BAYESIAN SURPLUS PRODUCTION MODELS (JABBA) APPLIED TO THE WESTERN ATLANTIC SKIPJACK TUNA STOCK ASSESSMENT

Rodrigo Sant'Ana
Universidade do Vale do Itajaí (Univali)

Eidi Kikuchi Santos
Universidade Federal do Rio Grande (FURG)

Bruno Mourato
Universidade Federal de São Paulo

Ai Kimoto
International Commission for the Conservation of Atlantic Tunas

SCRS/2022/099 Collect. Vol. Sci. Pap. ICCAT, 79(1): 649-674 (2022)

BAYESIAN SURPLUS PRODUCTION MODELS (JABBA) APPLIED TO THE WESTERN ATLANTIC SKIPJACK TUNA STOCK ASSESSMENT

Sant'Ana R.[1], Kikuchi E.[2], Mourato B.L. [3], Kimoto A.[4], Ortiz M.[4], and Cardoso L.G.[2]

SUMMARY

Bayesian State-Space Surplus Production Models were fitted to Western Atlantic skipjack tuna catch and CPUE data using the 'JABBA' R package. The ten scenarios were based on the previous assessment and on uncertainty grid proposed during the 2022 SKJ Data Preparatory Meeting, which in summary corresponded to nine runs based on variations in growth parameters and steepness. To implement these scenarios in a Bayesian surplus production model, a Pella-Tomlinson production function was used and priors for r and $B_{MSY}/B0$ was derived using the concept called Age-Structured Equilibrium Model (ASEM). All scenarios showed similar trends for the trajectories of B/B_{MSY} and F/F_{MSY} over time.

RÉSUMÉ

Les modèles de production excédentaire état-espace de type bayésien ont été ajustés aux données de capture et de CPUE du listao de l'Atlantique Ouest au moyen du progiciel JABBA R. Les dix scénarios étaient basés sur l'évaluation précédente et sur la grille d'incertitude proposée lors de la réunion de préparation des données sur le listao de 2022, ce qui correspondait en résumé à neuf scénarios basés sur des variations des paramètres de croissance et de la steepness. Pour mettre en œuvre ces scénarios dans un modèle bayésien de production excédentaire, une fonction de production Pella-Tomlinson a été utilisée et des priors pour r et BPME/B0 ont été dérivés en utilisant le concept appelé modèle structuré par l'âge en conditions d'équilibre (ASEM). Tous les scénarios ont montré une tendance similaire pour les trajectoires de B/BPME et F/FPME au fil du temps.

RESUMEN

Los modelos de producción excedente bayesianos de estado espacio se ajustaron a los datos de captura y CPUE de listado del Atlántico occidental utilizando el paquete R de 'JABBA'. Los diez escenarios se basaron en la evaluación anterior y en la matriz de incertidumbre propuesta durante la reunión de preparación de datos de listado de 2022, que en resumen correspondían a nueve ensayos basados en variaciones de los parámetros de crecimiento y la inclinación. Para implementar estos escenarios en un modelo de producción excedente bayesiano, se utilizó una función de producción de Pella-Tomlinson y se derivaron las distribuciones previas para r y BRMS/B0 utilizando el concepto denominado Modelo de equilibrio estructurado por edad (ASEM). Todos los escenarios mostraron una tendencia similar para las trayectorias de B/BRMS y F/FRMS a lo largo del tiempo.

KEYWORDS

Skipjack, stock status, CPUE fits, hindcast, life history priors

[1] Universidade do Vale do Itajaí, Escola do Mar, Ciência e Tecnologia, Laboratório de Estudos Marinhos Aplicados. Rua Uruguai, 458, Itajaí, SC, Brazil.

[2]Universidade Federal de Rio Grande, Laboratório de Recursos Pesqueiros Demersais, Fundação Universidade Federal do Rio Grande (FURG), Av. Itália km 8, Campus Carreiros, Rio Grande, Brazil.

[3]Universidade Federal de São Paulo, Instituto do Mar, Rua Carvalho de Mendonça, 144, Encruzilhada - Santos/SP, 11070-100

[4]ICCAT Secretariat. Calle Corazón de Maria 8, Madrid Spain 28002.

1. Introduction

The skipjack tuna (*Katsuwonus pelamis*) is widely distributed in the tropical and subtropical waters of the Atlantic, Indian and Pacific Oceans (ICCAT, 2006). The species has habitat preferences for an epipelagic realm, generally inhabiting open waters with optimum temperature range varying between 20 ºC and 30 ºC (ICCAT, 2006). As a function of its wide distribution, skipjack tuna has been intensively exploited by various fisheries around the world (ICCAT, 2006). For management purposes, the International Commission for the Conservation of Atlantic Tunas (ICCAT) considers two distinct stock units in the Atlantic Ocean, East and West stocks (ICCAT, 2006). In this study, we will foccus on the West stock of the skipjack only. This stock is mainly exploited by baitboat fleet, 83% of the total catch in average for the all-time series. The purse seiners had subsequently contributed to 12.4% on average, and the other fleets had contributed with less than 5%.

The last Western Atlantic skipjack tuna stock assessment was carried out in 2014 (ICCAT, 2014) and included outputs from four distinct models, (a) Mean length-based mortality estimator; (b) Catch-only model; (c) Bayesian Surplus Production model (BSP), and; (d) Stock Production Model Incorporating Covariates model (ASPIC). All detailed descriptions and concerns about the results of each model can be observed in the Report of the 2014 ICCAT East and West Atlantic Skipjack stock assessment meeting (ICCAT, 2014). The final summary of the models indicated that the West Atlantic skipjack tuna stock was not overfished and not experiencing overfishing. Reference points resulted from those models were indicates an MSY around 32,000 metric tons, F_{2013}/F_{MSY} close to 0.7 and B_{2013}/B_{MSY} probably close to 1.3 (ICCAT, 2014).

Here, we present the 2022 preliminary stock assessment results for West Atlantic skipjack tuna stock based on the Bayesian State-Space Surplus Production Model framework, JABBA (Just Another Bayesian Biomass Assessment; https://github.com/jabbamodel/JABBA; Winker *et al.*, 2018). The JABBA model is a fully documented, open-source R package (https://github.com/JABBAmodel) that has been formally included in the ICCAT stock catalogue (https://github.com/ICCAT/software/wiki/2.8-JABBA) and has been widely applied in a number of recent ICCAT stock assessments, including: South Atlantic blue shark (ICCAT, 2016b), Mediterranean albacore (ICCAT, 2017c), South Atlantic swordfish (ICCAT, 2017a; Winker *et al*., 2018), Atlantic shortfin mako shark stocks (south and north) (ICCAT, 2017d; Winker *et al.*, 2017, 2019a), Atlantic blue marlin (Mourato *et al.*, 2019), Atlantic bigeye tuna (Winker *et al.*, 2019b), Atlantic white marlin (Mourato *et al.*, 2020), Atlantic yellowfin tuna (Sant'Ana *et al.*, 2020), Mediterranean swordfish (Winker *et al.* 2020; ICCAT, 2017b) and South Atlantic albacore (Winker *et al.*, 2020c).

This preliminary assessment of the West Atlantic skipjack tuna stock is guided by the SCRS work plan. A grid scenario was built based on the discussions and recommendations that were raised during the 2022 Skipjack Data Preparatory Meeting. In this way, extensive model diagnostics, retrospective pattern analysis and model prediction skillness were provided to evaluate the fitted models. In addition, this document explores the sensitivity of the base case scenario to the inclusion of alternative and additional standardized CPUE indices that have been made available for this assessment.

2. Material and Methods

2.1. JABBA inputs

This stock assessment is implemented using the Bayesian state-space surplus production model framework called JABBA (Winker *et al.*, 2018), which is now available as '*R package*' that can be installed from github.com/jabbamodel/JABBA. JABBA's inbuilt options include: (1) automatic fitting of multiple CPUE time series and associated standard errors; (2) estimating or fixing the process variance, (3) optional estimation of additional observation variance for individual or grouped CPUE time series, and (4) specifying a Fox, Schaefer or Pella-Tomlinson production function by setting the inflection point B_{MSY}/K and converting this ratio into a shape parameter *m*, (5) extensive diagnostic procedures and associated plots (e.g. residual run tests) and (6) a routine to conduct retrospective analysis. A full JABBA model description, including formulation and state-space implementation, prior specification options and diagnostic tools is available in Winker *et al.* (2018).

2.2. *Fishery data*

The ICCAT Secretariat provided fishery catch data for West Atlantic skipjack tuna from 1952 to 2020 (Figure 1). Relative abundance indices were made available, principally, in the form of joint standardized CPUE time series. These indices cover various periods and represent the distinct fishing gears and fleets that operate over the W-SKJ stock. A summary of the available indices is described below:

- BRA BB Past (1981 – 1999) used in the 2014 assessment.
- BRA BB Present (2000 - 2020);

- BRA HL (2010 – 2016);
- USA LL (1993 – 2020);
- VEN PS (1987 – 2020).

The CV's for all indices were scaled to an 0.25 average.

2.3. Model specifications

The model specifications were based on uncertainty grid defined in the Skipjack Data Preparation Meeting that resulted in nine distinct scenarios. These scenarios incorporate three variations in growth parameters as provided in Anon (2022) and three variations of steepness (0.7, 0.8, and 0.9). All models were implemented using a Pella and Tomlinson production function (**Table 1**).

The priors of K was kept uninformative similar to those used in the last assessment of the species. For K, a lognormal distribution was implemented using JABBA "range" option. Lower and upper values ranged from 50,000 t to 200,000 t, which resulted in an approximated mean value of 106,190 t and a CV of 36%. For r, were developed priors distribution with an associated shape parameter of a Pella-Tomlinson production function from an Age-Structured Equilibrium Model (ASEM) approach with Monte-Carlo simulations (Winker *et al.*, 2019b). The stock parameters used as inputs for the ASEM models included the uncertainty grid configuration citet before and presented in **Table 1**.

For all scenarios, the same initial depletion prior ($\varphi = B_{1952}/K$) was defined by a beta distribution with mean = 0.93 and CV of 5%. All catchability parameters were formulated as uninformative uniform priors. Even as, the process error of $\log(B_y)$ in year y for all scenarios were defined by an inverse-gamma distribution with shape parameter equal to 0.01 and rate parameter equal to 0.01.

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 R package. In this study, three 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 (1992), Geweke (1992), and Gelman and Rubin (1992) diagnostics as implemented in the coda package (Plummer *et al.*, 2006).

2.4. Model diagnostics and sensitivity runs

To evaluate CPUE fits, the model predicted CPUE indices were compared to the observed CPUE. JABBA-residual plots were used to examine (1) colour-coded lognormal residuals of observed versus predicted CPUE indices for all fleet together with (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), and (3) a loess smoother through all residuals aids to detect the presence systematic residual patterns. In addition, it depicts the root-mean-squared-error (RMSE) as a goodness-of-fit statistic. We conducted a runs test to quantitatively evaluate the randomness of residuals (Carvalho *et al.*, 2017). The runs test diagnostic was applied to residuals of the CPUE fit on log-scale using the function runs.test in the R package tseries, considering the 2-sided p-value of the Wald-Wolfowitz runs test. The runs test results can be visualized within JABBA using a specifically designed plot function that illustrates which time series passed or failed the runs test and highlights individual data points that fall outside the three-sigma limits (e.g. Anhøj and Olesen, 2014).

To check for systematic bias in the stock status estimates, we also performed a retrospective analysis for central reference scenario (S05: ASEM h = 0.8 Pella m), by sequentially removing one year of data at a time over a period of eight years (n = 8), refitting the model after each data removal and comparing quantities of interest (*i.e.* biomass, fishing mortality, B/B_{MSY}, F/F_{MSY}, B/B_0 and MSY) to the reference model that is fitted to full data time series. To compare retrospective bias between the models, we computed Mohn's (1999) rho (ρ) statistic, specifically the commonly used formulation defined by Hurtado-Ferro *et al.* (2014).

Although the above model diagnostics are important to evaluate the goodness of fit to the data and the consistency of benchmarking retrospectively, providing scientific advice should also involve checking that the model has prediction skill of future states under alternative management scenarios. To do this, the model-free hindcasting cross-validation (HCXval) technique by Kell *et al.* (2016) was applied, where observations are compared to their predicted future values. The HCXval algorithm has in common with retrospective analysis that requires the same two routine procedures of sequential removal the observations and re-fitting the model to the so truncated data series, but HCXval involves the additional steps of projecting ahead over the missing years and then cross-validating these forecasts against observations to assess the model's prediction skill. A robust statistic for evaluating prediction skill is the Mean Absolute Scaled Error (MASE) proposed by Hyndman and Koehler (2006), which scales the mean absolute error of prediction residuals to a naïve baseline prediction, where a 'prediction' is said to have 'skill' if it improves the model forecast when compared to the naïve baseline. A widely used baseline forecast for time series is the 'persistence algorithm' that takes the value at the previous time step to predict the expected outcome at the next time step as a naïve in-sample prediction, e.g., tomorrow's weather will be the same as today's. The MASE score scales the mean absolute error of the prediction residuals to the mean absolute error of a naïve in-sample prediction. A MASE score higher than one can then be interpreted such that the average model forecasts are no better than a random walk. Conversely, a MASE score of 0.5 indicates that the model forecasts twice as accurately as a naïve baseline prediction; thus, the model has prediction skill.

Finally, the analysis included sensitivity model runs based on forward stepwise inclusion of each index one-by-one in the model. Taking as prior indices the BRA BB Past and BRA BB Present as default indices in the small model. The general idea with this comparative analysis was to evaluate the possible effects of the inclusion of each index over estimated biomass dynamic of this stock.

3. Results and Discussion

The MCMC convergence tests by Heidelberger and Welch (1992), Geweke (1992), and Gelman and Rubin (1992) were passed by all estimable key parameters for all models. Adequate convergence of the MCMC chains was also corroborated by visual inspection of trace plots (results available on request), which showed good mixing in general (*i.e.*, moving around the parameter space).

The model fits to each of the five standardized CPUE's indices are shown in **Figure 2** for each of the nine uncertainty grid scenarios. For all scenarios, the behavior of the model's fits appeared to be led by the pattern observed in the VEN PS index more than the other indices. Some variations and mainly, the decrease trends observed in the recent years of the BRA BB Present index and USA LL index tend to be not well interpreted by the model. This behavior are common when the trends observed along all time-series have poor signal and that can be corroborated by the presence of long and relatively flat time-series index.

The results of the log-residuals runs tests for each CPUE and each scenario are shown in **Figure 3**. Green panels indicate CPUE indices that passed the runs test with no evidence of a non-random residual pattern ($p > 0.05$) and red panels indicating a failed runs test. In addition, the inner shaded area shows 3-sigma limits around the overall mean as proposed by Anhøj and Olesen (2014) and the red circles identify each specific year where the residuals are larger than the threshold limit. In all scenarios were observed a same pattern, with a failed behavior in the runs test diagnostic procedure for BRA BB Present, BRA HL and VEN PS indices. The other two indices (USA LL and BRA BB Past) have passed completely in this diagnostic for all scenarios (**Figure 3**). The goodness-of-fit were comparable among all scenarios, in general, the RMSE statistics were consistent ranging from 42.1% to 42.7% (**Figure 4**). This pattern shown some conflicting between indices, mainly at the beginning and final of the time-series. The annual process error deviation estimated for all scenarios shown a similar stochastic pattern with a constant average centered around the zero and 95% credibility intervals always covering the zero value (**Figure 5**), which suggest no evidence of structural model misspecifications.

The medians of the marginal posteriors for K ranged between 121,544 t (S03) and 208,597 t (S07) (**Table 3**). The values estimated for posterior to prior median (PPMR) and variance (PPVR) ratios estimated to K indicates that this parameter have been mixed informed by data and priors for almost all scenarios (S04 – S09). The scenarios S01 ~ S03 had been well informed by the priors. Additionaly, there was not observed expressive reductions in the precision of the posteriors in relation to the priors defined to this parameter. For the r, the medians of the marginal posteriors ranged between 0.443 (S07) and 1.054 (S03). The values of PPMR and PPVR estimated for r, in general, show that the priors used have contributed to define the behavior of the posteriors, but it is also possible to realize an influence of the data in the posteriors (**Figure 6**). The initial depletion ($\varphi = B_{1952}/K$) marginal posteriors for each scenario were also similar and largely informed by the priors distributions.

The range of *MSY* median estimates were narrow between all nine scenarios, reaching the lower value in the S07 scenario (32,716 metric tons) and the higher value in the S03 scenario (40,152 metric tons) (**Table 3**). Furthermore, the marginal posterior medians for B_{MSY} varied between 50,945 (S02) and 79,276 (S07) metric tons, and estimates of F_{MSY} showed a small variation between the nine scenarios with median values varying from 0.414 (S07) to 0.799 (S03) (**Table 3**).

In general, all scenarios showed similar trends for the trajectories of B/B_{MSY} and F/F_{MSY} over time (**Figure 7**; **Figure 8**). The trajectory of B/B_{MSY} showed a sharp decrease after the year 1980 and a subsequently stable trend from 1984 to 2020. This stability in the second period, between 1984 and 2020, possibly are linked to the hyper-stable pattern observed on the most longer index used in the model (VEN PS index). As commented before, this hyper-stable behaviour in most longer time-series index allied to general poor trends in the other indices used in this assessment could be an explanatory hypothesis for this stable trend. The F/F_{MSY} trajectory shows a sharply increasing trend at the same year that was observed a decrease in B/B_{MSY} trajectory, and after that soft decrease trend from 1984 onwards (**Figure 7**; **Figure 8**). This abrupt increase pattern observed after 1980 marks the beginning of the operations of Brazilian baitboat fleet over this stock. For all scenarios evaluated here, the models do not evidenced periods of overfishing ($F/F_{MSY} > 1$) or even the stock are being overfished ($B/B_{MSY} < 1$) (**Figure 7**; **Figure 8**). In general, the B/B_0 trajectory showed a similar trend for all nine scenarios, with a decrease marked at the beginning of 1980's and a subsequent stable period after that (**Figure 9**).

The results of an eight year retrospective analysis applied to scenario S05 are depicted in **Figure 10**, respectively. In general, the Base Case scenario (S05) shows a negligible retrospective pattern. The estimated Mohn's rho for all stock quantities fell within the acceptable range of -0.15 and 0.20 (Hurtado-Ferro *et al.*, 2014; Carvalho *et al.*, 2017) and these results confirm the absence of an undesirable retrospective pattern (**Table 4**). The hindcasting cross-validation results for all updated indices show predictions within limits of the 95% CRI's suggesting good prediction skills for S05 scenario (**Figure 11**). Except for the BRA HL last time point exclusion that shown a prediction outside of the 95% CRI's limits. However, the mean absolute scaled error (MASE) estimated were slightly above of the reference level (MASE > 1) for BRA BB Present, BRA HL and VEN PS indices, which indicates that the average model forecasts are not better than a naïve baseline prediction – like a random walk process (Carvalho *et al.*, 2021). Nonetheless, for the index with a flat trend with low variation at the end of the time series is expected that the MASE estimation will be close to reference level one.

The results of the sensitivity analysis based on forward stepwise indices in models (**Table 2**) are shown in **Figure 12**. These results show that when excluded VEN PS index the answer from the model tends to capture the signal passed from the other indices, building a more pessimistic scenario for the recent status of the West Atlantic skipjack stock.

The Kobe biplots for all scenarios were shown in **Figure 13**. All scenarios show optimistic status with probabilities of the stock being stable on green area (**Figure 13**). However, if considered the exclusion of the hyper-stable effect observed in VEN PS index, this status will probably show a small change as observed in the sensitivity analysis. Although, these results are preliminary and they were more explored during the skipjack stock assessment meeting and the final model was presented in the report of the meeting.

References

Anon, 2022. Data input and assessment models settings for the evaluation of East and West Atlantic skipjack tuna stocks. Collect. Vol. Sci. Pap. SCRS/2022/093.

Anhøj, J., Olesen, A.V., 2014. Run charts revisited: A simulation study of run chart rules for detection of non-random variation in health care processes. PLoS One 9, 1–13. https://doi.org/10.1371/journal.pone.0113825

Carvalho, F., Punt, A.E., Chang, Y.J., Maunder, M.N., Piner, K.R., 2017. Can diagnostic tests help identify model misspecification in integrated stock assessments? Fish. Res. 192, 28–40.

Carvalho, F., Winker, H., Courtney, D., Kapur, M., Kell, L., Cardinale, M., Schirripa, M., Kitakado, T., Yemane, D., Piner, K. R., Maunder, M. N., Taylor, I., Wetzel, C. R., Doering, K., Johnson, K. F., Methot, R. D. 2021. A cookbook for using model diagnostics in integrated stock assessments. Fisheries Research, 240: 105959.

Gelman, A., Rubin, D.B., 1992. Inference from Iterative Simulation Using Multiple Sequences. Stat. Sci. 7, 457–472. https://doi.org/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. https://doi.org/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. https://doi.org/10.1093/icesjms/fsu198

Hyndman and Koehler, 2006. Another look at measures of forecast accuracy. International Journal of Forecasting 22(4), 679-688

ICCAT, 2006. ICCAT Manual – Species: 2.1.3. Skipjack tuna. Available on-line at https://iccat.int/Documents/SCRS/Manual/CH2/2_1_3_SKJ_ENG.pdf.

ICCAT, 2014. ICCAT Skipjack Stock Assessment Report. Available on-line at https://iccat.int/Documents/Meetings/Docs/2014_SKJ_ASSESS_ENG.pdf.

ICCAT, 2016. Report of the 2015 ICCAT blue shark stock assessment session. Collect. Vol. Sci. Pap. ICCAT 72, 866–1019.

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

ICCAT, 2017b. Report of the 2016 Mediterranean swordfish stock assessment meeting. Col. Vol. Sci. Pap. ICCAT 73, 1005–1096.

ICCAT, 2017c. Report of the 2017 ICCAT albacore species group intersessional meeting (including assessment of Mediterranean albacore). Collect. Vol. Sci. Pap. ICCAT 74, 45.

ICCAT, 2017d. Report of the 2017 ICCAT shortfin mako assessment meeting. Collect. Vol. Sci. Pap. ICCAT 74, 1465–1561.

Kell, L. T., Kimoto, A. and Kitakado., T. 2016. Evaluation of the prediction skill of stock assessment using hindcasting. Fisheries Research, 183:119–127

Mourato, B.L., Winker, H., Carvalho, F., Ortiz, M., 2019. Stock Assessment of blue marlin (Makaira nigricans) using a Bayesian State-Space Surplus Production Model JABBA. Collect. Vol. Sci. Pap. ICCAT 75, 1003–1025.

Mourato, B.L., Winker, H., Carvalho, F., Kimoto, A., Ortiz, M., 2020. Developing of Bayesian State-Space Surplus Production JABBA for Assessing Atlantic white marlin (Kajikia albida) stock. Col. Vol. Sci. Pap. ICCAT 76, 235–254.

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.

Plummer, M., Nicky Best, Cowles, K., Vines, K., 2006. CODA: Convergence Diagnosis and Output Analysis for MCMC. R News 6, 7–11.

Sant'Ana, R., Mourato, B., Kimoto, A., Walter, J., Winker, H., 2020. Atlantic Yellowfin tuna stock assessment: An Implementation of a Bayesian State-Space Surplus Production Model using JABBA. Col. Vol. Sci. Pap. ICCAT 76, 699–724.

Su and Yajima, 2012. R2jags-a package for running jags from R. https://cran.r-project.org/web/packages/R2jags/index.html

Winker, H., Carvalho, F., Sharma, R., Parker, D., Kerwath, S., 2017. Initial results for North and South Atlantic shortfin mako (Isurus oxyrinchus) stock assessments using the Bayesian surplus production model JABBA and the catch-resiliance method CMSY 74, 1836–1866.

Winker, H., Carvalho, F., Kapur, M., 2018. JABBA: Just Another Bayesian Biomass Assessment. Fish. Res. 204, 275–288. https://doi.org/http://doi.org/10.1016/j.fishres.2018.03.01.

Winker, H., Kerwath, S., Merino, G. Ortiz, M. 2018b. Bayesian state-space surplus production model JABBA assessment of Atlantic bigeye tuna (Thunnus obesus) stock. Collect. Vol. Sci. Pap. ICCAT, 75(7): 2129-2168.

Winker, H., Carvalho, F., Kerwath, S., 2019a. Age-structured biomass dynamics of north Atlantic shortfin mako with implications for the interpretation of surplus production models. ICCAT-SCRS 098, 1–19.

Winker, H., Mourato, B., Chang, Y. 2019b. Unifying parametrizations between age-structured and surplus production models: an application to Atlantic white marlin (Kajika albida) with simulation testing. Col. Vol. Sci. Pap. ICCAT SCRC/2019/103.

Winker, H., Kerwath, S.E., Merino, G., Ortiz, M., 2019c. Bayesian State-Space Surplus Production Model JABBA of Atlantic bigeye tuna (Thunnus obesus) stock. Col. Vol. Sci. Pap. ICCAT 75, 2129–2168.

Winker, H., Kimoto, A., Mourato, B., Tserpes, G., and Ortiz, M. 2020. Development of Bayesian state-space surplus production model JABBA for assessing the Mediterranean swordfish (Xiphias gladius) stock. SCRS/2020/082.

Table 1. Summary of the uncertainty grid scenarios for West Atlantic skipjack tuna.

Scenario	Model	r	B_{MSY}/K (m)
S01	ASEM h = 0.7 Pella m	Lognormal (0.545, 0.284)	0.40
S02	ASEM h = 0.8 Pella m	Lognormal (0.607, 0.318)	0.41
S03	ASEM h = 0.9 Pella m	Lognormal (0.668, 0.330)	0.42
S04	ASEM h = 0.7 Pella m	Lognormal (0.416, 0.148)	0.38
S05	ASEM h = 0.8 Pella m	Lognormal (0.440, 0.184)	0.37
S06	ASEM h = 0.9 Pella m	Lognormal (0.466, 0.219)	0.36
S07	ASEM h = 0.7 Pella m	Lognormal (0.366, 0.142)	0.38
S08	ASEM h = 0.8 Pella m	Lognormal (0.385, 0.172)	0.36
S09	ASEM h = 0.9 Pella m	Lognormal (0.402, 0.206)	0.35

Table 2. Summary of sensitivity analysis runs for West Atlantic skipjack tuna (*Katsuwonus pelamis*).

Scenario	Model	Type	Indices
S05	Pella m	ASEM h = 0.8	+ BRA BB Past + BRA BB Present
S05	Pella m	ASEM h = 0.8	+ BRA BB Past + BRA BB Present + USA LL
S05	Pella m	ASEM h = 0.8	+ BRA BB Past + BRA BB Present + USA LL + BRA HL
S05	Pella m	ASEM h = 0.8	+ BRA BB Past + BRA BB Present + USA LL + BRA HL + VEN PS

Table 3. Summary of posterior quantiles presented in the form of marginal posterior medians and associated the 95% credibility intervals of parameters for the Bayesian state-space surplus production models for West Atlantic skipjack tuna.

S01				S02			
Estimates	**Median**	**LCI (2.50%)**	**UCI (97.50%)**	**Estimates**	**Median**	**LCI (2.50%)**	**UCI (97.50%)**
K	135,554	89,686	223,440	K	124,239	81,155	190,556
r	0.861	0.552	1.321	r	0.980	0.618	1.551
ψ(psi)	0.940	0.815	0.991	ψ(psi)	0.939	0.816	0.991
σ_{proc}	0.103	0.056	0.166	σ_{proc}	0.101	0.056	0.164
F_{MSY}	0.724	0.465	1.112	F_{MSY}	0.782	0.493	1.238
B_{MSY}	54,219	35,872	89,372	B_{MSY}	50,945	33,278	78,138
MSY	38,457	29,754	59,238	MSY	39,119	30,300	59,165
B_{1952}/K	0.931	0.726	1.164	B_{1952}/K	0.930	0.724	1.166
B_{2020}/K	0.734	0.532	0.922	B_{2020}/K	0.752	0.546	0.931
B_{2020}/B_{MSY}	1.836	1.330	2.306	B_{2020}/B_{MSY}	1.834	1.330	2.271
F_{2020}/F_{MSY}	0.257	0.143	0.440	F_{2020}/F_{MSY}	0.253	0.145	0.427

S03				S04			
Estimates	**Median**	**LCI (2.50%)**	**UCI (97.50%)**	**Estimates**	**Median**	**LCI (2.50%)**	**UCI (97.50%)**
K	121,544	79,093	194,144	K	188,042	140,788	269,550
r	1.054	0.651	1.654	r	0.506	0.384	0.662
ψ(psi)	0.940	0.815	0.991	ψ(psi)	0.940	0.817	0.991
σ_{proc}	0.098	0.054	0.162	σ_{proc}	0.106	0.061	0.169
F_{MSY}	0.799	0.493	1.253	F_{MSY}	0.474	0.360	0.620
B_{MSY}	51,043	33,216	81,532	B_{MSY}	71,464	53,506	102,441
MSY	40,152	30,630	61,185	MSY	33,621	27,008	47,088
B_{1952}/K	0.932	0.730	1.166	B_{1952}/K	0.931	0.722	1.162
B_{2020}/K	0.769	0.566	0.943	B_{2020}/K	0.641	0.438	0.847
B_{2020}/B_{MSY}	1.832	1.347	2.245	B_{2020}/B_{MSY}	1.687	1.153	2.229
F_{2020}/F_{MSY}	0.246	0.142	0.417	F_{2020}/F_{MSY}	0.319	0.185	0.556

S05				S06			
Estimates	**Median**	**LCI (2.50%)**	**UCI (97.50%)**	**Estimates**	**Median**	**LCI (2.50%)**	**UCI (97.50%)**
K	172,595	122,341	261,704	K	155,467	107,402	238,505
r	0.575	0.408	0.800	r	0.651	0.447	0.936
ψ(psi)	0.939	0.815	0.990	ψ(psi)	0.939	0.819	0.991
σ_{proc}	0.104	0.059	0.167	σ_{proc}	0.105	0.059	0.167
F_{MSY}	0.568	0.403	0.790	F_{MSY}	0.680	0.467	0.977
B_{MSY}	63,873	45,275	96,850	B_{MSY}	55,971	38,667	85,865
MSY	36,040	28,110	51,995	MSY	37,617	29,019	55,466
B_{1952}/K	0.929	0.724	1.171	B_{1952}/K	0.931	0.721	1.169
B_{2020}/K	0.673	0.463	0.871	B_{2020}/K	0.689	0.477	0.885
B_{2020}/B_{MSY}	1.819	1.252	2.353	B_{2020}/B_{MSY}	1.914	1.324	2.458
F_{2020}/F_{MSY}	0.276	0.161	0.496	F_{2020}/F_{MSY}	0.251	0.143	0.453

Continues on the next page…

Table 3. Continued from the previous page.

S07				S08			
Estimates	**Median**	**LCI (2.50%)**	**UCI (97.50%)**	**Estimates**	**Median**	**LCI (2.50%)**	**UCI (97.50%)**
K	208,597	156,839	303,053	K	185,436	134,087	270,092
r	0.443	0.339	0.581	r	0.500	0.362	0.682
$\psi_{(psi)}$	0.939	0.814	0.991	$\psi_{(psi)}$	0.939	0.817	0.991
σ_{proc}	0.106	0.062	0.168	σ_{proc}	0.107	0.062	0.171
F_{MSY}	0.414	0.317	0.544	F_{MSY}	0.522	0.378	0.712
B_{MSY}	79,276	59,606	115,174	B_{MSY}	66,760	48,273	97,237
MSY	32,716	26,300	45,689	MSY	34,376	27,248	49,174
B_{1952}/K	0.927	0.720	1.163	B_{1952}/K	0.931	0.719	1.172
B_{2020}/K	0.628	0.432	0.833	B_{2020}/K	0.637	0.433	0.847
B_{2020}/B_{MSY}	1.651	1.136	2.192	B_{2020}/B_{MSY}	1.770	1.203	2.354
F_{2020}/F_{MSY}	0.335	0.195	0.576	F_{2020}/F_{MSY}	0.296	0.169	0.527

S09			
Estimates	**Median**	**LCI (2.50%)**	**UCI (97.50%)**
K	172,008	119,107	263,847
r	0.561	0.386	0.806
$\psi_{(psi)}$	0.940	0.814	0.990
σ_{proc}	0.104	0.059	0.167
F_{MSY}	0.618	0.426	0.888
B_{MSY}	60,216	41,697	92,367
MSY	36,731	28,686	54,241
B_{1952}/K	0.930	0.728	1.167
B_{2020}/K	0.668	0.465	0.871
B_{2020}/B_{MSY}	1.909	1.327	2.489
F_{2020}/F_{MSY}	0.259	0.145	0.455

Table 4. Summary Mohn's rho statistic computed for a retrospective evaluation period of eight years for the S05 scenario. The more the values diverge from zero, the stronger is the retrospective bias. Values falling between -0.15 and 0.2 are widely deemed as acceptable retrospective bias (Huerto *et al.*, 2014).

Scenario	***Stock Quantity***					
	B	F	B/B_{MSY}	F/F_{MSY}	B/K	MSY
S05	-0.1374	0.1661	-0.0685	0.1439	-0.0025	-0.0559

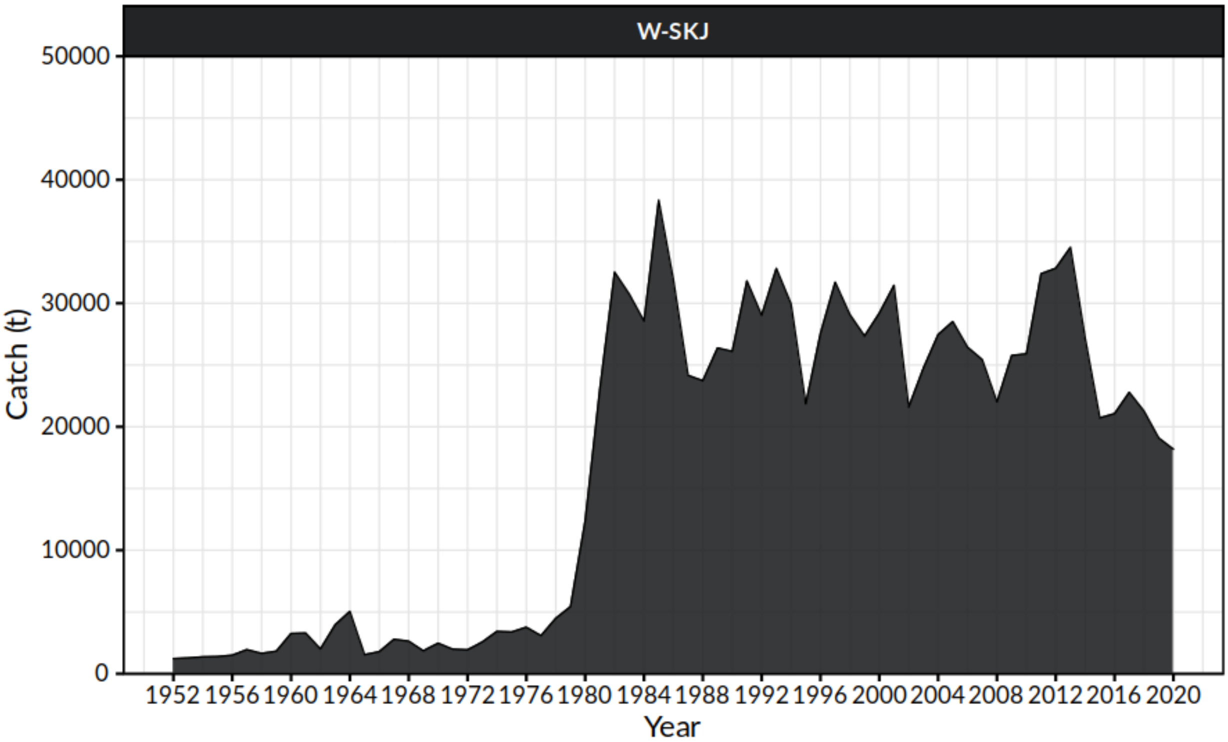


Figure 1. Catch time series in metric tons (t) between 1952 and 2020 for West Atlantic skipjack tuna.

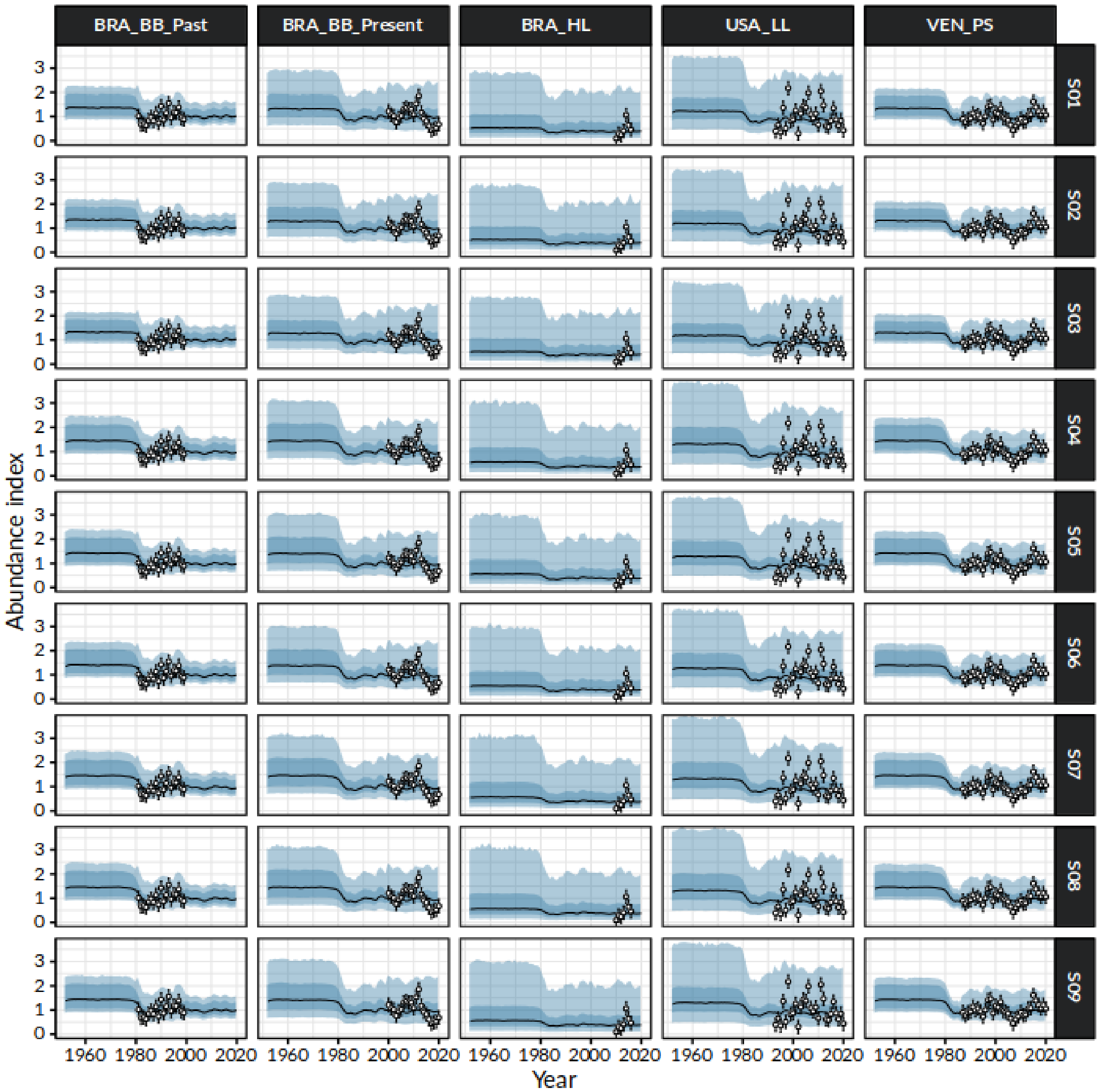


Figure 2. Time series of observed (circle) with error 95% Cis (error bars) and predicted (solid line) CPUE of West Atlantic skipjack tuna for the Bayesian state-space surplus production model JABBA for each scenario fitted. Dark shaded blue areas show 95% credibility intervals of the expected mean CPUE and light shaded blue areas denote the 95% posterior predictive distribution intervals.

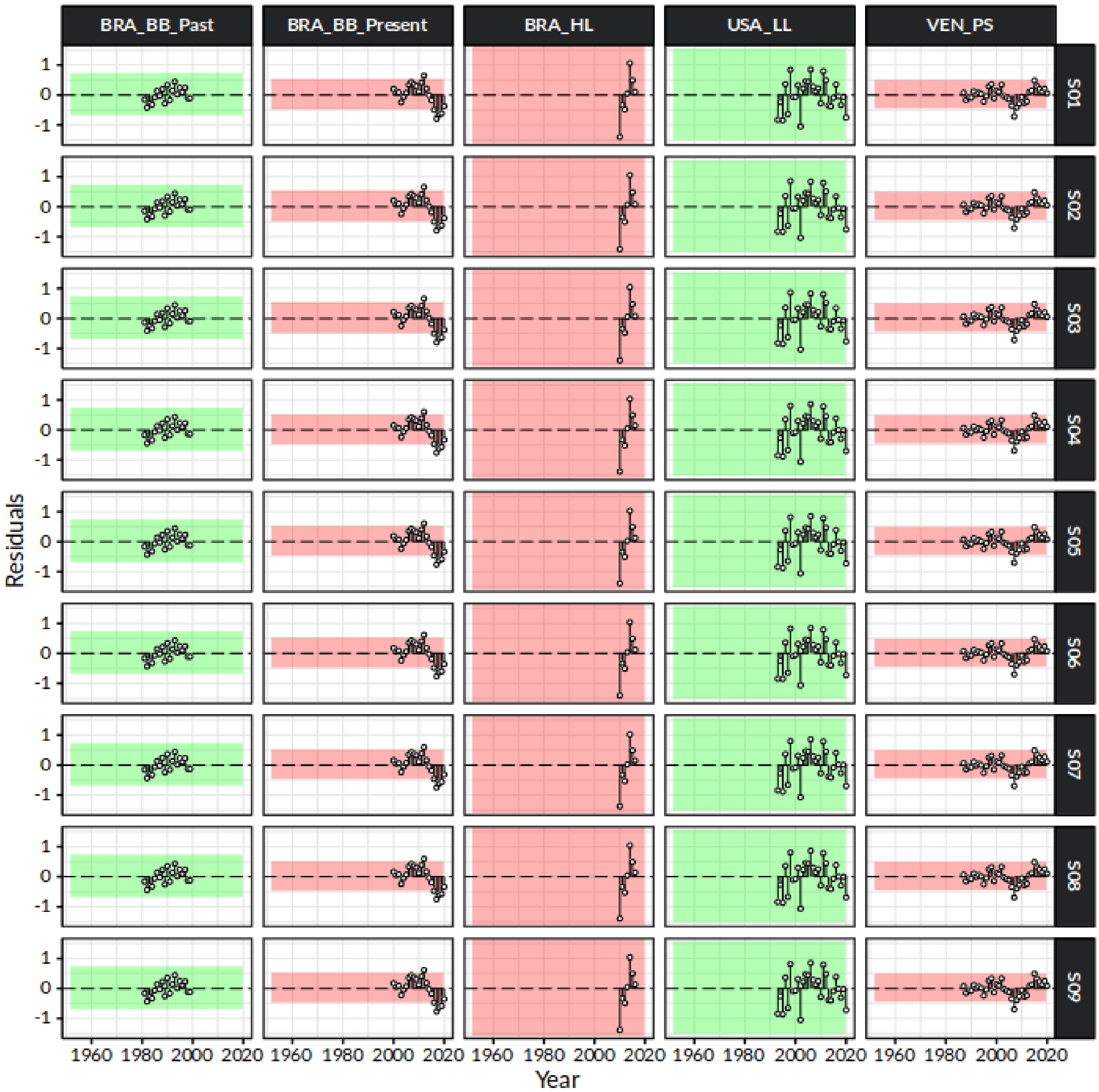


Figure 3. Runs tests to quantitatively evaluate the randomness of the time series of CPUE residuals for each scenario fitted for the West Atlantic skipjack tuna. Green panels indicate no evidence of lack of randomness of time-series residuals (p>0.05) while red panels indicate the opposite. The inner shaded area shows three standard errors from the overall mean and red circles identify a specific year with residuals greater than this threshold value (3x sigma rule).

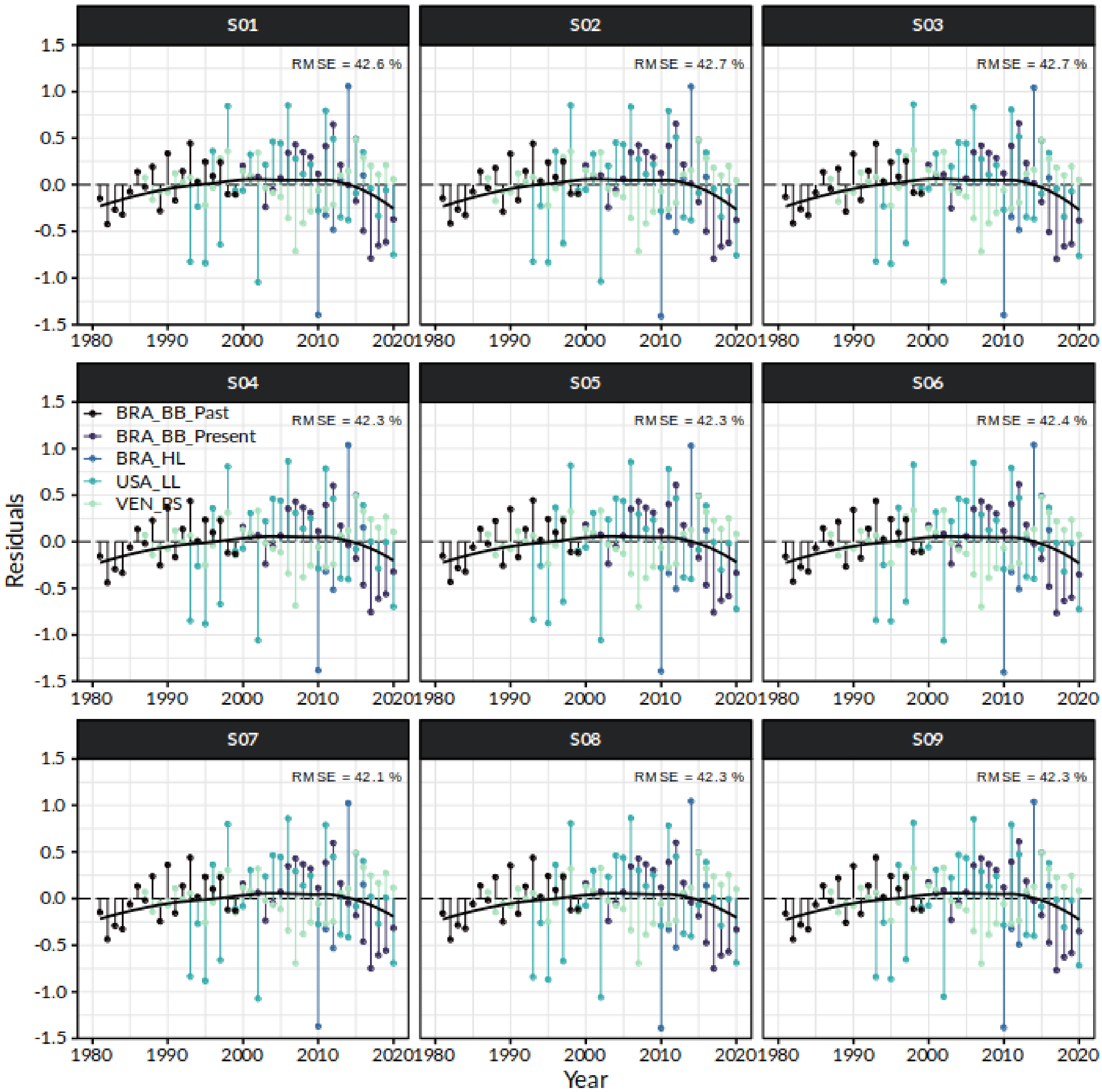


Figure 4. JABBA residual diagnostic plots for alternative sets of CPUE indices examined for each scenario fitted for the West Atlantic skipjack tuna. Solid black lines indicate a loess smoother through all residuals.

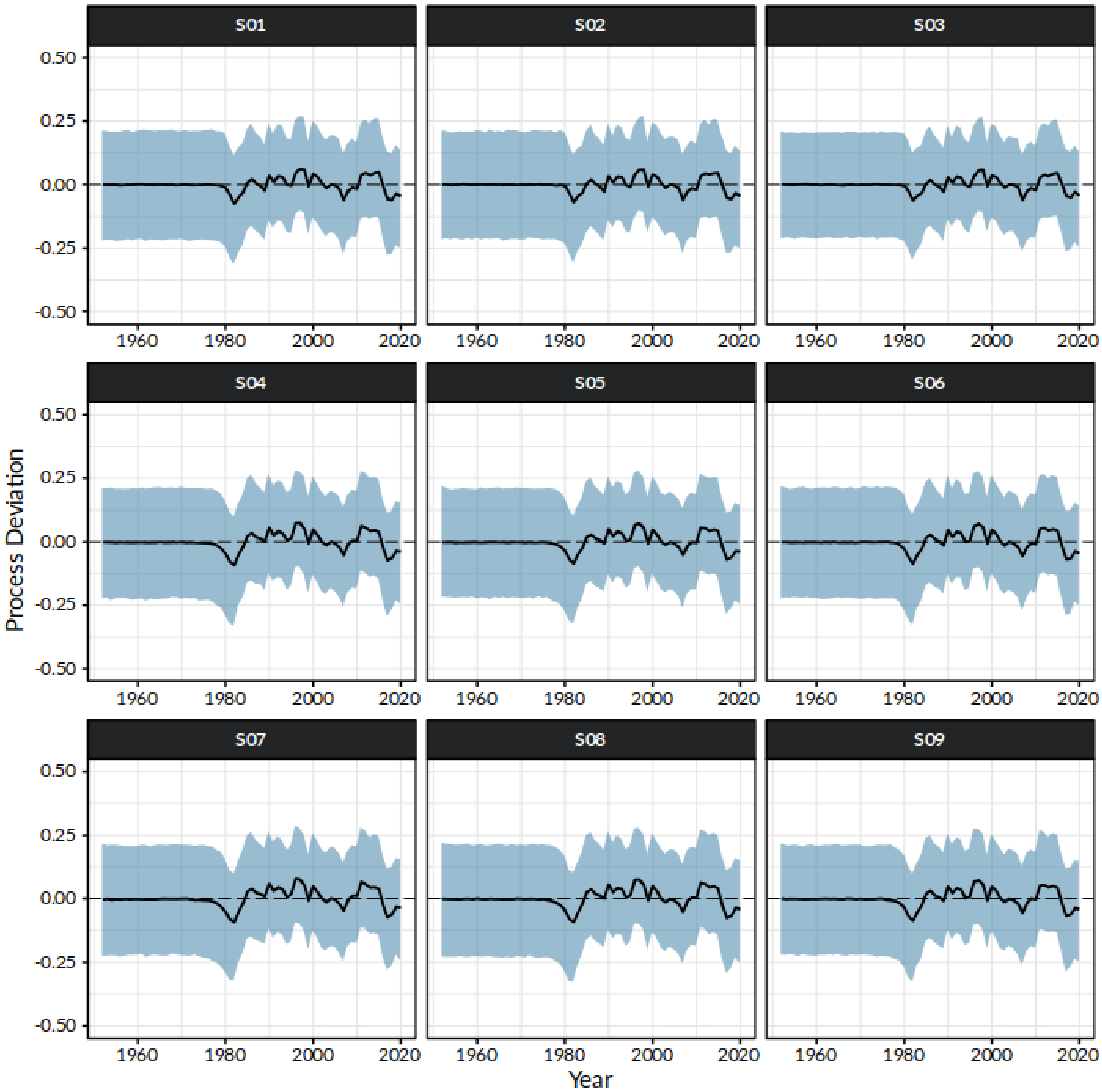
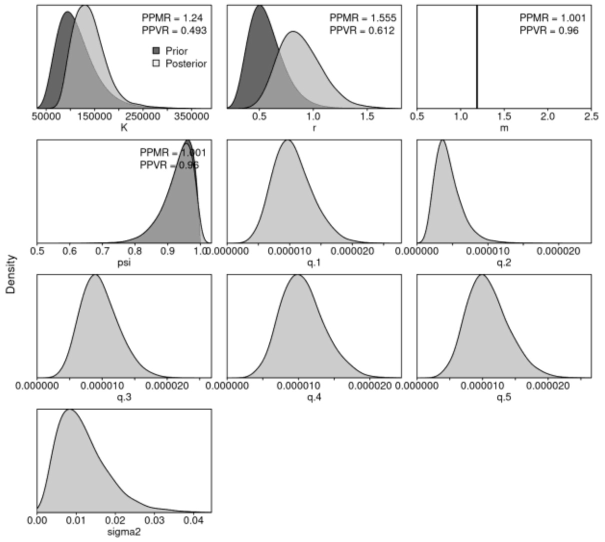
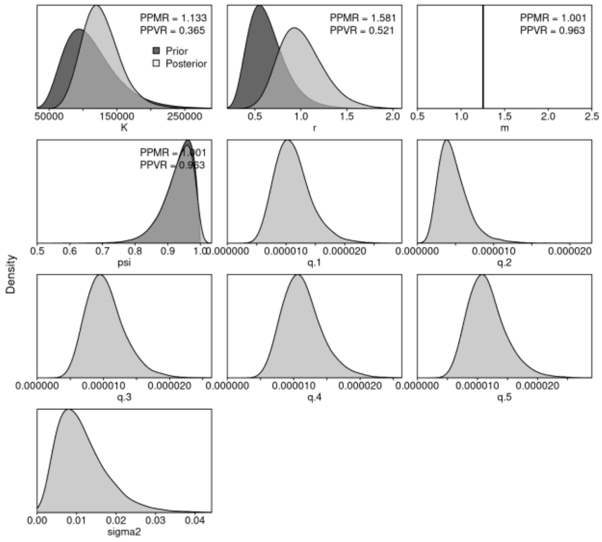


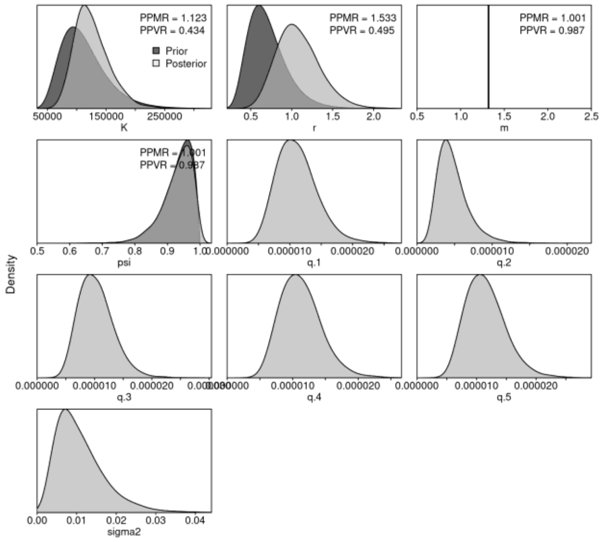
Figure 5. JABBA residual diagnostic plots for alternative sets of CPUE indices examined for each scenario fitted for the West Atlantic skipack tuna. Process error deviates (median: solid line) with shaded blue area indicating 95% credibility intervals.

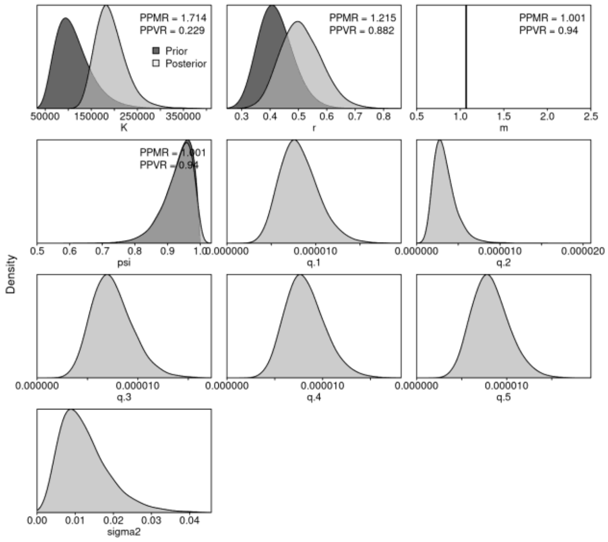
S01 **S02**





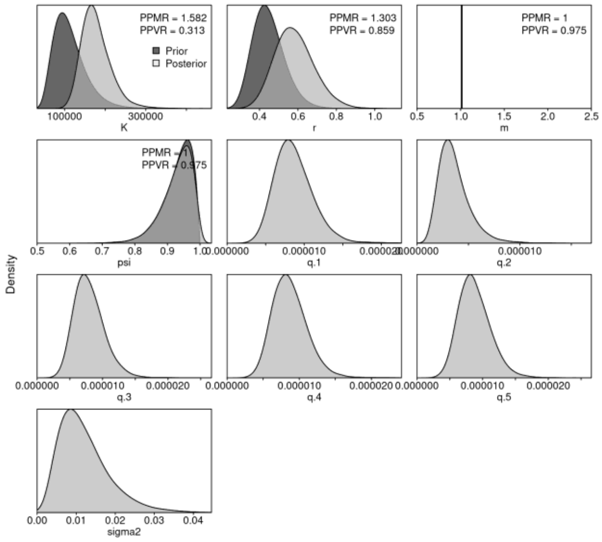
S03 **S04**



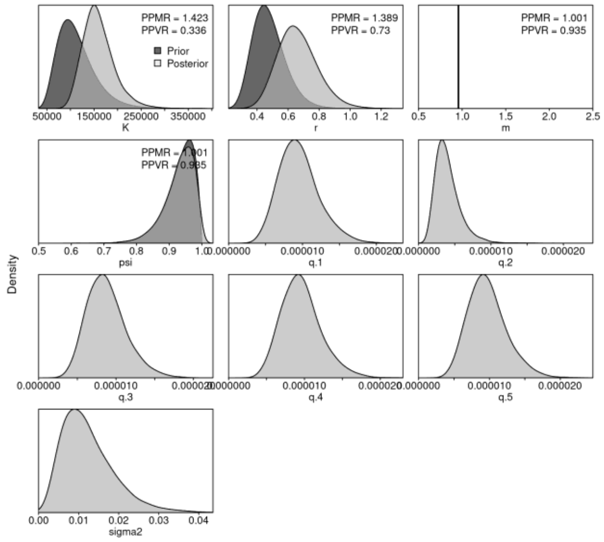


Continues on the next page…

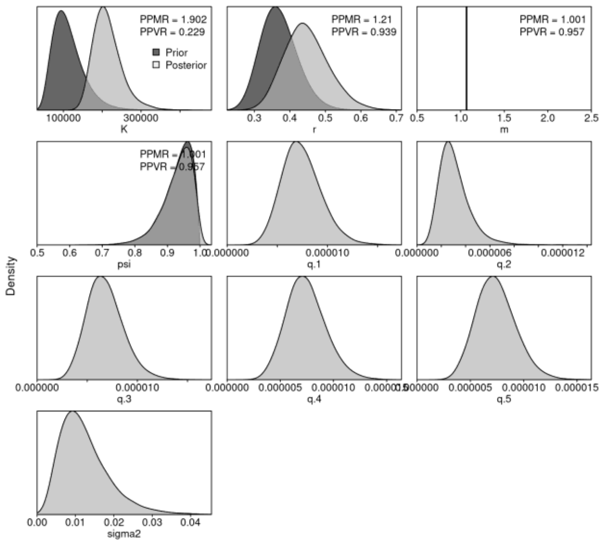
S05



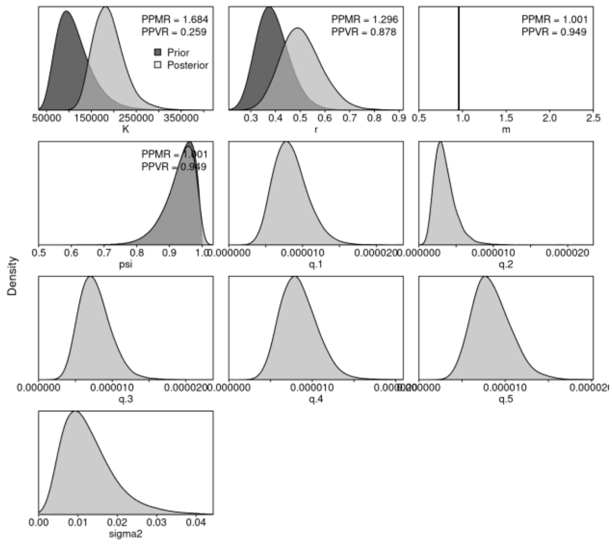
S06



S07



S08



Continues on the next page…

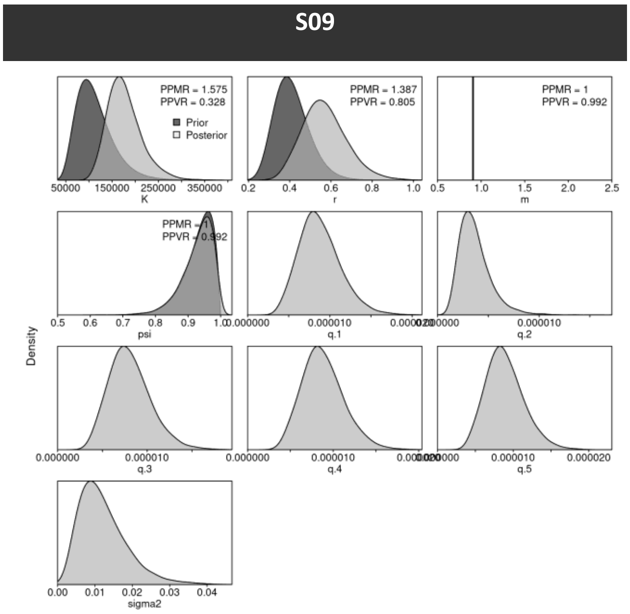


Figure 6. Prior and posterior distributions of various models and management parameters for the Bayesian state-space surplus production fitted for the West Atlantic skipjack tuna. PPRM: Posterior to Prior Ratio of Medians; PPRV: Posterior to Prior Ratio of Variances.

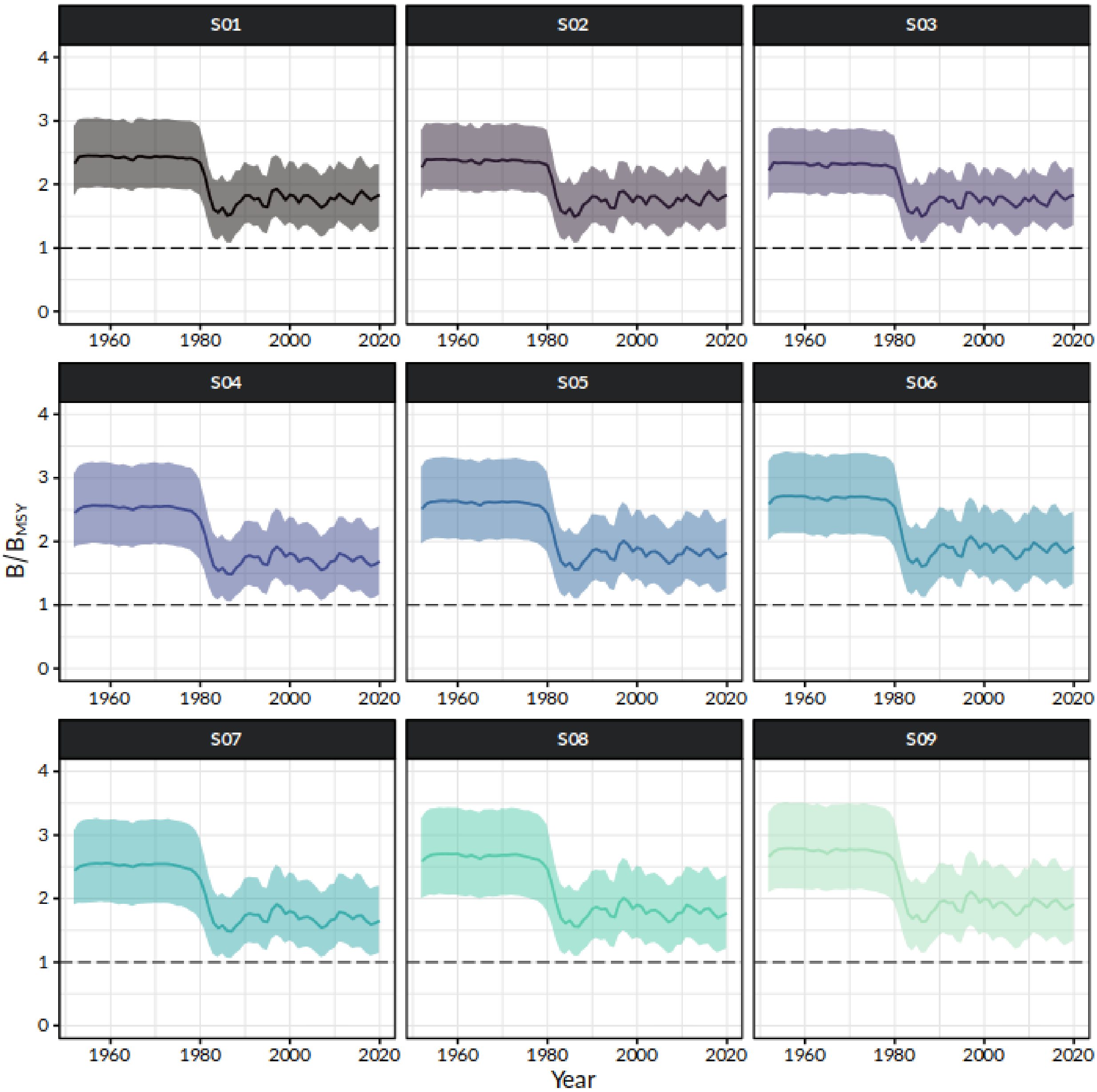


Figure 7. Trends in biomass relative to B_{MSY} (B/B_{MSY}) for each scenario from the Bayesian state-space surplus production JABBA model fits to West Atlantic skipjack tuna.

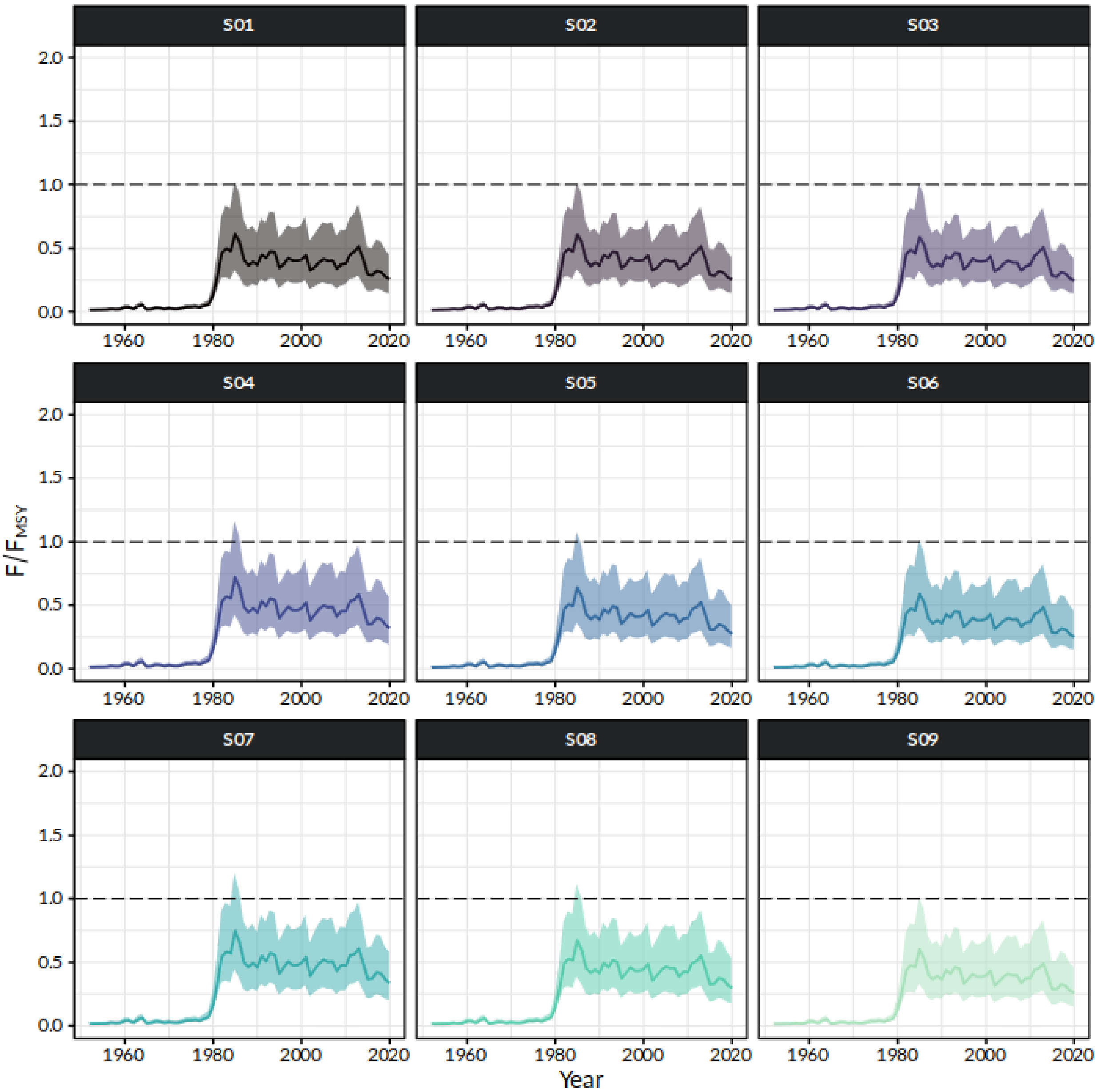


Figure 8. Trends in biomass relative to F_{MSY} (F/F_{MSY}) for each scenario from the Bayesian state-space surplus production JABBA model fits to West Atlantic skipjack tuna.

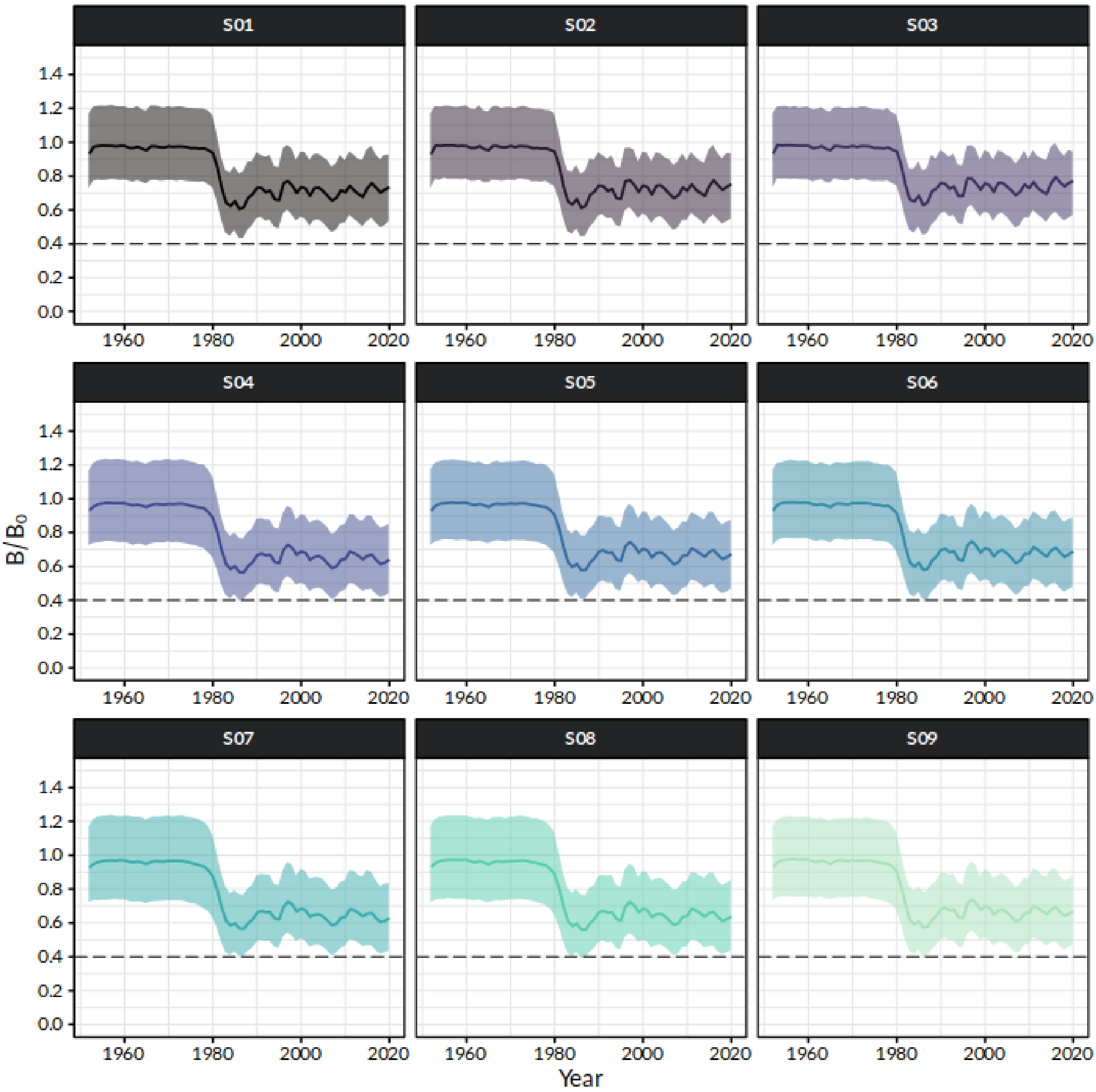


Figure 9. Trends in biomass relative to B_0 (B/B_0) for each scenario from the Bayesian state-space surplus production JABBA model fits to West Atlantic skipjack tuna.

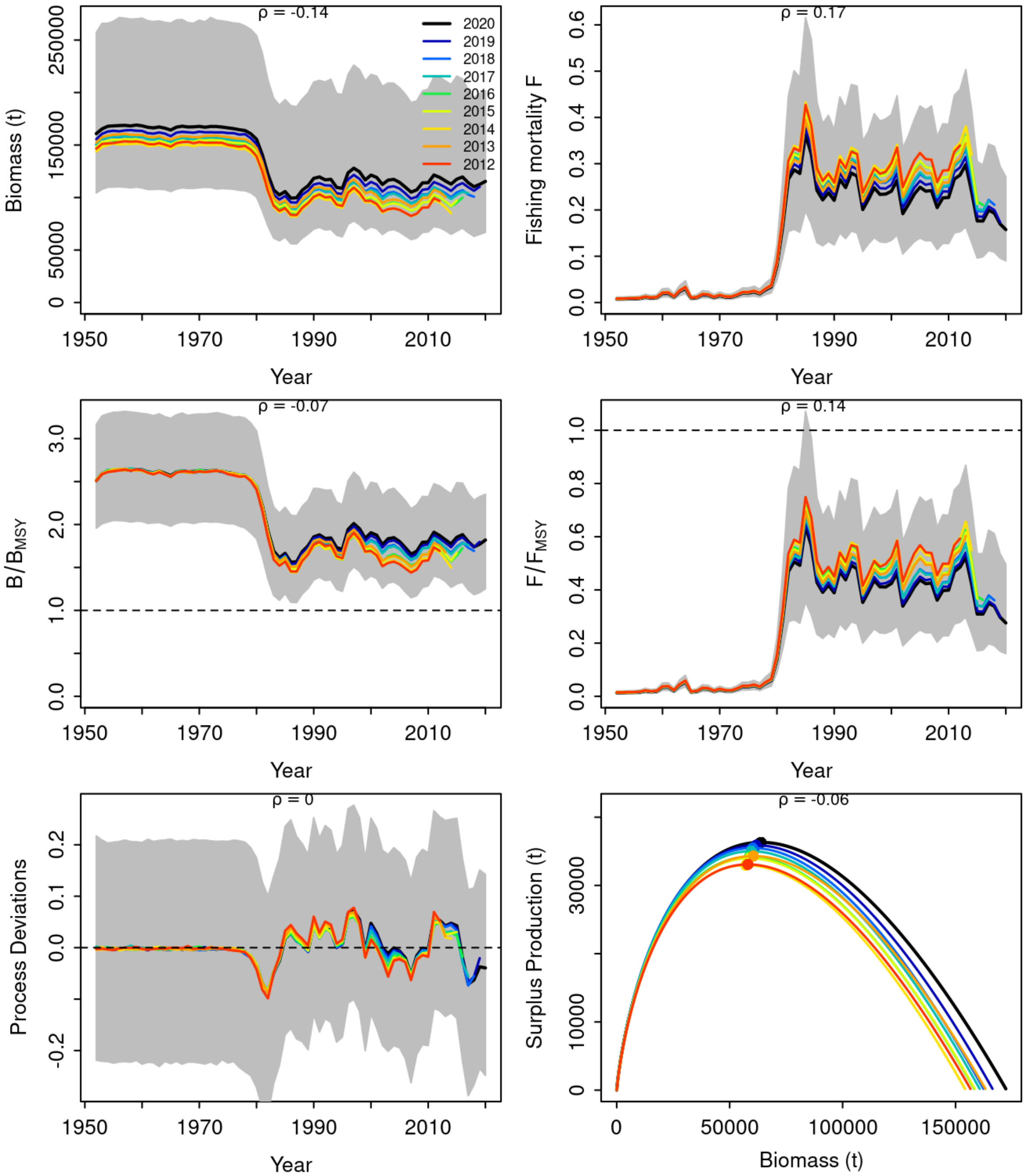


Figure 10. Retrospective analysis conducted for scenario S05 for West Atlantic skipjack tuna, by removing one year at a time sequentially (n=8) and predicting the trends in biomass and fishing mortality (upper panels), biomass relative to B_{MSY} (B/B_{MSY}) and fishing mortality relative to F_{MSY} (F/F_{MSY}) (middle panels) and process deviations and surplus production curve (bottom panels) from the Bayesian state-space surplus production model fits.

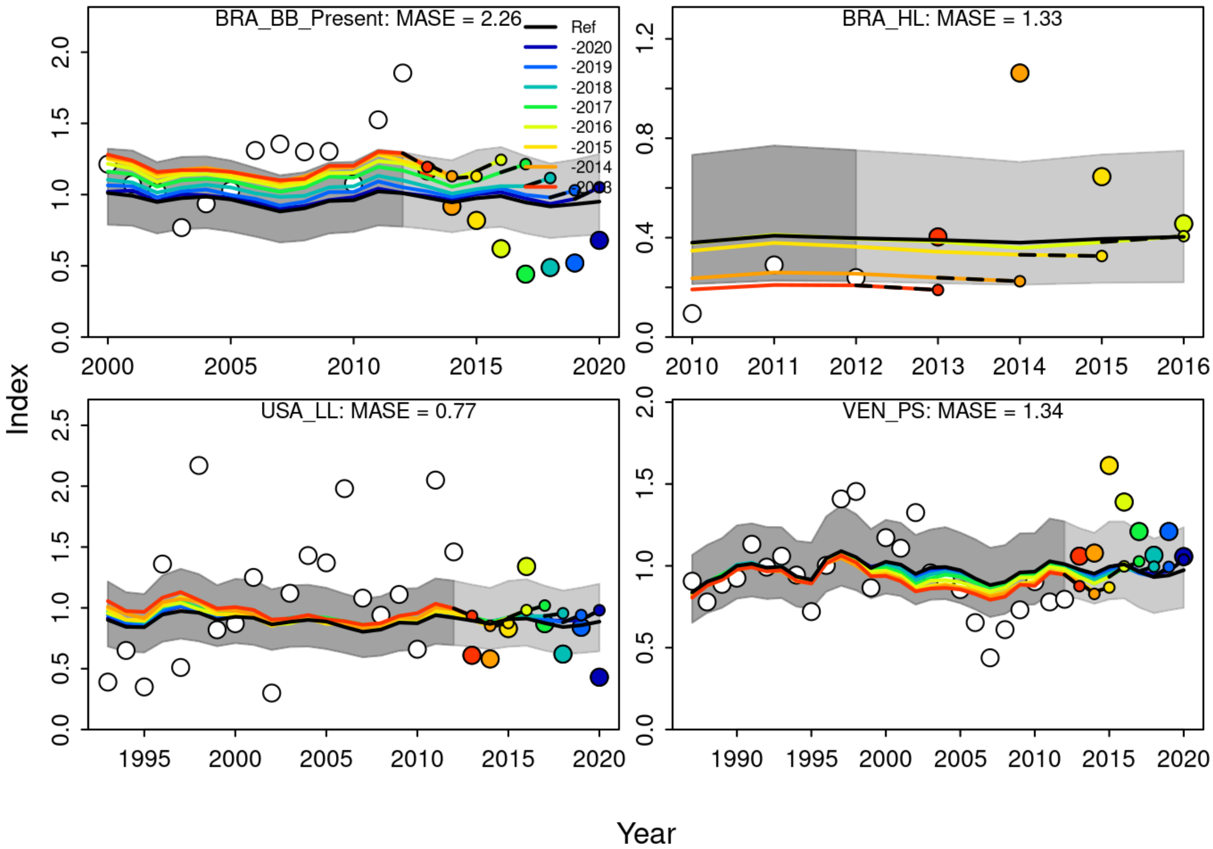


Figure 11. Hindcasting cross-validation results (HCxval) for the two scenarios S05 for West Atlantic skipjack tuna, showing one-year-ahead forecasts of CPUE values (2011-2019), performed with eight hindcast model runs relative to the expected CPUE. The CPUE observations, used for cross-validation, are highlighted as color-coded solid circles with associated light-grey shaded 95% confidence interval. The model reference year refers to the end points of each one-year-ahead forecast and the corresponding observation (i.e. year of peel + 1).

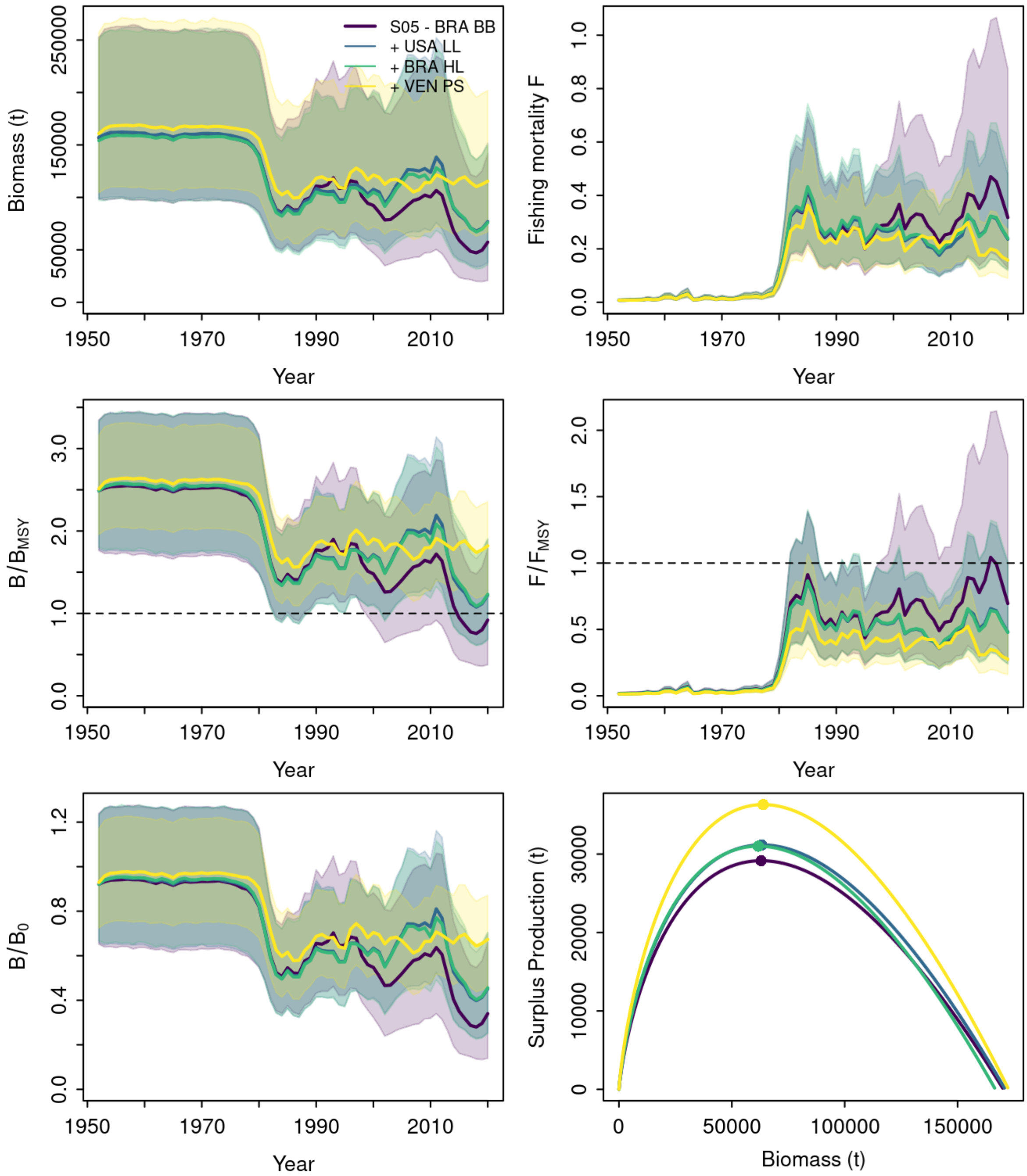


Figure 12. Sensitivity analysis performed for scenarios S05 showing the trends in biomass and fishing mortality (upper panels), biomass relative to B_{MSY} (B/B_{MSY}) and fishing mortality relative to F_{MSY} (F/F_{MSY}) (middle panels) and biomass relative to K (B/K) and surplus production curve (bottom panels) for the West Atlantic skipjack tuna.

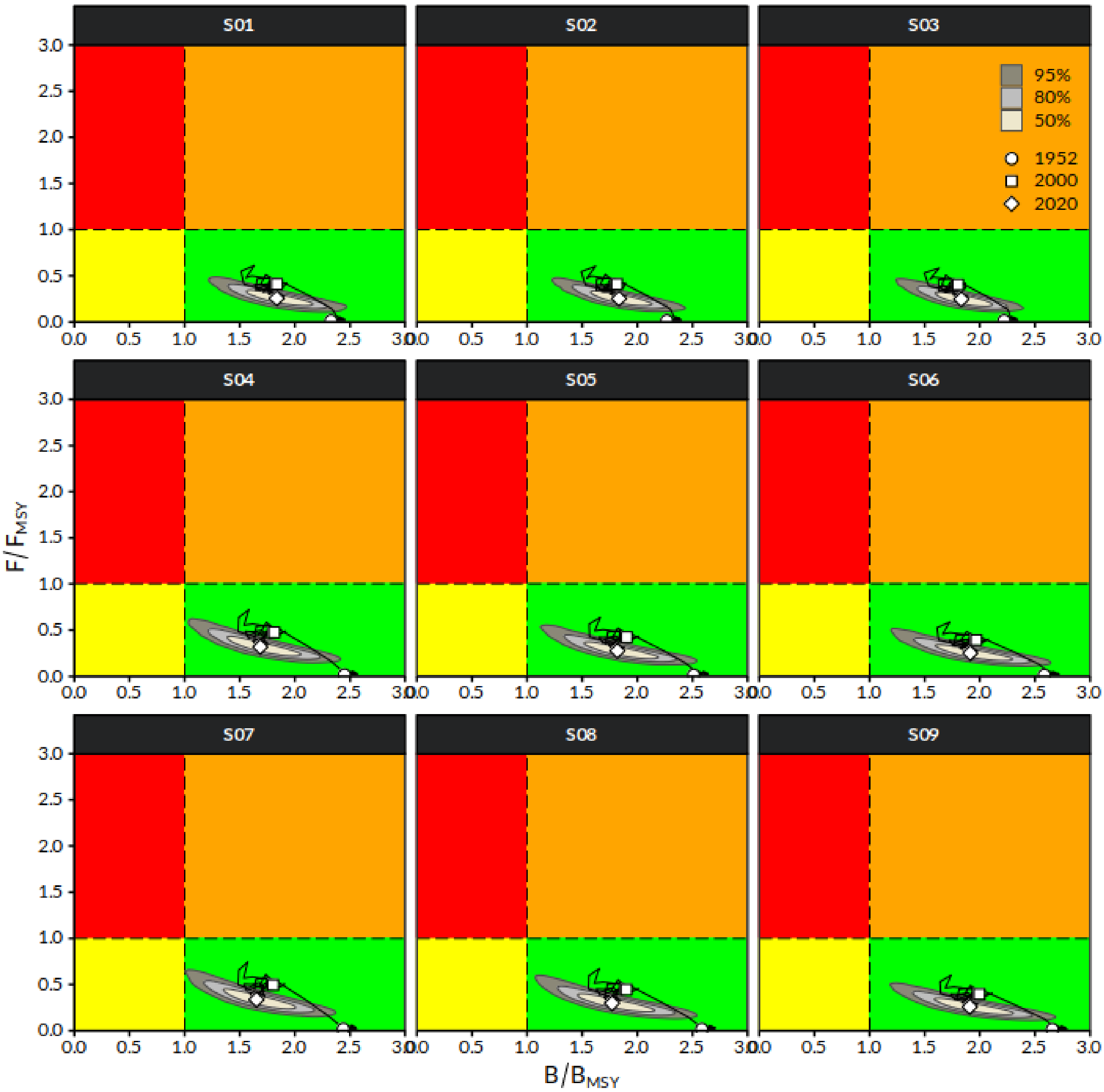


Figure 13. Kobe phase plot showing estimated trajectories (1952-2020) of B/B_{MSY} and F/F_{MSY} for the Bayesian state-space surplus production model for the West Atlantic skipjack tuna. Different grey shaded areas denote the 50%, 80%, and 95% credibility interval for the terminal assessment year. The probability of terminal year points falling within each quadrant is indicated in the figure legend.