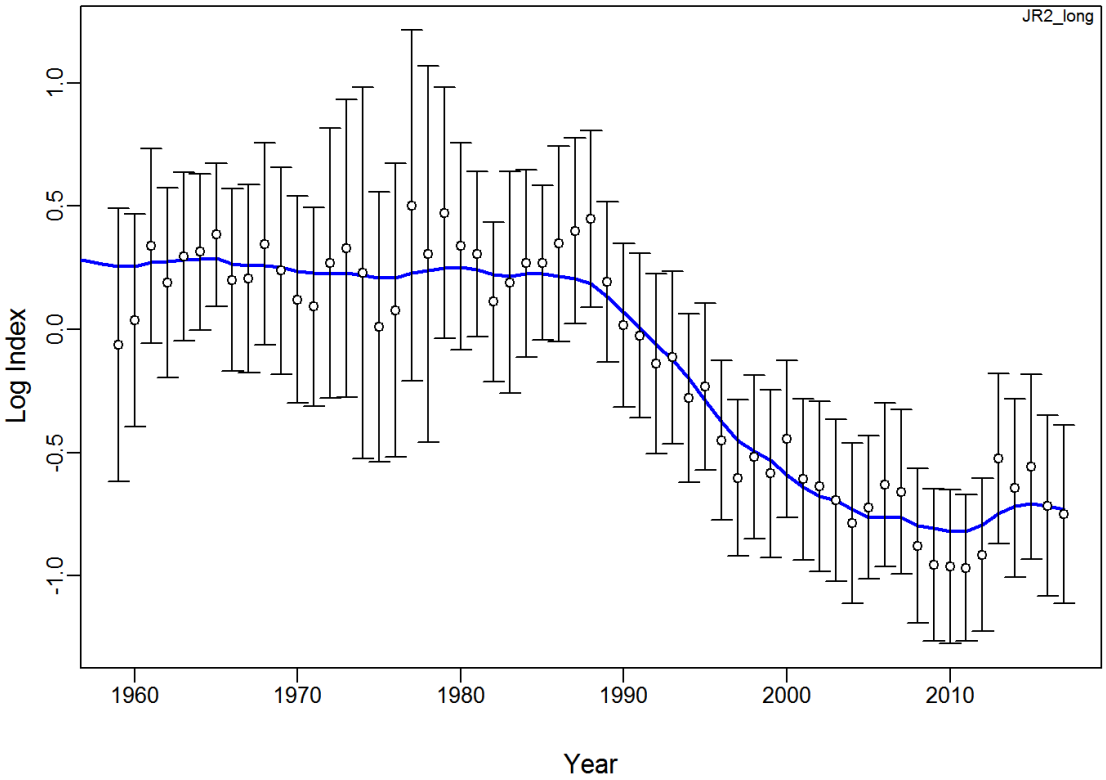

Figure 9. JABBA fits to the standardized JR2_long CPUE (in log scale) for scenario S2. The solid blue line is the model predicted value and the circles are observed data values. Vertical black lines represent the estimated 95% confidence intervals around the CPUE values.

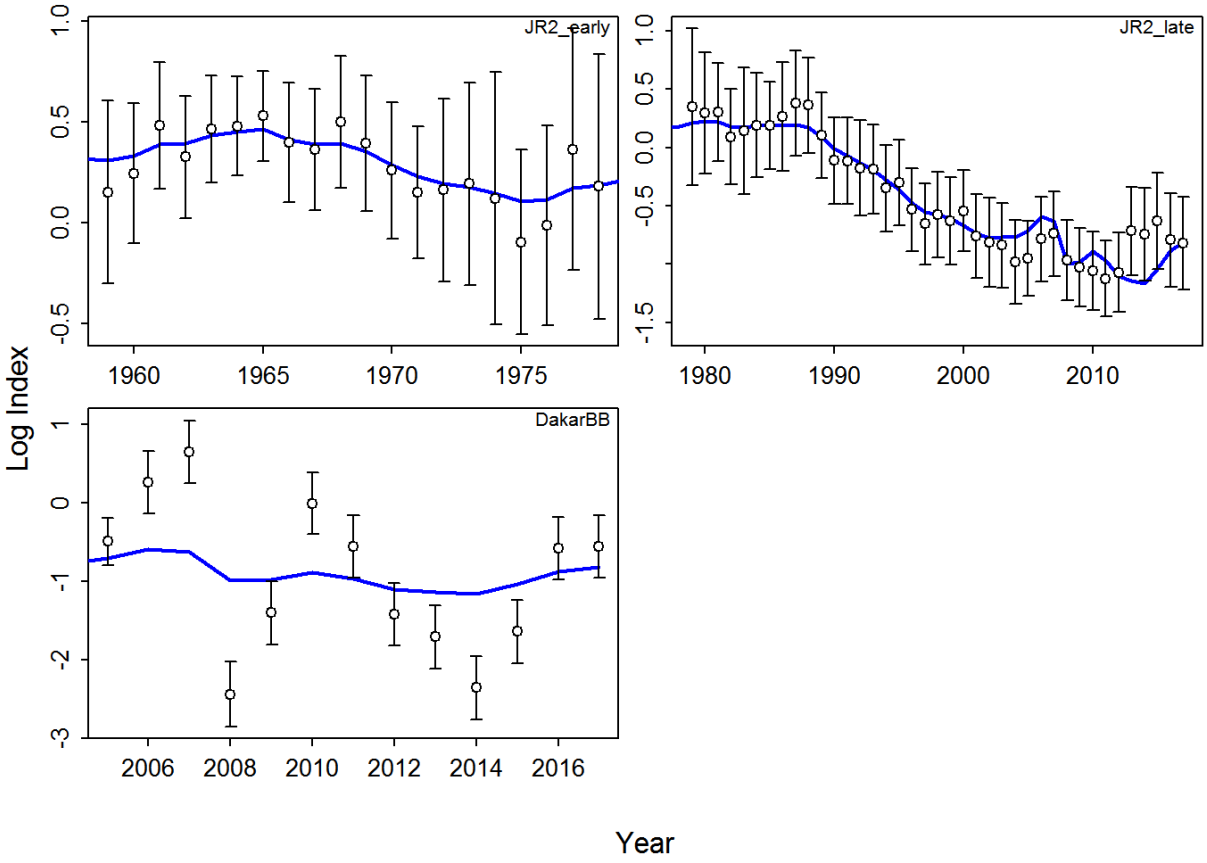

Figure 10. JABBA fits to the standardized JR2_early, JR2_later and DAK_BB CPUE (in log scale) for scenario S3. The solid blue line is the model predicted value and the circles are observed data values. Vertical black lines represent the estimated 95% confidence intervals around the CPUE values.

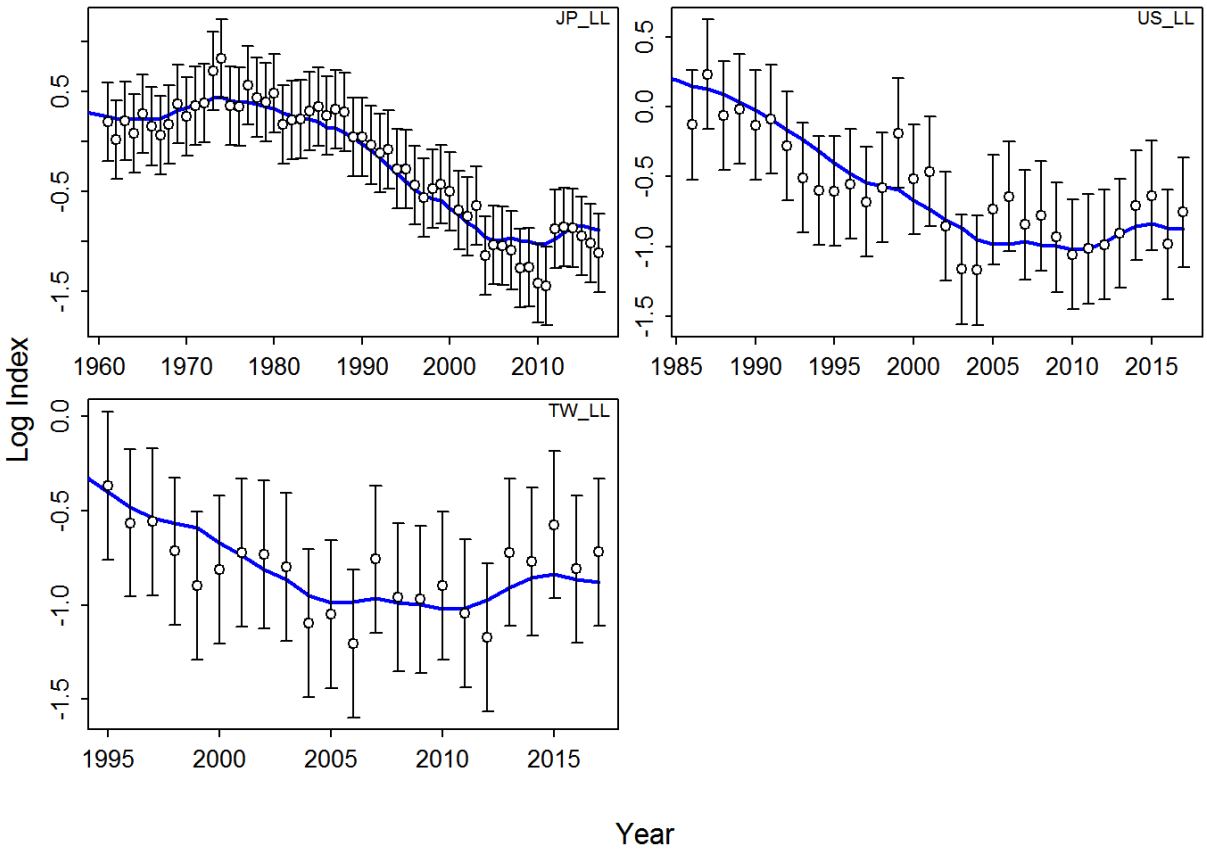

Figure 11. JABBA fits to the standardized JP_LL, TW_LL and US_LL (in log scale) for scenario S4. The solid blue line is the model predicted value and the circles are observed data values. Vertical black lines represent the estimated 95% confidence intervals around the CPUE values.

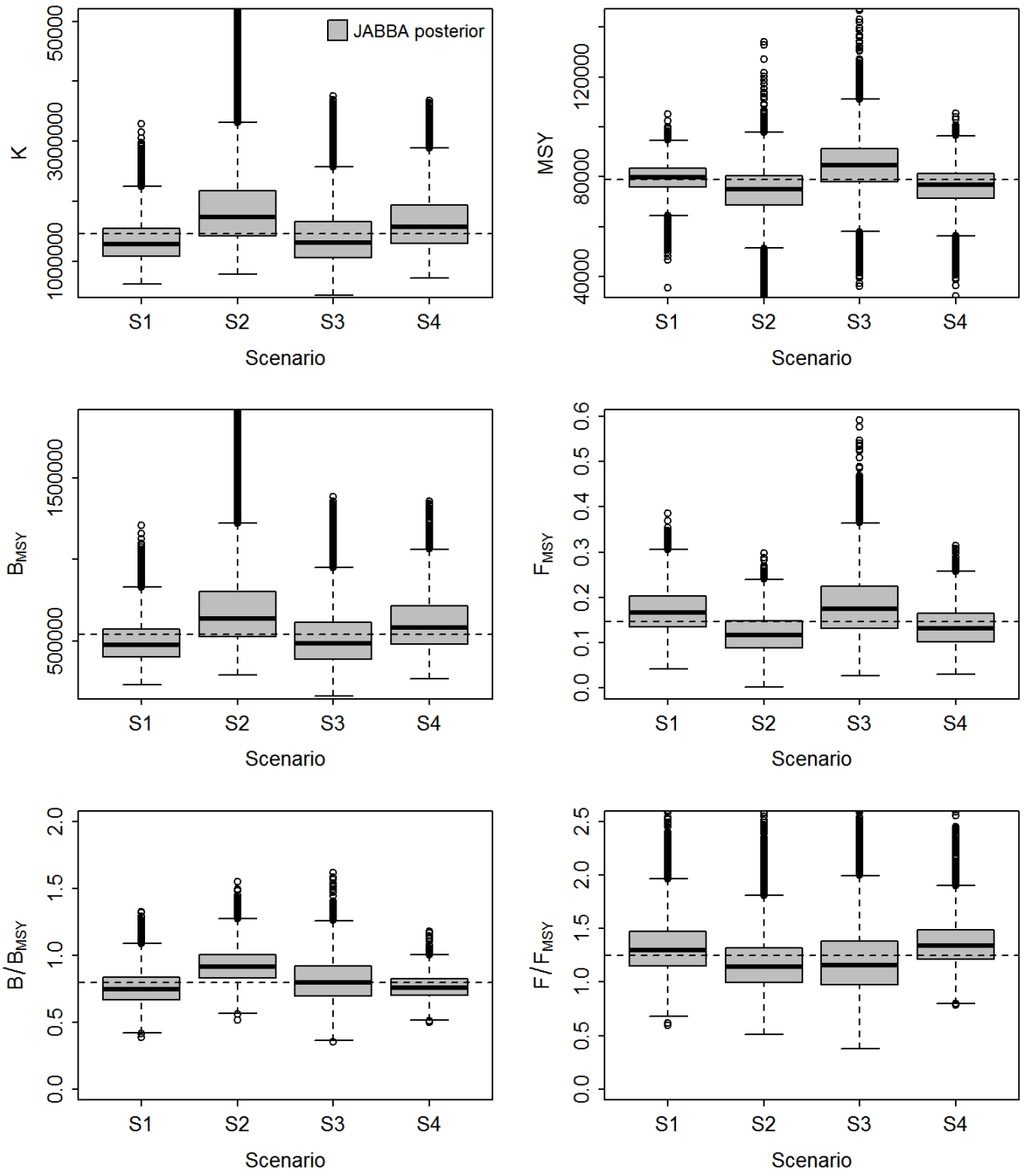

Figure 12. Boxplots summarizing the postirior distributions depicting the stock stutus for scenarios S1-S4, where B/B_{MSY} and F/F_{MSY} are presented for the final assessment year 2017. Dashed lines denote means across the for sceanrios.

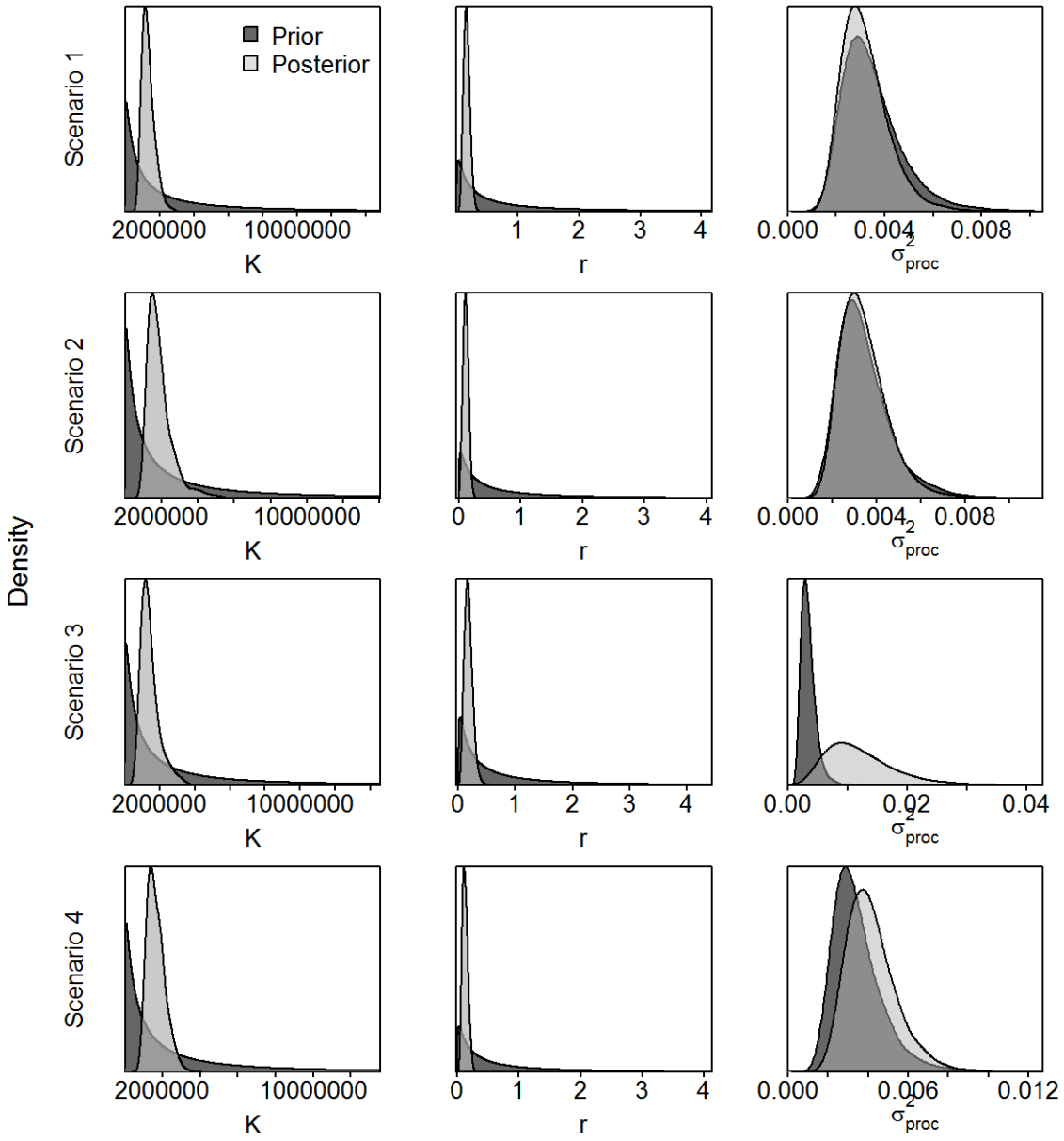

Figure 13. Prior and posterior distributions of K, r and process error variance σ^2_{proc} for scenarios S1-S4. Posteriors distributions are plotted using generic kernel densities.

.

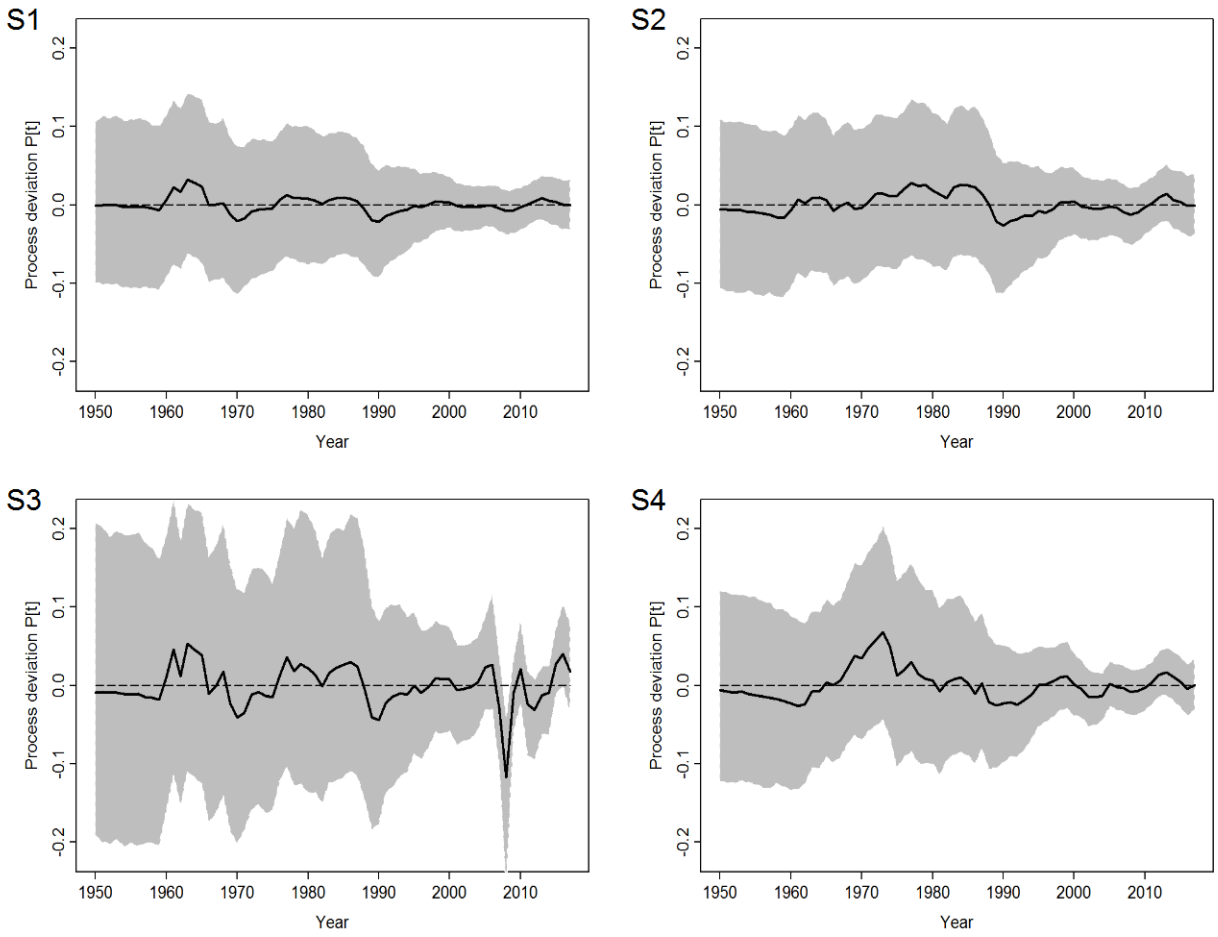

Figure 14. Process error deviation (η_y) plots, shown for scenarios S1=S4. Solid black lines denote the median process error η_y with associated 95% CIs illustrated by great shaded areas.

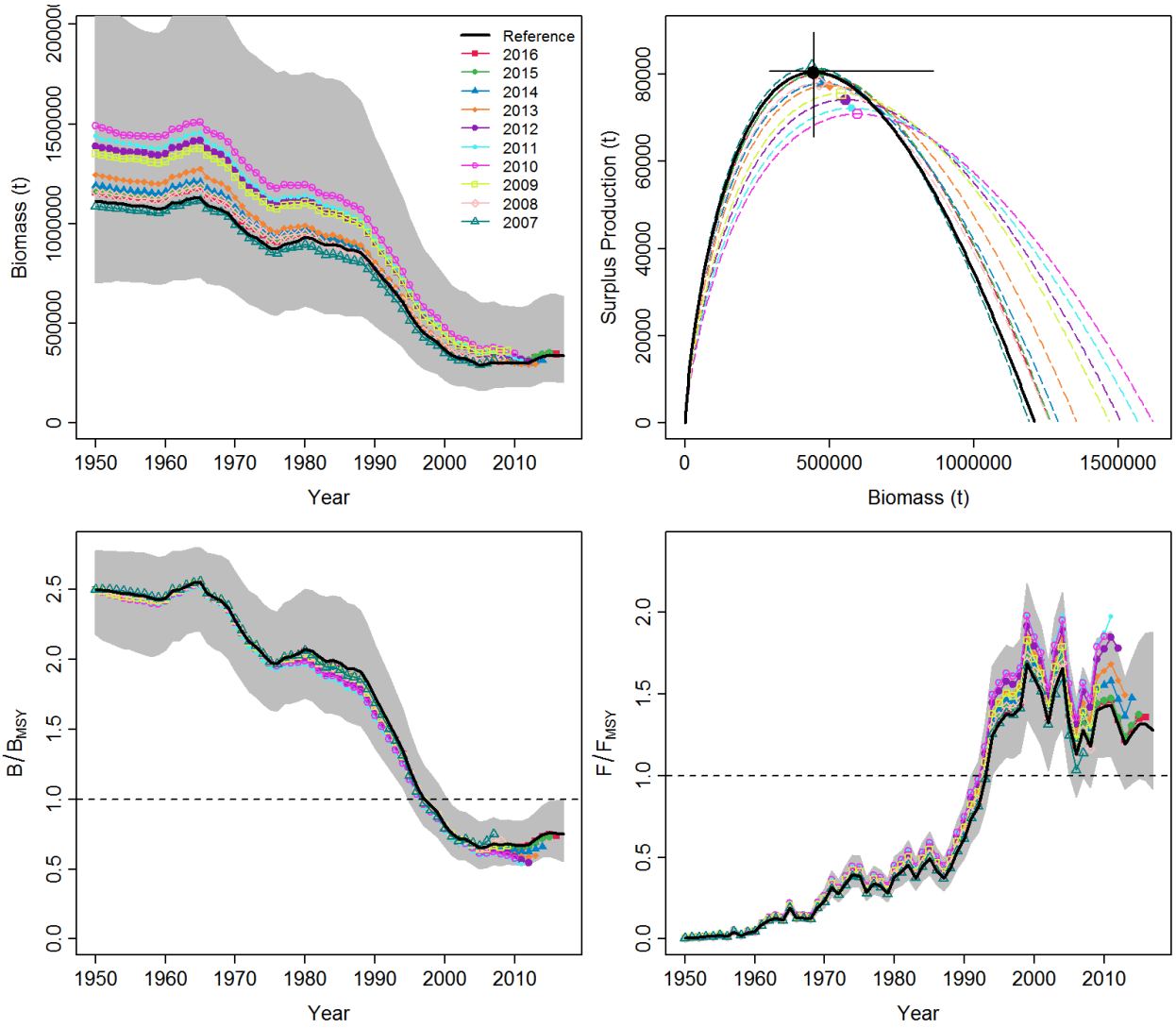

Figure 15. Retrospective analysis for stock biomass (t), surplus production function (maximum = MSY), B/B_{MSY} and F/F_{MSY} for the Atlantic bigeye tuna JABBA base case scenario S1. The label "Reference" indicates the reference case model fits and associated 95% CIs to the entire time series 1950-2017. The numeric year label indicates the retrospective results from the retrospective 'peel', sequentially excluding CPUE data back to 2007. Grey shaded areas denote the 95% CIs, which are indicated by crosshair for B_{MSY} and MSY defining the maximum of the surplus production curve.

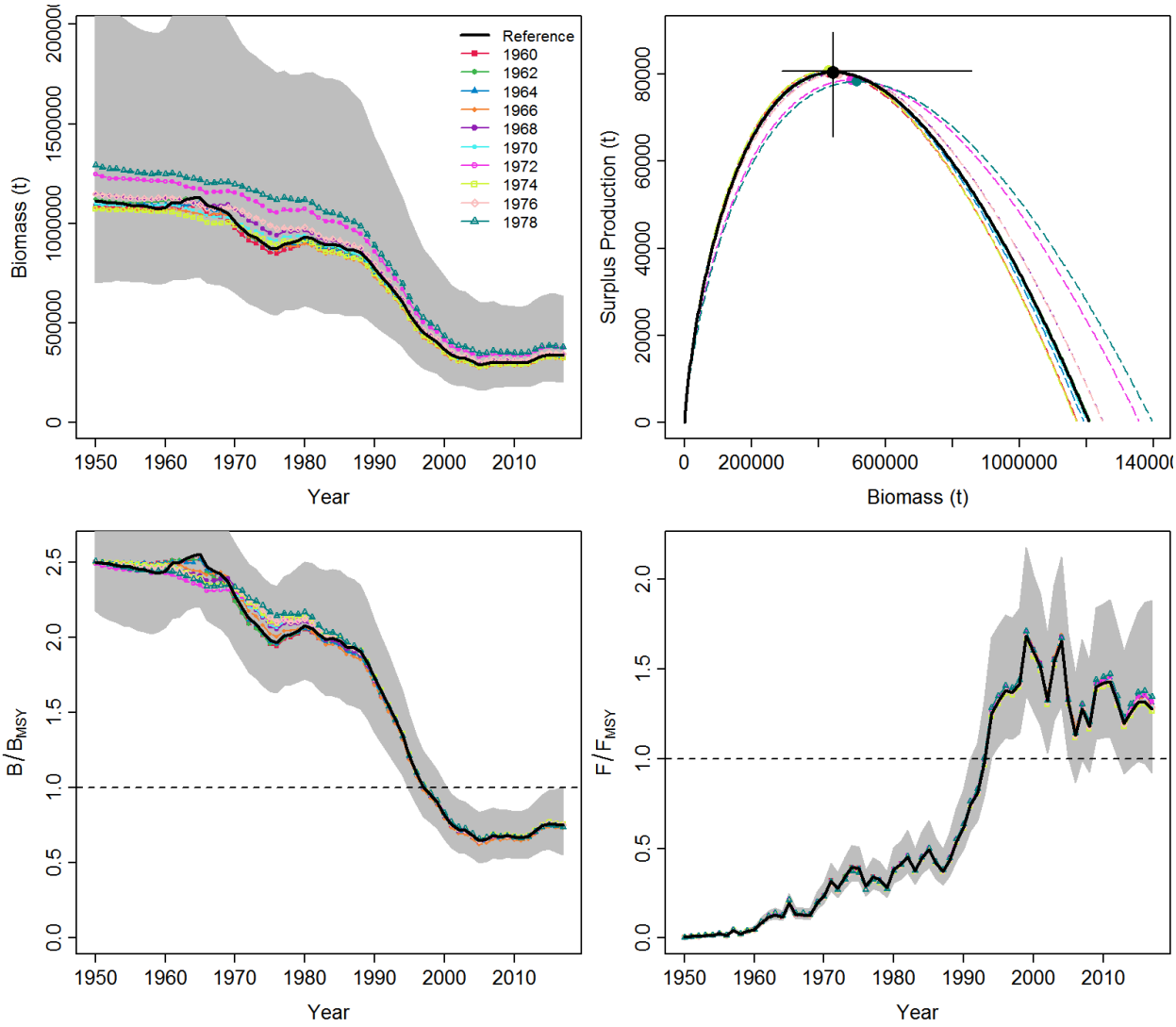

Figure 16. Progressive pattern analysis for stock biomass (t), surplus production function (maximum = MSY), B/B_{MSY} and F/F_{MSY} for the Atlantic bigeye tuna JABBA base case scenario S1. The label "Reference" indicates the reference case model fitted to the entire time series 1950-2017. The numeric year label indicates the retrospective results from the retrospective 'peel', sequentially excluding CPUE data back to 2007. Grey shaded areas denote the 95% CIs, which are indicated by crosshair for B_{MSY} and MSY defining the maximum of the surplus production curve.

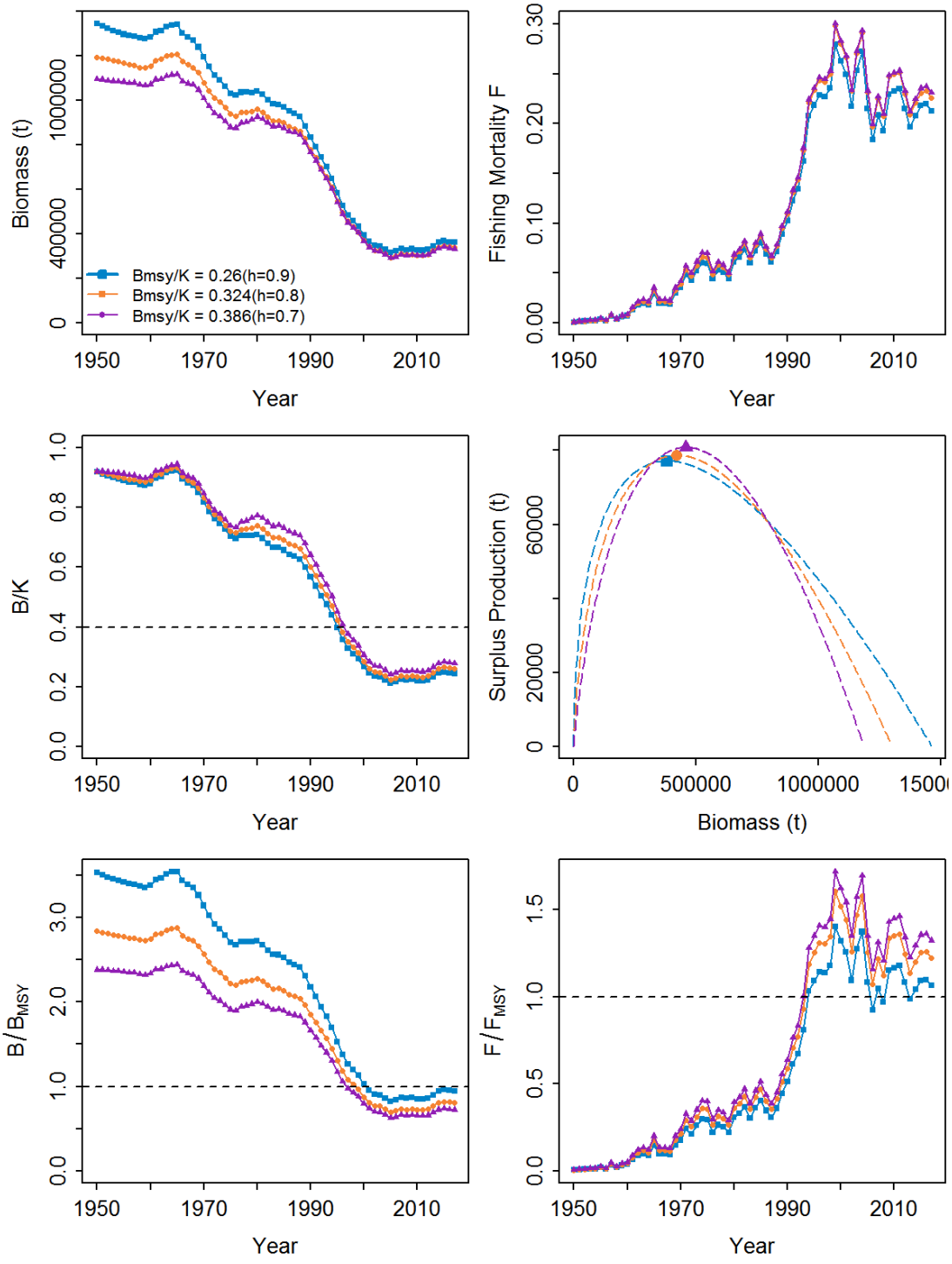

Figure 17. JABBA reference case model runs for three alternative B_{MSY}/K input values showing predicted biomass (t), fishing mortality (F), biomass to unfished biomass (B/K), surplus production function (maximum = MSY), B/B_{MSY} and F/F_{MSY} for the Atlantic bigeye tuna JABBA base case scenario S1. The value of B_{MSY}/K= 0.324 corresponds a steepness of h = 0.8 assumed for the stock synthesis (s33) reference case 2015 and 2018.

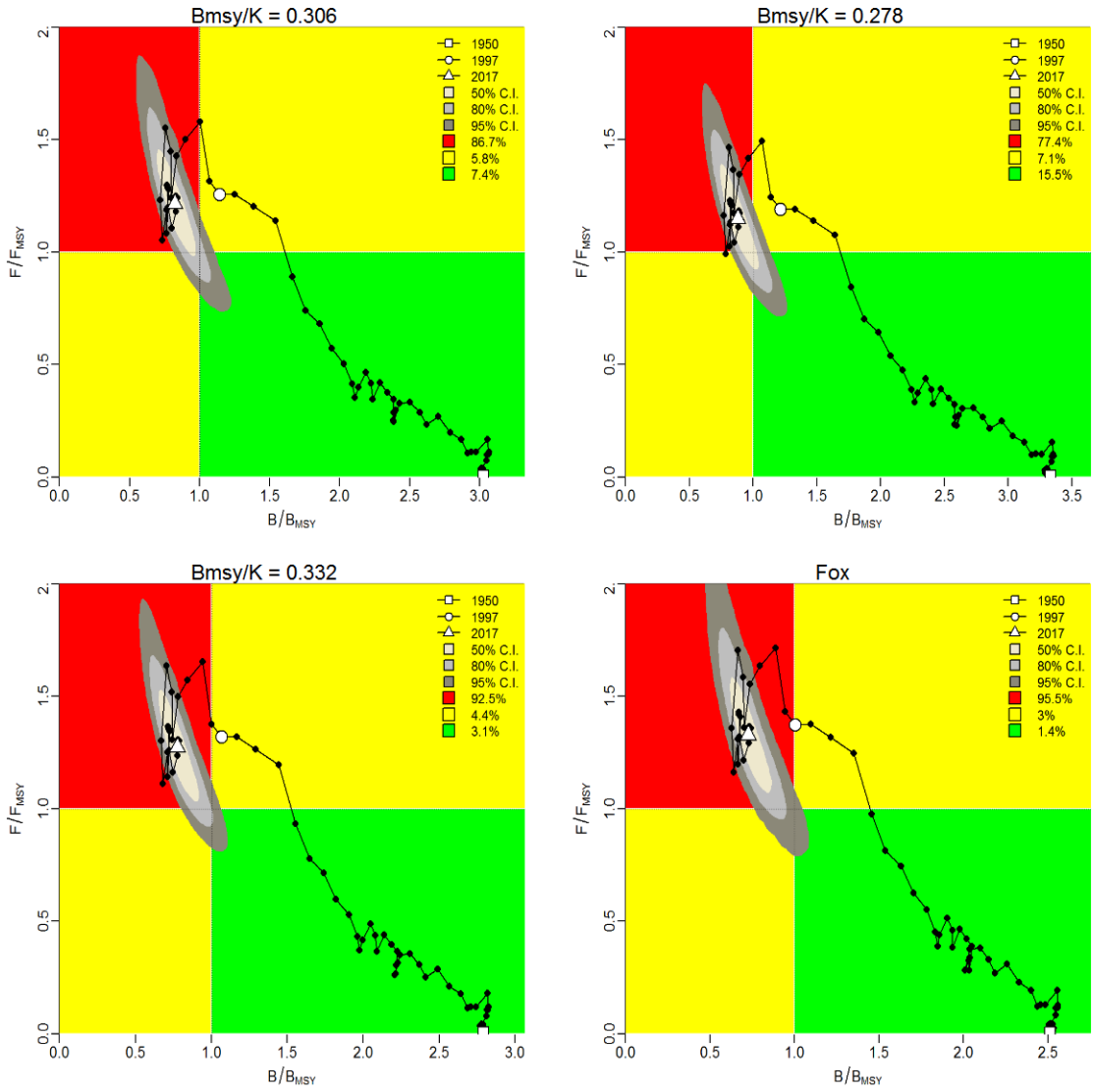

Figure 18. Kobe phase plot showing estimated trajectories (1950-2017) of B/B_{MSY} and F/F_{MSY} for the three alternative B_{MSY}/K input values. The value of B_{MSY}/K= 0.306 corresponds a steepness of h = 0.8 assumed for the stock synthesis (s33) reference case 2015 and 2018. Different grey shaded areas denote the 50%, 80%, and 95% credibility interval for the terminal assessment year. The probability of terminal year points falling within each quadrant is indicated in the figure legend.

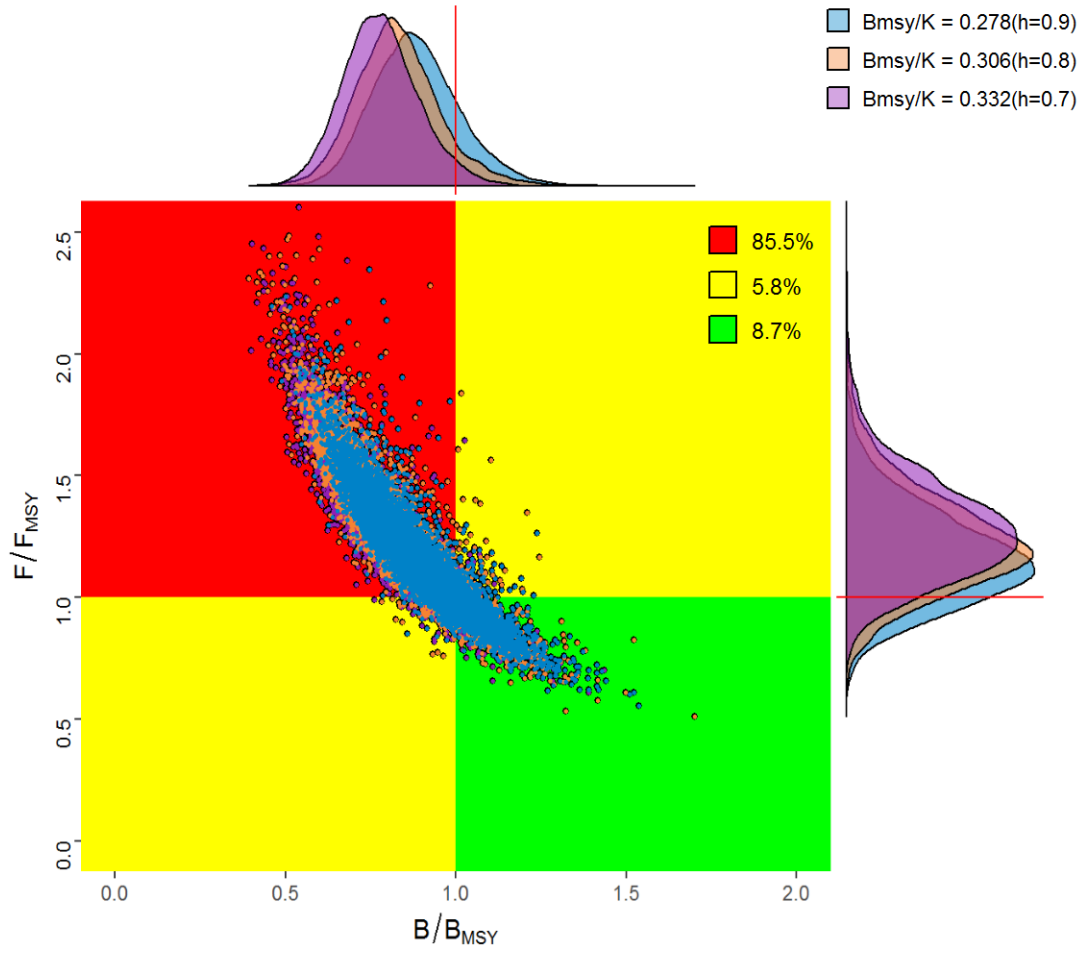

Figure 19. Kobe phase plot showing the combined posteriors of B/B_{MSY} and F/F_{MSY} (1950-2017) from reference case runs for the three alternative B_{MSY}/K input values using the 'Kobe' library in FLR (Kell *et al*., 2007b). The value of B_{MSY}/K= 0.306 corresponds a steepness of h = 0.8 assumed for the stock synthesis (s33) reference case 2018. The probability of terminal year points falling within each quadrant is indicated in the figure legend.

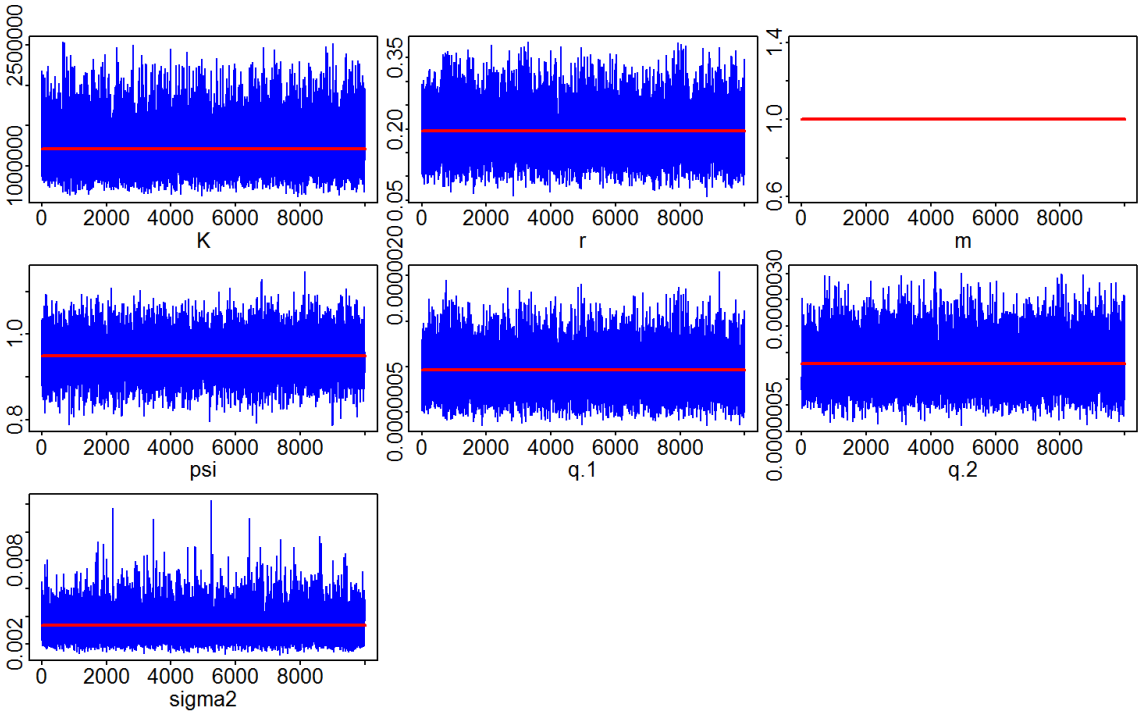

Figure A1. Trace plots for the model (scenario S1) parameter drawn from MCMC samples in the Bayesian state-surplus production model for the Atlantic bigeye tuna.

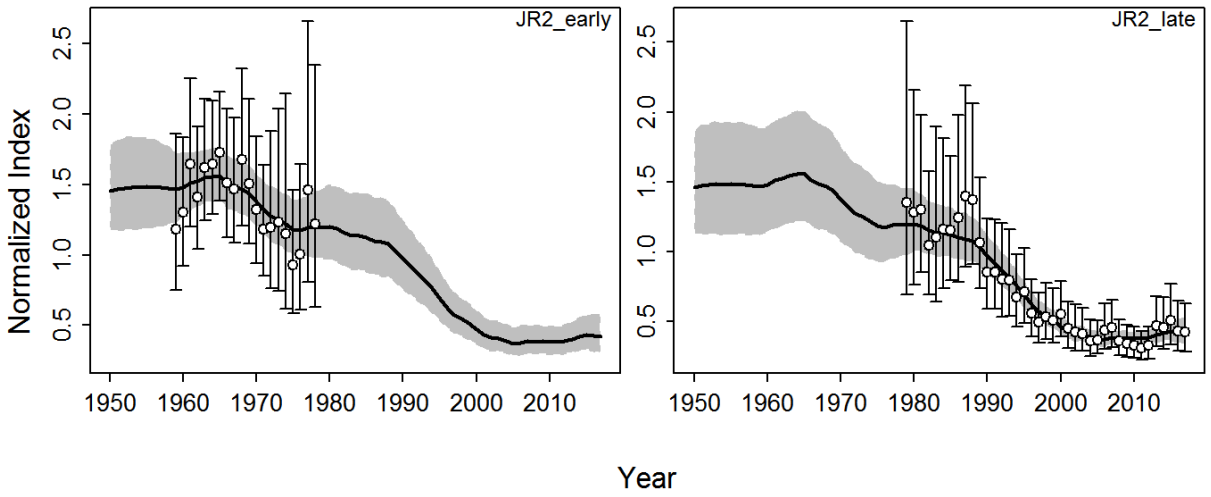

Figure A2. Time-series of observed (circle and SE error bars) and predicted (solid line) CPUE of Atlantic bigeye tuna for scenario S1 (JR2 early + late). Shaded grey area indicates 95% C.I.

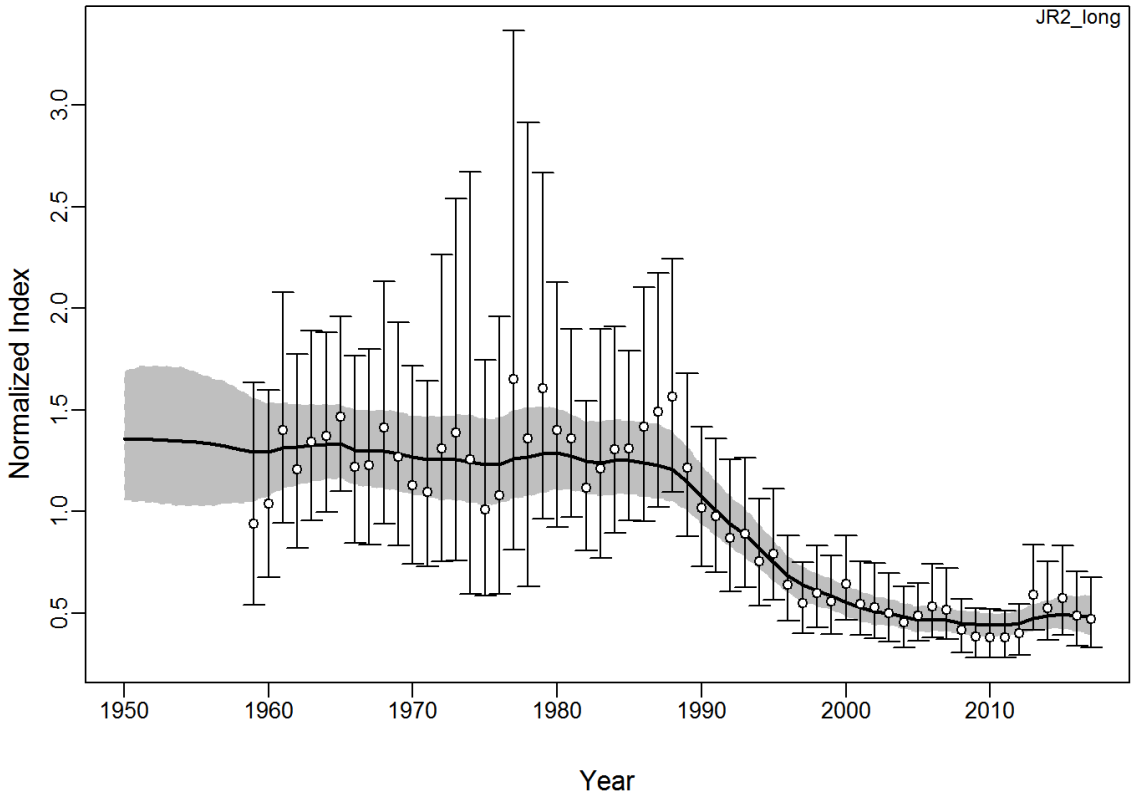

Figure A3. Time-series of observed (circle and SE error bars) and predicted (solid line) CPUE of Atlantic bigeye tuna for scenario S2 (JR2 long). Shaded grey area indicates 95% C.I.

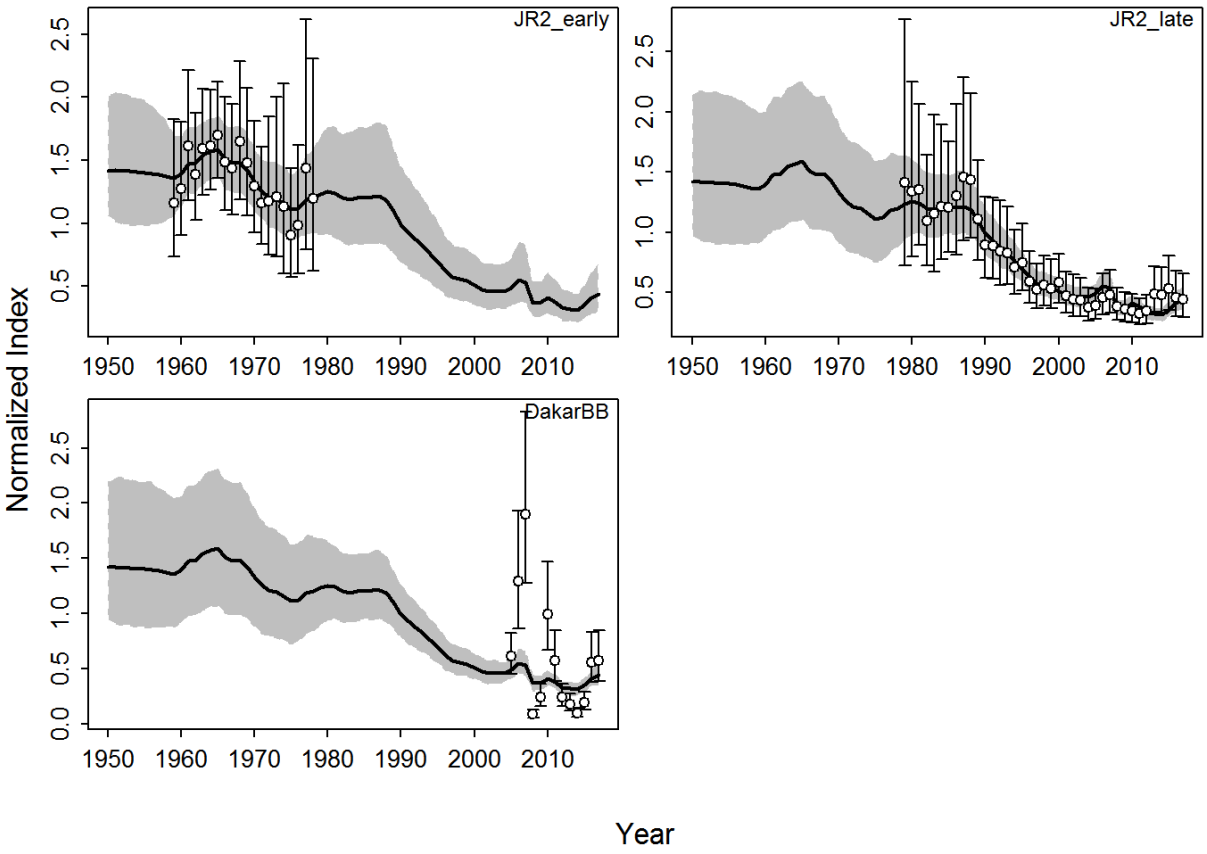

Figure A4. Time-series of observed (circle and SE error bars) and predicted (solid line) CPUE of Atlantic bigeye tuna for scenario S3 (JR2 early + late, DAK_BB). Shaded grey area indicates 95% C.I.

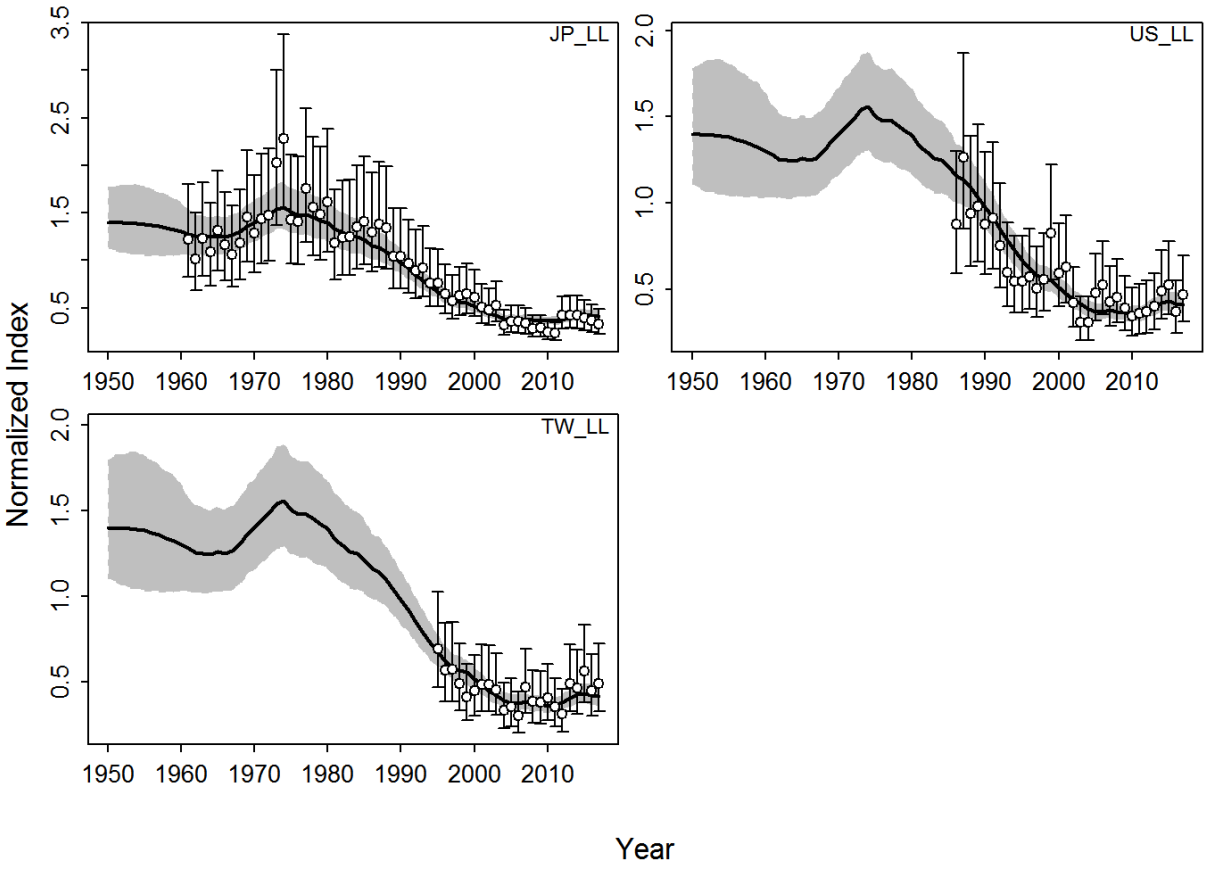

Figure A5. Time-series of observed (circle and SE error bars) and predicted (solid line) CPUE of Atlantic bigeye tuna for scenario S4 (JP_LL, US_LL, TW_LL). Shaded grey area indicates 95% C.I.

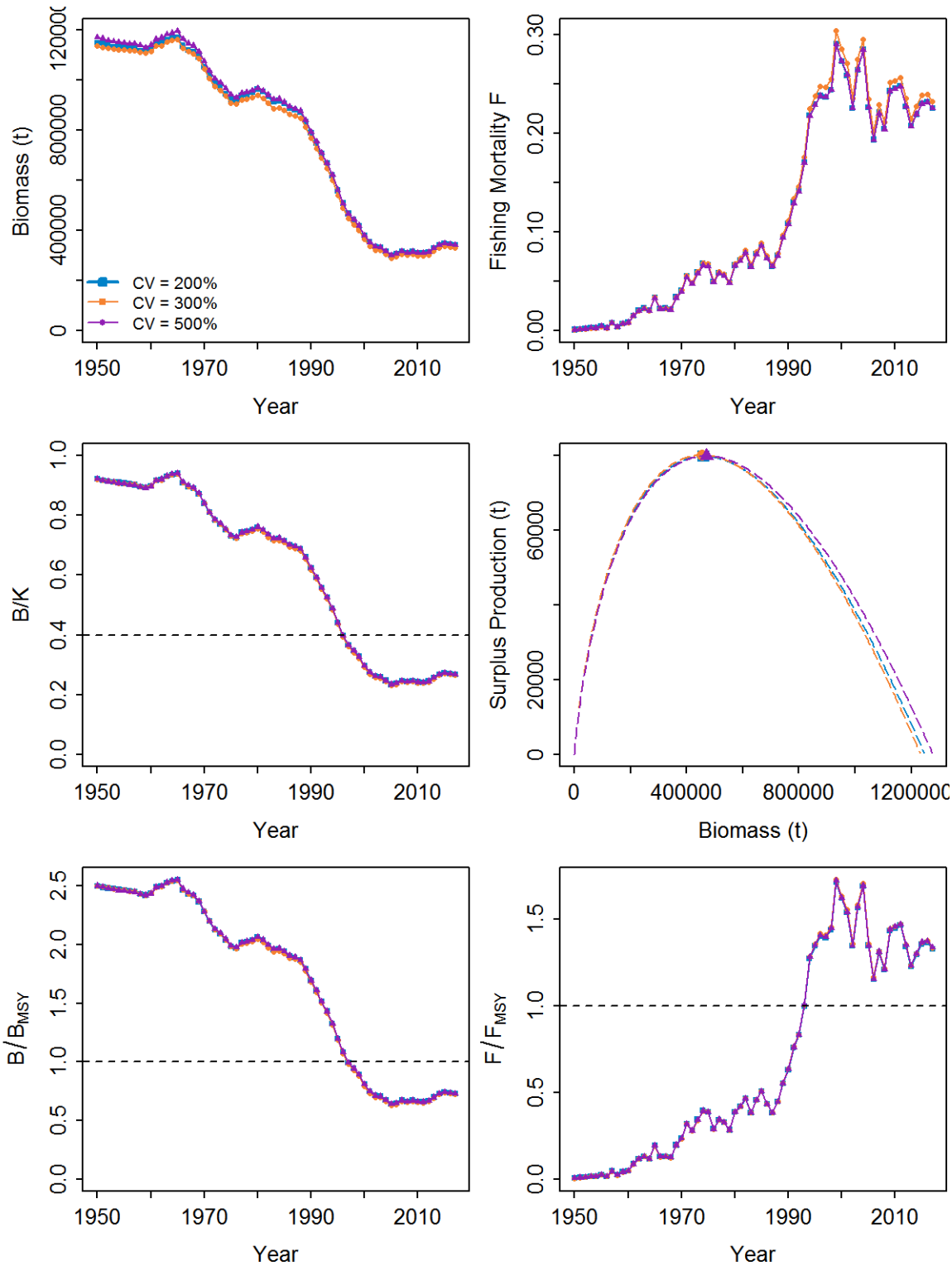

Figure A6. Sensitivity analysis illustrating the effect of simultaneously increasing the prior CVs for *r* and *K* from 200% to 500% on the predicted biomass (t), fishing mortality (*F*), biomass to unfished biomass (*B*/*K*), surplus production function (maximum = MSY), B/B_{MSY} and F/F_{MSY} for Atlantic bigeye tuna.

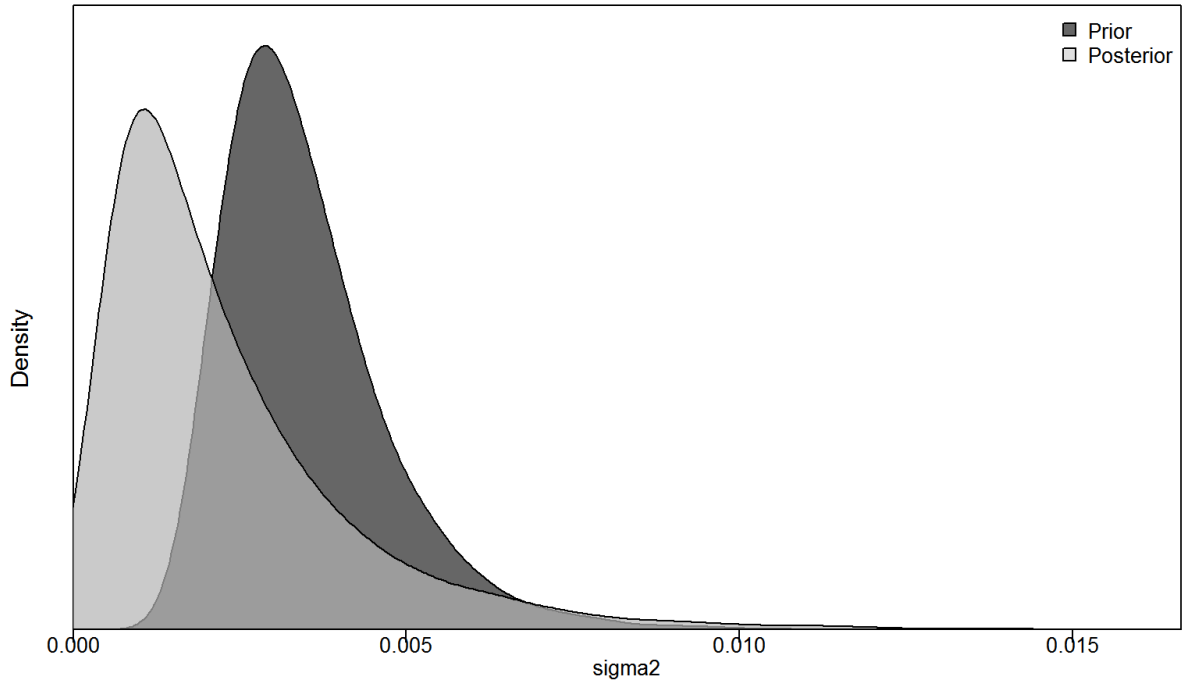

Figure A7. Comparison of the "freely" estimated process variance posterior, using an uninformative prior with σ^2_{proc}~1/gamma(0.001,0.001) and the initially developed informative prior based on age-structured simulations, with σ^2_{proc}~1/gamma(9.606,0.03).