PRELIMINARY STOCK ASSESSMENT OF NORTHEASTERN ATLANTIC PORBEAGLE (*LAMNA NASUS*) USING THE BAYESIAN STATE-SPACE SURPLUS PRODUCTION MODEL *JABBA*

M. Ortiz[1], N. Taylor[1], A. Kimoto[1] and R. Forselledo[2]

SUMMARY

*Bayesian State-Space Surplus Production Models were fitted to Northeastern Atlantic porbeagle shark (*Lamna nasus*) catch and relative abundance indices using the 'JABBA' R package. This document presents details on the model diagnostics and stock status estimates for preliminary scenarios. A ref model was fitted to three indices reviewed by the ICES WKELASMO in 2022 and the full model also included a fourth historical index applied in 2009 stock assessment. The prior assumptions in the surplus production function were kept consistent with the ICES WKELASMO assessment presented in 2022. We evaluated model plausibility using four objective model diagnostics: (1) model convergence, (2) fits to the data, (3) consistency (e.g. retrospective patterns) and (4) prediction skill. Our results suggest that the full model represents the most plausible candidate model that incorporates all available relative indices of abundance. Results are consistent with the SPiCT model runs, indicating that the stock is currently overfished (0.45 B/B_{MSY}) but no experiencing overfishing (0.01 F/F_{MSY}). Additional sensitivity runs indicated that the full model was robust to alternative productivity and variance assumptions.*

RÉSUMÉ

*Les modèles de production excédentaire état-espace de type bayésien ont été ajustés aux données de capture et aux indices d'abondance relative du requin-taupe commun (*Lamna nasus*) de l'Atlantique Nord-Est en utilisant le progiciel « JABBA » R. Ce document présente des détails sur les diagnostics du modèle et les estimations de l'état des stocks pour des scénarios préliminaires. Un modèle de référence a été ajusté à trois indices examinés par le WKELASMO du CIEM en 2022 et le modèle complet comprenait également un quatrième indice historique appliqué à l'évaluation du stock de 2009. Les postulats a priori dans la fonction de production excédentaire correspondaient à la dernière évaluation du WKELASMO du CIEM présentée en 2022. Nous avons évalué la plausibilité des modèles en utilisant quatre diagnostics de modèle objectifs : (1) la convergence des modèles, (2) les ajustements aux données, (3) la cohérence (p.ex. les schémas rétrospectifs) et (4) la capacité de prédiction. Nos résultats suggèrent que le modèle complet représente le modèle potentiel le plus plausible qui incorpore tous les indices d'abondance relative disponibles. Les résultats concordent avec les scénarios du modèle SPiCT, indiquant que le stock est actuellement surexploité (0,45 B/BPME) mais ne fait l'objet de surpêche (0,01 F/FPME). Des analyses de sensibilité supplémentaires indiquaient que le modèle complet était robuste aux postulats alternatifs de productivité et de variance.*

RESUMEN

*Se ajustaron modelos bayesianos de producción excedente estado-espacio a los índices de captura y abundancia relativa del marrajo sardinero (*Lamna nasus*) del Atlántico nororiental utilizando el paquete R "JABBA". Este documento presenta detalles sobre los diagnósticos del modelo y las estimaciones del estado del stock para los escenarios preliminares. Se ajustó un modelo de referencia a tres índices revisados por el WKELASMO de ICES en 2022 y el modelo completo también incluyó un cuarto índice histórico aplicado en la evaluación de stock de 2009. Los supuestos previos en la función de producción excedente se mantuvieron coherentes con la última evaluación del WKELASMO de ICES, presentada en 2022. La plausibilidad de estos*

[1] ICCAT Secretariat. Calle Corazón de Maria 8, Madrid Spain 28002.
[2] Dirección Nacional de Recursos Acuáticos DINARA, Laboratorio de Recursos Pelágicos, Constituyente 1497, CP 11200 Montevideo Uruguay.

modelos se evaluó mediante cuatro diagnósticos objetivos del modelo: (1) la convergencia del modelo, (2) el ajuste a los datos, (3) la coherencia (por ejemplo, patrones retrospectivos) y (4) la capacidad de predicción. Nuestros resultados sugieren que el modelo completo representa el modelo candidato más plausible que incorpora todos los índices de abundancia relativa disponibles. Los resultados son coherentes con los ensayos del modelo SPiCT, e indican que e stock está actualmente sobrepescado /0,45 B/B_{RMS}) pero no experimentando sobrepesca (0,01 F/F_{RMS}). Los ensayos de sensibilidad adicionales indicaron que el modelo completo era robusto ante las hipótesis alternativas de productividad y varianza.

KEYWORDS

Northeastern Atlantic porbeagle, stock status, biomass dynamic model,
Model diagnostics, hindcasting

1. Introduction

The porbeagle (*Lamna nasus*) is a widely distributed shark in cold and temperate waters of the Atlantic Ocean, including the Mediterranean Sea (Castro 1983, Compagno 1984 and 2001). Porbeagle inhabits the open ocean and continental shelves but is also found close to the coast. For management purposes, the International Commission for the Conservation of Atlantic Tunas (ICCAT) considers there to be four stocks, northeast and northwest Atlantic, and a South Atlantic east and west stocks.

The previous stock assessment for the Northeastern Atlantic porbeagle stock was carried out by ICES/ICCAT in 2009 (Anon., 2010). The 2009 assessment presented the stock status estimates using a version of the Bayesian surplus production model BSP (Winker *et al*., 2018). The model was fitted to a catch series (1926 to 2008) and two CPUE indices that started in 1972 (French longline) and 1986 (Spanish longline).

This document presents the stock assessment results for Northeastern Atlantic porbeagle stock based on the Bayesian State-Space Surplus Production Model software, Just Another Bayesian Biomass Assessment (JABBA, Winker 2018), using updated catch (1926 to 2020) and four standardized indices of abundance, three of them reviewed at the Benchmark Workshop for selected elasmobranch stocks (WKELASMO) data preparatory meeting; i) a French longline CPUE series (1972 – 2009) (Biais 2022a, ii) a Norway longline CPUE (1950-1964, 1968-1972) (Biais 2022b), and iii) a longline survey index (2018-2019) that has been extended back in time with commercial longline data (2000-2009) (Biais 2022). The fourth index included in this analysis is the historical Spanish longline index (1986-2007) that was reviewed and applied in the 2009 stock assessment (Mejuto *et al*., 2010). The main improvements compared to the 2009 assessment include new historical indices back in 1950, just after the main catches of this stock, a composite survey index that although limited to two years has been extended back in time (2000 to 2009) by combining with an index from a single commercial longline vessel by set fishing operations (Biais 2022).

2. Material and Methods

This preliminary stock assessment is implemented using the Bayesian state-space surplus production model framework JABBA (Winker *et al*., 2018). JABBA's built-in options include: automatic fitting of multiple CPUE time series and associated standard errors; estimating or fixing the process variance, optional estimation of additional observation variance for individual or grouped CPUE time series, and specifying a Fox, Schaefer or Pella-Tomlinson production function by setting the inflection point B_{MSY}/K and converting this ratio into shape a parameter m. JABBA also provides a comprehensive toolbox to conduct model diagnostics to objectively evaluate the four model plausible criteria recommended in Carvalho *et al*. (2021): (1) model convergence (2) fit to the data, (3) model consistency (retrospective pattern) and (4) prediction skill through hindcast cross-validation (Kell *et al*. 2016; 2021). The full JABBA model description, including formulation and state-space implementation, prior specification options, and diagnostic tools is available in Winker *et al*. (2018). Following its first application to Mediterranean albacore in 2017 based on an early development version in 2017 (ICCAT, 2017a), JABBA has evolved into a fully documented, open-source R package (https://github.com/JABBAmodel/JABBA), which has been included in the ICCAT stock catalogue (https://github.com/ICCAT/software/wiki/2.8-JABBA). JABBA has subsequently been applied in a number of recent ICCAT stock assessments south Atlantic swordfish (ICCAT, 2017b; Winker *et al*., 2018), Atlantic shortfin mako shark stocks (south and north) (ICCAT, 2017c; Winker *et al*., 2020, 2017), Atlantic blue marlin (Mourato *et al*., 2019), Atlantic bigeye tuna (Winker *et al*., 2019), Atlantic White marlin (Mourato *et al*., 2020), Atlantic yellowfin tuna (Sant'Ana *et al*., 2020), Mediterranean swordfish (Winker *et al*. 2020a), South Atlantic albacore (Winker *et al*. 2020b), and Mediterranean albacore (Winker *et al*. 2021).

2.1. Fishery data

Fishery catch data for northeastern porbeagle were made available by ICES/ICCAT and reviewed by WKELASMO during the data preparatory meeting (ICES 2022 WKELASMO, Ortiz *et al*., 2022) for the period 1926-2020 (**Figure 1**). During the data preparatory meeting, the Group reviewed three indices of abundance, including two new series, the Norway historic longline series from 1950 to 1972 (missing 3 years 1965-67) that covers a wider spatial area of the stock and the main fishery during this period, just after the major peak of catches of porbeagle in 1947 (**Figure 2**). And a longline survey carry out in 2018 and 2019, the only index of abundance available after the full no-retention management policy implemented after 2010. And the third index reviewed was the updated French longline CPUE that targets porbeagle from 1972 – 2009 and covers the Bay of Biscay and the Celtic Sea, this index was available also at the 2009 stock assessment.

For this preliminary assessment evaluation, it was included also the Spanish longline index as it was used in the 2009 assessment with a larger geographical coverture, but for which porbeagle is a non-target species (Mejuto *et al*., 2010). For a reference scenario (ref), the CPUEs reviewed by the WKELASMO group at the data preparatory meeting were included the French longline (FRA-LL), the Norway longline (NOR-LL), and the Survey composite index (Survey), all converted to biomass units as agreed by the Group. The corresponding standard errors and coefficient of variance from the standardization of the index in biomass units were also included.

In addition, we considered an alternative scenario (Scenario 2 Full), where JABBA was fitted to the additional historical Spanish longline index as presented in 2009. The indices used in this assessment were provided in mass per unit effort and assumed to be proportional to biomass.

2.2. Model specifications and sensitivity runs

Initially, two candidate model scenarios were considered:

- Ref: a reference scenario, fitted to the three indices (FRA-LL, Survey-Index, and NOR-LL) reviewed in the data preparatory meeting.

- Full: a model, fitted to all four available indices including the SPA-LL index from the 2009 stock assessment.

For the unfished equilibrium biomass K, we used default settings of the JABBA R package in the form of vaguely informative lognormal prior with a large CV of 100% and a central value that corresponds to eight times the maximum total catch which is consistent with other platforms, such as Catch-MSY (Martell and Froese, 2013) or the initial value for K in SPiCT (Pederson and Berg 2017). We assumed a Fox production function setting the inflection point at B_{MSY}/K = 0.37, a lognormal prior distribution for r with mean of 0.059 and a standard deviation for log(r) of 0.457 and initial beta prior for the relative biomass ($\varphi = B_{1926}/K$) with mean = 0.90 and CV of 10% (or alpha = 9.1, beta = 1.011). All catchability parameters were formulated as uninformative uniform priors, while the process error of log(B_y) in year y was estimated "freely" by the model using an uninformative inverse-gamma distribution with both scaling parameters set at 0.001. The prior for r (0.059) is the same as used in the SPiCT model and agreed by the WKELASMO group, however, the standard deviation for log(r) is greater than the one used in the SPiCT model, providing a less restrictive r prior in the JABBA runs.

Initial trials indicated that it was challenging to reliably estimate observation errors using an additional variance approach for model internal weighting (e.g. Winker *et al*. 2020), because several of the indices covered only a few years and were subject to missing values and irregular spacing. To address this, it was initially considered a fixed observation error approach by assuming a standard error for log(CPUE) of 0.25 for all indices. Thereafter, this restriction was relaxed and the model used the coefficient of variance of the indices as provided in the biomass units. To explore sensitivity, additional tests were conducted for alternative observation and process error variance settings (**Table 1**) as well sensitivity analyses exploring alternative weighting for the indices of abundance, as is common practice in many age-structured tuna assessments. The sensitivity tests also included an alternative assumption of a higher prior mean for r (3*0.059) (**Table 1**). To examine the sensitivity of the assessment results to the inclusion of individual CPUE indices, we iteratively re-fitted the models while excluding one index at the time and refitting the model (i.e. Jackknife index analysis).

2.3. Model diagnostics

The evaluation of model diagnostics follows the principles in Carvalho *et al*. (2021), who recommended objectively evaluating the base-case candidate model based on the following four model plausible criteria: (1) model convergence (2) fit to the data, (3) model consistency (retrospective pattern) and (4) prediction skill through hindcast cross-validation (Kell *et al*., 2017; 2021).

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 using a Markov Chains Monte Carlo (MCMC) simulation. 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 (Heidelberger and Welch, 1992) and Geweke (1992) and Gelman and Rubin (1992) diagnostics as implemented in the coda package (Plummer *et al*., 2006).

To evaluate the JABBA fit to the abundance index data, the model predicted values were compared to the observed indices. JABBA-residual plots were used to examine (1) color-coded lognormal residuals of observed versus predicted CPUE indices by 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 which highlights systematically auto-correlated residual patterns to evaluate the randomness of model residuals. In addition, it provides the root-mean-squared-error (RMSE) as a goodness-of-fit statistic. We conducted run tests 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 1-sided *p*-value of the Wald-Wolfowitz runs test (Carvalho *et al*. 2021). The run 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 time-series data points that fall outside the three-sigma limits (Anhøj and Olesen, 2014).

To check for model consistency with respect to the stock status estimates, it was also performed a retrospective analysis by removing one year of data at a time sequentially (n = 5), refitting the model 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 time series. To compare the bias between the models, we computed Mohn's (Mohn, 1999) rho (ρ) statistic and specifically the commonly used formulation Hurtado-Ferro *et al*. (2015).

Although the above model diagnostics are important to evaluate model convergence, the fit to the data and retrospective consistency, providing scientific advice should also involve checking that the model has prediction skill of future states under alternative management scenarios (Carvalho *et al*., 2021). To validate a model's prediction skill requires that the system be observable and measurable (Kell *et al*., 2021). Therefore, we applied a hindcasting cross-validation (HCXval) technique (Kell *et al*., 2016), where observations are compared to their predicted future values. HCxval is a form of cross-validation where, like retrospective analysis, recent data are removed, and the model is refitted with the remaining data, 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), 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 (Kell *et al*. 2021). 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. Unfortunately, in the case of northeastern porbeagle there is only one index of abundance in the last decade, the survey index conducted in 2018 and 2019, therefore the hindcasting cross-validation has limited used when evaluating the model with the end year of 2020.

3. Results and Discussion

The MCMC convergence tests by Heidelberger and Welch (Heidelberger and Welch, 1992) and Geweke (1992) and Gelman and Rubin (1992) were passed for all key parameters for both the reference run and the full model. 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 relative abundance indices, comprising the standardized CPUE indices FRA-LL, NOR-LL, and SPA-LL, and the composite Survey-Index, are shown for the Ref and full models in **Figure 3** and **Figure 4**, respectively. Both models appeared to fit the abundance trends reasonably well for the FRA-LL, for the other indices, the plots show some patterns of mostly positive residuals, although the statistical of normality distribution was not rejected (e.g. green color in residual plots). It was noticable the large residuals for the SPA_LL index, which were expected due to the large variance of this index as the average CV is about 89%, roughly twice the variance compared to the other indices. Nonetheless, run tests conducted on the log-residuals provided no evidence to reject the hypothesis of randomly distributed residual patterns for the three indices used in the ref and full models (**Figures 3** and **4**). The overall goodness-of-fit indicated a moderate precision of the fits from the reference runs (RMSE = 38.1%), which degraded for the "full" model (RMSE = 73.6%**)**, that incorporated the historical SPA-LL index (**Figure 5**). The residual patterns of the years 1987-2009 indicated a conflict between residuals from the FRA-LL and the Survey-Index plus the SPA-LL indices (**Figure 5**). This still resulted in an, on average, a negative trend in the residual pattern for most recent years, which is probably due only two years with observations in 2018 and 2019 and the contradictory pattern of the FRA-LL and SPA-LL at the end of 2008. The estimated process error deviations had a similar trend for the ref and full models, showing particularly negative variations in the last decade, likely associated with the almost null catches after 2010. The process deviations for the terminal year are close to zero and therefore to average expectation (**Figure 6**).

The median of marginal posteriors for r was estimated to be lower for the full model at 0.052 than for the ref model at 0.054 (**Table 2**). The scale of absolute estimates for K and B_{MSY} was similar for both scenarios (**Table 2**), which was also associated with a comparable posterior to the prior ratio of variance for K (PPRV = 0.116) of the ref model compared to the full model (PPRV = 0.064) as shown in **Figures 6-7**. This indicates that the indices may hold information about the total biomass to effectively update the posterior of K given the relatively vague prior. Estimates of the median MSY of 1286 (t) and 1166 (t) were close between the ref and full models. The posterior median of B_{2020}/B_{MSY} was slightly higher for the full model (0.511) compared to the ref model (0.397), while the posterior medians of F_{2020}/F_{MSY} (**Table 2**) were similar at about 1% for the current fishing mortality about the fishing mortality at MSY. The most notable differences between the two scenarios were therefore the increased uncertainty about the fit to the historical SPA-LL index and the slightly lower total surplus biomass scale for the full model (1,166 t) but the confidence bounds overlap completely between model scenarios (**Figure 8**).

The sensitivity runs indicated that the full model was largely robust to alternative assumptions about r, the terminal year (2020 or 2015), and the observation and process errors (**Figures 9** and **10**). Assuming a higher prior mean for r showed a limited effect on the current stock status in terms of B_{2020}/B_{MSY} and F_{2020}/F_{MSY}, and resulted in similar estimates of MSY (**Figure 10**). The only effects on stock status estimates were observed when the terminal year of the model was changed to 2010 or 2015, however, the confidence bounds of the estimates of MSY, B_{MSY}, and K have substantial overlap. Decreasing the process error to 0.1 resulted in a slightly higher r.

The Jackknife index analysis, applied to the full model by removing one index at a time, showed that removing the 'FRA-LL' was the most influential with regard to the stock status trajectories and fishing mortality trend particularly in the period 1970 to 2000. (**Figure 11**). Next, the NOR-LL index exclusion also had an intermediate effect on the trajectories of biomass and fishing mortality. If the survey index is excluded, the trend of biomass since 2010 changes, with a perception of a more rapid biomass increase in recent years (**Figure 11**).

The retrospective analysis applied over a horizon of five years to the ref and full models **(Figures 12-13)** indicated a better retrospective pattern for the ref model (**Figure 12)**, however, Mohn's estimates for both models fell within the acceptable threshold of -0.15 and 0.2 for long live species (Huerto-Ferro *et al.*, 2015). Except for Mohn's estimate of F/F_{MSY}, which was slightly lower than -0.15 in the full model (**Table 3**, **Figure 12**).

Hindcasting cross-validation results were limited as there is only one index in the last decade (the composite survey-index 2018/19) and suggested that the Survey-Index has some prediction skill for both the ref and full model scenarios as judged by the MASE scores < 1 (**Figure 14**), which provides a means to validate that short-term forecast are consistent with the 'future' observations that were unknown to the model (Kell *et al.*, 2021). The

MASE scores for the Survey index were close to 0.8 the prediction residuals appeared relatively small. Generally, MASE scores were marginally better in the ref model.

The surplus production phase plots show similar trends for both the ref and full scenarios, suggesting that the stock has been in an overfished state since the 1950s, but with fishing pressure dropping below F_{MSY} in the last decade (**Figures 8** and **15**). Catches have exceeded MSY for several years while biomass remained below B_{MSY} and from 1946 onwards and these high catches were no longer sustainable and decreased to less than 400 t at the beginning of the 1960s, however, new fisheries were developed, and by the mid-1970s catches were close to 2000 t, however, catches decreased even more and since the 1980s they oscillating around 500 t until 2010 when due to the ban on porbeagle retention, the catch has almost disappeared. The northeast porbeagle stock was overfished and experience overfishing (e.g. red quadrant) for most of the time series (1960-2010), which brought the biomass to very low levels (0.14 B/B_{MSY} in 2005). After 2010, with the implementation of strong measures, fishing pressure had substantially decreased, and the stock biomass has been able to increase albeit at a slow pace given the biological characteristics of the porbeagle. At present (2020), the probability that the stock is overfished and that overfishing is currently occurring (e.g. red quadrant) is estimated to be 0% for both the ref and full models (**Figure 15**). However, both scenarios indicate that the stock is still overfished (~ 0.45 B/B_{MSY}) but fishing mortality is currently well below F_{MSY}, (~ 0.01 F/F_{MSY}) (e.g. yellow quadrant) with a 98% probability. Given that in the last decade commercial catches stopped hence no fishery-dependent indices of abundance are available, and only 2 observations (2018/19) are available from a Survey in the Bay of Biscay area, there is high uncertainty in the recovery trend of this stock. Nonetheless, the sensitivity runs with terminal years of 2010 and 2015, showed consistent trends of biomass and fishing mortality compared to the ref and full model runs, as well the estimated reference points, indirectly supporting the stock status determination and conclusions in 2020.

In line with the recommendations by the 2021 ICCAT Working Group of Stock Assessment Methods (WGSAM), we evaluated the plausibility of two alternative JABBA model scenarios for the northeast porbeagle based on best practices in using model diagnostics (Carvalho *et al*., 2021). These criteria are: (1) model convergence, (2) fits to the data, (3) model consistency (e.g. retrospective patterns), and (4) prediction skill. Our results suggest that full represents the most plausible candidate model for the northeast porbeagle stock status. Specifically, the full model converged adequately, provided a robust fit to the data, and is largely consistent retrospectively, while including all available information on indices of abundance. However, due to the limited number of indices in the last decade, the hindcasting cross-validation test was limited, and no conclusion was reached about desirable prediction skill.

Acknowledgments

We would like to thank the ICES WKELASMO scientists for providing the input data and we greatly appreciate the collaboration of Dr. H. Winker for making available the JABBA software, R scripts, and its useful suggestions and comments for improving this work.

References

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

Anonymous. 2010. Report of the 2009 porbeagle stock assessments meeting. (Copenhagen, Denmark, June 22 to 27, 2009). Collect. Vol. Sci. Pap. ICCAT 65(6): 1909-2005.

Biais, G. 2022. Porbeagle abundance survey in the Bay of Biscay and in the Celtic Sea in 2018 and 2019. Working document ICES WKELASMO workshop.

Biais, G. 2022a. Standardized catch rates of Porbeagle in the northeast Atlantic ocean from French longliner data by trip. Working document ICES WKELASMO workshop.

Biais, G. 2022b. Standardized catch rates of Porbeagle in the northeast Atlantic ocean from Norwegian longliner logbooks. Working document ICES WKELASMO workshop.

Biais, G. 2022c. SPiCT runs for the northeast Atlantic porbeagle. Working document ICES WKELASMO workshop.

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. https://doi.org/10.1016/j.fishres.2016.09.018

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., and Methot R. D.2021. A cookbook for using model diagnostics in integrated stock assessments. Fisheries Research Volume 204. https://doi.org/10.1016/j.fishres.2021.105959

Castro, J.I. 1983. The sharks of North American waters. Texas A&M Univ. Press, College Station, TX, 180 pp.

Collette, B. B. and C. E. Nauen. 1983. FAO species catalogue, vol. 2, Scombrids of the world. FAO. Fisheries synopsis 125 (2): 137 p.

Compagno, L. J. V., 2001. Sharks of the world. An annotated and illustrated catalogue of shark species known to date. Volume 2. Bullhead, mackerel and carpet sharks (Heterodontiformes, Lamniformes and Orectolobiformes). FAO Species Catalogue for Fishery Purposes. No. 1, Vol. 2. Rome, FAO. 2001. 269p.

Compagno, L.J.V. 1984. FAO species catalogue. Vol 4. Sharks of the world. An annotated and illustrated catalogue of shark species known to date. Part 2. Carcharhiniformes. FAO Fish. Synop. (125, Vol. 4, Part 2),655 p.

Gelman, A., Rubin, D.B., 1992. lnference 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., 2015. 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

ICES 2022. Benchmark Workshop for selected elasmobranch stocks (WKELASMO). ICES Scientific Reports. Report. https://doi.org/10.17895/ices.pub.21025021.v1

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

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

ICCAT, 2017c. 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. http://www.sciencedirect.com/science/article/pii/S0165783616301540.

Kell, L. T., Sharma, R., Kitakado, T., Mosqueira, I., Winker, H., Cardinale, M., and Fu, D. 2021. Validation of stock assessment models using prediction skill: Is it me or my model talking? ICES Journal of Marine Science: https://doi.org/10.1093/icesjms/fsab104.

Martell, S., Froese, R., 2013. A simple method for estimating MSY from catch and resilience 504–514. https://doi.org/10.1111/j.1467-2979.2012.00485.x

Mejuto, J., M. Ortiz, B. Garcia-Cortes, J. Ortiz de Urbina, A.M. Ramos-Cartelle. 2010. Historical data and standardized catch rates of porbeagle (*Lamna nasus*) caught as by-catch of the Spanish surface longline fishery targeting swordfish (*Xiphias gladius*) in the Atlantic Ocean. Collect. Vol. Sci.Pap. ICCAT 65(6): 2006-2030.

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. https://doi.org/10.1006/jmsc.1999.0481

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.

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.

Ortiz, M., Mayor, C., Palma, C., and Taylor N.G. 2022. Review of the catch series for northeast porbeagle (*Lamna nasus*) as input for stock assessment. SCRS/2022/022.

Pedersen, M.W. and Berg, C.W. 2017. A stochastic surplus production model in continuous time. Fish and Fisheries 18(2): 226-243. https://doi.org/10.1111/faf.12174

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.

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., Carvalho, F., Kerwath, S., 2020. Age-structured biomass dynamics of north Atlantic shortfin mako with implications for the interpretation of surplus production models. Col. Vol. Sci. Pap. ICCAT 76, 316–336.

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-resilience method CMSY 74, 1836–1866.

Winker, H., Kerwath, S.E., Merino, G., Ortiz, M., 2019. 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. L., Tserpes, G., and Ortiz, M. 2020a. Development of Bayesian state-space surplus production model JABBA for assessing the Mediterranean swordfish (*Xiphias gladius*) stock. ICCAT Coll.Vol.Sci.Pap., 77: 508–536.

Winker, H., Mourato, B., Parker, D., R, S. A., Kimoto, A., and Ortiz, M. 2020b. Preliminary stock assessment of South Atlantic albacore tuna (*Thunnus alalunga*) using the Bayesian state-space surplus production model JABBA. Collective Volume of Scientific Papers ICCAT, 77: 352–376.

Winker H., Pinto C., and Kimoto A. 2021. Preliminary stock assessment of Mediterranean albacore (Thunnus alalunga) using the Bayesian State-Space Surplus Production Model JABBA. Collective Volume of Scientific Papers ICCAT, 78: 245-277.

Table 1. Model specifications of alternative productivity and variance parameters used in the sensitivity analysis for the full model. r-prior: mean value of the prior for r, σ_{obs_index}: observation error assumption for the indices, σ_{proc}: process error assumption.

Run	r-prior	σ_{obs_index}	σ_{proc}	Description
End year 2000	0.059	CV indx	est	Catch and index final year 2000
End year 2015	0.059	CV indx	est	Catch and index final year 2015
Fixed Process error	0.059	CV indx	0.1	Fixed process error
High r prior	0.177	CV indx	est	Increased r prior mean by factor of 3
low se Survey	0.059	CV indx	est	lower fixed observation error Survey

Table 2. 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 northeast porbeagle reference and the "full" (all 4 indices of abundance) model runs.

	Reference			**Full**		
Estimates	Median	2.50%	97.50%	Median	2.50%	97.50%
K	64,247	40,803	108,337	61,275	38,247	126,950
r	0.05434	0.03218	0.08449	0.05169	0.02883	0.07972
ψ (psi)	0.92452	0.67969	0.99702	0.92869	0.68636	0.99724
σ_{proc}	0.11800	0.05200	0.19400	0.11800	0.05600	0.19300
F_{MSY}	0.054	0.032	0.083	0.051	0.028	0.079
B_{MSY}	23,776	15,100	40,093	22,676	14,154	46,981
MSY	1,286.4	825.6	1,849.4	1,166.0	721.9	1,870.0
B_{1926}/K	0.901	0.629	1.162	0.91	0.64	1.16
B_{2020}/K	0.147	0.063	0.309	0.189	0.074	0.391
B_{2020}/B_{MSY}	0.397	0.17	0.836	0.511	0.201	1.057
F_{2020}/F_{MSY}	0.014	0.007	0.033	0.012	0.006	0.028

Table 3. Summary Mohn's rho statistic for the reference (Ref) and full models, computed for a retrospective evaluation period of five years. Estimates in green are within the suggested range of acceptable values for long live species (-0.15 to 0.20, Hurtado-Ferro et al., 2015).

	Stock Quantity					
Scenario	B	F	B/B_{MSY}	F/F_{MSY}	B/K	MSY
Ref	0.096	-0.086	0.144	-0.147	0.144	0.033
Full	0.120	-0.105	0.139	-0.173	0.139	0.060

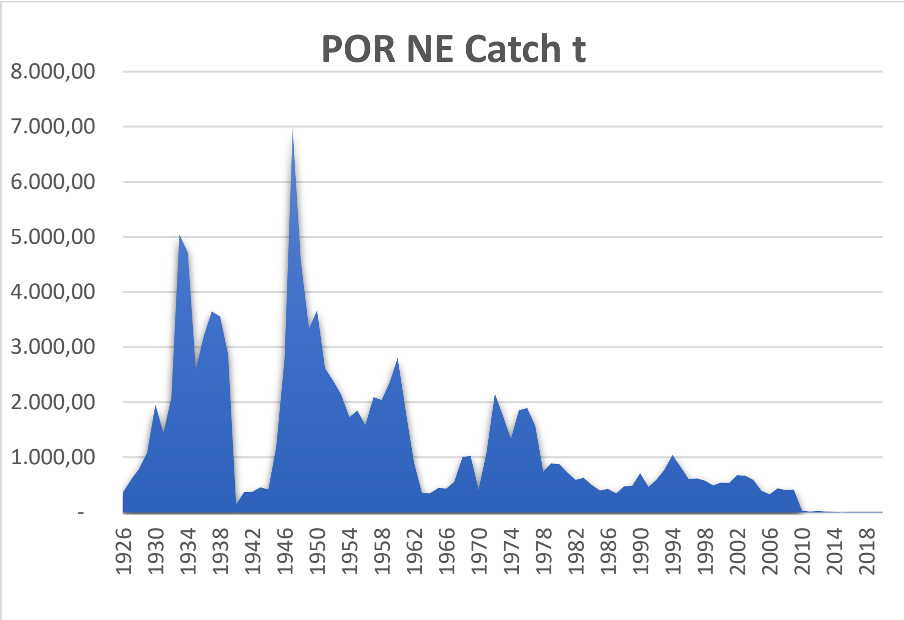


Figure 1. Catch time series 1926 – 2020 in metric tons (t) for the northeast Atlantic porbeagle.

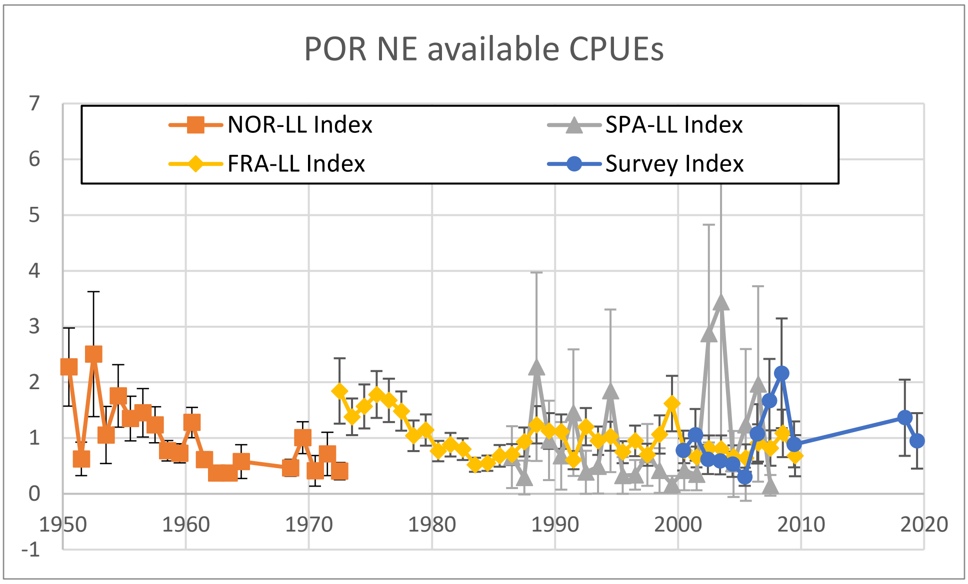


Figure 2. Time-series of relative indices of abundance scaled to their mean considered in the JABBA stock assessment for northeast porbeagle. Error bars represent the $\pm$ one standard error of the annual index estimates.

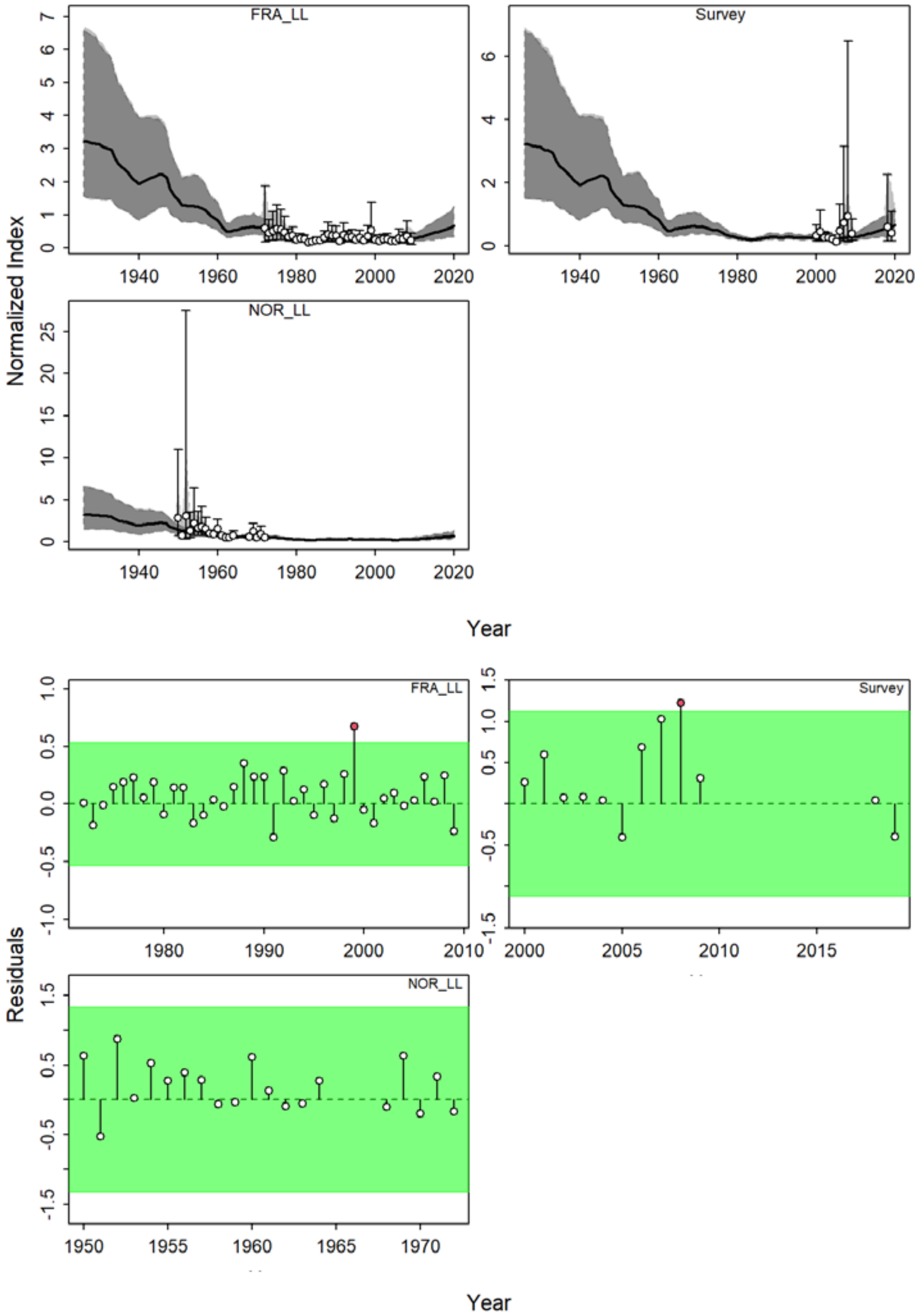


Figure 3. Time-series of observed (circle) with error 95% CIs (error bars) and predicted (solid line) CPUE (top) and Runs tests to quantitatively evaluate the randomness of the time series of CPUE residuals (bottom) for the northeast porbeagle reference scenario (Ref). On the top panel, the Dark shaded grey areas show 95% credibility intervals of the expected mean CPUE and light-shaded grey areas denote the 95% posterior predictive distribution intervals. On the bottom panel, green areas indicate no evidence of lack of randomness of time-series residuals (p>0.05) while red panels would 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 (3- sigma rule).

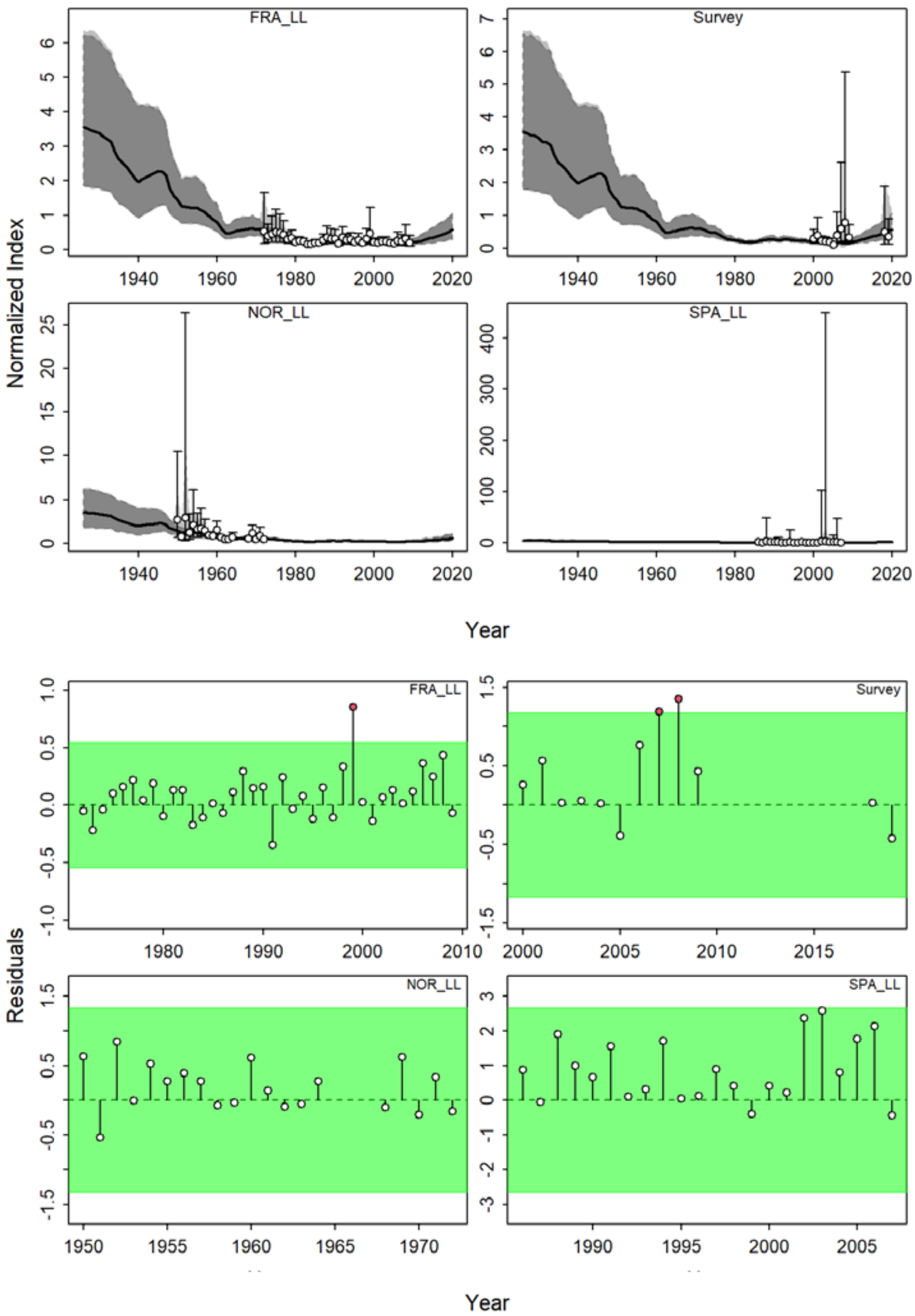


Figure 4. Time-series of observed (circle) with error 95% CIs (error bars) and predicted (solid line) CPUE (top) and Runs tests to quantitatively evaluate the randomness of the time series of CPUE residuals (bottom) for the northeast porbeagle "full" model. On the top panel, the Dark shaded grey areas show 95% credibility intervals of the expected mean CPUE and light-shaded grey areas denote the 95% posterior predictive distribution intervals. On the bottom panel, green areas indicate no evidence of lack of randomness of time-series residuals ($p>0.05$) while red panels would 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 (3- sigma rule).

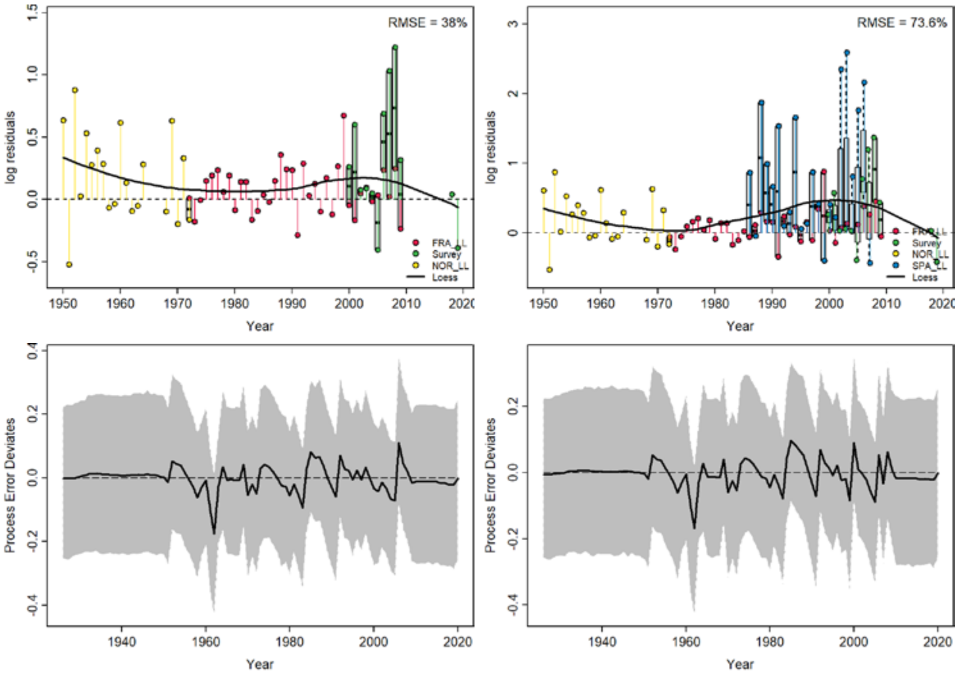


Figure 5. JABBA residual diagnostic plots for alternative sets of relative abundance indices examined for each scenario (Left column: reference scenario Ref; Right: “full” model) for northeast porbeagle. Top panels: Boxplots indicating the median and quantiles of all residuals available for any given year, and solid black line show the loess smoother through all residuals. Bottom panels: Process error deviates (median: solid line) with shaded grey area indicating 95% credibility intervals.

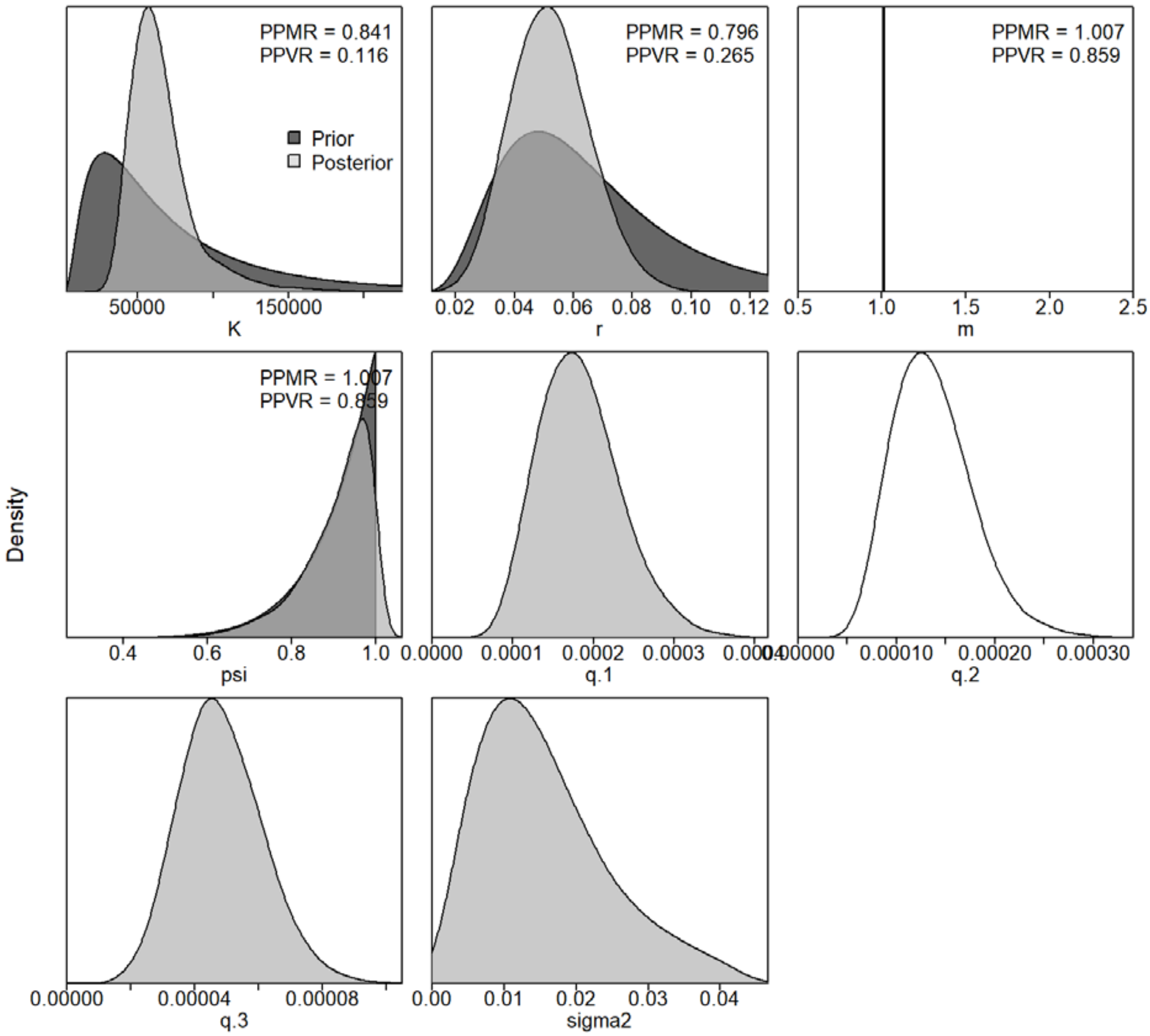


Figure 6. Prior and posterior distributions of various models and management parameters for the reference scenario for northeast porbeagle. PPRM: Posterior to Prior ratio of means; PPRV: Posterior to Prior ratio of variances.

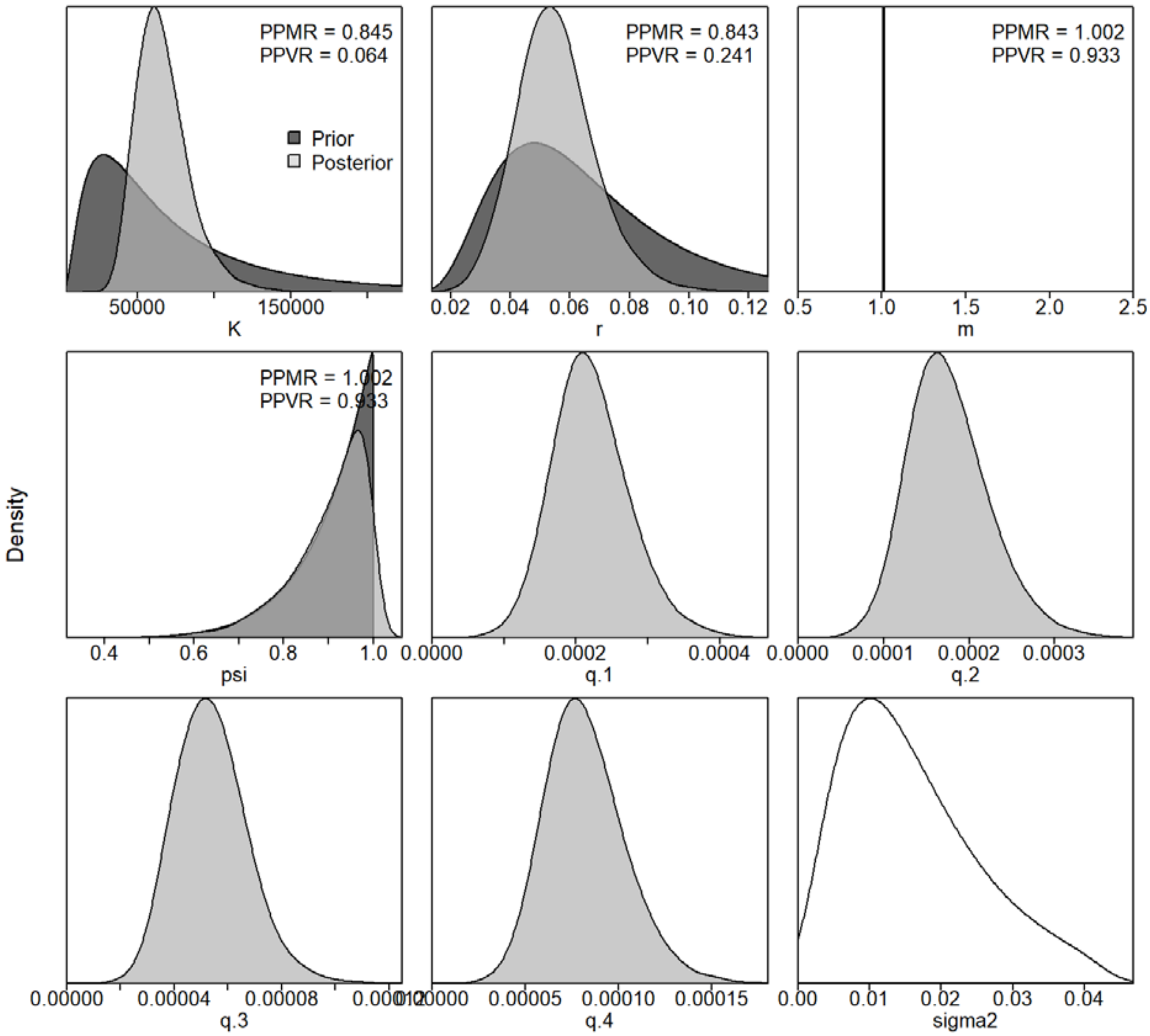


Figure 7. Prior and posterior distributions of various models and management parameters for the full model for northeast porbeagle. PPRM: Posterior to Prior ratio of means; PPRV: Posterior to Prior ratio of variances.

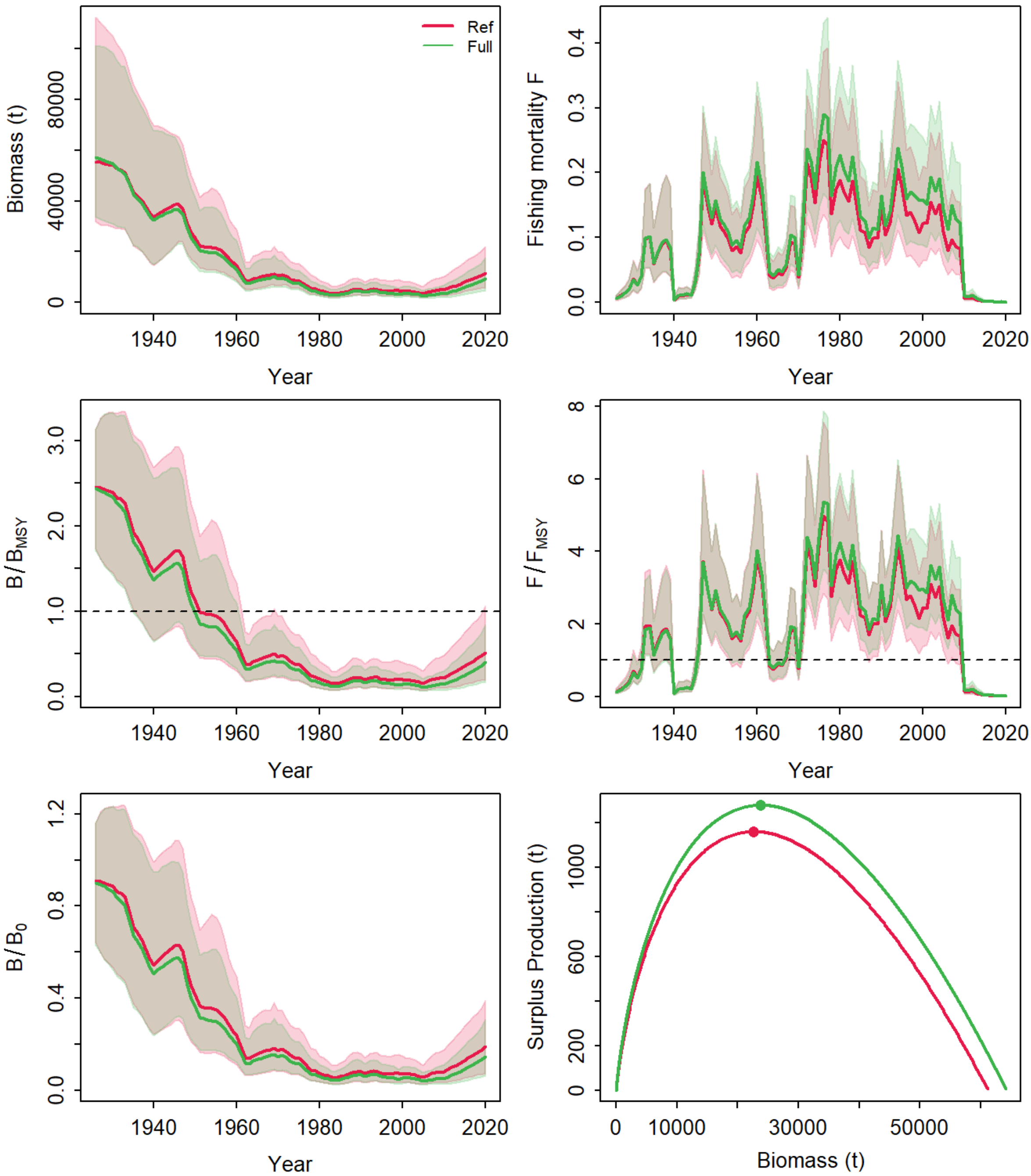


Figure 8. Comparison stock trajectory estimates for the northeast porbeagle reference scenario Ref (red) and Full model (green), showing trends in biomass and fishing mortality (upper panels), biomass relative to B_{MSY} (B/B_{MSY}), and fishing mortality relative to F_{MSY} (F/F_{MSY}) (middle panels) and biomass relative to K (B/K) and the surplus production curve (bottom panels).

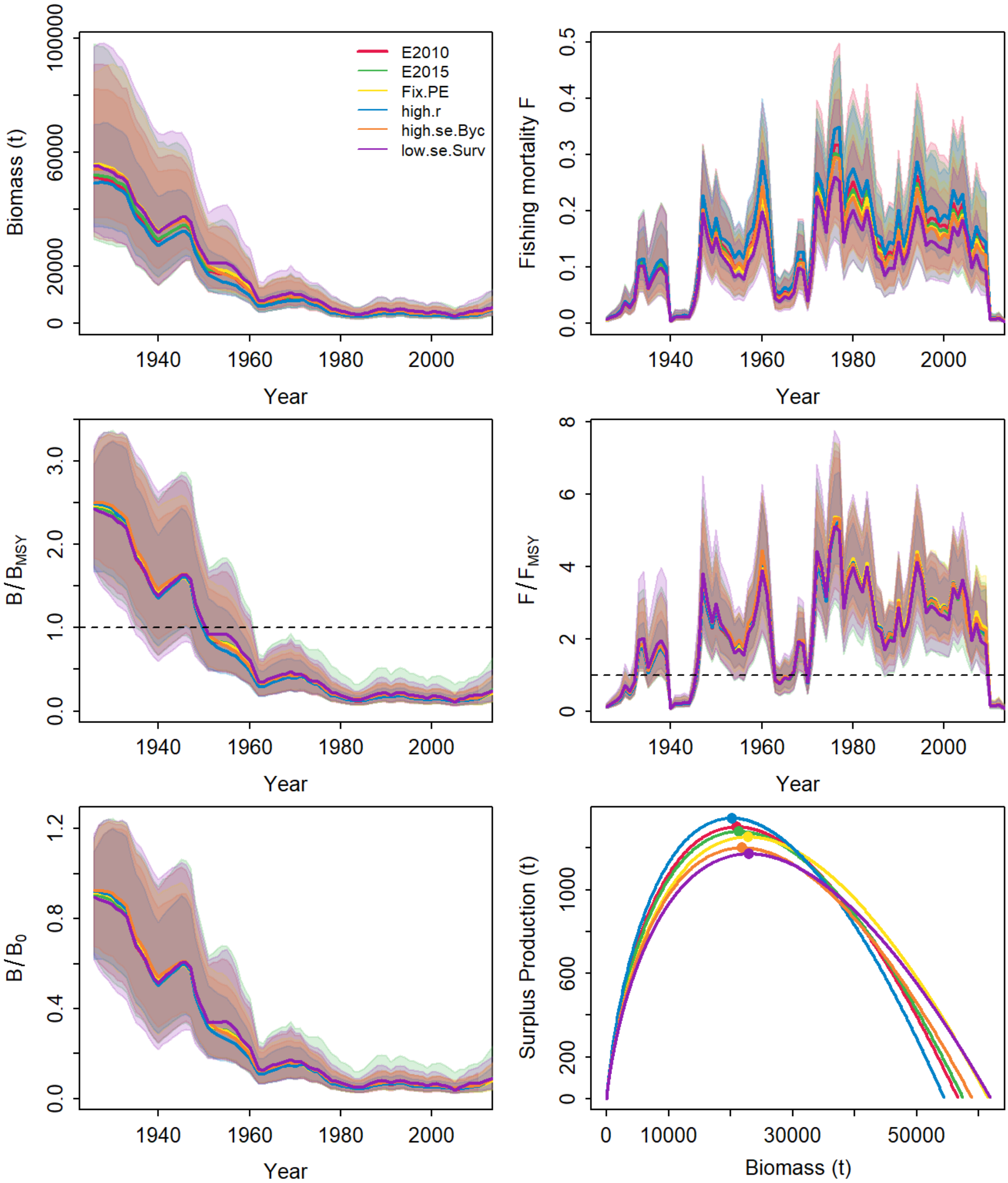


Figure 9. Sensitivity analysis performed on the "full" model for the terminal year, alternative productivity (*r*), and variance parameter assumptions described in **Table 2**. E2010, E2015: terminal year 2010 and 2015 respectively, high-r: increased *r* prior mean by a factor of 3, fixed.pe: process error fixed to 0.1, and low.se.Surv: lower fixed observation error for the Survey-Index. For comparison values of the reference (ref) model are also plotted.

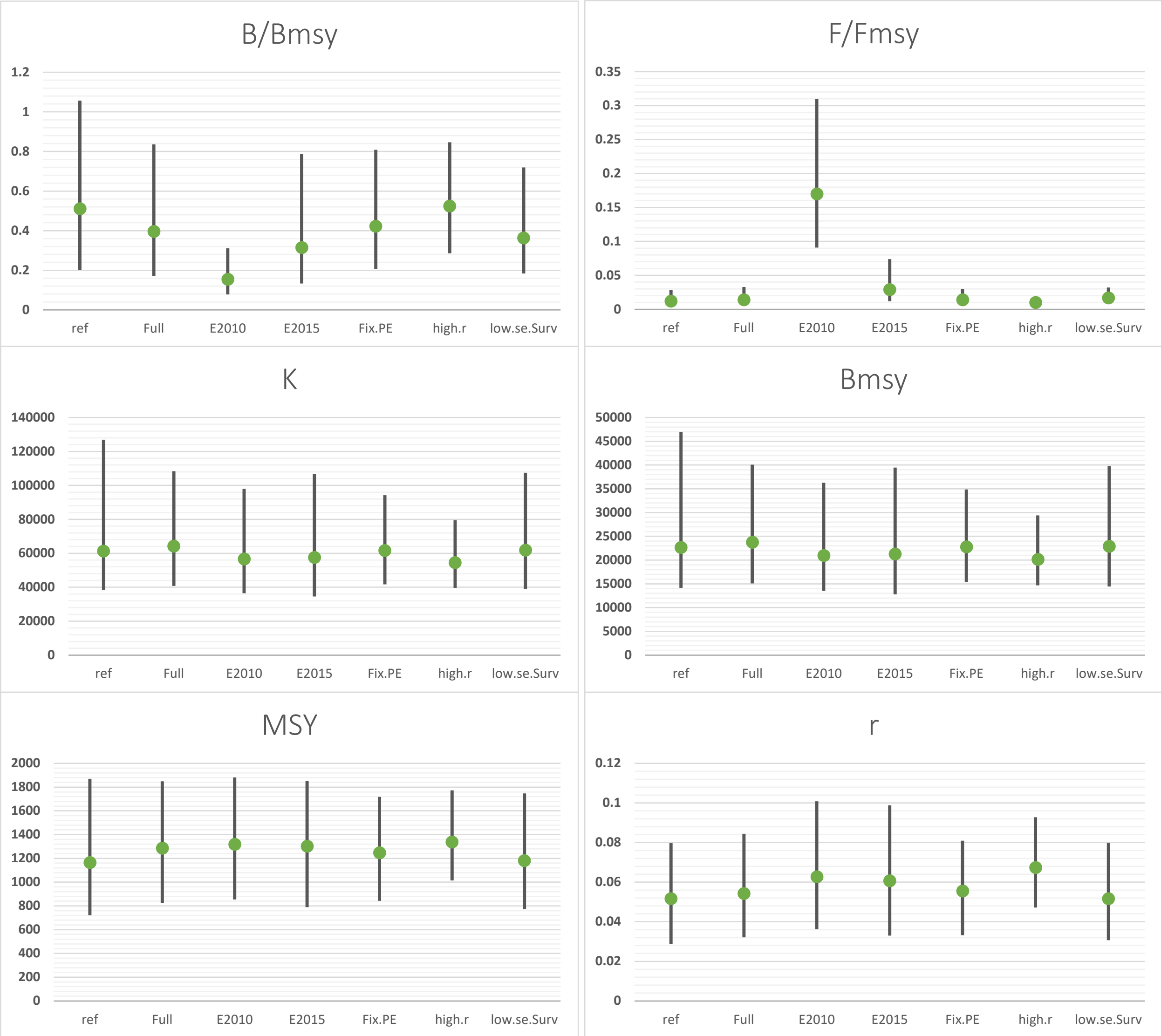


Figure 10. Sensitivity analysis done on the full model, plotted estimates of final year biomass over B_{MSY} (B/B_{MSY}), final year fishing mortality over F_{MSY} (F/F_{MSY}), Carrying capacity (*K*), Biomass at MSY (B_{MSY}), MSY, and *r* for the northeast porbeagle (see **Table 2** for details). Markers indicate the estimated median of the posterior with the upper and lower 95% confidence bounds. E2010, E2015: terminal year 2010 and 2015 respectively, high-r: increased r prior mean by a factor of 3, fixed.pe: process error fixed to 0.1, and low.se.Surv: lower fixed observation error for the Survey-Index. For comparison values of the reference (ref) model are also plotted

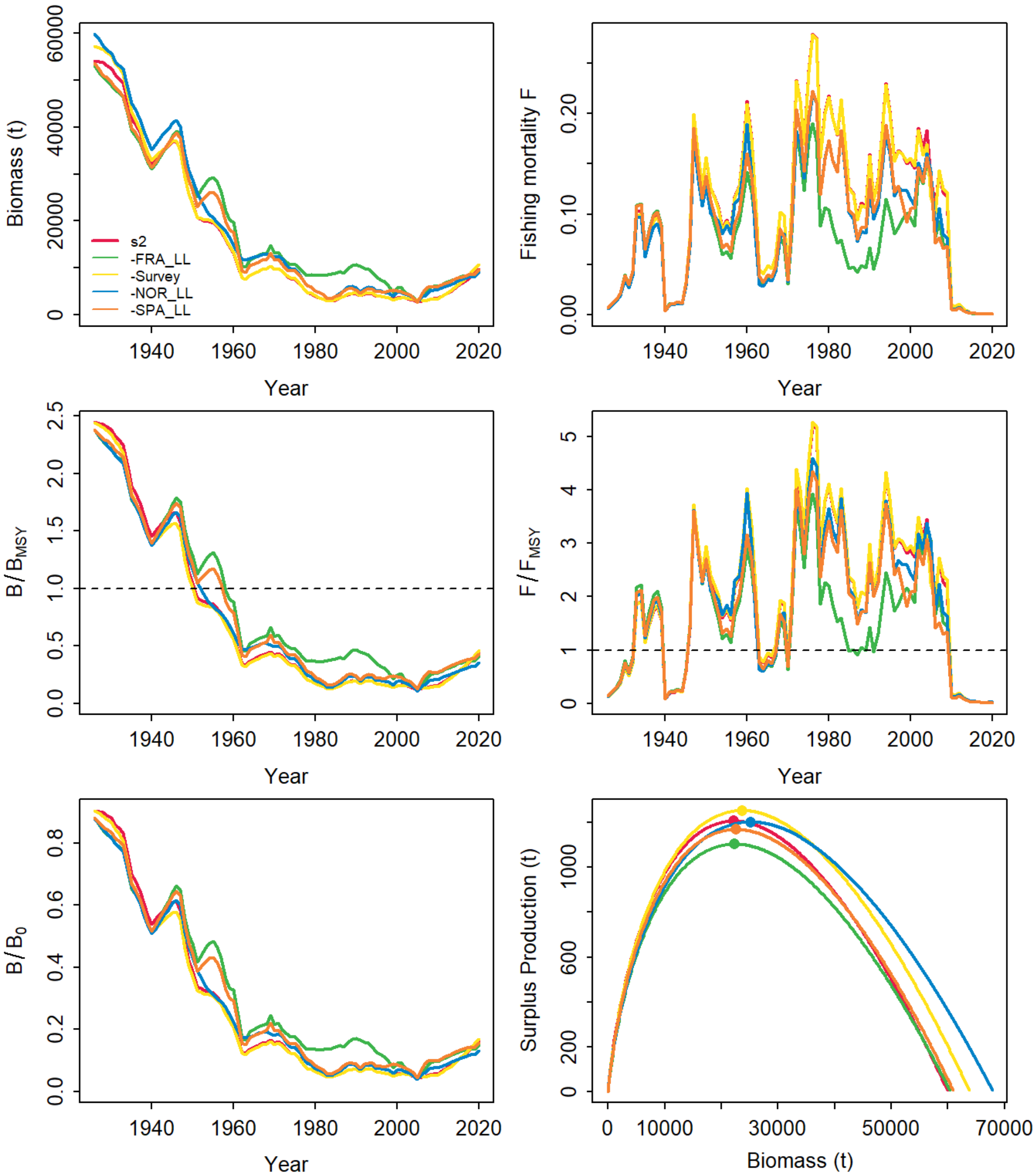


Figure 11. Jackknife index analysis performed on the full model, by removing one CPUE index at a time and predicting the trends in biomass and fishing mortality (top row), biomass relative to BMSY (B/BMSY), and fishing mortality relative to F_{MSY} (F/F_{MSY}) (middle row) and biomass relative to K (B/B_0) and surplus production curve (bottom row) from the Bayesian state-space surplus production model fitting to the northeast porbeagle catch and index series.

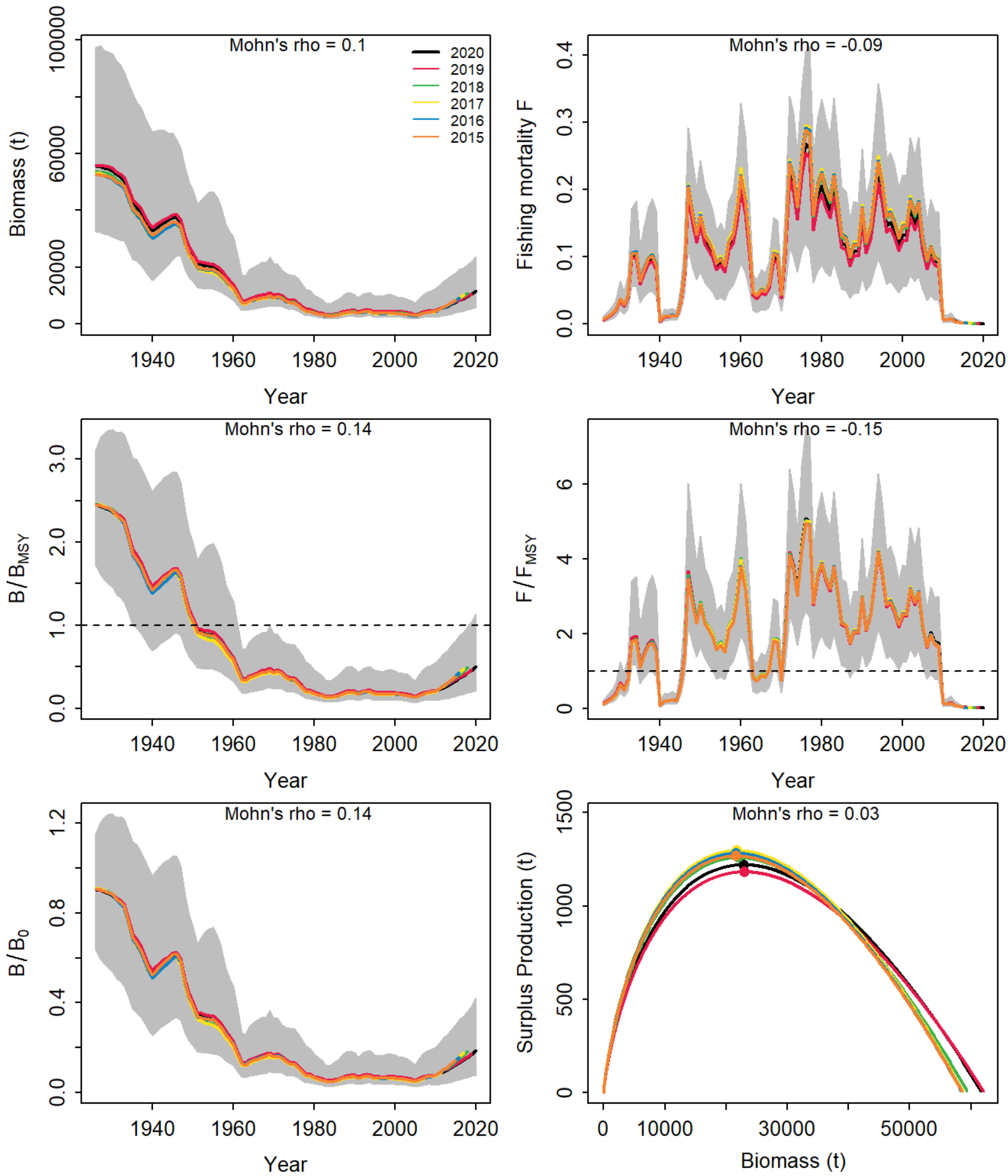


Figure 12. Retrospective analysis performed for the reference scenario, by removing one year at a time sequentially (n=5) and predicting the trends in biomass and fishing mortality (upper panels), biomass relative to B_{MSY} (B/B_{MSY}) and fishing mortality relative to F_{MSY} (F/F_{MSY}) (middle panels) and biomass relative to K (B/K) and surplus production curve (bottom panels) for each scenario from the Bayesian state-space surplus production model fits to northeast porbeagle.

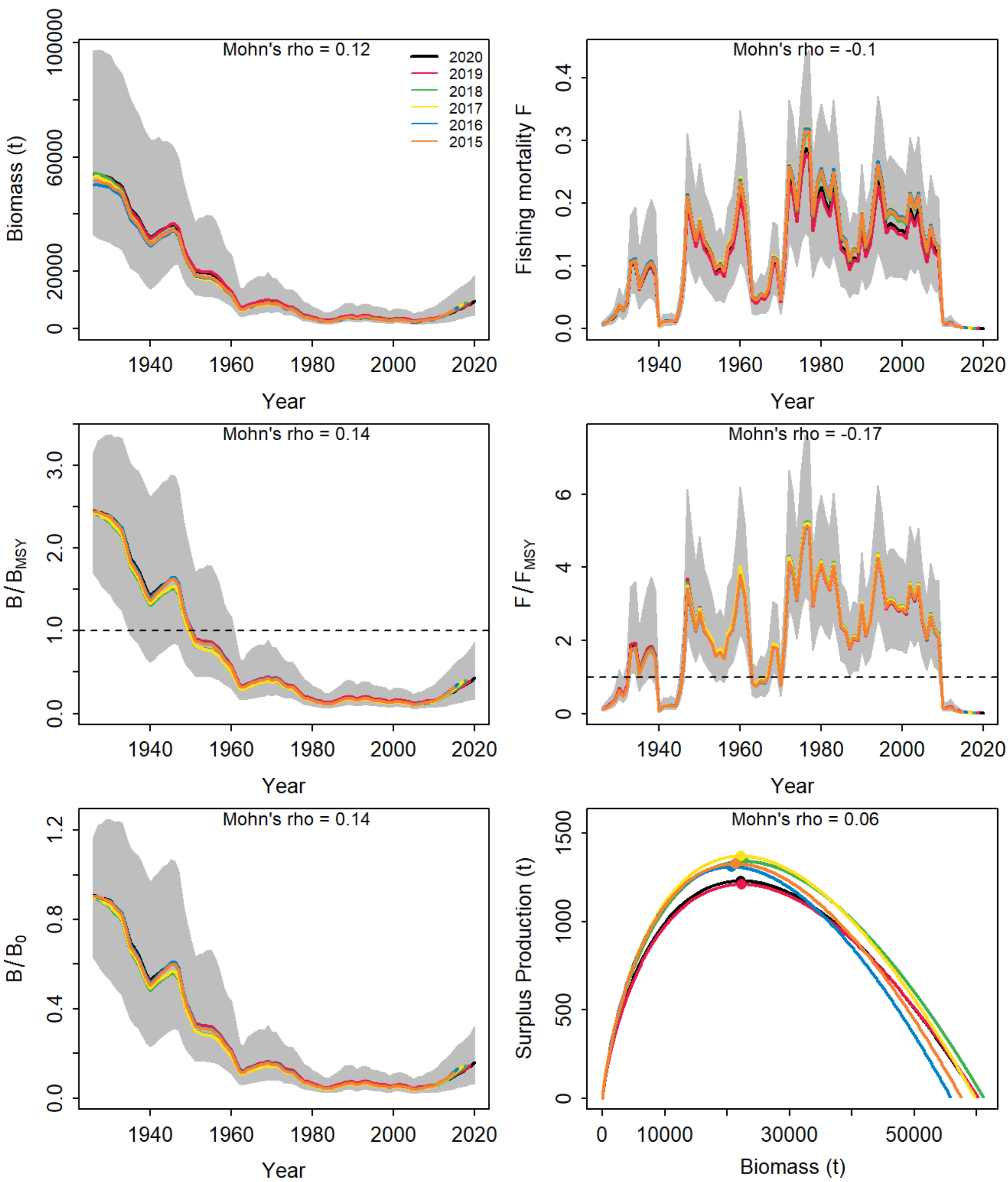
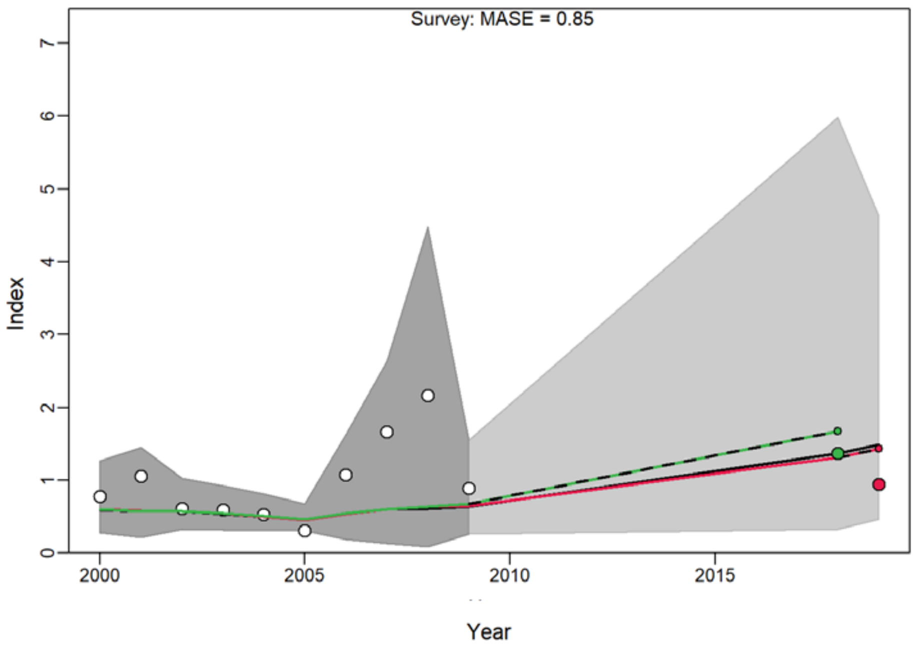


Figure 13. Retrospective analysis performed for the "full" model, by removing one year at a time sequentially (n=5) and predicting the trends in biomass and fishing mortality (upper panels), biomass relative to B_{MSY} (B/B_{MSY}) and fishing mortality relative to F_{MSY} (F/F_{MSY}) (middle panels) and biomass relative to K (B/K) and surplus production curve (bottom panels) for each scenario from the Bayesian state-space surplus production model fits to northeast porbeagle.



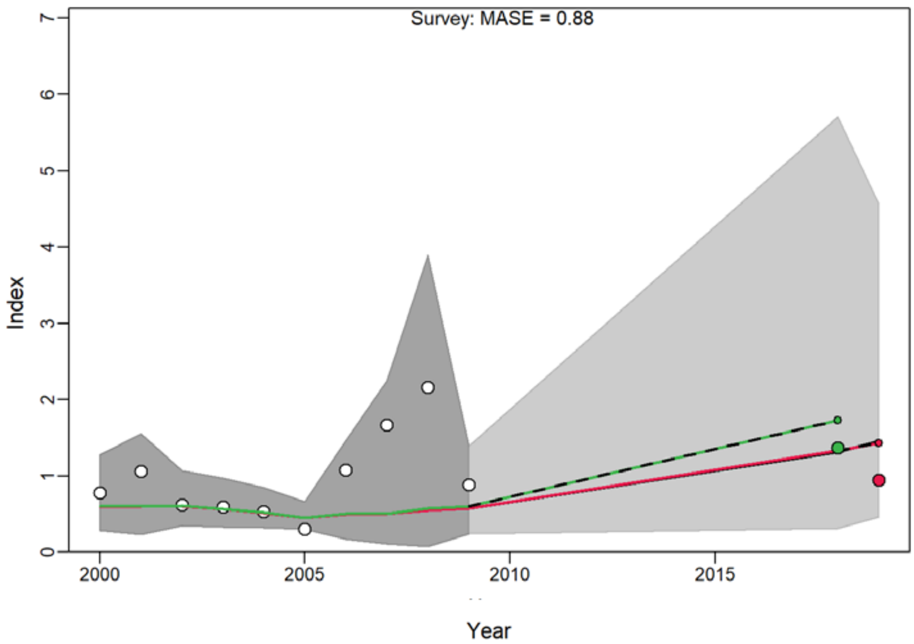


Figure 14. Hindcasting cross-validation results (HCxval) for the reference scenario (top) and the “full” model (bottom) for northeast porbeagle, showing one-year-ahead forecasts of the Survey CPUE values (2018-2019), performed with five model hindcast runs. The CPUE observations, used for cross-validation as prediction residuals, 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).

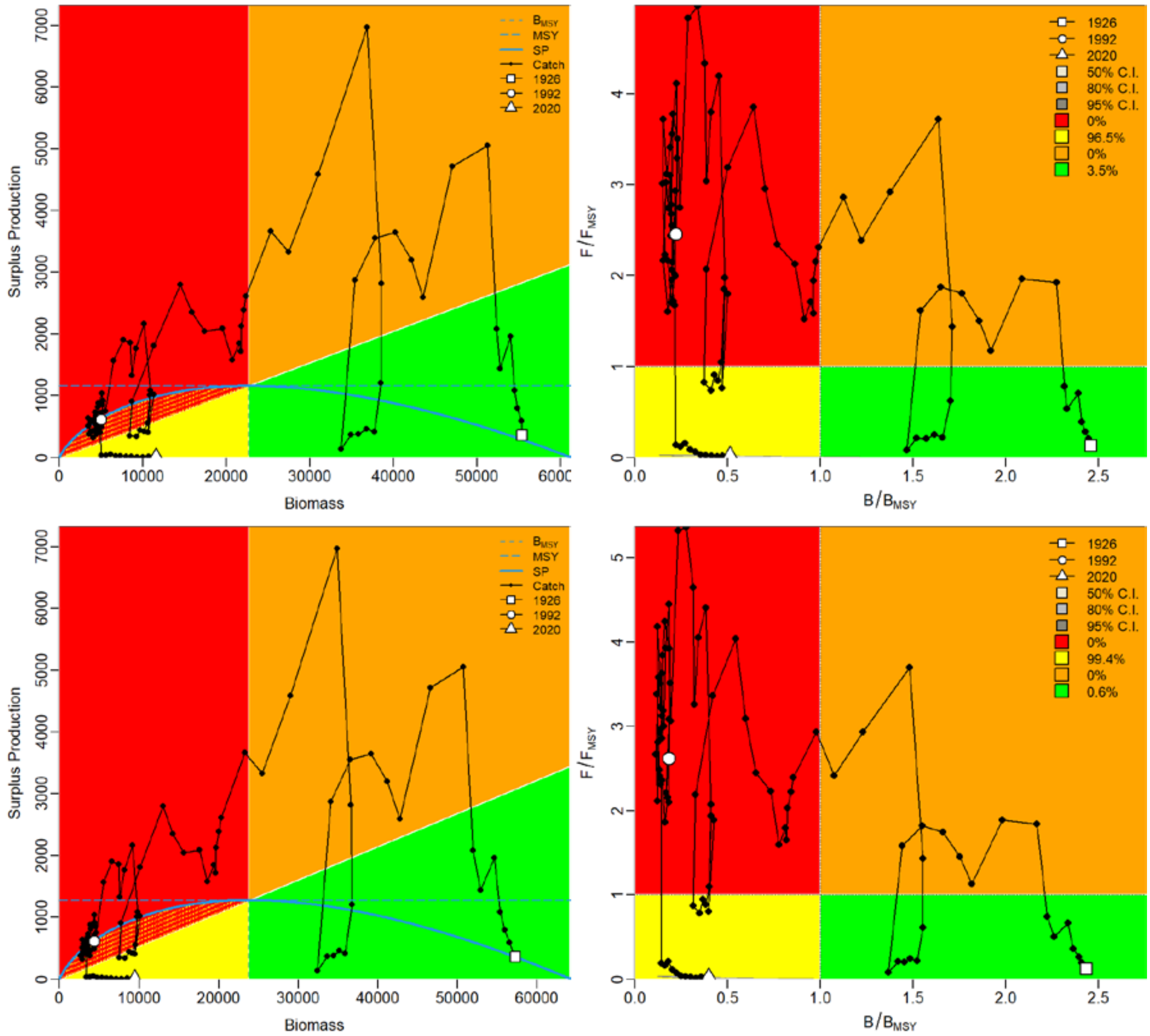


Figure 15. JABBA surplus production phase plot for the reference run (top row) and the "full" model (bottom row) showing trajectories of the catches in relation to the B_{MSY} and MSY (left column) and Kobe phase plot showing estimated trajectories (1926-2020) of B/B_{MSY} and F/F_{MSY} for the Bayesian state-space surplus production model for the northeast porbeagle (right column). The probability of terminal year points falling within each quadrant is indicated in the figure legend.