North Pacific Fisheries Commission

NPFC-2022-TWG CMSA05-WP06

Update of Virtual Population Analysis and State-Space Assessment Model for Operating Models of Chub Mackerel Stock Assessment in NPFC

Shota Nishijima, Kazuhiro Oshima, and Momoko Ichinokawa, Shinohara Naoto

Fisheries Resources Institute,
Japan Fisheries Research and Education Agency (FRA)
(Corresponding author: Shota Nishijima, nishijimash@affrc.go.jp)

Summary

The Technical Working Group for Chub Mackerel Stock Assessment (TWG CMSA) in NPFC has decided to use an operating model (OM) for comparing the performance of different four assessment model candidates. In this paper, we report the updated results of tuned virtual population analysis (VPA) and state-space assessment model (SAM), candidate stock assessment models proposed by Japan, under the determined scenarios to include biological uncertainties on natural mortality, weight, and maturity. We changed a few model configurations from the previous analysis to avoid overfitting and stabilize parameter estimation, which will be useful to the application of these models to pseudo-data generated from OM.

Note: Working document will be submitted to the NPFC 5th Meeting of Technical Working Group on Chub Mackerel Stock Assessment. This manuscript is an earlier draft version submitted to CMSA SWG OM02. Document not to be cited without author's permission.

Introduction

The Technical Working Group on Chub Mackerel Stock Assessment (TWG CMSA) in NPFC determined that (1) the candidates of stock assessment models (VPA, ASAP, KAFKA, and SAM) would be compared by an operating model (OM), and (2) the operating model would be based on POPSIM-A (NPFC 2019). POPSIM-A uses a stock assessment model as an operating model and, therefore, input data are needed for the development of operating models by fitting stock assessment model candidates (Deroba et al. 2014). At the TWG CMSA03 and CMSA04, we showed the preliminary results of tuned virtual population analysis (VPA) and state-space assessment model (SAM), which are candidate models proposed by Japan (Nishijima et al. 2020, 2021a). Although Japan updated natural mortality coefficients (M) based on the re-estimation of von-Bertalanffy growth curve (Nishijima et al. 2021b), it has been agreed to use the previous M estimators with no update for the stock assessments in the OM development (NPFC 2021). Moreover, the weights to the respective scenarios and the performance measures for the comparison of stock assessment models have also been determined at the TWG CMSA04 (NPFC 2021). Here, we report the updated results of VPA and SAM with a few changes in model configurations under the determined scenarios.

Model

Virtual population analysis

The VPA assumes no error in catch-at-age and conducts a backward calculation of population dynamics. We assumed that the age structure was from 0 to 6+ and used the Pope's approximation (Pope 1972) to estimate fish numbers and fishing mortality coefficients:

$$N_{a,y} = N_{a+1,y+1}\exp(M_a) + C_{a,y}\exp\left(\frac{M_a}{2}\right), \quad \text{if } a \leq 4 \qquad (1)$$

$$N_{5,y} = \frac{C_{5,y}}{C_{5,y} + C_{6+,y}} N_{6+}\exp(M_5) + C_{5,y}\exp\left(\frac{M_5}{2}\right), \qquad (2)$$

$$N_{6+,y} = \frac{C_{6+,y}}{C_{5,y} + C_{6+,y}} N_{6+}\exp(M_6) + C_{6+,y}\exp\left(\frac{M_6}{2}\right), \qquad (3)$$

where $N_{a,y}$ is the fish number at age a in year y and $C_{a,y}$ is the catch at age at age a in year y, and M_a is the natural mortality coefficients at age a. We here used $M_a = 0.41$ for all age classes under the scenarios A, C, and E, but instead used M_a = (0.57, 0.47, 0.38, 0.32, 0.28, 0.26, 0.24) from age 0 to 6+ for the scenarios B, D, and F, because the TWG CMSA has agreed to turn M values back to

previous ones (NPFC 2021). Three types of weight-at-age and maturity-at-age were used: (weighted-)average for the scenarios A and B, the highest for the scenarios C and D, and the lowest for the scenarios E and F (Table 1). The fish numbers in the terminal year (2019) were calculated from the fishing morality coefficients in the terminal year:

$$N_{a,2019} = \frac{C_{a,2019} \exp\left(\frac{M_a}{2}\right)}{1 - \exp\left(-F_{a,2019}\right)}. \quad (4)$$

The fishing mortality coefficients except for the terminal year were computed from

$$F_{a,y} = -\log\left\{1 - \frac{C_{a,y}}{N_{a,y}} \exp\left(\frac{M_a}{2}\right)\right\}. \quad (5)$$

We also assumed that the fishing mortality coefficient of plus group (A+) were identical to that of A−1:

$$F_{6+,y} = F_{5,y}. \quad (6)$$

We used 'ridge VPA' to stabilize the terminal F estimates, which included a ridge penalty (squared term of estimated parameters) in the optimization, i.e., penalized likelihood (Okamura et al. 2017):

$$\text{minimize} \quad (1-\lambda)\sum_k \sum_y \left[\frac{\ln(2\pi v_k^2)}{2} + \frac{\left\{\ln\left(I_{k,y}\right) - \ln\left(q_k X_{k,y}^{b_k}\right)\right\}^2}{2v_k^2}\right] + \lambda \sum_{a=0}^{5} F_{a,2019}^2, \quad (7)$$

where λ is the penalty coefficient ($0 < \lambda < 1$), $I_{k,y}$ is the value of index k in year y, v_k^2 is the variance of index k, q_k is the proportionality constant, and b_k is the nonlinear coefficient between index k and its associated estimates X_k. We used all six abundance indices from Japan (fleets 2-5), China (fleet 7) and Russia (fleet 9) following the agreement at the TWG CMSA03 (NPFC 2020). The Japanese abundance indices are of recruitment numbers (i.e. $X_{k,y} = N_{0,y}$) (summer survey index for fleet 2 and autumn survey index for fleet 3) and of spawning stock biomass (i.e. $X_{k,y} = SSB_y$) (dip-net fishery index for fleets 4 and egg survey index for fleet 5), while the Chinese and Russian indices were used by assuming $X_{k,y} = \sum_{a=0}^{6+} S_{a,y} \times B_{a,y}$ (i.e., vulnerable stock size) based on the intersessional agreement (SWG_OM01 Summary), where $S_{a,y}$ is the selectivity at age in year y and $B_{a,y}$ is the biomass at age in year y. The selectivity at age was calculated so that the maximum fishing mortality coefficient was equal to one: $S_{a,y} = F_{a,y}/\max(\boldsymbol{F_y})$. We estimated $b_{k,y}$ to treat hyperstability or hyperdepletion for all of the six abundance indices. To avoid overfitting and stabilize parameter estimation, we assumed the variances v_k^2 were equal between the two recruitment indices, between

the two SSB indices, and between the Chinese and Russian fishery indices. We selected the penalty coefficient so as to minimize the absolute value of Mohn's rho (Mohn 1999) of average fishing mortality coefficient in the five-year retrospective analysis:

$$\rho = \frac{1}{5}\sum_{i=1}^{5}\left(\frac{\hat{F}^{R}_{2019-i} - F_{2019-i}}{F_{2019-i}}\right), \tag{8}$$

where $\hat{F}_{2019-i}$ is the estimate of average fishing mortality coefficients in year $2019-i$ using the full data. Therefore, the ridge VPA can reduce a retrospective bias to some extent. The selected λ were 0.77 for the scenario A, 0.48 for B, 0.76 for C, 0.75 for D, 0.76 for E, and 0.39 for F.

State-space assessment model

The basic model structure of SAM followed the original one (Nielsen and Berg 2014). Numbers at age a in year y are described as:

$$\log(N_{0,y}) = \log(N_{0,y-1}) + \eta_{0,y}, \tag{9}$$

$$\log(N_{a,y}) = \log(N_{a-1,y-1}) - F_{a-1,y-1} - M_{a-1,y-1} + \eta_{a,y}, \qquad 1 \le a \le 5 \tag{10}$$

$$\log(N_{6+,y}) = \log\left(N_{5,y-1}e^{-F_{5,y-1}-M_{5,y-1}} + N_{6+,y-1}e^{-F_{6+,y-1}-M_{6+,y-1}}\right) + \eta_{6+,y}\,, \tag{11}$$

where $\eta_{a,y}$ is the process error at age a in year y. Although we applied the hockey-stick (HS) stock-recruit relationship at the previous time (Nishijima et al. 2021b), estimation of HS stock-recruitment relationship makes parameter estimation unstable, which will pose difficulty to the application of SAM to pseudo data generated from OM. Therefore, we instead used a random walk of recruitment (eqn. 9) and applied a post-hoc analysis of stock-recruitment relationship in the same way as VPA (see below). We used the same six datasets (scenarios) of different natural mortality coefficients (M) and biological parameters as in VPA (Table 1). We assumed different magnitudes of the process errors for age 0 and older: $\eta_{0,y} \sim N(0, \omega_R^2)$, $\eta_{a,y} \sim N(0, \omega_{S,a}^2)$ ($a > 0$). We fixed the variance for the ages older than 0 at a small value ($\omega_{S,a}^2 = 0.0001$) because it was unlikely to converge when estimating this parameter. (Table 1). The fishing mortality coefficient was assumed to follow the multivariate random walk:

$$\log(\boldsymbol{F_y}) = \log(\boldsymbol{F_{y-1}}) + \boldsymbol{\xi}_y\,, \qquad \text{if } y \neq 2011 \tag{12}$$

where $\boldsymbol{F_y} = (F_{1,y}, \ldots, F_{A-1,y})^T$, $\boldsymbol{\xi_y} \sim \mathrm{MVN}(0, \boldsymbol{\Sigma})$, and $\boldsymbol{\Sigma}$ is the variance-covariance matrix of

multivariate normal distribution (MVN). The diagonal elements of matrix $\mathbf{\Sigma}$ were σ_a^2, while off-diagonal elements were assumed to be $\rho^{|a-a'|}\sigma_a\sigma_{a'}$ ($a \neq a$'). $\rho^{|a-a'|}$ corresponded to the correlation coefficient of F between ages a and a', and this assumption reflected the decrease in correlation with increasing age difference. In addition, we assumed $F_{6+,y} = F_{5,y}$ in accordance with tuned VPA. The random walk was omitted in 2011 because the fishing effort on chub mackerel possibly greatly decreased since the previous year because of the Great East Japan Earthquake and tsunami in March 2011. We found positive retrospective bias in stock abundance and negative bias in fishing mortality if assuming a random walk in 2011 (Nishijima et al. 2020).

The SAM estimated the errors in catch-at-age in a lognormal fashion:

$$\log(C_{a,y}) = \log\left(\frac{F_{a,y}}{F_{a,y}+M_{a,y}}\left(1-\exp\left(-F_{a,y}-M_{a,y}\right)\right)N_{a,y}\right)+\varepsilon_{a,y} \quad (13)$$

where $\varepsilon_{a,y}$~N$(0,\tau_a^2)$. We used the six indices in the same way as the VPA:

$$\log(I_{k,y}) = \log\left(q_r X_y^{b_k}\right)+\eta_{k,y}\,, \quad (14)$$

where $\eta_{k,y}$ is the measurement error of index k in year y: $\eta_{k,y}$~$N(0,\ \nu_k^2)$.

The SAM has to estimate many parameters. We then imposed the following constraints to stabilize estimation and avoid overfitting:

$$\omega_{S,a} = \omega_S\ \left(\forall a\ (a>0)\right), \quad (15)$$

$$\sigma_0 = \sigma_1, \sigma_2 = \sigma_3 = \cdots = \sigma_A\,, \quad (16)$$

$$\tau_2 = \tau_3, \tau_5 = \tau_{6+}. \quad (17)$$

In addition, we added two constraints and assumptions on the fitting to the abundance indices so that the likelihood that a convergence error will be unlikely to occur when we will apply the SAM to pseudo data generated from OM. First, we assumed the variances ν_k^2 were equal between the two recruitment indices, between the two SSB indices, and between the Chinese and Russian fishery indices, as in VPA. Second, we fixed at b_k = 1 for the two SSB indices because these two indices were relatively proportional to SSB although either index (fleet 4, egg survey) were significantly nonlinear under some scenarios. As a sensitivity trial we also analyzed the models with b_k = 1 for all abundance indices under the basecase scenarios (A and B), and showed its results in Figs. S1-2.

In contrast to VPA, SAM regards state variables, such as numbers at age and F at age, as latent random variables, which requires complex, difficult numerical integral calculation for many

random effects. We therefore used Template Model Builder (TMB: Kristensen, Nielsen, Berg, Skaug, & Bell, 2016), an R package which enables fast computation for latent variable models. We also applied a bias correction method of mean values because random effects were estimated by logarithmic scale (Thorson and Kristensen 2016).

Retrospective analysis

We conducted a retrospective analysis as a diagnostic of robustness and estimation bias. However, the Chinese and Russian abundance indices have short years (five and four years, respectively), which led to an estimation error while removing recent-year data. We removed these indices when the available years became less than three in the retrospective analysis. We fixed λ in the equation 7 for the retrospective analysis. We calculated Mohn's rho for biomass, SSB, recruitment, and average fishing mortality coefficient. We showed results of other diagnostics regarding residuals in fitting of prediction to samples in supplementary figures (Figs. S3-14).

Basic biological parameters and biological reference point

We first calculated F%SPR and F0.1 as biological reference points that do not use a stock-recruitment relationship. We estimated them using M at age, weight at age, maturity at age, and estimated F at age of each year. We also tried to calculate Fmax but could not find a solution for some years. We therefore excluded Fmax in this document.

We then computed basic biological parameters and biological reference points that are based on a stock-recruitment relationship. As both VPA and SAM did not internally estimate a stock-recruitment relationship, we made a post-hoc analysis of the HS stock-recruitment relationship , using SSB and the number of recruits estimated in VPA and SAM:

$$\log(R_y) = \begin{cases} \alpha \times SSB_y, & SSB_y < \beta \\ \alpha \times \beta, & SSB_y \geq \beta \end{cases}, \quad (18)$$

where α is the slope at the origin and β is the breaking point at SSB. We estimated these parameters by assuming log-normal error in the number of recruits. We used the averages of M at age, weight at age, maturity at age, and estimated F over years to derive SSB0 (SSB at F=0), R0 (recruitment at F=0), steepness (h), MSY, SSBmsy (SSB that allows for MSY), %SPRmsy (percentage of spawner per recruit at the MSY level relative to that at F=0), SSBmsy/SSB0, and F/Fmsy. The definition of steepness in the HS function depended on that by Punt et al. (2014): $h = 1 - \beta/SSB0$. The MSY-based reference points were obtained by assuming that selectivity at age is the one that obtained by the average of F at age over years. Although we here estimated the

'deterministic' MSY-based reference points for simplicity, it is forewarned that the deterministic MSY-based reference points are necessarily more optimistic than 'stochastic' MSY-based reference points that are computed by a population dynamics simulation incorporating stochasticity including recruitment variability (Okamura et al. 2020). As the biological parameters related to growth and maturity in chub mackerel are much time-varying due to density dependence (Watanabe and Yatsu 2004, 2006, Kamimura et al. 2021), the biological reference points would be time-varying (Miller and Brooks 2021). Therefore, we also calculated per-year reference points using each year's data and estimates.

Results

Estimates of abundances and fishing mortalities

In SAM, the past estimates of total biomass were almost the same among scenarios, whereas the past SSB and recruitment were, respectively, higher and lower in the scenarios A, C, and E (age common *M*) than in the scenarios B, D, and F (age-specific M) (Table 2; Fig. 1, left panels). The recent estimates were relatively different among scenarios: especially, SSB were estimated to be much higher in the scenarios C and D (highest weight and maturity) than the other scenarios (Fig. 1, left and second top panel). This is because although the total fish numbers were the highest under the scenarios E and F (lowest weight and maturity), the much higher weight- and maturity-at-age under the scenarios C and D led to the highest SSB (Table 2). AIC values were 1002.9 for the scenario A, 1004.4 for B, 1017.2 for C, 1018.6 for D, 1000.2 for E, and 1002.9 for F.

In VPA, the past estimates of abundances and exploitation rates were not so different from those in SAM (Table 2). However, the abundance estimates (biomass, SSB and recruitment) in recent years were much higher in VPA than in SAM, while the exploitation rates in recent years were lower in VPA than in SAM (Fig. 1). Exceptionally, VPA under the scenarios E and F estimated a similar trend of SSB as SAM. The scenarios C and D (the highest weight and maturity) estimated higher total biomass and SSB than the other scenarios (Fig. 1, right panels). This is also because the effects of the highest weight- and maturity-at-age with the scenarios C and D increased biomass and SSB (Table 2). The ridge VPA, unlike SAM, cannot be compared by AIC because of using a penalized likelihood rather than a marginal likelihood.

In addition, the selectivity at age was moderately different between models especially for recent years: SAM estimated relatively higher selectivity for ages 1 and 2, but VPA estimated relatively higher selectivity for ages 4 and 5 in the latest two years (Fig. 2).

Fitting to the abundance indices

The patterns of fitting to the abundance indices were different between SAM and VPA with all scenarios. VPA fitted better to the higher recruitment indices (fleets 2 and 3) in 2016 and 2018 than SAM (Figs. 3-8). By contrast, SAM fitted better to a SSB index (fleet 4) in recent years than VPA (Figs. 3-8). These differences possibly caused much higher abundances of VPA estimates than those of SAM estimates in recent years (Fig. 1).

The two recruitment indices (fleets 2 and 3) were significantly hyper-stable ($b_k < 1$) under most scenarios in SAM and VPA, and had more nonlinear relationships with estimated abundances in SAM than in VPA (Table3, Figs. 9-14). By contrast, the SSB index of fleet 5 (egg survey index) was significantly hyper-depleted ($b_k > 1$) under the scenarios A-F in VPA whereas it was assumed to be proportional in SAM (Table 3). The SSB index of fleet 4 (dip-net fishery index) showed significantly nonlinear relationships under the scenarios C, D, and F in VPA (Table 3, Figs. 9-14). It is noteworthy that the indices of the fleets 7 and 9 were extremely hyper-stable except for the scenario D in VPA, although their p values were not statistically significant for SAM because of large standard error due to small sample sizes (Table 3, Figs. 9-14). This indicates that these indices are not informative of abundances at least with this usage.

Retrospective analysis

In the retrospective analysis with SAM, there were no serious biases under all scenarios (Table 4, Figs. 15-20), although a positive bias in the fishing mortality coefficient was found under the scenario C (Table 3, Fig.17). In the retrospective analysis of VPA, positive biases in biomass and SSB and negative biases in the fishing mortality coefficient were observed except for the scenario D (Table 4, Figs. 21-26).

Basic biological parameters and biological reference points

When using the averages of biological parameters and fishing mortality coefficients over years, the basic biological parameters (SSB0, R0) and MSY-based reference points (SSBmsy, MSY) that were associated with absolute values related to abundances were much larger in VPA than in SAM for the scenarios except for E and F (lowest weight and maturity) because higher estimates in recruitment and SSB led to higher parameter values of the breaking point (β) (Table 5, Fig. 27). The parameter α (slope of HS relationship at the origin), steepness, and the relative reference points

(%SPRmsy, SSBmsy/SSB0, F/Fmsy) were almost same between SAM and VPA, while the difference of M (age-common or age-specific) had a larger impact on these values (Table 5), suggesting that the setting of M will be important for evaluating the productivity of this stock and considering biological reference points.

The per-year analysis of %SPR and F relative to F0.1 revealed that the fishing impacts had been generally high until the 2000s, but decreased in the 2010s (Fig. 28, top and middle panels). F was generally higher than Fmsy until the 1990s but became lower than Fmsy thereafter in both SAM and VPA (Fig. 28, bottom panels). However, F relative to Fmsy has moderately increased since 2015 for all scenarios in SAM and the scenario E and F in VPA (Fig. 28, bottom panels). The recent increase in F/Fmsy is closely related to the declines of maturity at age and weight at age, which reduces SPR0 and %SPRmsy (Fig. 29).

The temporal trends of SSB/SSB0 were totally different between SAM and VPA and among the six scenarios as well (Fig. 30, top panels). SSB/SSBmsy in the past years (e.g., 1970s) were also much different between SAM and VPA, but interestingly its recent estimates were relatively similar: in both SAM and VPA, SSB are around SSBmsy under the scenario A, B, E and F or over SSBmsy under the scenarios C and D (Fig. 20, bottom panels). This implies that we should take care of not only recent estimates but also past estimates when we will use a depletion statistic such as SSB/SSB0 and SSB/SSBmsy as a performance measure for evaluating the stock assessment models. The past estimates SSB/SSBmsy in SAM were more robust among the scenarios than those in VPA.

Discussion

We showed that there is a large difference of abundance estimates for recent years between VPA and SAM (Fig. 1). The reason for this is that VPA allows a flexible change in annual selectivity at age, whilst SAM estimates a gradual change in selectivity at age under the assumption of random walk of F at age. This difference of model configuration caused lower selectivities for ages 1-3 in VPA than in SAM (Fig. 2), leading to the recent inflation of recruitment estimates in VPA compared to SAM. In addition, SAM is less likely to estimate outliers of recruitment than in VPA because of random-walk recruitment. VPA tended to (over-)fit to the high values of recruitment indices (fleets 2 and 3) in 2016 and 2018 (Figs. 3-8). By contrast, SAM fitted less to the recruitment indices, but fitted better to a SSB index (fleet 4). The difference of model configuration and the data conflict between the recruitment indices and the SSB index caused the large difference of abundance estimates between SAM and VPA in recent years.

The abundance indices from the fleets 7 and 9 exhibited extreme hyperstability in both SAM and VPA (Table 3, Figs. 9-14). This is because those index values little change during the short time period when the data are available. The hyperstability can be a cause of overfishing because the abundance index is kept at a certain level even the actual abundance declines (Hutchings 1996, Rose and Kulka 1999), and therefore, we should discuss the way to resolve this problem towards the benchmark stock assessments. Estimating the nonlinearity coefficients between abundance and their index like this document is a simple effective way to improve the accuracy of stock sizes under hyperstability or hyperdepletion (Hashimoto et al.2018). At the same time, there are strong needs to improve tuning method and to standardize those indices as well as to check meta data of those indices.

The two settings of M had little impact on absolute values of abundances and exploitation rates (Fig. 1), but relatively large influences on %SPR and F relative to F0.1 from the past to latest years (Fig. 28, top and middle panels): the scenarios with age-specific M had larger fishing impacts than the scenarios with age-common M. By contrast, F relative to Fmsy were greatly affected by the difference of maturity and weight in recent years rather than by the difference of M (Fig. 28, bottom panels). These biological parameters and data (M, maturity, and weight) will be important for assessing stock status relative to biological reference points.

Time-varying life-history parameters related to maturation and growth are one of the key characteristics for chub mackerel (Watanabe and Yatsu 2004, 2006, Kamimura et al. 2021). This indicates that basic biological parameters such as steepness and biological reference points such as SSBmsy change dynamically over years (Miller and Brooks 2021). We, hence, suggest that attention to the dynamic nature of chub mackerel should be paid not only in stock assessment but also future prediction and stock management. Even if we can correctly conduct stock assessment regarding past estimates, it is adequately possible that future prediction and management advice could be biased due to the misspecification of future biological parameters. Although %SPRmsy has sometimes been used as a proxy of Fmsy (Zhou et al. 2020), fixing %SPRmsy at a value as a proxy of Fmsy may not be effective for chub mackerel because %SPRmsy changed drastically in recent years (Fig. 29, bottom panels). This change of Fmsy caused the rapid increase F/Fmsy in recent years under some scenarios (Fig. 28, bottom panels) despite the fact that the exploitation rate was kept at a low level (Fig. 1, bottom panels). The time-varying biological parameters and their uncertainties into future prediction and stock management will be an important topic, which should be discussed towards the benchmark stock assessment work of chub mackerel in Northwestern

Pacific.

References

Deroba, J. J., D. S. Butterworth, R. D. Methot, J. A. A. DeOliveira, C. Fernandez, A. Nielsen, S. X. Cadrin, M. Dickey-Collas, C. M. Legault, J. Ianelli, J. L. Valero, C. L. Needle, J. M. O'Malley, Y. J. Chang, G. G. Thompson, C. Canales, D. P. Swain, D. C. M. Miller, N. T. Hintzen, M. Bertignac, L. Ibaibarriaga, A. Silva, A. Murta, L. T. Kell, C. L. DeMoor, A. M. Parma, C. M. Dichmont, V. R. Restrepo, Y. Ye, E. Jardim, P. D. Spencer, D. H. Hanselman, J. Blaylock, M. Mood, and P. J. F. Hulson. 2014. Simulation testing the robustness of stock assessment models to error: Some results from the ICES strategic initiative on stock assessment methods. ICES Journal of Marine Science 72:19–30.

Hashimoto, M., Okamura, H., Ichinokawa, M., Hiramatsu, K., and Yamakawa, T. (2018) Impacts of nonlinear relationship between abundance and its index in a tuned virtual population analysis. Fisheries Science 84: 335–347.

Hutchings, J. A. 1996. Spatial and temporal variation in the density of northern cod and a review of hypotheses for the stock's collapse. Canadian Journal of Fisheries and Aquatic Sciences 53:943–962.

Kamimura, Y., M. Taga, R. Yukami, C. Watanabe, and S. Furuichi. 2021. Intra-and inter-specific density dependence of body condition, growth, and habitat temperature in chub mackerel (*Scomber japonicus*). bioRxiv:2021.03.25.436928.

Kristensen, K., A. Nielsen, C. W. Berg, H. Skaug, and B. M. Bell. 2016. TMB: Automatic differentiation and laplace approximation. Journal of Statistical Software 70:1–21.

Manabe, A., Kamimura, Ichinokawa, M., and Oshima, K. 2021. Maturity at age of chub mackerel under different stock level in the northwestern Paficic Ocean. NPFC-2021-TWG CMSA04-WP07.

Miller, T. J., and E. N. Brooks. 2021. Steepness is a slippery slope. Fish and Fisheries 22:634–645.

Mohn, R. 1999. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. ICES Journal of Marine Science 56:473–488.

Nielsen, A., and C. W. Berg. 2014. Estimation of time-varying selectivity in stock assessments using state-space models. Fisheries Research 158:96–101.

Nishijima, S., R. Yukami, K. Oshima, and M. Ichinokawa. 2020. Application of Virtual Population Analysis (VPA) and State-space Assessment Model (SAM) to the Shared Data of Chub Mackerel in the Northwest Pacific. NPFC-2020-TWG CMSA03-WP05.

Nishijima, S. 2020. Compilation and summary of shared data for operating models of the chub mackerel in the Northwestern Pacific. NPFC-2020-TWG CMSA03-WP04.

Nishijima, S., Y. Kamimura, R. Yukami, A. Manabe, K. Oshima, and M. Ichinokawa. 2021a.

Update on natural mortality estimators for chub mackerel in the Northwest Pacific Ocean. Page NPFC-2021-TWG CMSA04-WP05.

Nishijima, S., R. Yukami, K. Oshima, and M. Ichinokawa. 2021b. Re-analysis of Virtual Population Analysis (VPA) and State-space Assessment Model (SAM) for Operatinng Models of Chub Mackerel in NPFC. NPFC-2021-TWG CMSA04-WP08.

NPFC 2019. 2nd Meeting of the Technical Working Group on Chub Mackerel Stock Assessment. 2019. 2nd Meeting Report. NPFC-2019-TWG CMSA02-Final Report. (available at www.npfc.int)

NPFC. 2020. 3rd Meeting of the Technical Working Group on Chub Mackerel Stock Assessment. NPFC-2020-TWG CMSA03-Final Report. (available at www.npfc.int)

NPFC. 2022. 4th Meeting of the Technical Working Group on Chub Mackerel Stock Assessment. NPFC-2020-TWG CMSA03-Final Report. (in press)

Okamura, H., Y. Yamashita, and M. Ichinokawa. 2017. Ridge virtual population analysis to reduce the instability of fishing mortalities in the terminal year. ICES Journal of Marine Science 74:2427–2436.

Okamura, H., M. Ichinokawa, R. Hilborn. 2020. Evaluating a harvest control rule to improve the sustainability of Japanese fisheries. BioRxiv. https://doi.org/10.1101/2020.07.16.207282

Punt, A. E., A. D. M. Smith, D. C. Smith, G. N. Tuck, and N. L. Klaer. 2014. Selecting relative abundance proxies for BMSY and BMEY. ICES Journal of Marine Science 71:469–483.

Pope, J. G. 1972. An investigation in the accuracy of the virtual population analysis using cohort analysis. ICNAF Research Bulletin 9:65–74.

Rose, G., and D. Kulka. 1999. Hyperaggregation of fish and fisheries: how catch-per-unit-effort increased as the northern cod (*Gadus morhua*) declined. Canadian Journal of Fisheries and Aquatic Sciences 56:118–127.

Thorson, J. T., and K. Kristensen. 2016. Implementing a generic method for bias correction in statistical models using random effects, with spatial and population dynamics examples. Fisheries Research 175:66–74.

Watanabe, C., and A. Yatsu. 2004. Effects of density-dependence and sea surface temperature on interannual variation in length-at-age of chub mackerel (*Scomber japonicus*) in the Kuroshio-Oyashio area during 1970-1997. Fishery Bulletin 102:196–206.

Watanabe, C., and A. Yatsu. 2006. Long-term changes in maturity at age of chub mackerel (*Scomber japonicus*) in relation to population declines in the waters off northeastern Japan. Fisheries Research 78:323–332.

Zhou, S., A. E. Punt, Y. Lei, R. A. Deng, and S. D. Hoyle. 2020. Identifying spawner biomass per-recruit reference points from life-history parameters. Fish and Fisheries 21:760–773.

Tables

Table 1: Six scenarios to be used for the stock assessment analyses for the operating model development.

Scenario	Description	M	Weight-at-age	Maturity-at-age	Catch (at-age)	Abundance index	Fleet
A	Base-case 1	0.41	Average	Average	Average	All six	Single
B	Base-case 2	Gislason	Average	Average	Average	All six	Single
C	Highest weight and maturity	0.41	Highest	Highest	Average	All six	Single
D	Highest weight and maturity	Gislason	Highest	Highest	Average	All six	Single
E	Lowest weight and maturity	0.41	Lowest	Lowest	Average	All six	Single
F	Lowest weight and maturity	Gislason	Lowest	Lowest	Average	All six	Single

Table 2: Total numbers and SSB of summary statistics throughout the whole period (minimum, median, maximum, and mean), some years (1970, 1980, 1990, 2000, 2010, and 2019), and their ratios of the latest year (2019) to the summary statistics.

Model	Scenario	Min		Median	Max		Mean	1970	1980	1990	2000	2010	2019	2019/Min	2019/Median	2019/Max	2019/Mean
Total number (billion)																	
SAM	A	0.68	(2000)	6.22	42.59	(2018)	12.10	24.42	11.66	0.82	0.68	4.96	35.14	51.63	5.65	0.83	2.90
	B	0.73	(2000)	6.77	44.28	(2018)	12.89	27.64	12.66	0.89	0.73	4.80	32.74	44.84	4.84	0.74	2.54
	C	0.68	(2000)	6.18	40.22	(2018)	11.97	24.40	11.66	0.82	0.68	4.87	33.83	49.93	5.47	0.84	2.83
	D	0.73	(2000)	6.86	42.99	(2018)	12.89	27.62	12.65	0.89	0.73	4.77	32.51	44.68	4.74	0.76	2.52
	E	0.68	(2000)	6.72	48.87	(2018)	13.21	24.23	11.61	0.81	0.68	5.43	41.33	60.75	6.15	0.85	3.13
	F	0.73	(2000)	7.68	52.47	(2018)	14.18	27.57	12.55	0.89	0.73	5.63	40.71	55.69	5.30	0.78	2.87
VPA	A	0.76	(2001)	7.07	151.38	(2018)	17.77	26.34	11.99	0.97	1.05	4.25	101.16	132.50	14.30	0.67	5.69
	B	0.86	(2001)	7.74	141.67	(2018)	17.90	29.80	12.68	1.02	1.16	4.05	80.03	93.08	10.34	0.56	4.47
	C	0.76	(2001)	7.17	168.25	(2018)	18.58	26.34	11.99	0.97	1.05	4.30	111.80	146.44	15.60	0.66	6.02
	D	0.86	(2001)	7.82	157.50	(2018)	18.63	29.80	12.68	1.02	1.16	4.08	88.20	102.58	11.28	0.56	4.73
	E	0.76	(2001)	7.74	405.97	(2018)	29.37	26.34	11.99	0.97	1.05	5.55	261.70	342.78	33.80	0.64	8.91
	F	0.86	(2001)	8.32	187.61	(2018)	19.89	29.80	12.68	1.02	1.16	4.28	103.63	120.52	12.46	0.55	5.21
SSB (thousand ton)																	
SAM	A	43.9	(2002)	340.8	1300.9	(1979)	470.0	677.2	1024.3	86.5	62.1	166.7	852.1	19.39	2.50	0.66	1.81
	B	38.1	(2002)	293.6	1149.6	(1979)	410.7	605.6	884.0	68.9	53.1	136.6	754.8	19.78	2.57	0.66	1.84
	C	44.2	(2002)	341.7	1668.8	(2017)	528.4	680.4	1018.8	86.6	62.0	170.2	1309.1	29.64	3.83	0.78	2.48
	D	38.3	(2002)	285.6	1462.4	(2017)	461.9	608.1	880.4	69.0	53.1	133.4	1179.0	30.74	4.13	0.81	2.55
	E	43.8	(2002)	341.4	1300.6	(1979)	461.0	676.0	1024.5	86.5	62.2	187.5	842.5	19.23	2.47	0.65	1.83
	F	38.0	(2002)	299.9	1150.0	(1979)	394.6	604.2	884.9	68.8	53.1	142.4	706.6	18.59	2.36	0.61	1.79
VPA	A	44.5	(2002)	340.8	2431.3	(2019)	578.9	665.2	1093.3	99.4	64.9	120.1	2431.3	54.59	7.13	1.00	4.20

B	37.3	(2002)	283.9	1613.9	(2019)	464.9	583.7	943.9	74.5	53.5	102.0	1613.9	43.23	5.69	1.00	3.47
C	44.5	(2002)	342.8	6761.1	(2019)	817.1	665.2	1093.3	99.4	64.9	120.2	6761.1	151.80	19.72	1.00	8.27
D	37.3	(2002)	303.6	6754.8	(2019)	764.1	583.7	943.9	74.5	53.5	102.7	6754.8	180.93	22.25	1.00	8.84
E	44.5	(2002)	340.3	1418.1	(1978)	480.8	665.2	1093.3	99.4	64.9	120.0	1048.2	23.53	3.08	0.74	2.18
F	37.3	(2002)	277.0	1260.0	(1978)	396.7	583.7	943.9	74.5	53.5	101.9	666.0	17.84	2.40	0.53	1.68

Note: The numbers of brackets show the years when the minimum or maximum values recorded.

Table 3: The index-related parameters of b (nonlinearlity coefficient) and ν (standard deviation) estimated by SAM and VPA under the scenarios A to F.

Model	Scenario	***b* (nonlinearlity coefficient)**						***ν* (standard deviation)**					
		Fleet 2	Fleet 3	Fleet 4	Fleet 5	Fleet 7	Fleet 9	Fleet 2	Fleet 3	Fleet 4	Fleet 5	Fleet 7	Fleet 9
SAM	A	1.70*	2.01*	1.00	1.00	0.00	0.00	1.10	1.10	0.52	0.52	0.20	0.20
	B	1.73*	2.03*	1.00	1.00	0.00	0.00	1.08	1.08	0.53	0.53	0.20	0.20
	C	1.87*	2.20*	1.00	1.00	0.00	0.00	1.10	1.10	0.61	0.61	0.20	0.20
	D	1.88*	2.19*	1.00	1.00	0.00	0.00	1.08	1.08	0.64	0.64	0.20	0.20
	E	1.51*	1.82*	1.00	1.00	0.00	0.00	1.16	1.16	0.48	0.48	0.20	0.20
	F	1.57*	1.87*	1.00	1.00	0.00	0.00	1.13	1.13	0.51	0.51	0.20	0.20
VPA	A	1.34	1.52*	0.87	0.50*	0.49	0.27	1.01	1.01	0.56	0.56	0.18	0.18
	B	1.45*	1.62*	0.95	0.55*	0.32*	0.24	0.97	0.97	0.58	0.58	0.18	0.18
	C	1.31	1.48	0.64*	0.36*	0.33	0.40	1.00	1.00	0.60	0.60	0.18	0.18
	D	1.28	1.44	0.60*	0.34*	-1.08*	3.10*	1.01	1.01	0.62	0.62	0.08	0.08
	E	1.33	1.49*	1.22	0.72	0.45	0.24	1.00	1.00	0.53	0.53	0.18	0.18
	F	1.44*	1.60*	1.38*	0.84	0.30*	0.22	0.97	0.97	0.56	0.56	0.18	0.18

* The probability of $b < 1$ (hyperstability) or $b > 1$ (hyperdepletion) is $p < 0.05$.

Table 4: Mohn's rho of the scenarios A to F with SAM and VPA.

Model	Scenario	Biomass	SSB	Recruitment	F
SAM	A	-0.13	-0.09	-0.04	0.22
	B	0.02	0.08	0.12	-0.12
	C	-0.16	-0.15	-0.02	0.34
	D	-0.10	-0.07	0.02	0.01
	E	-0.12	-0.07	0.01	0.13
	F	-0.07	0.05	0.04	0.01
VPA	A	0.33	0.57	0.14	-0.23
	B	0.28	0.72	0.09	-0.24
	C	0.46	0.80	0.18	-0.24
	D	0.03	-0.03	0.01	0.26
	E	0.29	0.35	0.11	-0.21
	F	0.22	0.47	0.05	-0.21

Table 5: Basic biological parameters and biological reference points when using the averages of biological parameters and fishing mortality coefficients over years.

Model	Scen-ario	α (10^6/ton)	β (1000 ton)	σ_R	SPR0 (g)	SSB0 (1000 ton)	R0 (billion)	Steepness (h)	SSBmsy (1000 ton)	MSY (1000 ton)	%SPR msy	SSBmsy /SSB0	F/ Fmsy
SAM	A	0.0074	974	0.75	665	4760	7.16	0.8	970	873	20	0.2	1.02
	B	0.0093	914	0.76	828	7060	8.52	0.87	910	917	13	0.13	1.03
	C	0.0072	894	0.75	693	4440	6.41	0.8	890	808	20	0.2	0.99
	D	0.0093	777	0.76	853	6140	7.2	0.87	780	800	13	0.13	0.99
	E	0.0078	968	0.78	649	4910	7.56	0.8	970	908	20	0.2	0.99
	F	0.0101	785	0.79	815	6490	7.96	0.88	780	845	12	0.12	0.99
VPA	A	0.0075	2177	0.88	665	10840	16.31	0.8	2180	1942	20	0.2	1
	B	0.0099	1056	0.9	828	8670	10.47	0.88	1060	1103	12	0.12	0.99
	C	0.0071	2990	0.88	693	14800	21.37	0.8	2990	2624	20	0.2	0.97
	D	0.0093	2664	0.91	853	21140	24.8	0.87	2660	2687	13	0.13	0.97
	E	0.0086	917	0.94	649	5130	7.9	0.82	920	935	18	0.18	0.93
	F	0.0114	715	0.96	815	6620	8.12	0.89	720	846	11	0.11	0.94

,

Figures

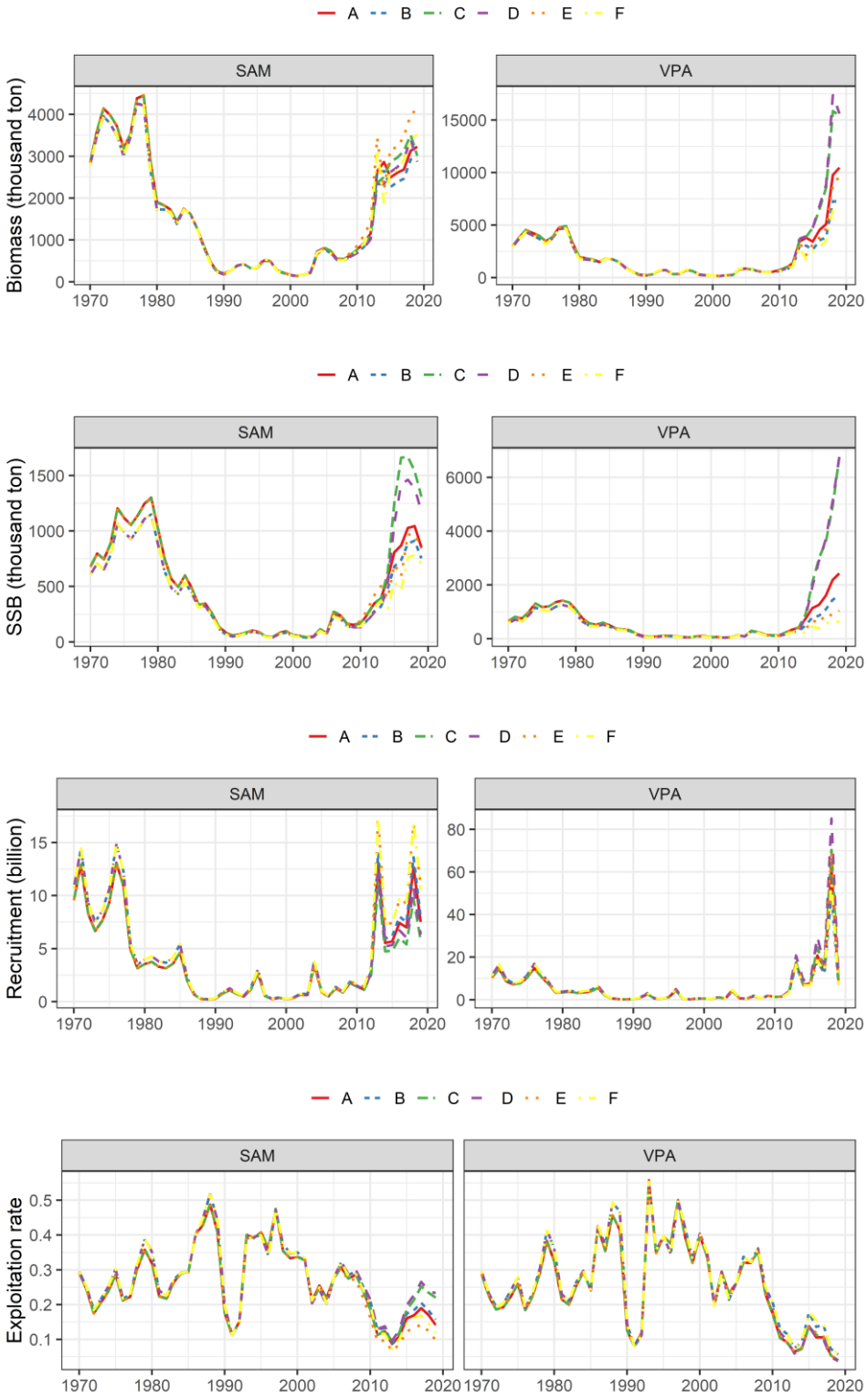

Figure 1: the estimates of total biomass (1st column), SSB (2nd column), recruitment number (3rd column), and exploitation rate (4th column) with SAM (left) and VPA (right) under the scenarios

A to F.

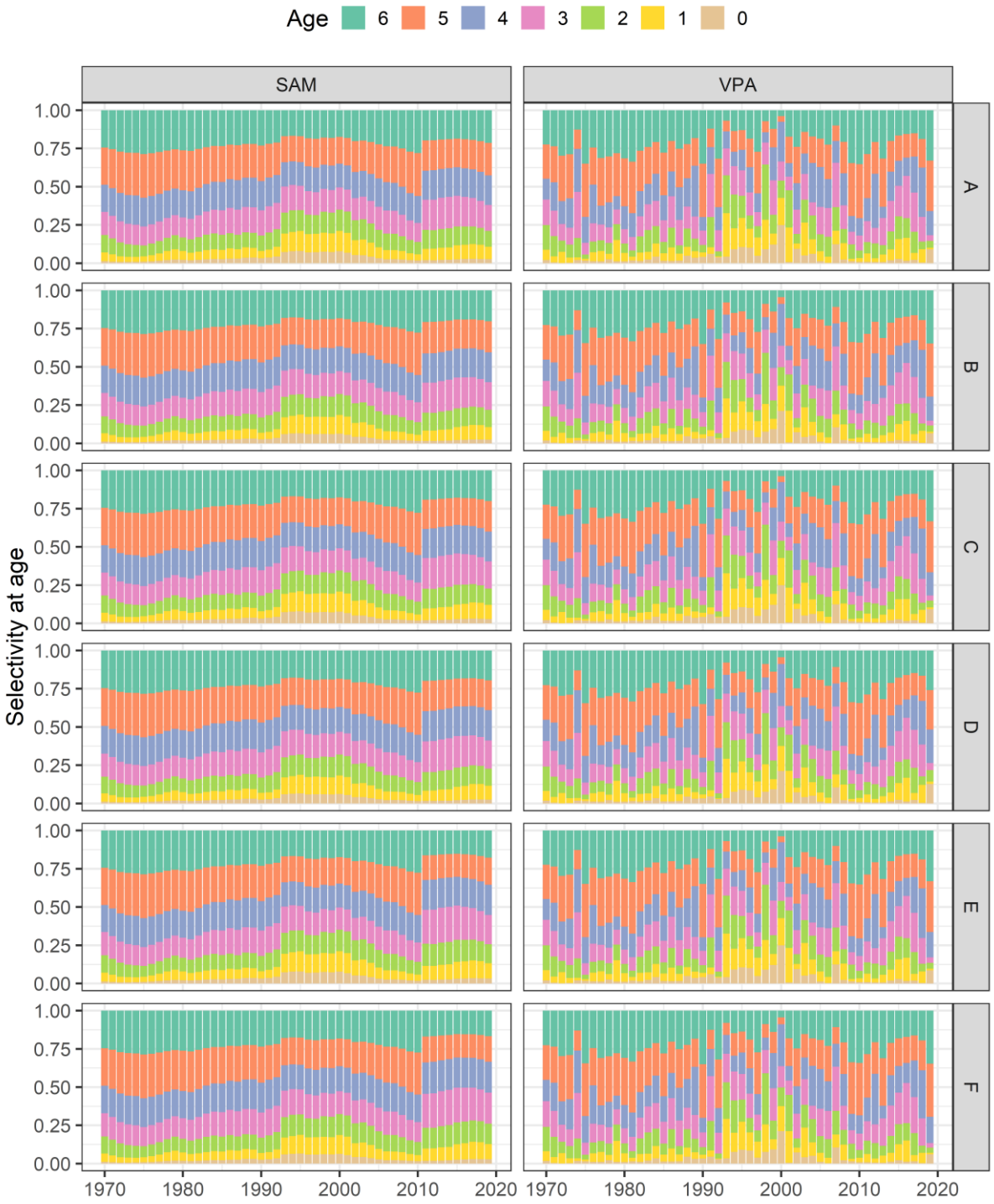

Figure 2: Selectivity at age in SAM (left) and VPA (right) under the scenario A to F. Selectivity is scaled so that its sum is equal to one.

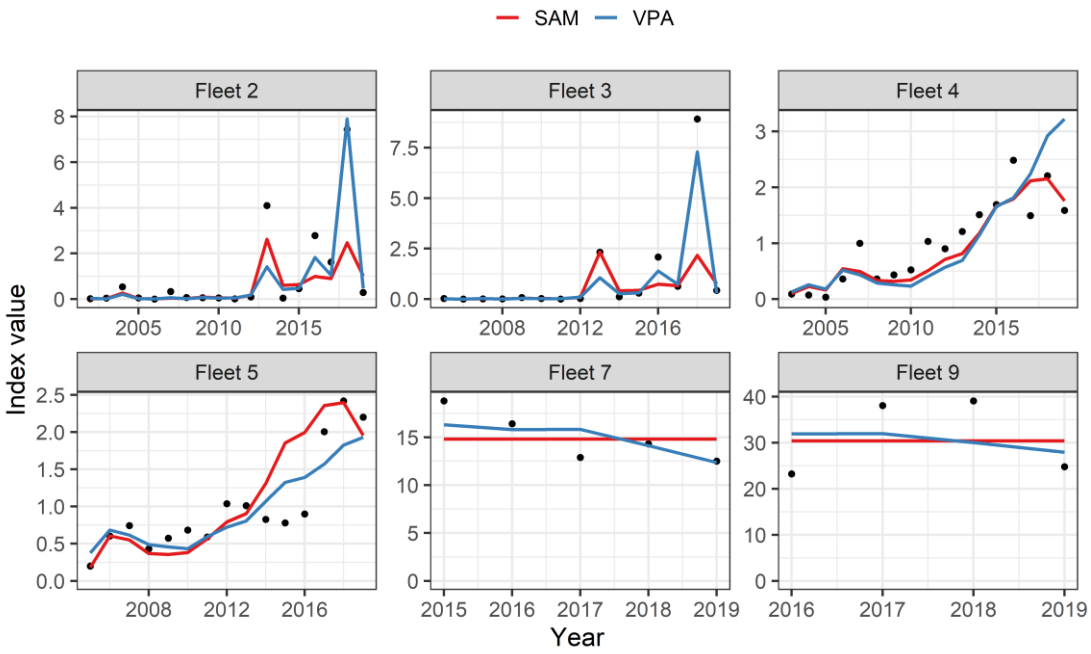

Figure 3: Index values observed (points) and their predicted values by SAM (red lines) and VPA (blue lines) under the scenario A.

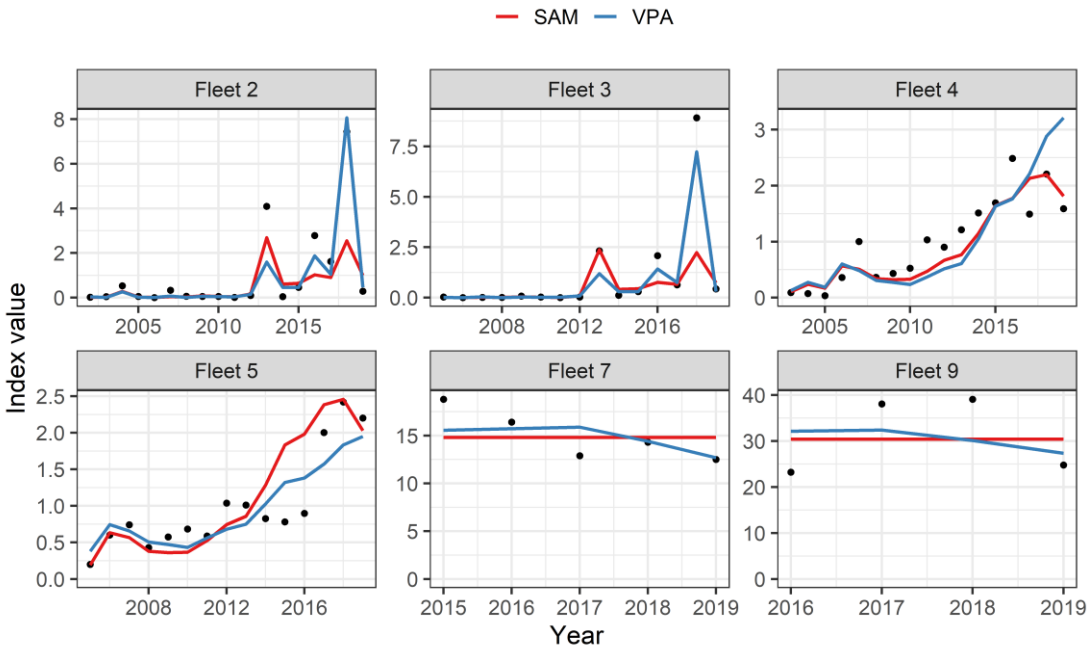

Figure 4: Index values observed (points) and their predicted values by SAM (red lines) and VPA (blue lines) under the scenario B.

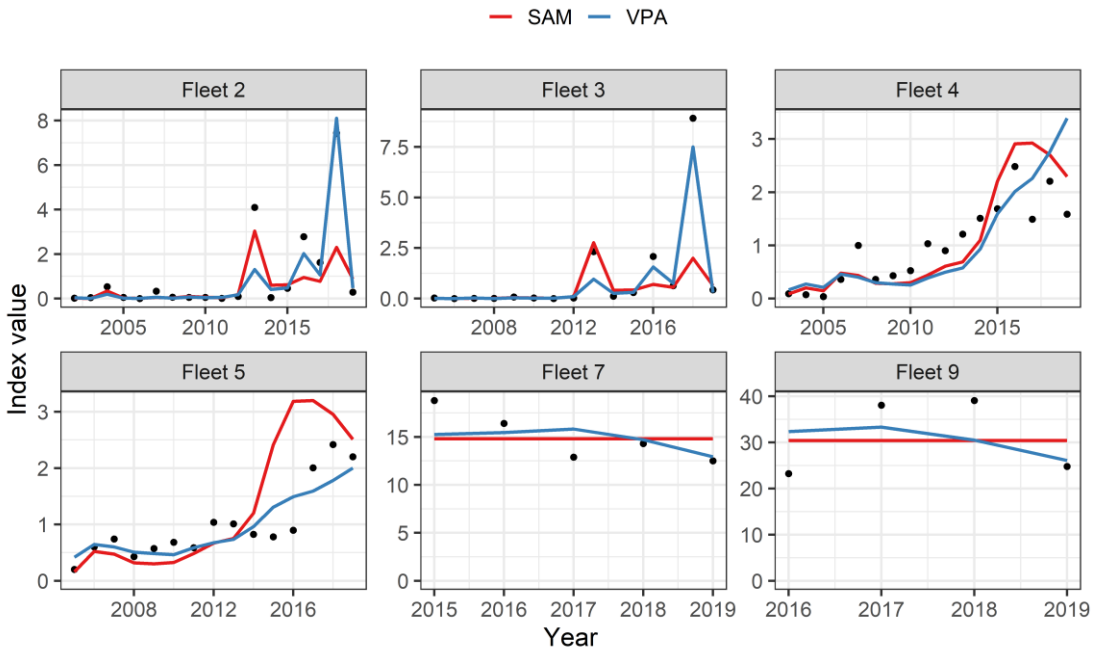

Figure 5: Index values observed (points) and their predicted values by SAM (red lines) and VPA (blue lines) under the scenario C.

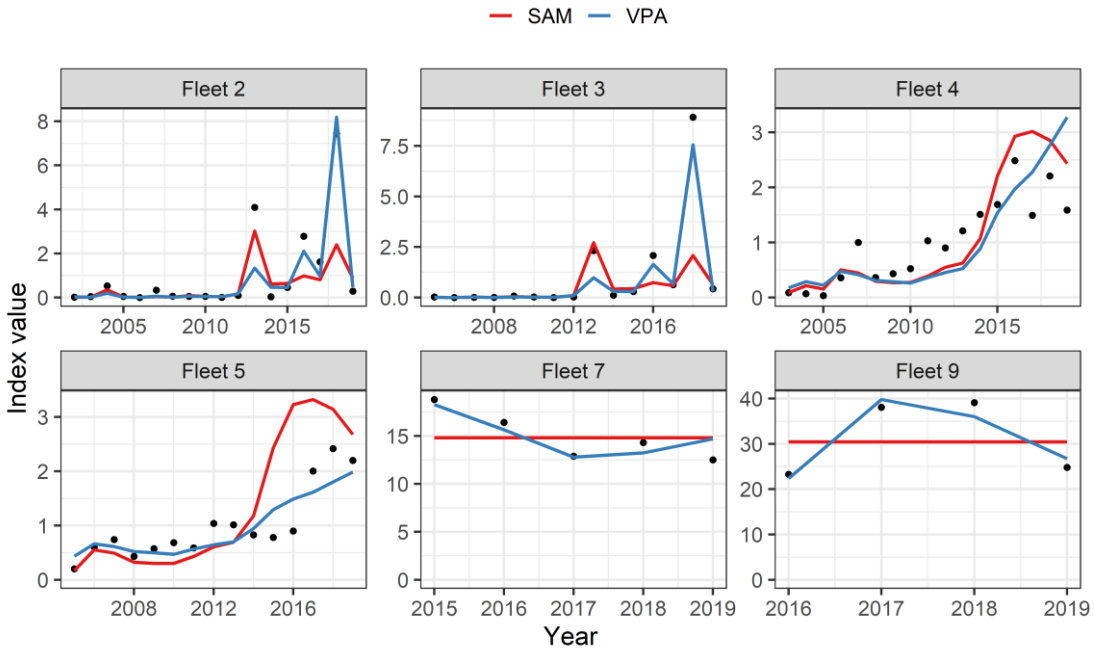

Figure 6: Index values observed (points) and their predicted values by SAM (red lines) and VPA (blue lines) under the scenario D.

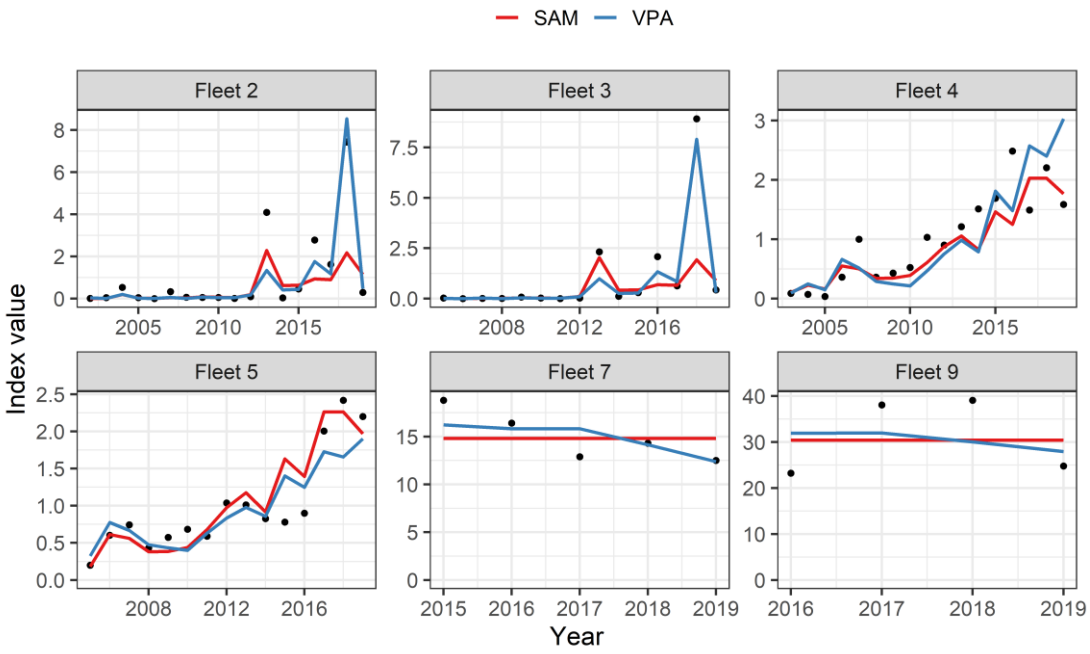

Figure 7: Index values observed (points) and their predicted values by SAM (red lines) and VPA (blue lines) under the scenario E.

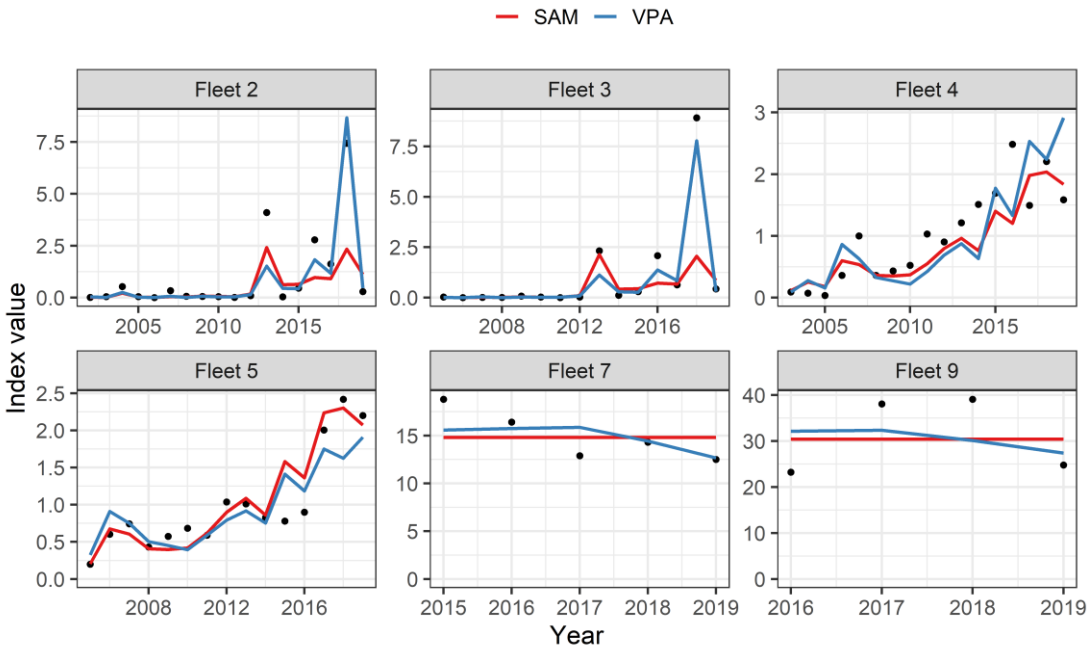

Figure 8: Index values observed (points) and their predicted values by SAM (red lines) and VPA (blue lines) under the scenario F.

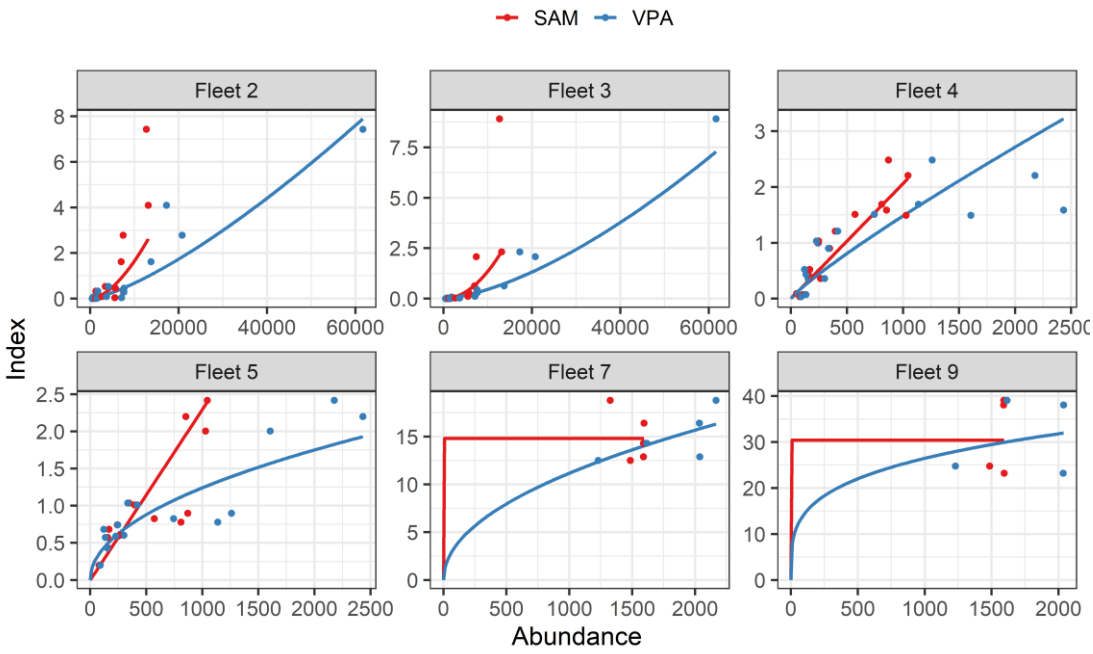

Figure 9: The relationships between abundance indices and their corresponding abundance estimates in SAM (red) and VPA (blue) under the scenario A.

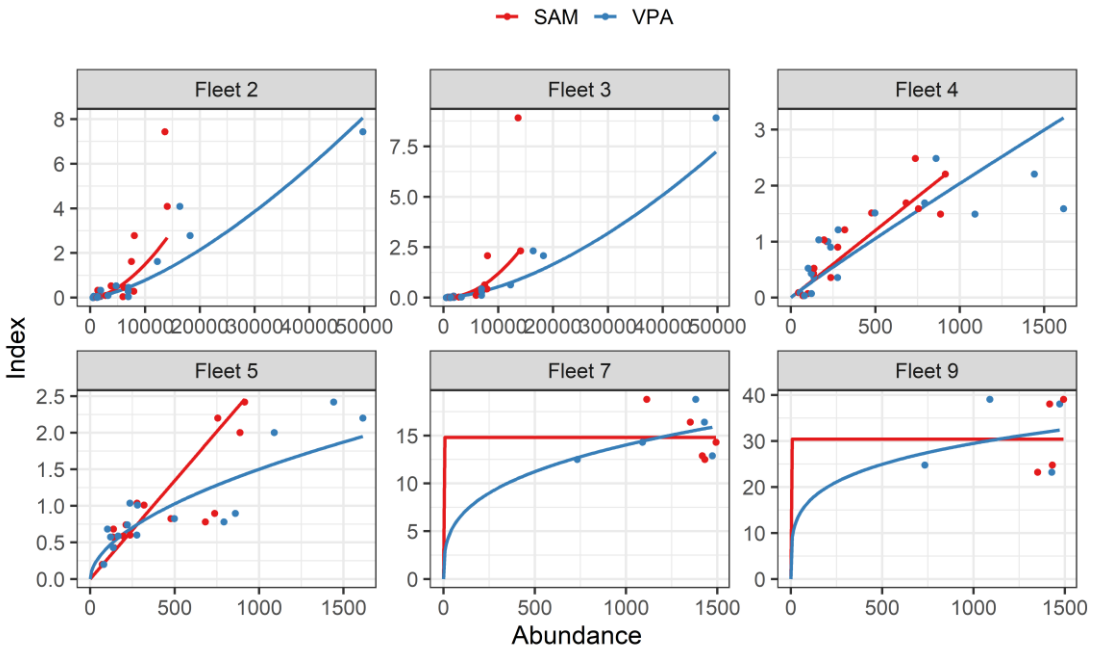

Figure 10: Relationships between abundance indices and their corresponding abundance estimates in SAM (red) and VPA (blue) under the scenario B.

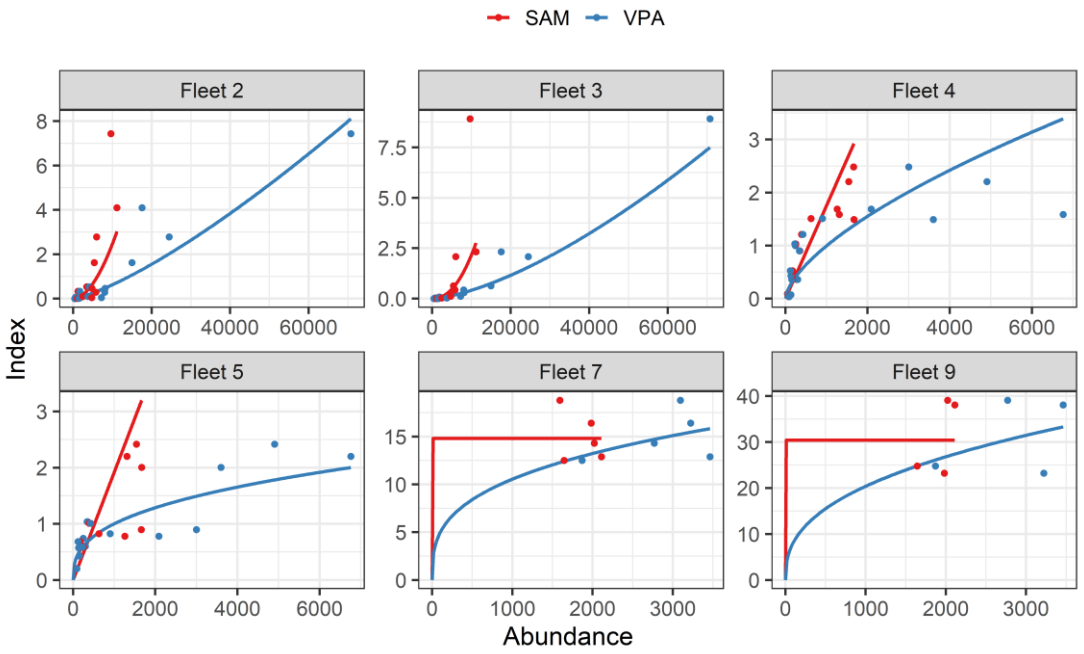

Figure 11: Relationships between abundance indices and their corresponding abundance estimates in SAM (red) and VPA (blue) under the scenario C.

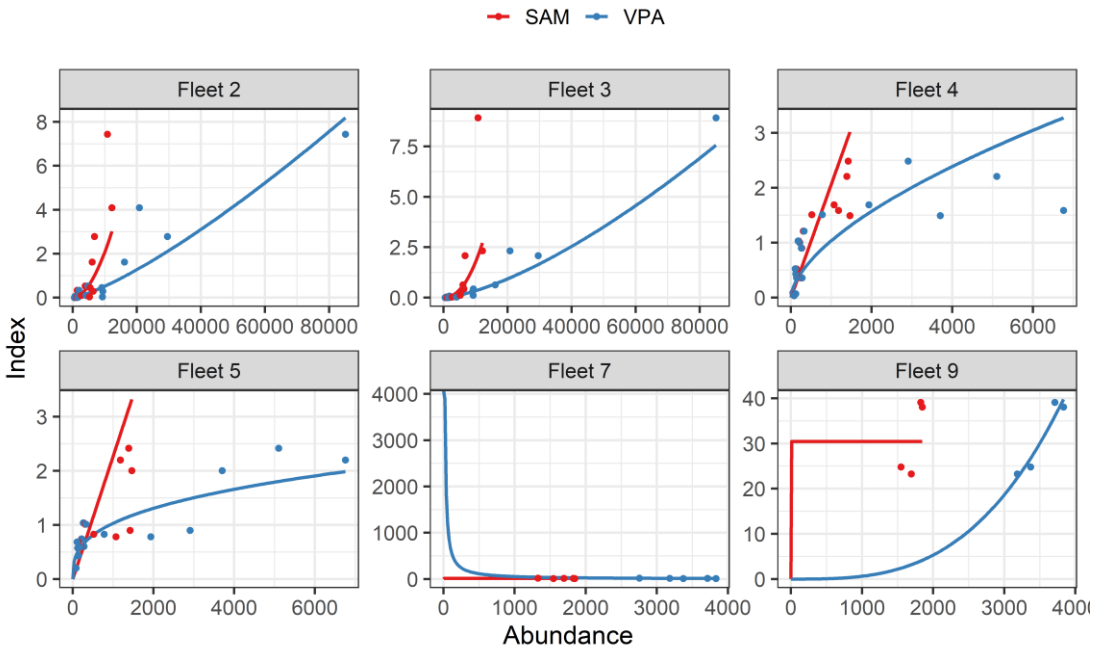

Figure 12: Relationships between abundance indices and their corresponding abundance estimates in SAM (red) and VPA (blue) under the scenario D.

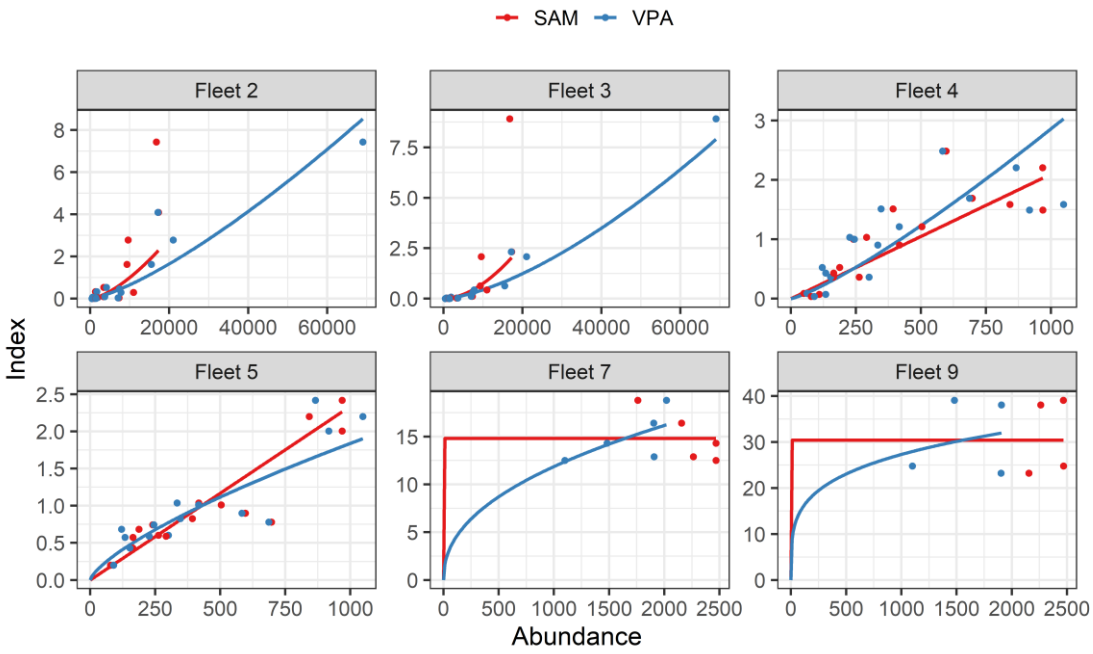

Figure 13: Relationships between abundance indices and their corresponding abundance estimates in SAM (red) and VPA (blue) under the scenario E.

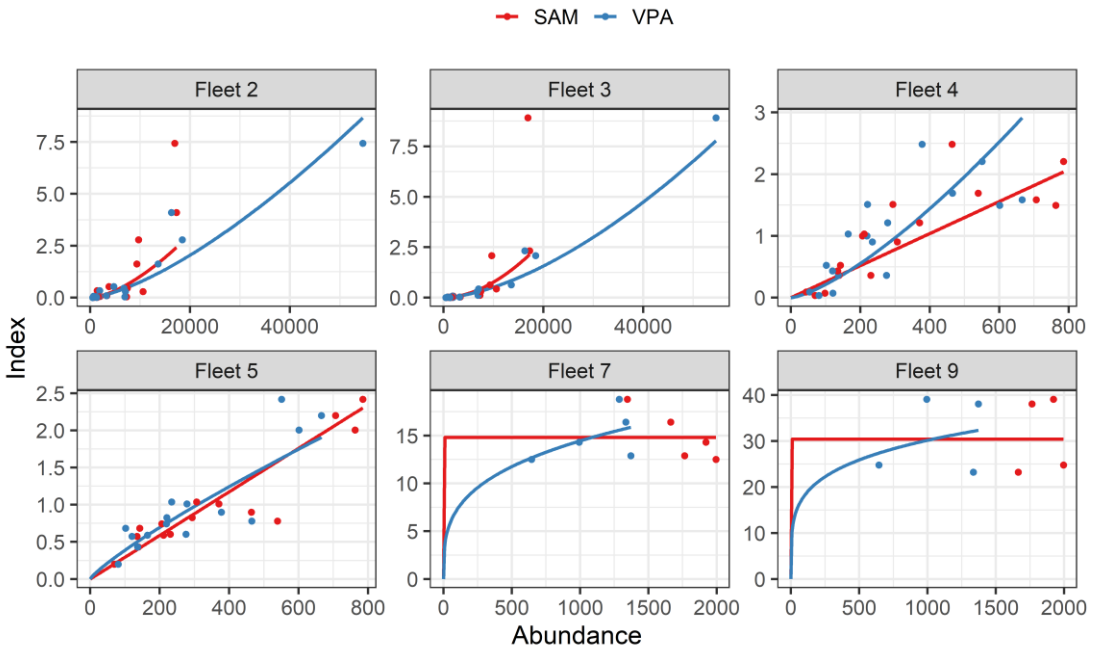

Figure 14: Relationships between abundance indices and their corresponding abundance estimates in SAM (red) and VPA (blue) under the scenario F.

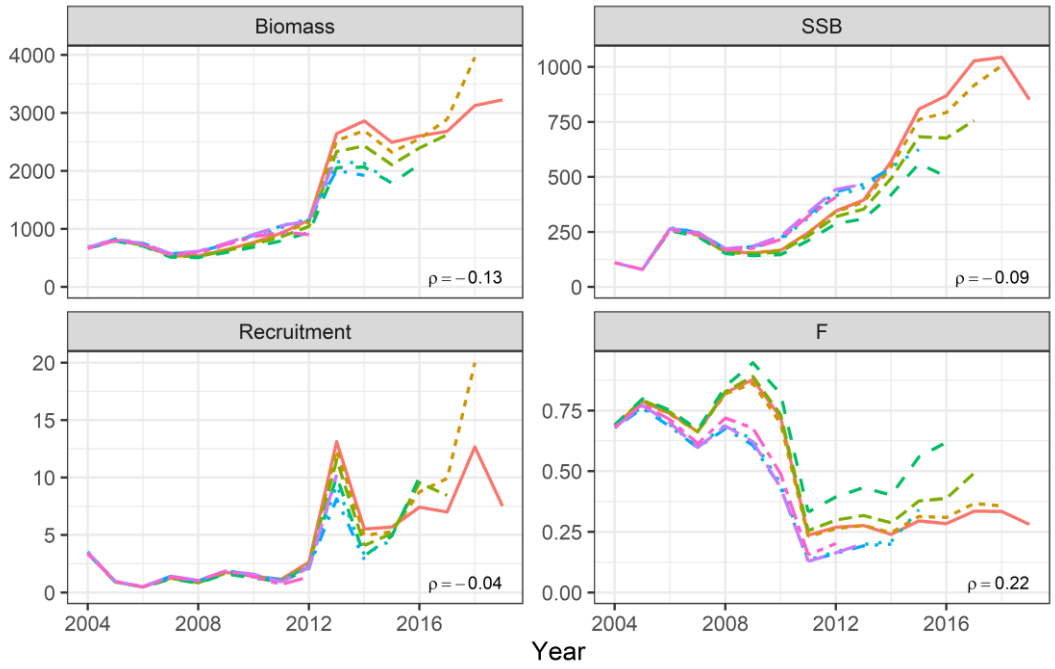

Figure 15: Retrospective pattens SAM under the scenario A.

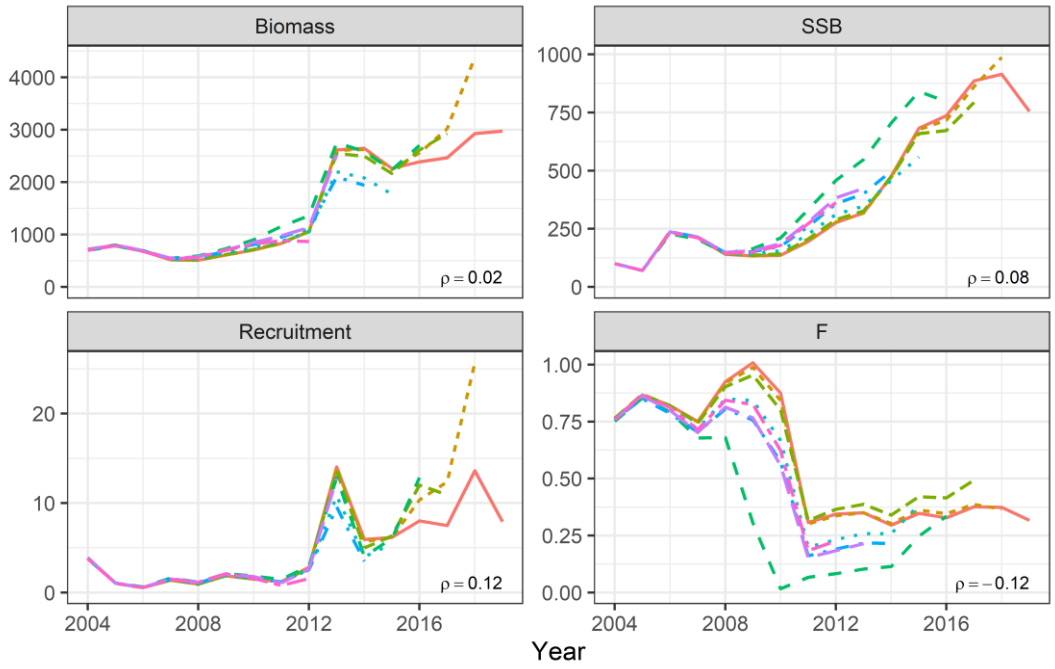

Figure 16: Retrospective pattens of SAM under the scenario B.

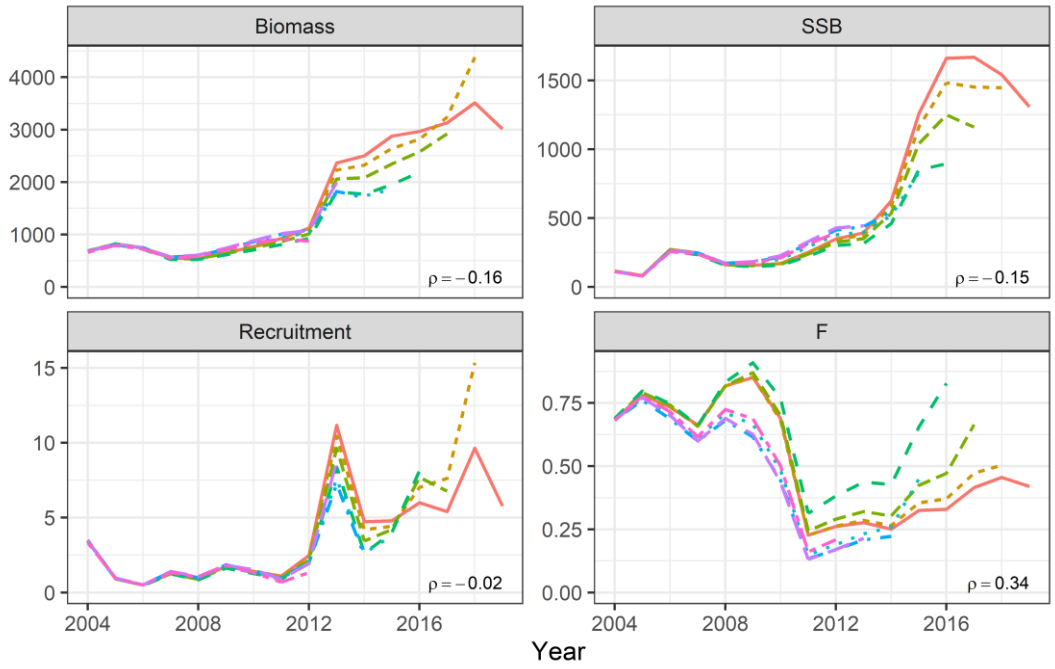

Figure 17: Retrospective pattens of SAM under the scenario C.

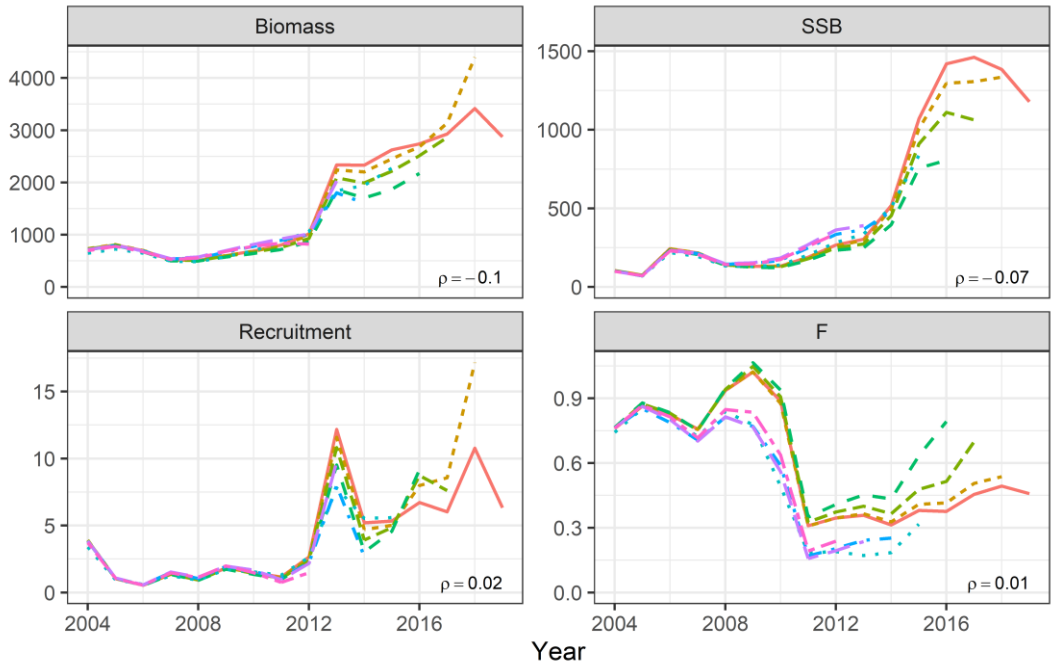

Figure 18: Retrospective pattens of SAM under the scenario D.

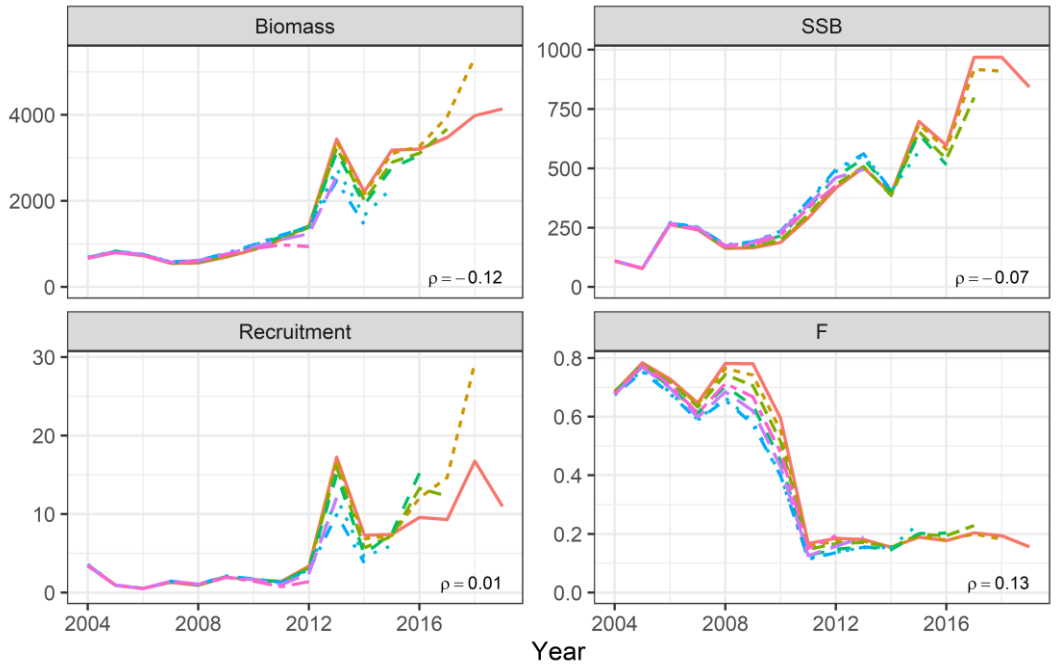

Figure 19: Retrospective pattens of SAM under the scenario E.

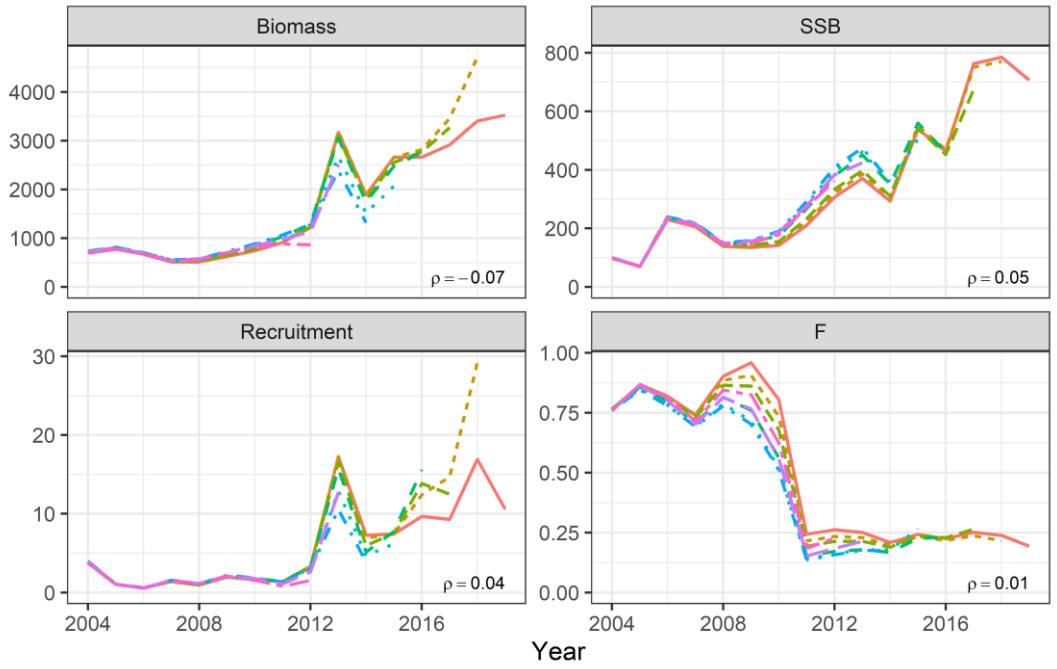

Figure 20: Retrospective pattens of SAM under the scenario F.

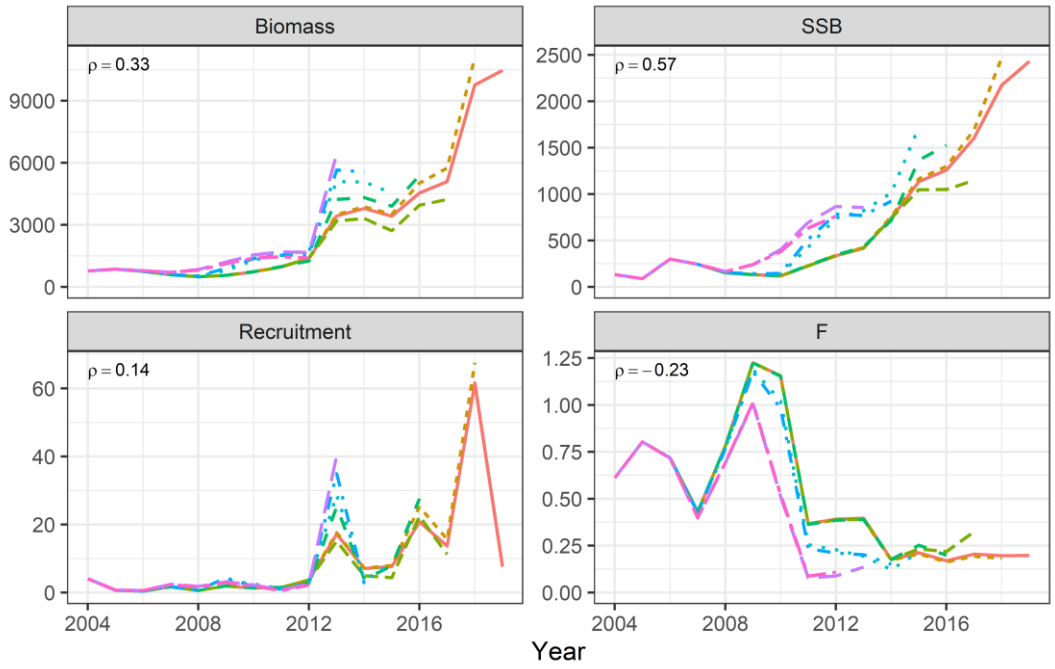

Figure 21: Retrospective pattens of VPA under the scenario A.

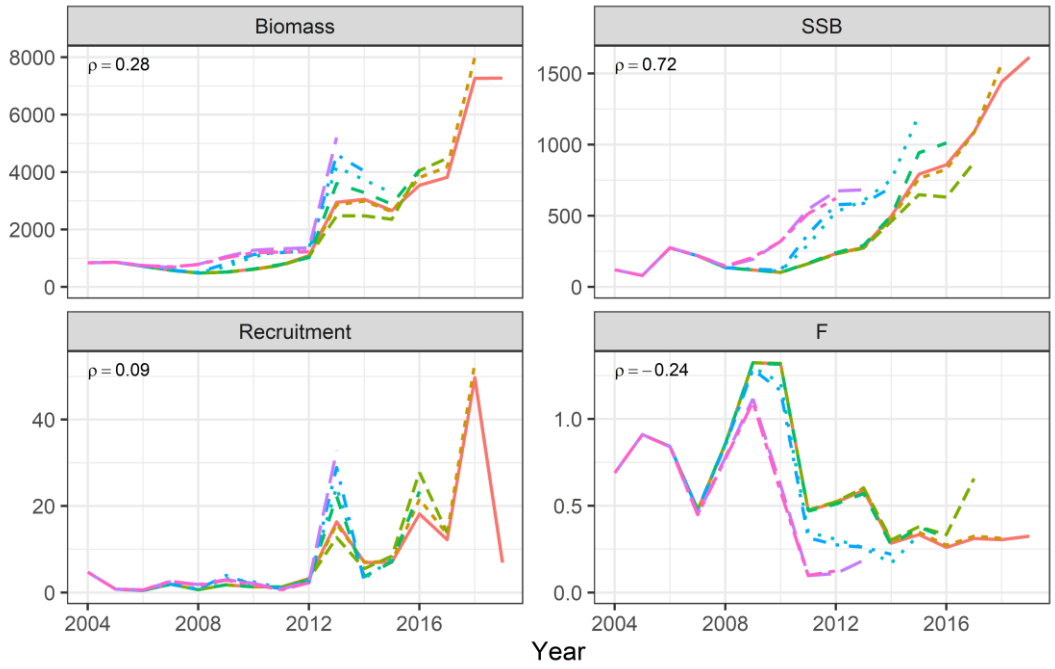

Figure 22: Retrospective pattens of VPA under the scenario B.

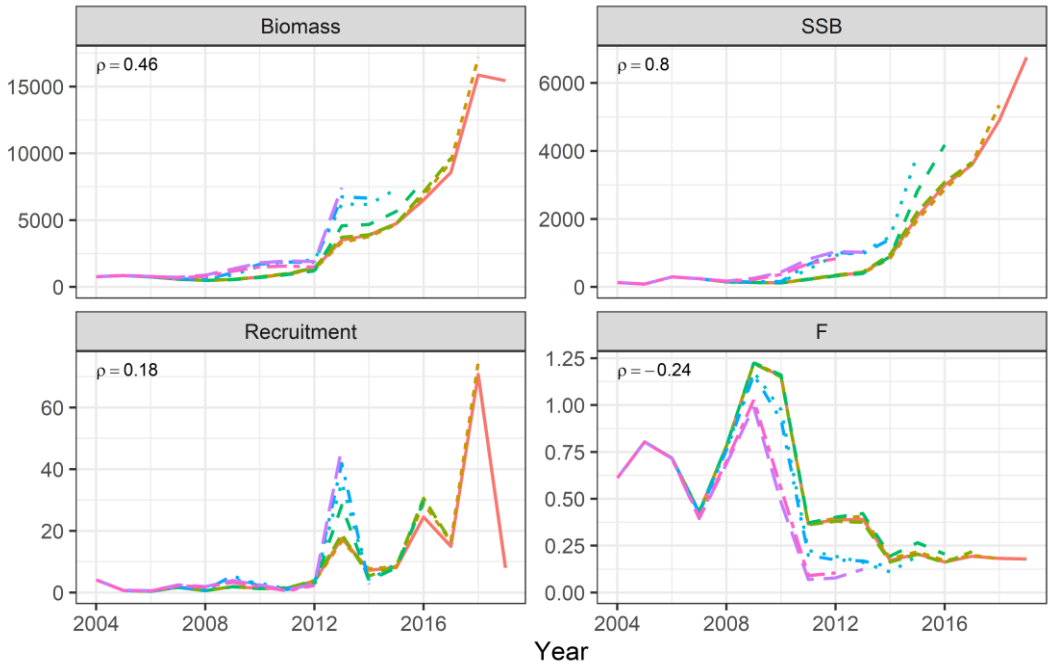

Figure 23: Retrospective pattens of VPA under the scenario C.

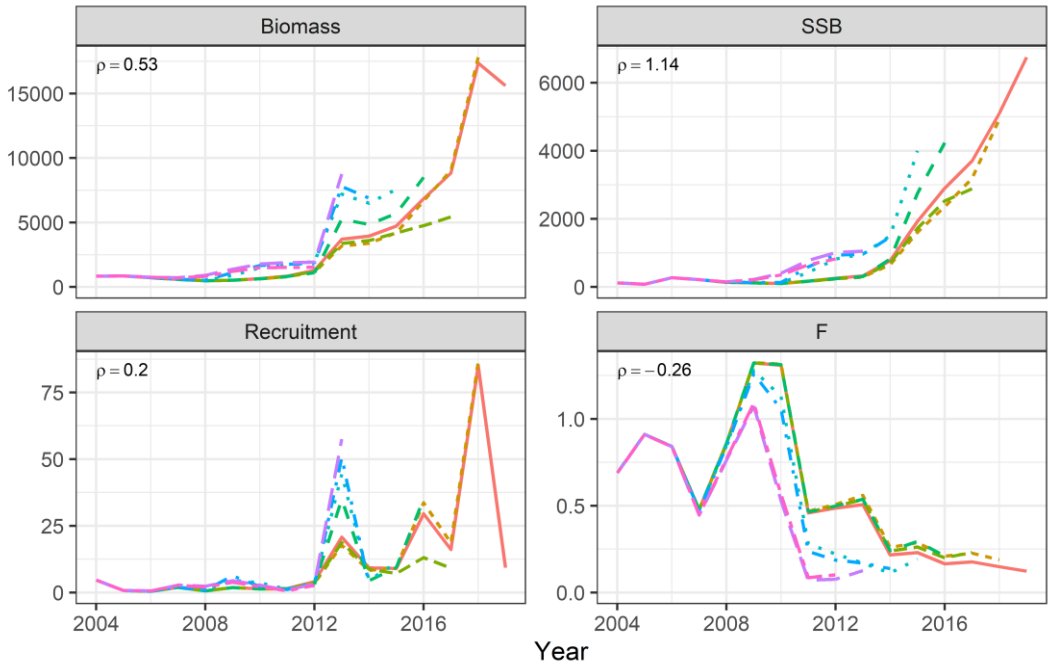

Figure 24: Retrospective pattens of VPA under the scenario D.

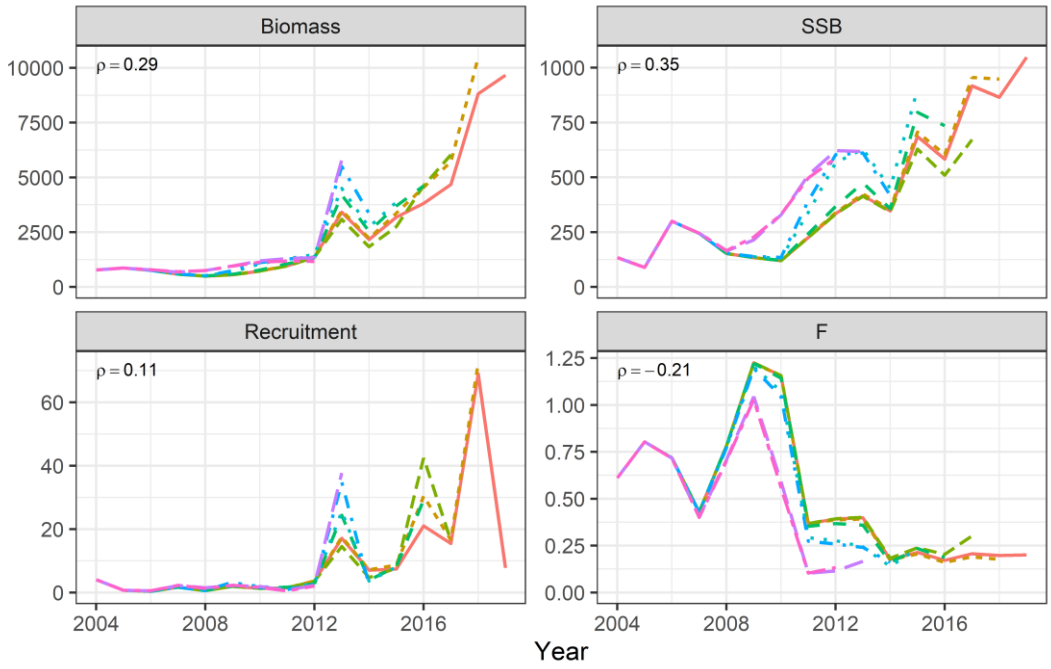

Figure 25: Retrospective pattens of VPA under the scenario E.

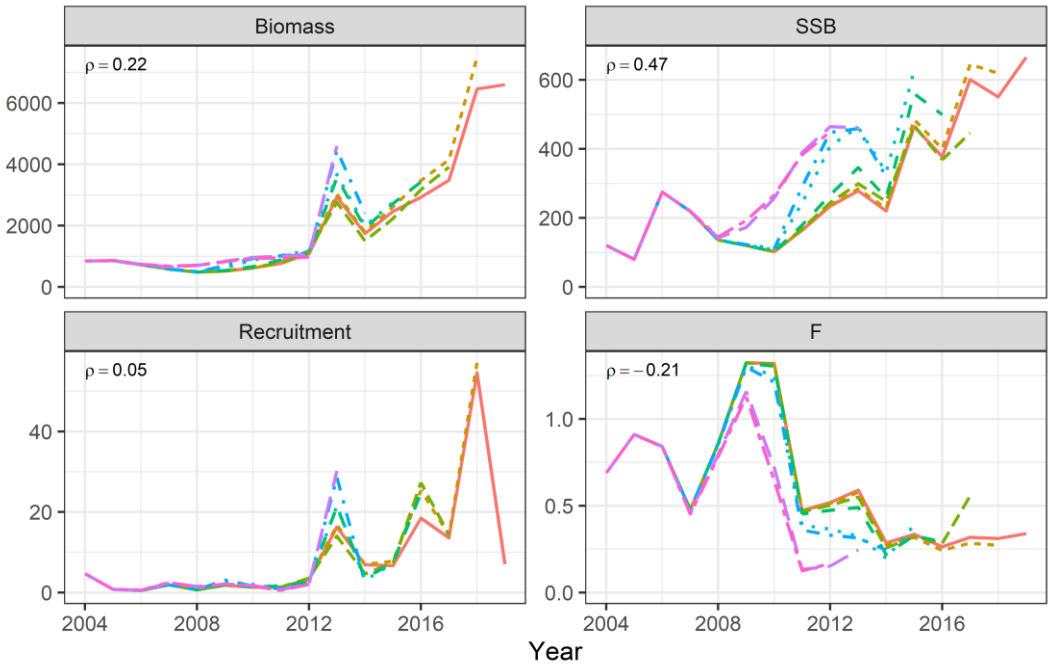

Figure 26: Retrospective pattens of VPA under the scenario F.

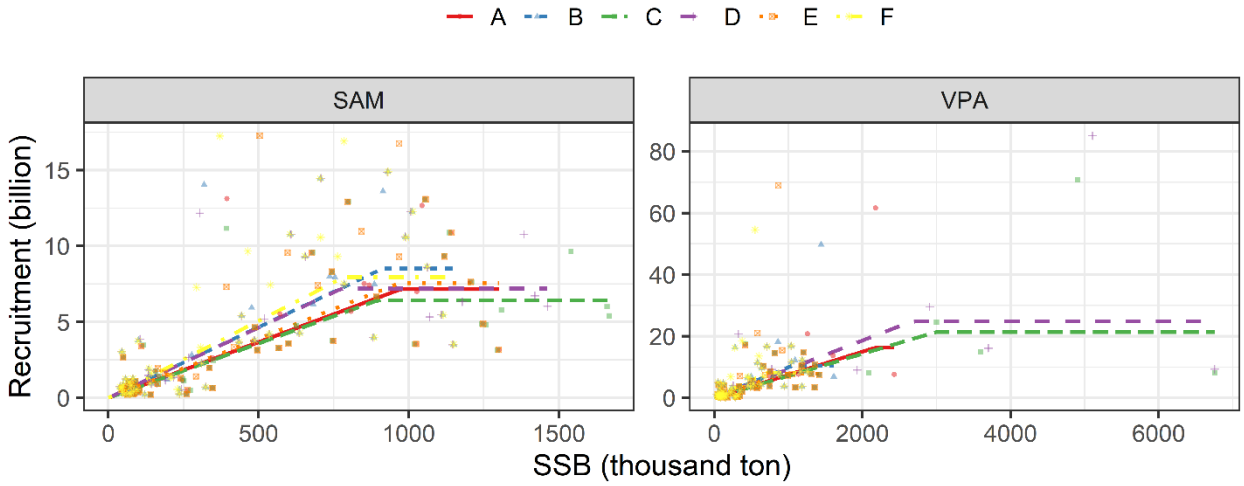

Figure 27: Continuous hockey-stick stock-recruit relationships in SAM and VPA with different scenarios.

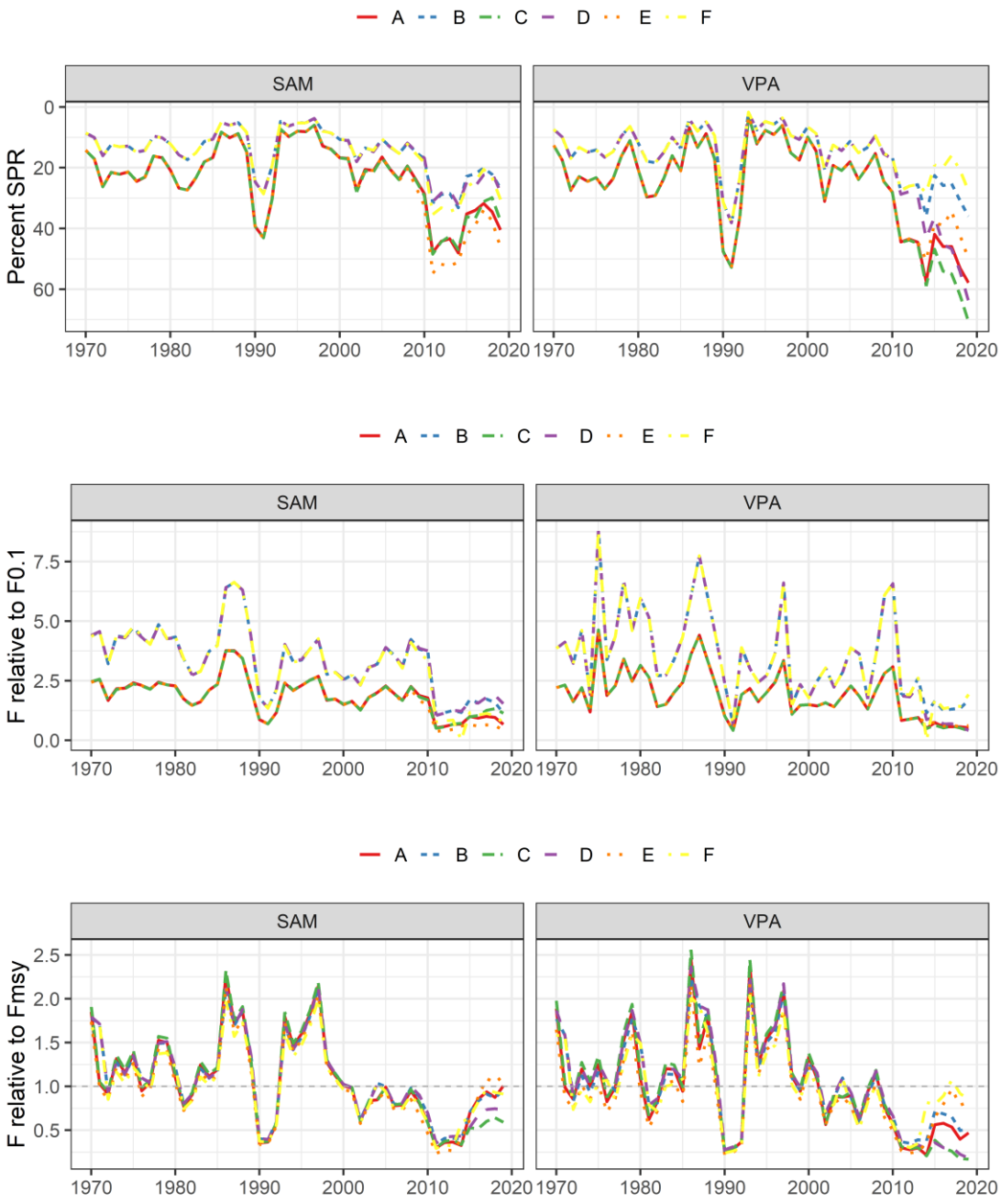

Figure 28: Temporal trends of percent SPR (top), F relative to F0.1 (middle), and F relative to Fmsy (bottom) in SAM (left) and VPA (right) under the scenarios A to F when using per-year biological parameters and F-at-age estimates. The values of Fmsy here is based on the time-varying estimates shown in Fig. 29.

Figure 29: Temporal trends of SPR0 (top) and %SPRmsy (bottom) in SAM (left) and VPA (right) under the scenarios A to F.

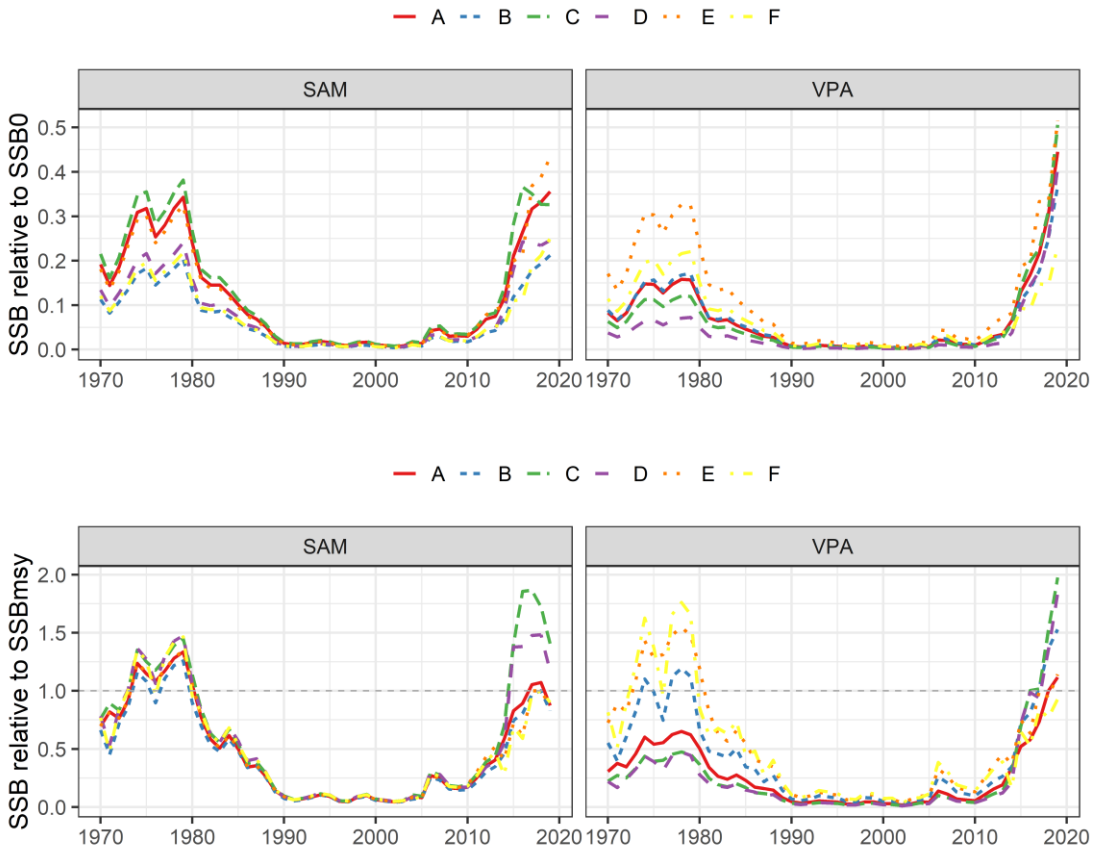

Figure 30: Temporal trends of SSB relative to SSB0 (top) and SSB relative to SSBmsy (bottom) in SAM (left) and VPA (right) under the scenario A to F.

Supplementary Figures

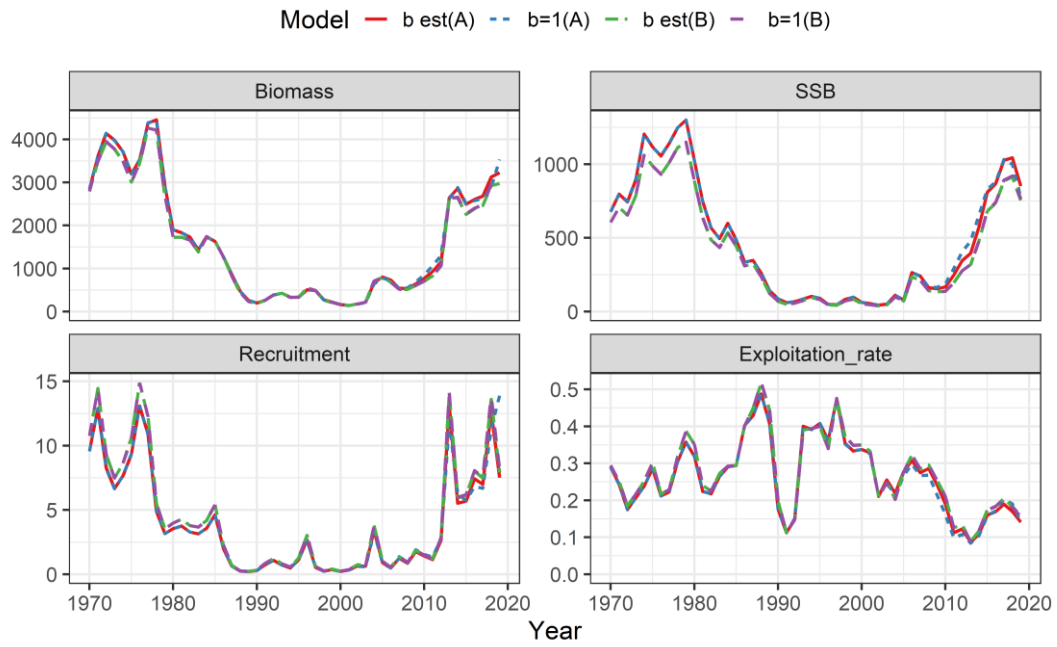

Figure S1: Comparison of estimates in biomass, SSB, recruitment, and exploitation rate between SAM with the nonlinear coefficients (*b*) estimated except for the SSB indices and SAM with b fixed at 1 under the scenarios A and B.

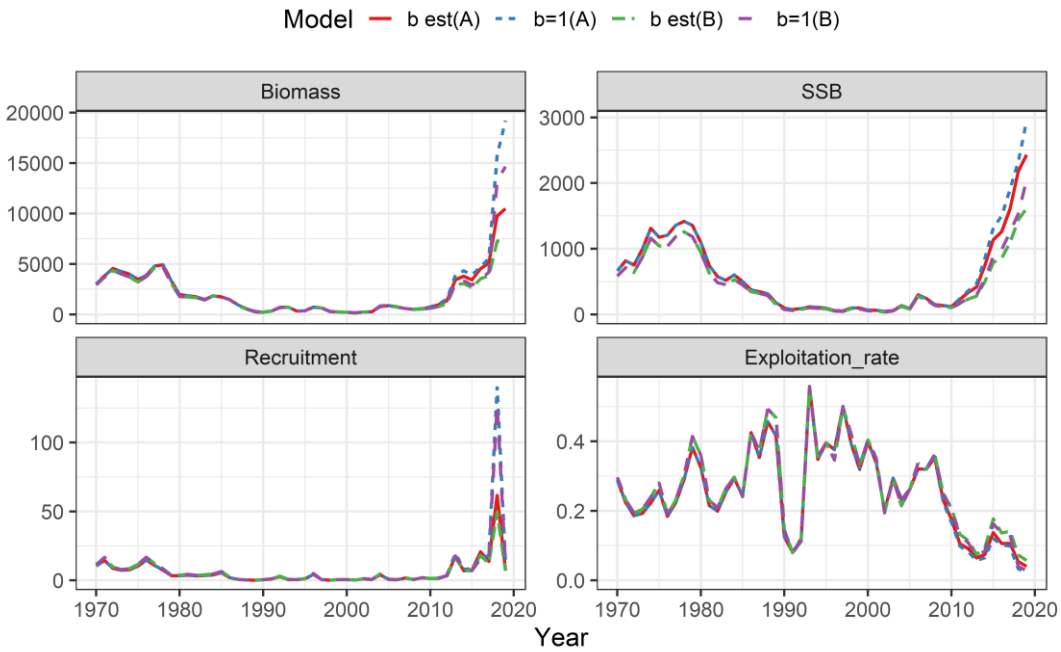

Figure S2: Comparison of estimates in biomass, SSB, recruitment, and exploitation rate between

VPA with the nonlinear coefficients (*b*) estimated except for the SSB indices and VPA with b fixed at 1 under the scenarios A and B.

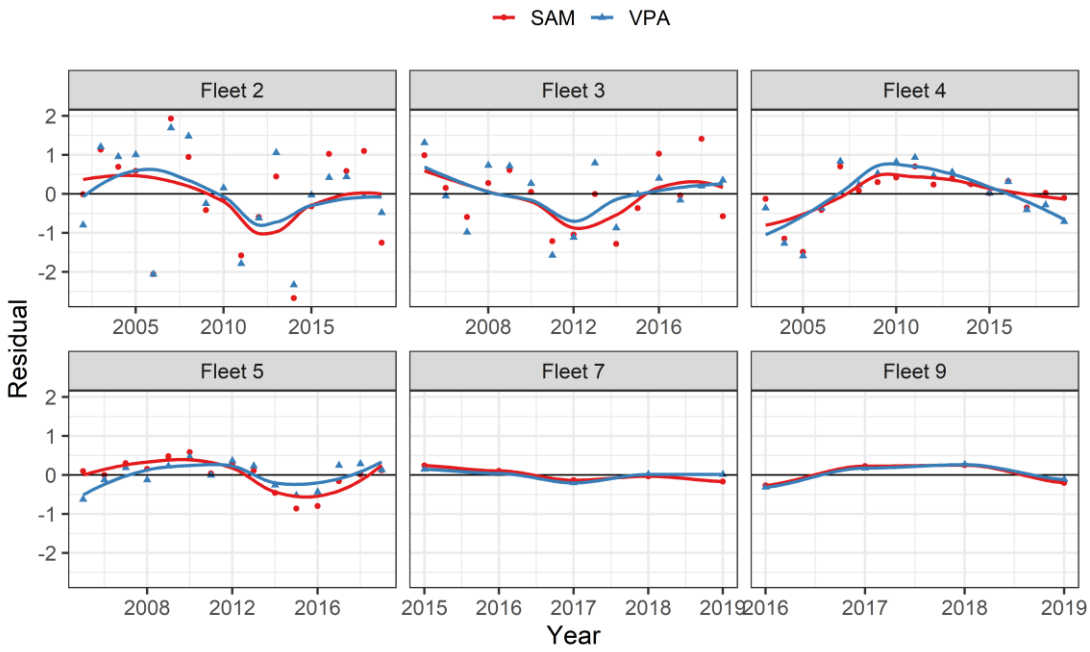

Figure S3: Residuals of abundance indices under the scenario A in SAM and VPA. The curves are the prediction by the LOESS (locally estimated scatterplot smoothing) regression.

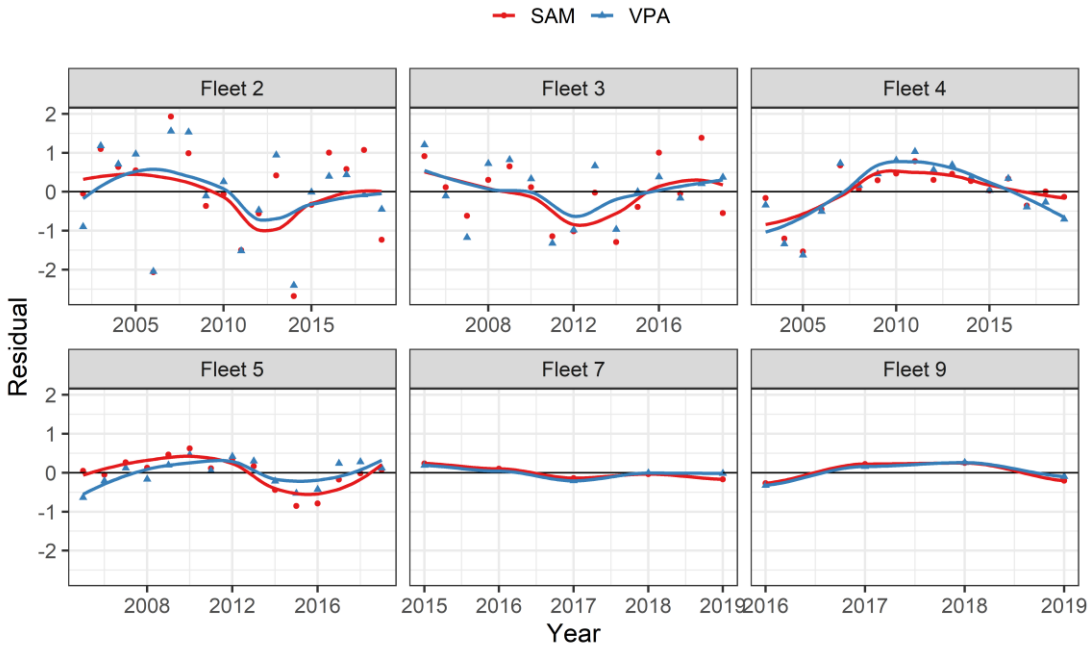

Figure S4: Residuals of abundance indices under the scenario B in SAM and VPA.

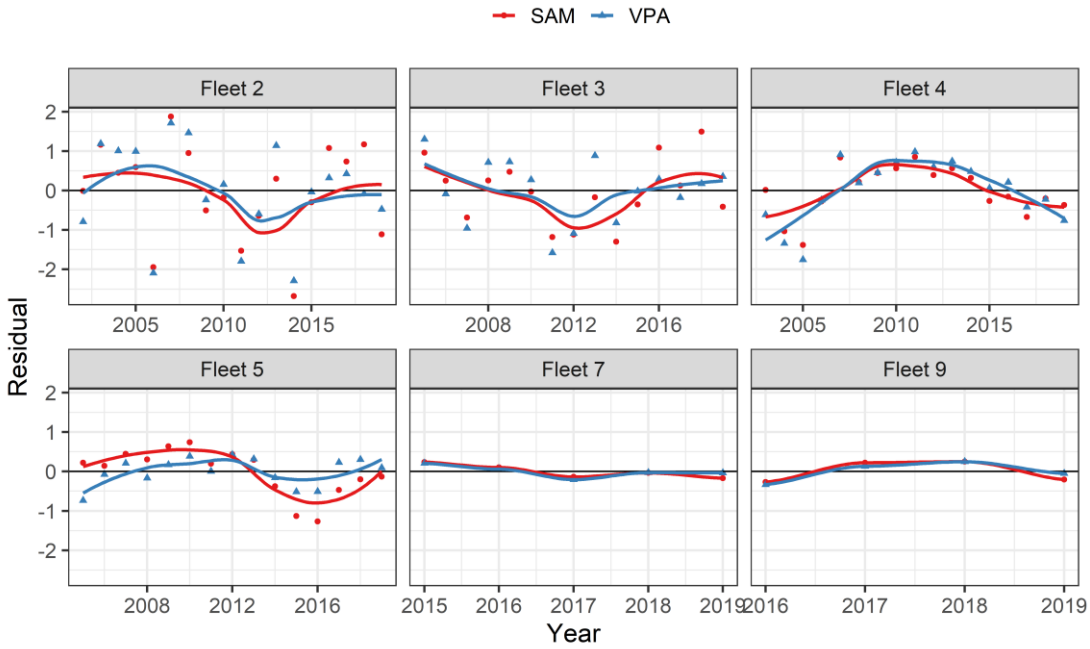

Figure S5: Residuals of abundance indices under the scenario C in SAM and VPA.

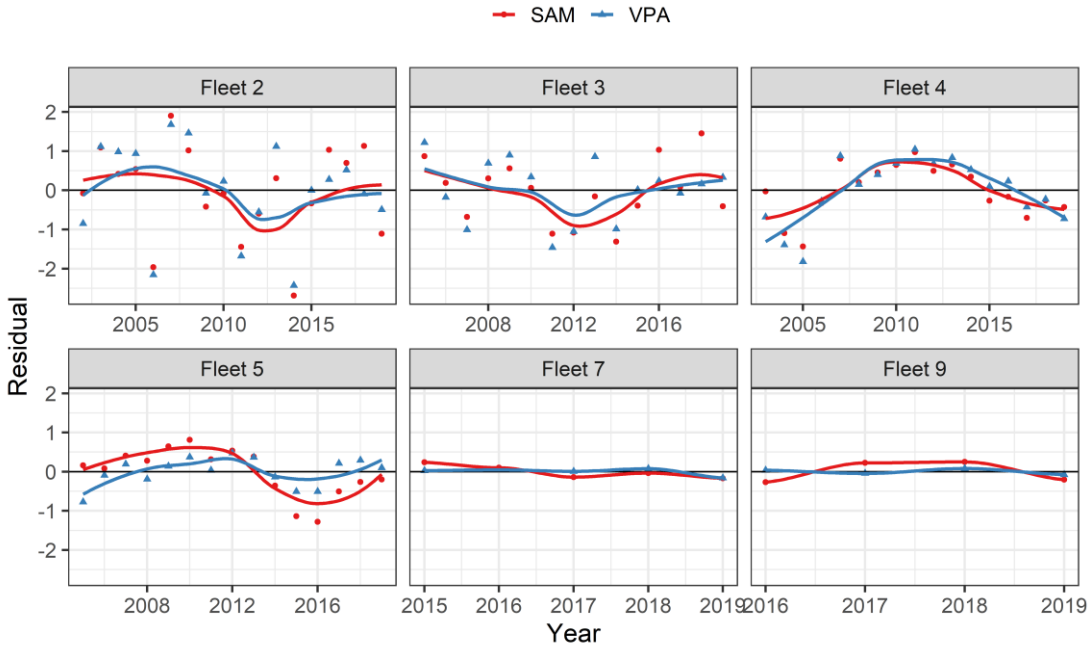

Figure S6: Residuals of abundance indices under the scenario D in SAM and VPA.

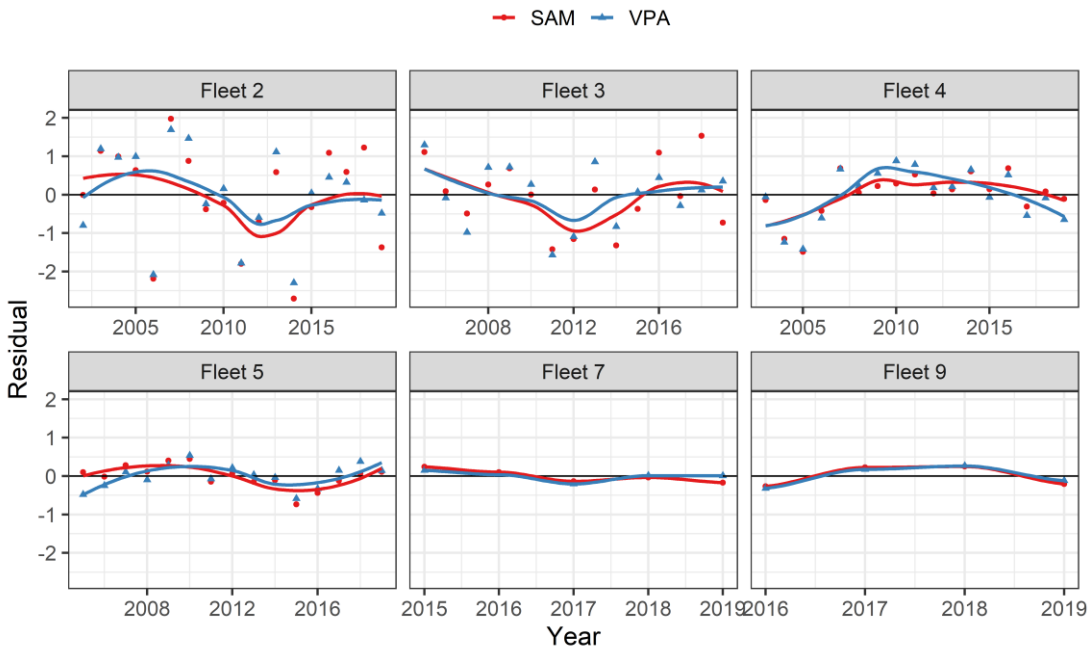

Figure S7: Residuals of abundance indices under the scenario E in SAM and VPA.

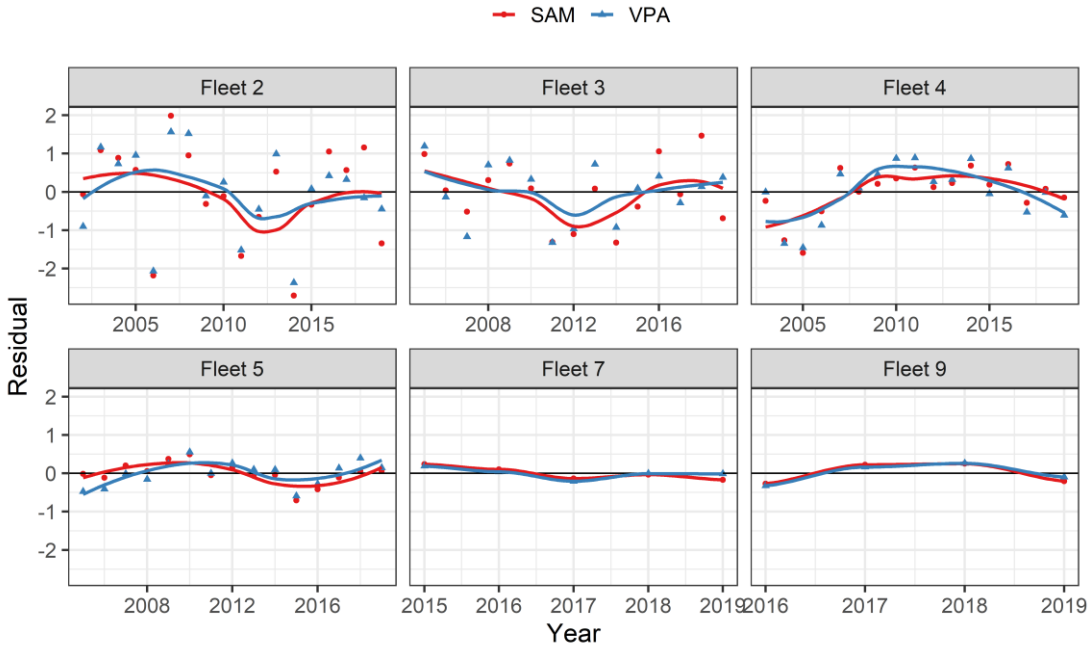

Figure S8: Residuals of abundance indices under the scenario F in SAM and VPA.

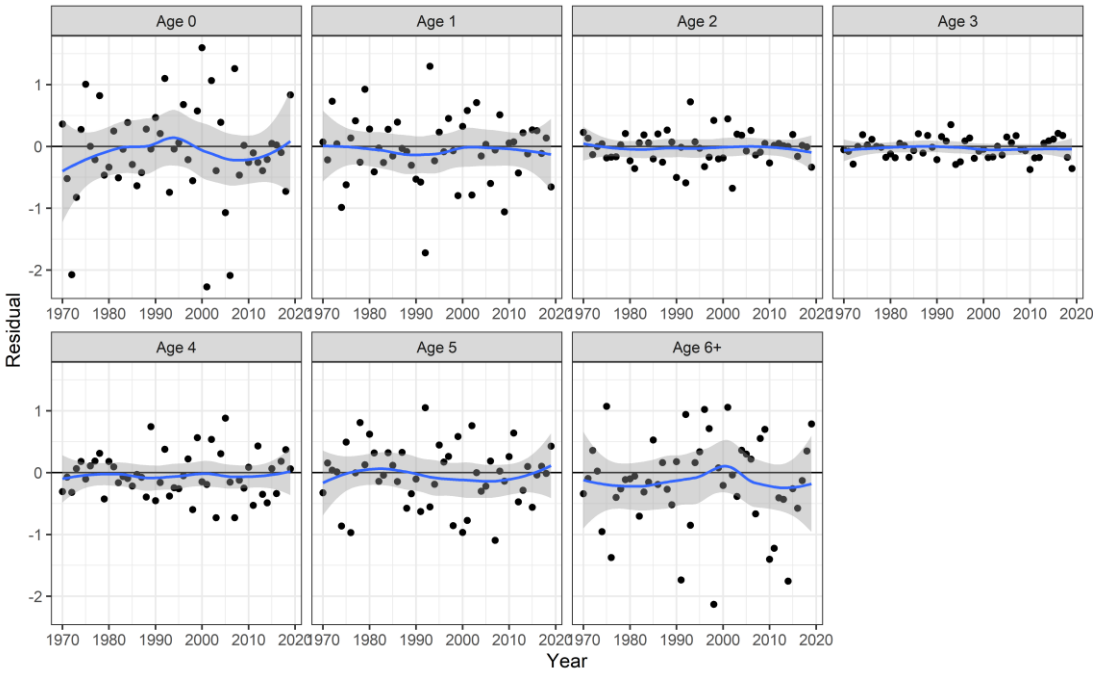

Figure S9: Residuals of catch-at-age under the scenario A in SAM. The blue curves are the prediction by the LOESS regression with 95% confidence intervals.

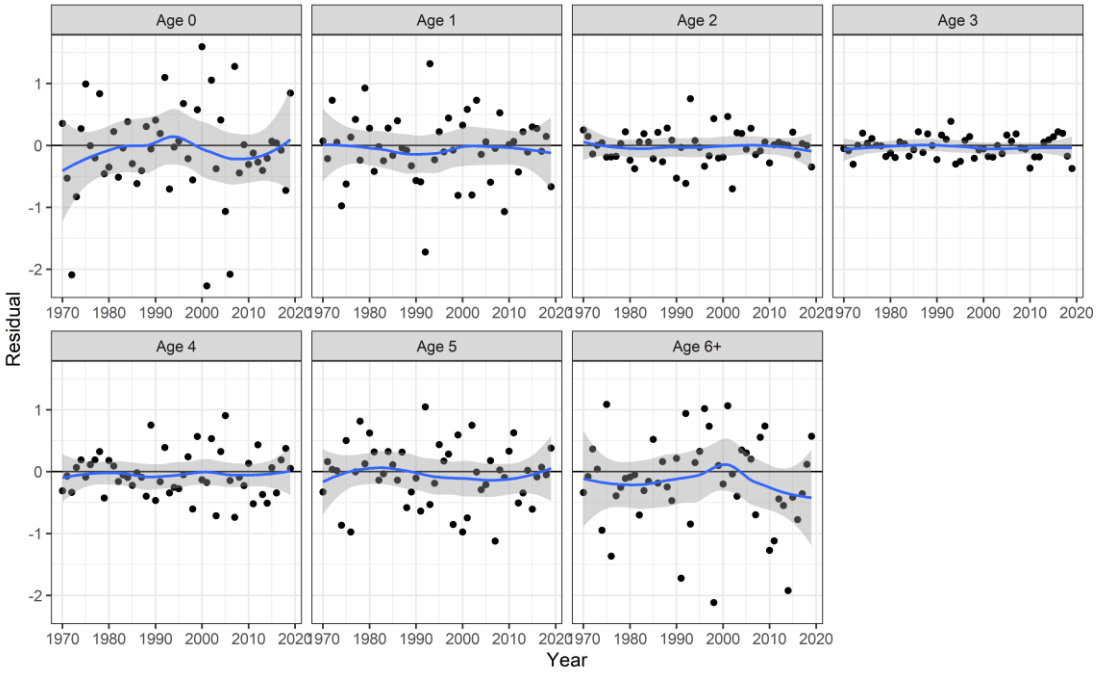

Figure S10: Residuals of catch-at-age under the scenario B in SAM.

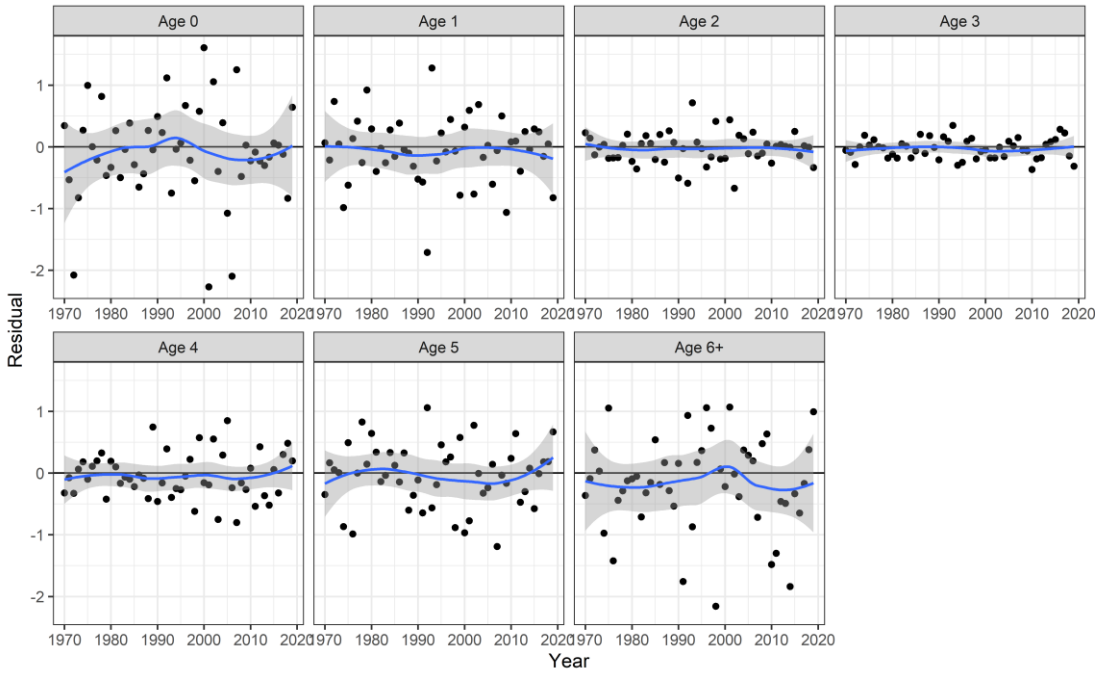

Figure S11: Residuals of catch-at-age under the scenario C in SAM.

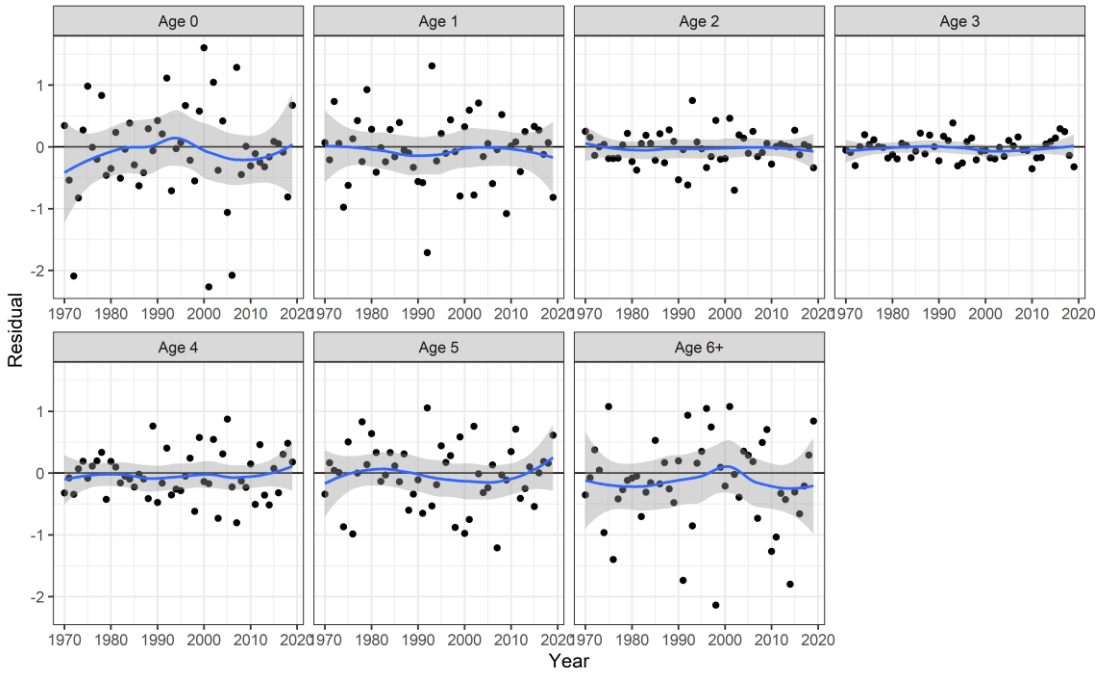

Figure S12: Residuals of catch-at-age under the scenario D in SAM.

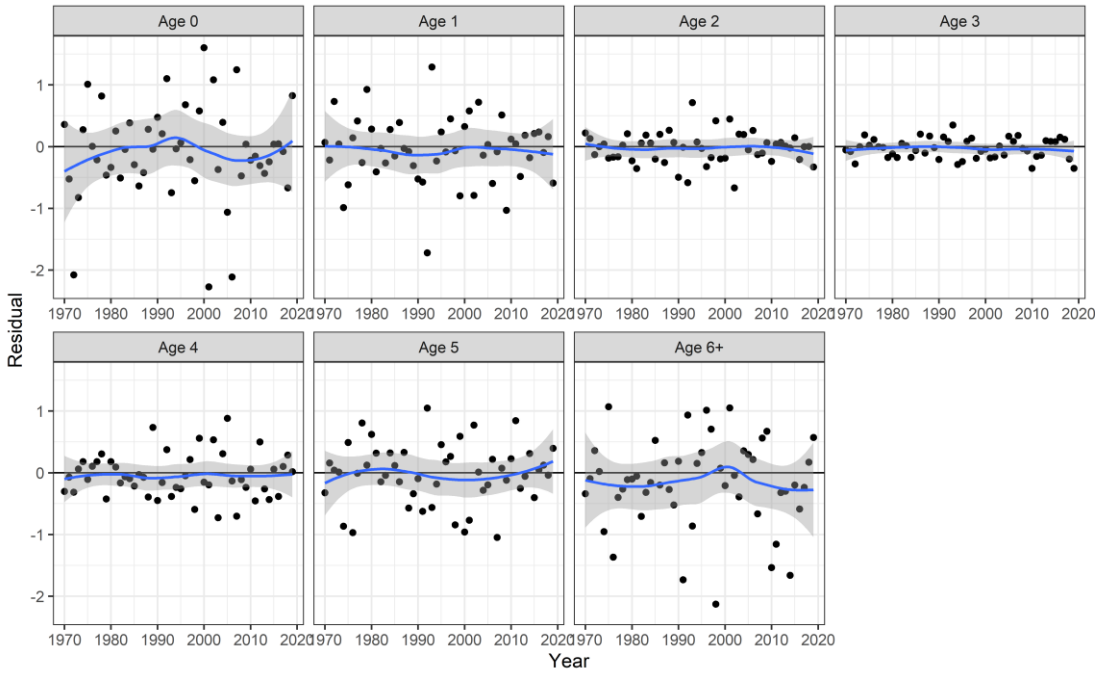

Figure S13: Residuals of catch-at-age under the scenario E in SAM.

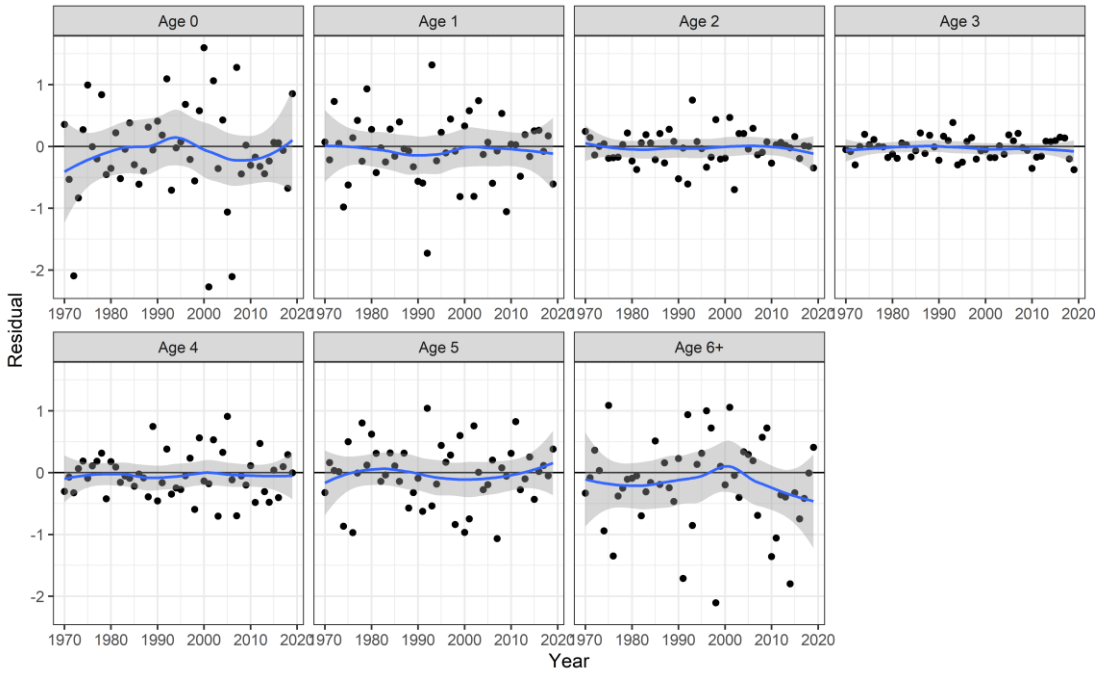

Figure S14: Residuals of catch-at-age under the scenario F in SAM.