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

Shota Nishijima, Ryuji Yukami, Kazuhiro Oshima, and Momoko Ichinokawa

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. The recent abundance estimates in VPA were much higher than those in the SAM and the recent fishing pressure was lower in the VPA. The scenarios under the highest maturity and weight estimated higher SSB in recent years and larger retrospective biases of SSB than the other scenarios in both SAM and VPA. The application of continuous hockey-stick stock-recruit model enabled to estimate feasible biological reference points. We found a few potential problems in both SAM and VPA: (1) the abundance indices from China and Russia were extremely hyper-stable against vulnerable stock size, (2) depletion statistics such as SSB/SSB0 and SSB/SSBmsy were highly sensitive to the choices of stock-recruit function, and (3) the MSY-based reference points were moderately sensitive to the biological parameters of maturity- and weight-at-age. How to treat these problems will be a matter of discussion in the TWG CMSA towards the construction of OM and the benchmark stock assessment.

Note: Working document submitted to the NPFC 4th Meeting of Technical Working Group on Chub Mackerel Stock Assessment, 22–25 June 2021. 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, 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, 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), using the shared data submitted from each Member (Nishijima 2020). At the same time, the TWG CMSA has determined to set six scenarios to include potential uncertainties of biological parameters and data (Table 1) and to rerun the stock assessment models using the determined scenarios by the next TWG CMSA meeting (NPFC 2020). Furthermore, Japan has intersessionally updated natural mortality coefficients (M) based on the re-estimation of von-Bertalanffy growth curve, resulting in higher estimates of *M* than the previous estimates (Nishijima et al. 2021). Here, we report the updated results of VPA and SAM by changing the following points from the previous analysis (Nishijima et al. 2020): (1) analyzing the six scenarios to evaluate the effect of uncertainties of biological parameters, (2) including the abundance indices from China and Russia, (3) estimating a continuous hockey-stick stock-recruit relationship, and (4) estimating some basic biological parameters, such as steepness and SSB0, and biological reference points, such as F0.1 and Fmsy, in the performance measures for evaluating the stock assessment models (NPFC 2020).

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:

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

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

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

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 $M_a = 0.53$ for all age classes under the scenarios A, C, and E, but instead used $M_a = (0.80, 0.60, 0.51, 0.46, 0.43, 0.41, 0.40)$ from age 0 to 6+ for the scenarios B, D, and F. In addition, 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. The fish numbers in the terminal year (2019) were calculated from the fishing morality coefficients in the terminal year:

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

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

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

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

$$F_{6+,y} = F_{5,y}. \tag{6}$$

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):

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

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: $S_{a,y} = F_{a,y}/\max(\mathbf{F_y})$. We estimated $b_{k,y}$ to treat hyperstability or hyperdepletion. We selected the penalty coefficient so as to minimize the absolute value of Mohn's rho (Mohn 1999) of spawning stock biomass in the five-year retrospective analysis:

$$\rho = \frac{1}{5}\sum_{i=1}^{5}\left(\frac{\widehat{SSB}^{R}_{2019-i} - \widehat{SSB}_{2019-i}}{SSB_{2019-i}}\right), \tag{8}$$

where $\hat{F}_{2019-i}$ is the estimate of average fishing mortality coefficients in year $2019-i$ using the full data and $\hat{B}^{R}_{2019-i}$ is the corresponding estimate when removing the data after $2019-i$. Therefore, the ridge VPA can reduce a retrospective bias to some extent. The selected λ were 0.20 for the scenario A, 0.01 for B, 0.30 for C, 0.08 for D, 0.93 for E, and 0.45 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:

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

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

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

where $\eta_{a,y}$ is the process error at age a in year y. The hockey-stick (HS) stock-recruit relationship has been recommended for the Japanese fisheries stocks in domestic stock assessments of Japan to obtain feasible MSY-based reference points (Ichinokawa et al. 2017) and the HS model has been used for this stock in Japan (Nishijima et al. 2019, Yukami et al. 2019). However, SAM is difficult to estimate the usual hockey-stick model owing to the indifferentiable nature of breaking point. Therefore, we alternatively applied a continuous hockey-stick model (Mesnil and Rochet 2010) that smooths the breaking point:

$$f(SSB_y) = \frac{\alpha}{2}\left\{SSB_y + \sqrt{\beta^2 + \frac{\gamma^2}{4}} - \sqrt{(SSB_y - \beta)^2 + \frac{\gamma^2}{4}}\right\}, \tag{12}$$

where α is the slope at the origin, β corresponds to the breaking point, and γ is the smoothing

parameter. We chose γ=100 thousand ton to stably estimate parameters. Using the Beverton-Holt (BH) model estimated an almost linear relationship between recruitment and SSB (Fig. 1), and thus unrealistically large spawning stock biomass without fishing (SSB0) and SSBmsy compared with the estimated historical SSB. The estimates of abundances and exploitation rates when using the BH model were little different from those when using the continuous HS model (Figs. S2, 3). 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 of a failure to converge when estimating this parameter.

The fishing mortality coefficient was assumed to follow the multivariate random walk:

$$\log\left(\boldsymbol{F_y}\right) = \log\left(\boldsymbol{F_{y-1}}\right) + \boldsymbol{\xi_y}, \qquad\qquad \text{if } y \neq 2011 \qquad (12)$$

where $\boldsymbol{F_y} = (F_{1,y}, \dots, F_{A+,y})^T$, $\boldsymbol{\xi_y} \sim \text{MVN}(0, \boldsymbol{\Sigma})$, and $\boldsymbol{\Sigma}$ is the variance-covariance matrix of multivariate normal distribution (MVN). The diagonal elements of matrix $\boldsymbol{\Sigma}$ 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:

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

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

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

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

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

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

$$\sigma_0 = \sigma_1, \sigma_2 = \sigma_3 = \cdots = \sigma_A \,, \tag{16}$$

$$\tau_2 = \tau_3, \tau_5 = \tau_{6+}. \tag{17}$$

These limitations were determined based on the Akaike information criteria (AIC).

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. Moreover, since SAM was difficult to converge in the retrospective analysis, we fixed the stock-recruit parameter β at the estimated value in the full data 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. S4-15).

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 VPA does not assume a stock-recruitment relationship when estimating stock abundances, we estimated the continuous HS relationship after VPA estimation, using SSB and the number of recruits estimated by VPA. 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): $h = 1 − β/SSB0$. 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 panels). 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 1008.7 for the scenario A, 1010.4 for B, 1015.1 for C, 1016.4 for D, 1009.8 for E, and 1009.8 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, right panels). Exceptionally, VPA under the scenario F estimated a similar trend of SSB as SAM. Considering the difference among scenarios within VPA, the scenario E (age common M, the lowest weight and maturity) led to the highest biomass and recruitment, and the lowest exploitation rates, whilst the scenarios C and D (the highest weight and maturity) estimated higher SSB than the other scenarios as in SAM (Fig. 1, right panels). This is

again because the effects of the highest weight- and maturity-at-age with the scenarios C and D exceeded the effects of increased total numbers under the other scenarios (Table 2). The ridge VPA, unlike SAM, cannot be compared by AIC because of use of 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), which was also reflected in smaller standard deviations (SDs) of VPA than those of SAM for the fleets 2 and 3 (Table 3). By contrast, SAM fitted better to a SSB index (fleet 4) in recent years than VPA (Figs. 3-8), which was associated with smaller SDs of SAM than those of VPA for the fleet 4. These differences are a cause for 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 ($b < 1$) 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 under the scenarios A-D in both SAM and VPA, while the SSB index of fleet 4 (dip-net fishery index) showed almost linear relationships except for the scenarios C and D in VPA (Table 3, Figs. 9-14). It is noteworthy that the indices of the fleets 7 and 9 were extremely hyper-stable in VPA with the scenarios with A-D and in SAM with all scenarios, 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 the base-case scenarios A and B (Table 3, Figs. 15-16). We found some positive biases in F and some negative biases in SSB under the scenarios between C and F (Table 3, Figs. 17-20). Accordingly, the scenarios A and B had higher performances than other scenarios in terms of retrospective bias.

VPA was also unlikely to cause serious retrospective biases in abundance estimates except for the biomass under the scenario E (Table 4, Figs. 21-26). However, VPA caused overestimation biases in F under the scenarios A to D, although the biases of F were related to the magnitude of penalty (λ in Equation 7).

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 because higher estimates in recruitment and SSB led to higher parameter values of breaking points (β) (Table 5, Fig. 27). The parameter α (slope of continuous HS relationship at the origin) and steepness were slightly lower in SAM than in VPA, while the relative reference points (%SPRmsy, SSBmsy/SSB0, F/Fmsy) were slightly higher in SAM than in VPA (Table 5), suggesting that VPA provides a more optimistic view to this stock. Comparing different scenarios, the age-common M (A, C, and E) led to the higher SSB-related values (SSB0, SSBmsy), MSY, the relative MSY-based reference points (%SPRmsy, SSBmsy/SSB0), but the lower steepness than the age-specific *M* (B, D and F) (Table 5).

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). Exceptionally, F relative to F0.1 in the terminal year (2019) in VPA under the scenario B became abruptly high. This is because the instability in estimation in VPA caused enormous F for ages 5 and 6+ in the terminal year (Fig. 2). %SPR was relatively robust against the choice of SAM or VPA (Fig. 28, top 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 increased since 2015 for the scenarios A, B, E and F in both SAM and VPA. In particular, SAM under the scenarios E and F (the lowest weight and maturity) estimated higher F than Fmsy in the latest three years. There are three reasons for this conspicuous pattern: (1) SAM estimated higher exploitation rates in recent years (Fig. 1, bottom panels), (2) the steepness (or resilience to fishing) was estimated to be lower in SAM than in VPA (Table 5), and (3) the scenarios E and E led to a steep decline of SPR0 due to the decreased weights and the delayed maturity, causing an increase in %SPRmsy especially for SAM (Fig. 29). In other words, F/Fmsy increased recently, because the changes in the biological parameters of maturity- and weight-at-age decreased Fmsy although the

exploitation rates were kept at low levels. Since the steepness was lower in the age-common M scenarios than in the age-specific M scenarios (Table 5), the scenario E in SAM had higher F/Fmsy than the scenario F (Fig. 28, bottom panels).

The values of SSB/SSB0 in recent years were robust between SAM and VPA, but the values in past years were quite different between the two models (Fig. 30, top panels): SSB/SSB0 in recent five years was estimated to be around 0.50 in both models, whilst SSB/SSB0 in the 1970s was around 0.30 in SAM but 0.20 in VPA. SSB in the recent five years was around SSBmsy for the scenarios A, B, E, and F, but over SSBmsy for the scenario D and E in both SAM and VPA, whereas in the 1970s SSB was around SSBmsy with all scenarios in SAM, but lower than SSBmsy with all scenarios except for the scenario F in VPA (Fig. 30, 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.

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 estimating stock-recruitment relationship. VPA, which does not have a stock-recruitment relationship internally, 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. Although a previous simulation study demonstrated the ridge VPA estimated more correctly recent abundances than SAM in a situation with few abundance indices (Okamura et al. 2018), this may not be applicable for chub mackerel because more abundance indices are available.

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 for 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. Moreover, these abundance indices elevated the difficulty in model estimation and convergence because of their short timeseries, which could prevent us from completing retrospective analysis. There, hence, deems to be some rooms to discuss about the application of those abundance indices to not only the benchmark stock assessment but also the model comparison via operating models. Consequently, at the present stage, we can suppose excluding these abundance indices from base-case scenarios but including them in a sensitivity trial as another scenario.

The scenarios with the highest maturity and weight (C and D) resulted in much higher SSB for recent years among the six scenarios in both SAM and VPA (Fig. 1, middle panels). This is because the setting of highest maturity and weight increased SSB (Table 2). In SAM, the values of AIC were higher in the scenario C and D and the retrospective biases in SSB and F were higher than the other scenarios (Table 4, Figs. 15-20). In VPA, the retrospective bias in SSB was also higher than the other scenarios (Table 4, Figs. 21-26). As the rates of maturity and growth certainly decreased along with the recent increase in stock size (Kamimura et al. 2021; Manabe et al. 2021), the retrospective analyses provide supportive evidence of the recent decline in weight- and maturity-at-age. Therefore, it may better to prioritize the other scenarios than the scenarios of highest weight and maturity when evaluating model performance by operating models. 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). This suggests that the uncertainties of maturity and weight will be problematic.

The MSY-based reference points were also sensitive to the choice of stock-recruit relationship. The Beverton-Holt stock-recruitment relationship is the most common but did not work well for

the this stock of chub mackerel because it estimated almost linear relationships and huge SSB0 and SSBmsy that were implausible for biological reference points (Fig. S1). We alternatively used the continuous hockey-stick relationship which provided plausible MSY-based reference points. A recent meta-analysis (Zhou et al. 2020) estimated the order of Perciformes estimated at 46.3 ± 17.1% (mean ± SD). Therefore, the range of estimated SPRmsy (22% to 33% in Table 4) falls within the 95% interval of this meta-analysis (mean ± 1.96SD). The high sensitivity of MSY-based reference points to the assumed type of stock-recruitment relationship will reduce the importance of MSY-based performance measures to be used in the stock assessment model competition through the operating model process due to much different MSY-based reference points even when similar abundance estimates were estimated: the comparison of MSY-based reference points is likely to make no sense. Considering the high sensitivity of MSY-based reference points, a possible option is to put less priority on (or not use daringly) the MSY-based reference points in performance measures of operation model testing. Instead, it will be better to focus on the performance measures that do not use a stock-recruit relationship, as the start point of our discussion. Candidates of the performance measures are absolute biomass and numbers at some historical benchmark years (first, middle, and last year of the stock assessment period), and the ratio showing the scale of the historical trends such as SSB/SSBmax (Table 2) as well as F0.1, and F%SPR.

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 in SAM (Fig. 28, bottom-left panel), although the exploitation rate was kept at a low level (Fig. 1, bottom-left panel). 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.

Ichinokawa, M., H. Okamura., H. Kurota. 2017. The status of Japanese fisheries relative to fisheries around the world. ICES Journal of Marine Science 74: 1277–1287.

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.

Mesnil, B., and M. J. Rochet. 2010. A continuous hockey stick stock-recruit model for estimating MSY reference points. ICES Journal of Marine Science 67:1780–1784.

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

Nishijima, S., R. Yukami, S. Isu, Y. Kamimura, and S. Furuichi. 2019. Research Institute Meeting Report on (Biological) Reference Points for the Pacific Stock of Chub Mackerel (*Scomber Japonicus*) in FY2019.

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. Page NPFC-2020-TWG CMSA03-WP05.

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. Page 25 NPFC-2020-TWG CMSA03-Final Report. (available at www.npfc.int)

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., Y. Yamashita, M. Ichinokawa, and S. Nishijima. 2018. Comparison of the performance of age-structured models with few survey indices. ICES Journal of Marine Science 75:2016–2024.

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.

Yukami, R., S. Nishijima, S. Isu, C. Watanabe, Y. Kamimura, and S. Furuichi. 2019. Stock assessment and evaluation for the Pacific stock of chub mackerel (fiscal year 2018). Pages 163-208 (in Japanese) Marine fisheries stock assessment and evaluation for Japanese waters (fiscal year 2018/2019). Fisheries Agency and Fisheries Research and Education Agency of Japan.

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[1]	Weight -at-age	Maturity -at-age	Catch (at-age)	Abundance index	Fleet
A	Base-case 1	0.41 → 0.53	Average	Average	Average	All six	Single
B	Base-case 2	Gislason	Average	Average	Average	All six	Single
C	Highest weight and maturity	0.41 → 0.53	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 → 0.53	Lowest	Lowest	Average	All six	Single
F	Lowest weight and maturity	Gislason	Lowest	Lowest	Average	All six	Single

1: The median values of natural mortality coefficients (*M*) have been changed from 0.41 to 0.53 according to the update on the von-Bertalanffy growth curve (Nishijima et al. 2021).

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)																	
	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
SAM	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
	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
VPA	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)																	
	A	54.2	(1997)	385.7	1528.3	(2018)	589.2	793.1	1159.4	110.7	72.6	255.6	1312.1	24.23	3.40	0.86	2.23
	B	48.7	(2002)	342.5	1361.4	(1979)	505.8	713.4	1072.5	90.0	62.9	187.4	1050.1	21.55	3.07	0.77	2.08
	C	54.1	(1997)	374.3	3027.2	(2018)	735.2	793.3	1163.0	110.7	72.4	233.4	2918.0	53.97	7.80	0.96	3.97
SAM	D	48.5	(2002)	343.2	2521.5	(2018)	631.8	713.6	1069.9	90.1	62.9	174.8	2415.5	49.78	7.04	0.96	3.82
	E	53.8	(1997)	401.7	1492.8	(1979)	558.7	787.6	1182.3	109.5	72.7	274.0	1035.0	19.26	2.58	0.69	1.85
	F	48.6	(2002)	349.3	1344.0	(1979)	493.7	711.5	1048.2	89.7	63.1	226.8	935.8	19.25	2.68	0.70	1.90
VPA	A	56.2	(2002)	405.1	3446.6	(2019)	699.1	774.8	1295.3	136.0	82.7	149.0	3446.6	61.35	8.51	1.00	4.93

B	48.0	(2002)	350.1	2509.4	(2019)	582.5	690.2	1135.8	104.2	68.8	126.8	2509.4	52.31	7.17	1.00	4.31
C	56.2	(2002)	405.1	9569.5	(2019)	998.5	774.8	1295.3	136.0	82.7	149.4	9569.5	170.35	23.62	1.00	9.58
D	48.0	(2002)	350.1	7015.4	(2019)	807.8	690.2	1135.8	104.2	68.8	126.9	7015.4	146.25	20.04	1.00	8.69
E	56.2	(2002)	415.0	2904.6	(2019)	670.8	774.8	1295.3	136.0	82.7	159.3	2904.6	51.71	7.00	1.00	4.33
F	48.0	(2002)	346.5	1493.2	(1978)	500.4	690.2	1135.8	104.2	68.8	127.9	1197.2	24.96	3.46	0.80	2.39

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	1.56*	1.91*	1.13	0.64*	0.00	0.00	1.24	0.84	0.52	0.31	0.15	0.24
	B	1.61*	1.94*	1.15	0.65*	0.00	0.00	1.22	0.81	0.56	0.32	0.15	0.24
	C	1.55*	1.89*	0.82	0.45*	0.00	0.00	1.26	0.88	0.63	0.34	0.15	0.24
	D	1.59*	1.91*	0.82	0.45*	0.00	0.00	1.24	0.85	0.67	0.35	0.15	0.24
	E	1.43	1.77*	1.25	0.76	0.00	0.00	1.31	0.95	0.51	0.27	0.15	0.24
	F	1.47*	1.81*	1.26	0.77	0.00	0.00	1.29	0.93	0.53	0.27	0.15	0.24
VPA	A	1.26*	1.45*	0.87	0.50*	0.21*	0.10*	1.18	0.75	0.70	0.31	0.11	0.23
	B	1.33*	1.51*	0.91	0.53*	0.15*	0.10*	1.15	0.73	0.72	0.31	0.11	0.23
	C	1.23	1.42*	0.64*	0.37*	0.19*	0.13*	1.18	0.76	0.75	0.33	0.12	0.23
	D	1.30*	1.48*	0.66*	0.38*	0.15*	0.12*	1.15	0.73	0.78	0.34	0.12	0.23
	E	1.03	1.21	0.95	0.55	0.62	0.17	1.26	0.83	0.65	0.30	0.09	0.24
	F	1.25	1.43*	1.25	0.74	0.34	0.17*	1.18	0.75	0.68	0.30	0.11	0.23

* 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	-0.10	-0.18	0.02	0.25
	B	-0.06	-0.14	0.03	0.17
	C	-0.15	-0.26	0.03	0.39
	D	-0.15	-0.28	0.00	0.67
	E	-0.19	-0.22	-0.04	0.37
	F	-0.15	-0.20	-0.02	0.37
VPA	A	0.14	0.00	-0.22	1.08
	B	0.08	0.02	-0.21	2.18
	C	0.15	0.16	-0.13	0.86
	D	0.12	0.18	-0.11	1.07
	E	0.39	0.00	-0.11	0.23
	F	0.14	0.00	-0.12	0.35

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	0.0084	1400	0.75	376	4420	12	0.68	1420	986	33	0.32	1.09
	B	0.0123	1000	0.76	329	4080	12	0.75	1040	825	26	0.25	1.08
	C	0.0083	1400	0.74	394	4580	12	0.69	1430	1023	32	0.31	1.02
	D	0.0123	1000	0.76	343	4220	12	0.76	1040	858	25	0.25	1.01
	E	0.0092	1130	0.79	364	3790	10	0.70	1160	865	31	0.31	1.04
	F	0.013	1000	0.79	320	4160	13	0.76	1030	848	25	0.25	1.06
VPA	A	0.0091	2960	0.87	376	10080	27	0.71	2990	2272	30	0.30	0.99
	B	0.0130	2160	0.88	329	9200	28	0.77	2200	1855	24	0.24	1.02
	C	0.0085	5180	0.85	394	17350	44	0.70	5210	3852	30	0.30	0.98
	D	0.0121	3730	0.87	343	15550	45	0.76	3770	3110	24	0.24	1.00
	E	0.0102	2410	0.99	364	8960	25	0.73	2450	2083	27	0.27	0.91
	F	0.0151	950	0.97	320	4600	14	0.79	990	945	22	0.22	0.95

Figures
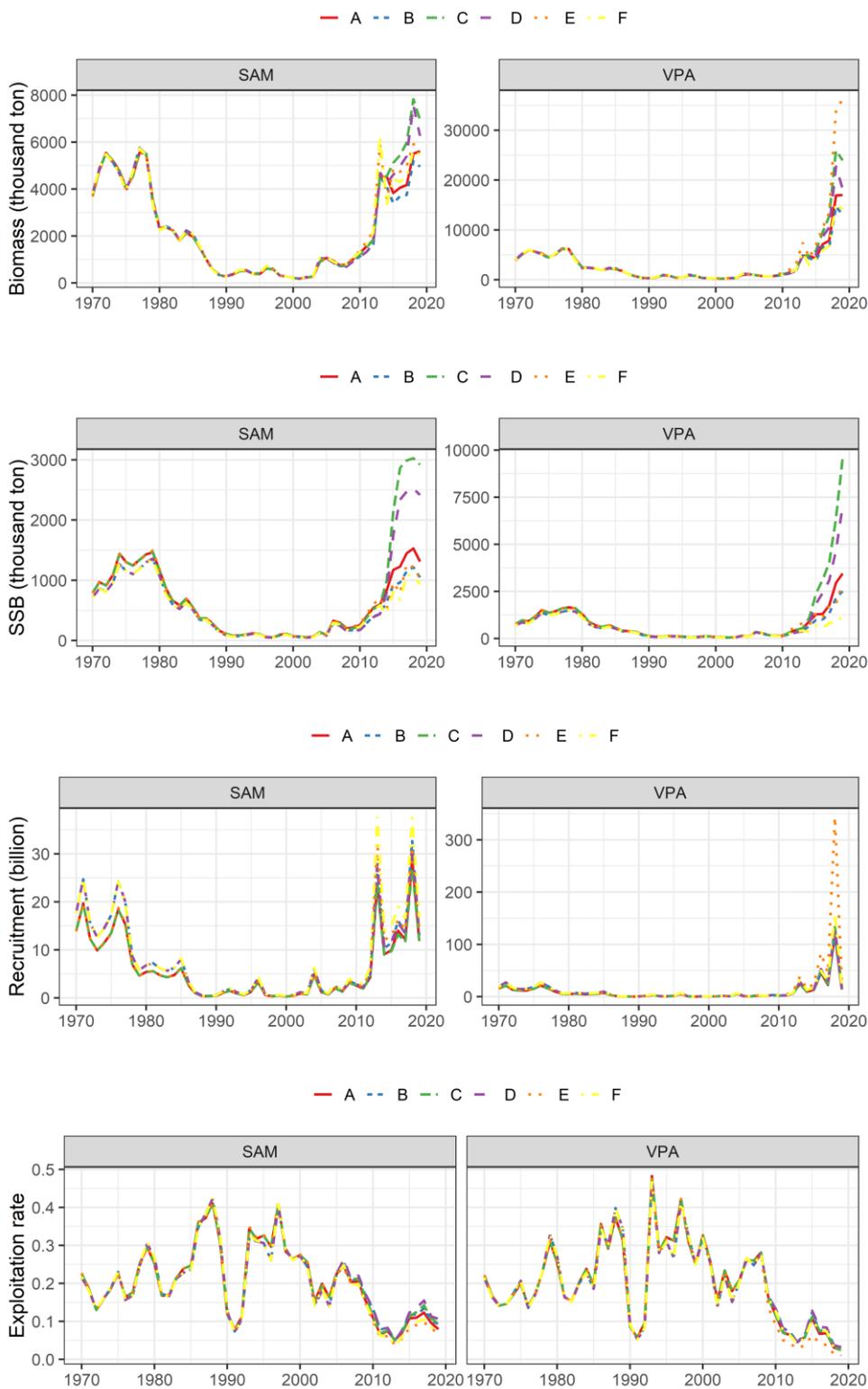
Figure 1: the estimates of total biomass (1st column), SSB (2nd column), recruitment number (3rd column), and exploitation rate (4th column) with SAM (left) and VPA (right) under the scenarios A to F.
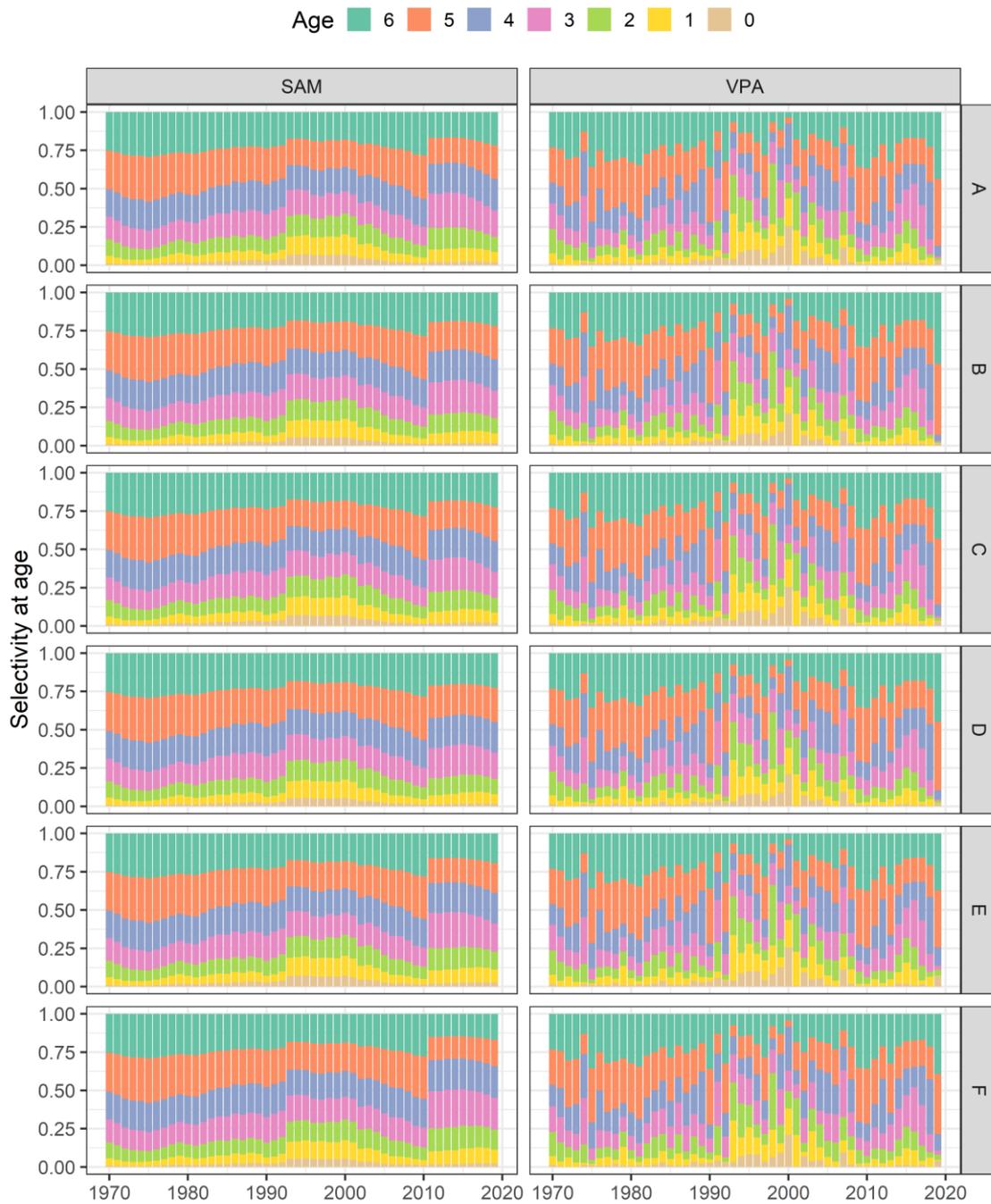
Figure 2: Selectivity at age in SAM (left) and VPA (right) under the scenario A to F. Selectivity is scaled so that its sum is equal to one.
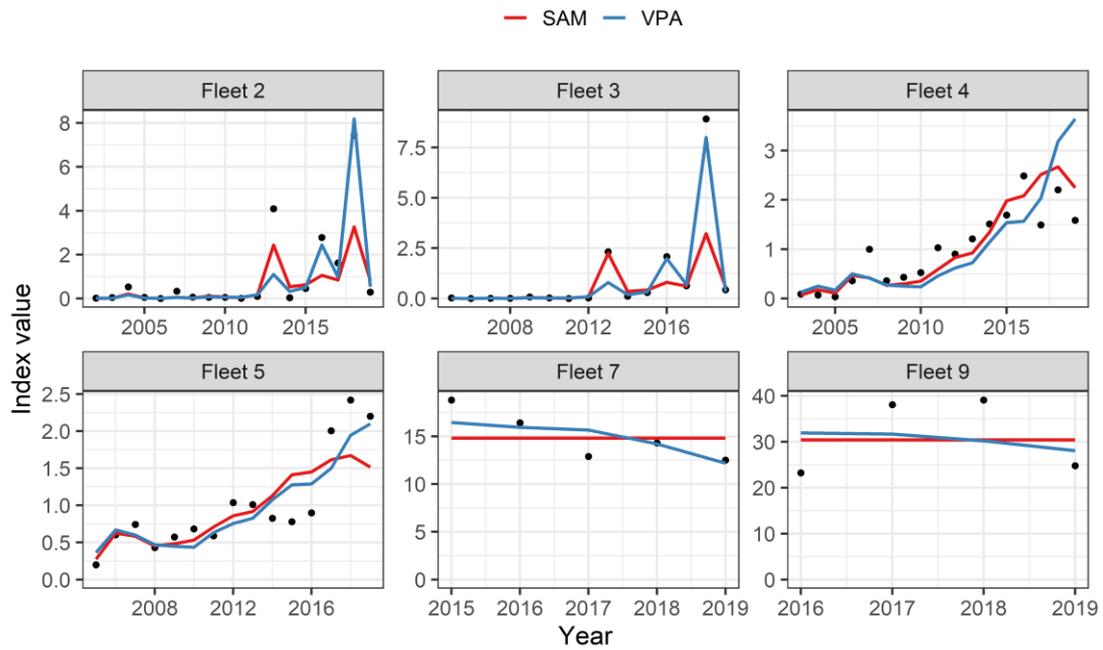
Figure 3: Index values observed (points) and their predicted values by SAM (red lines) and VPA (blue lines) under the scenario A.
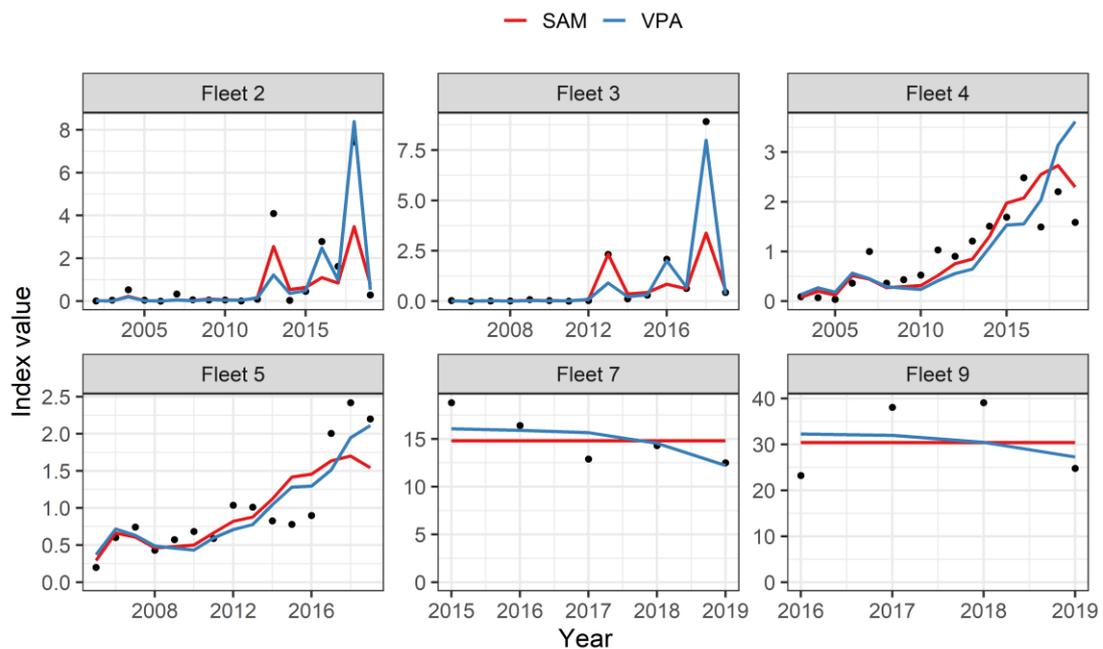
Figure 4: Index values observed (points) and their predicted values by SAM (red lines) and VPA (blue lines) under the scenario B.
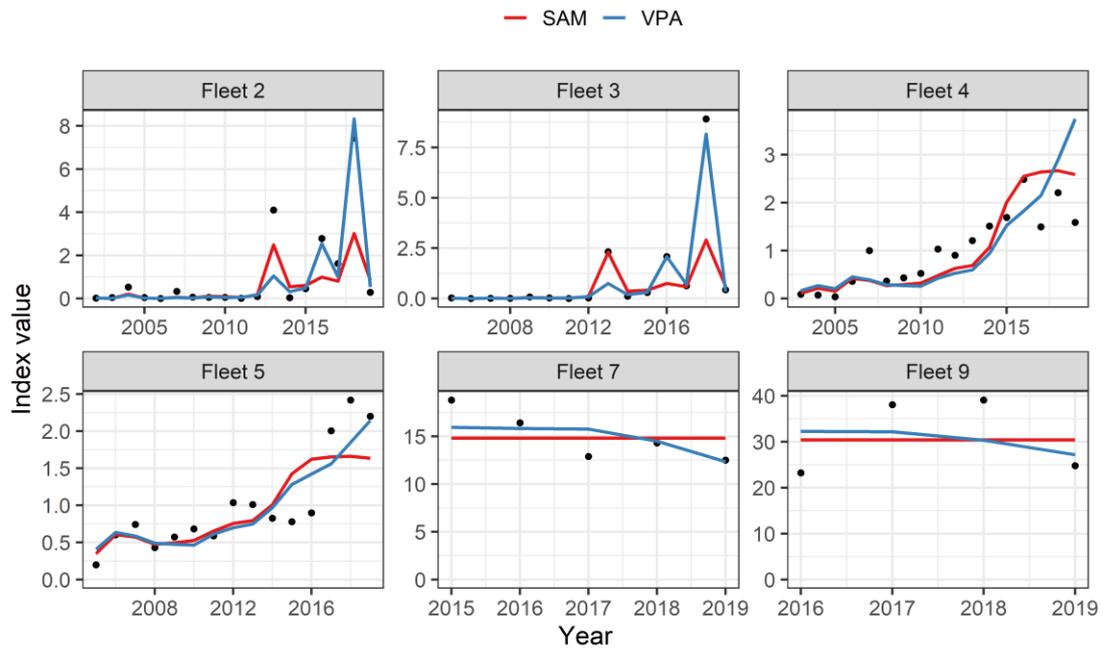
Figure 5: Index values observed (points) and their predicted values by SAM (red lines) and VPA (blue lines) under the scenario C.
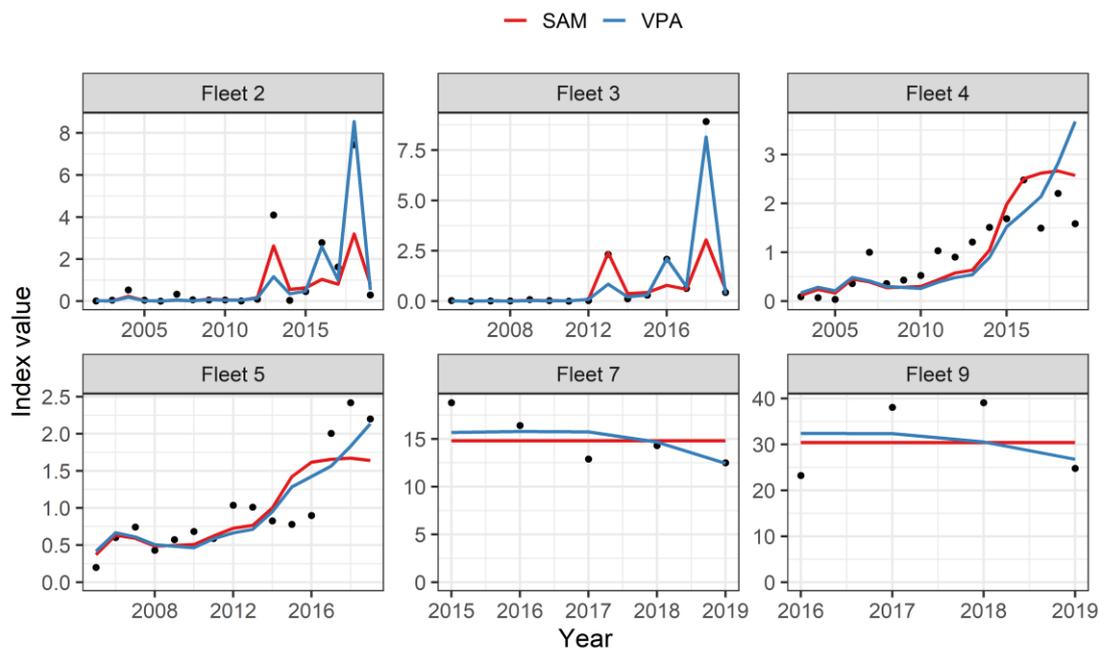
Figure 6: Index values observed (points) and their predicted values by SAM (red lines) and VPA (blue lines) under the scenario D.
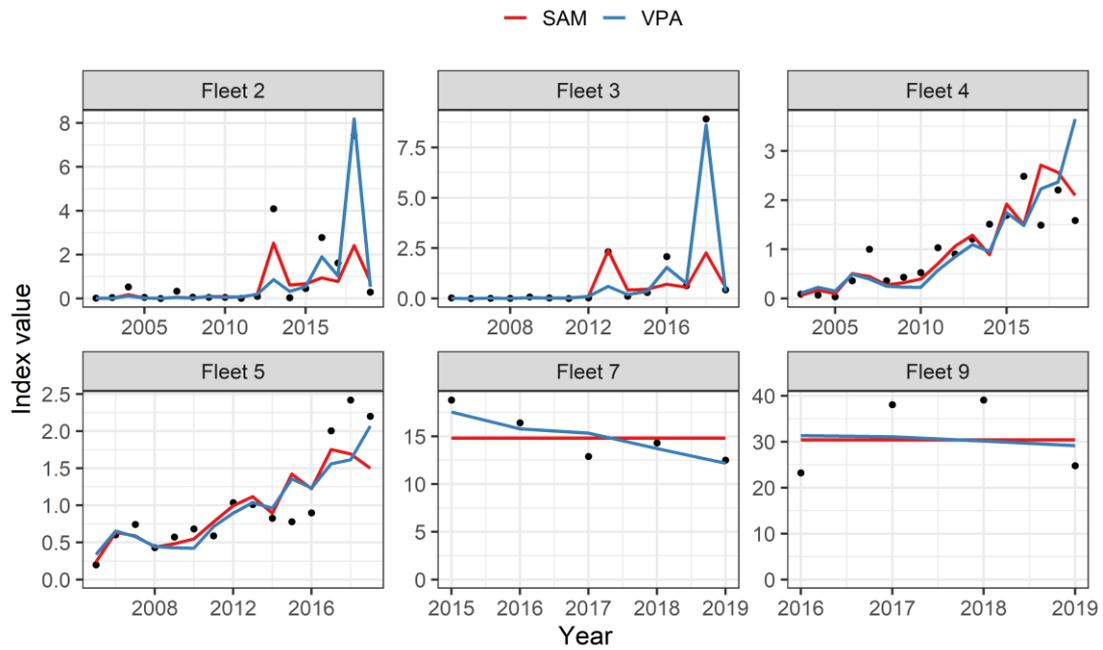
Figure 7: Index values observed (points) and their predicted values by SAM (red lines) and VPA (blue lines) under the scenario E.
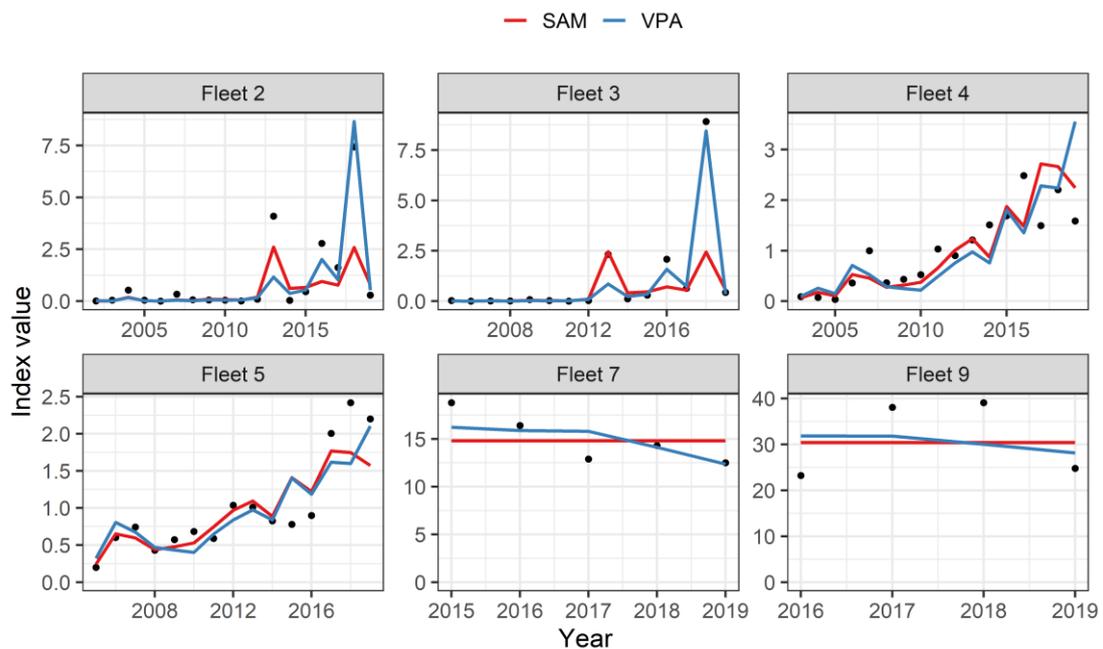
Figure 8: Index values observed (points) and their predicted values by SAM (red lines) and VPA (blue lines) under the scenario F.
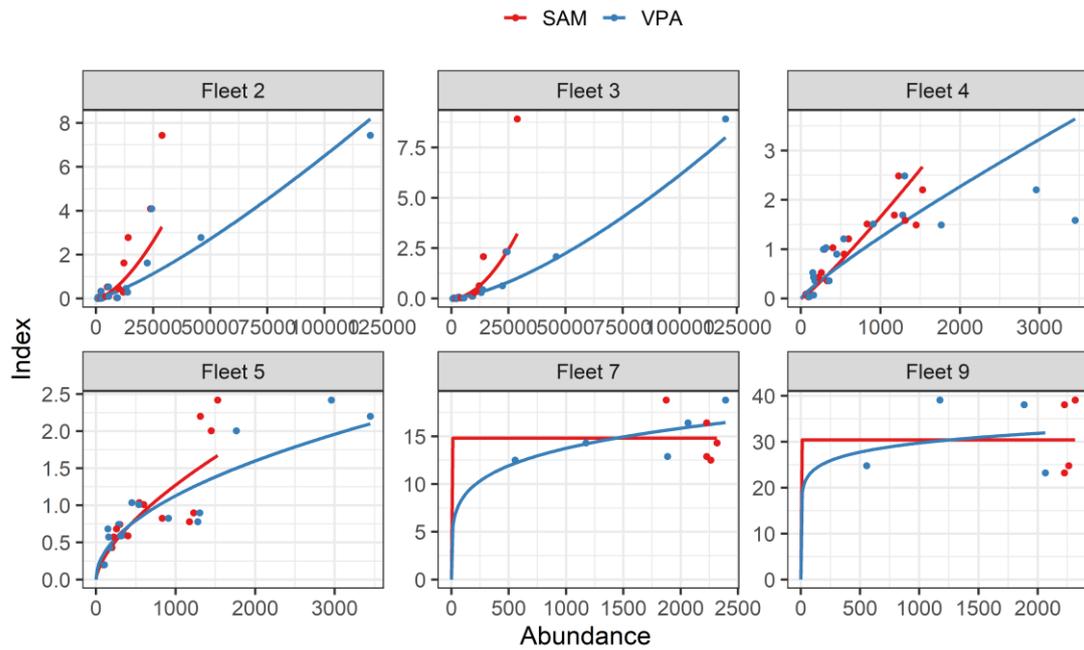
Figure 9: The relationships between abundance indices and their corresponding abundance estimates in SAM (red) and VPA (blue) under the scenario A.
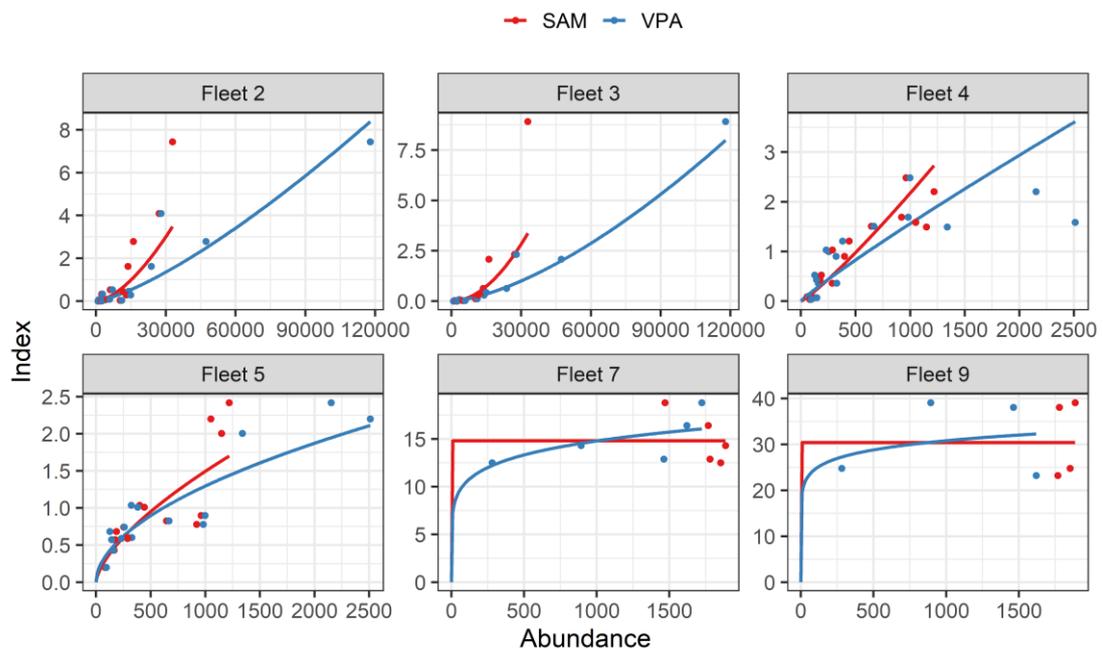
Figure 10: Relationships between abundance indices and their corresponding abundance estimates in SAM (red) and VPA (blue) under the scenario B.
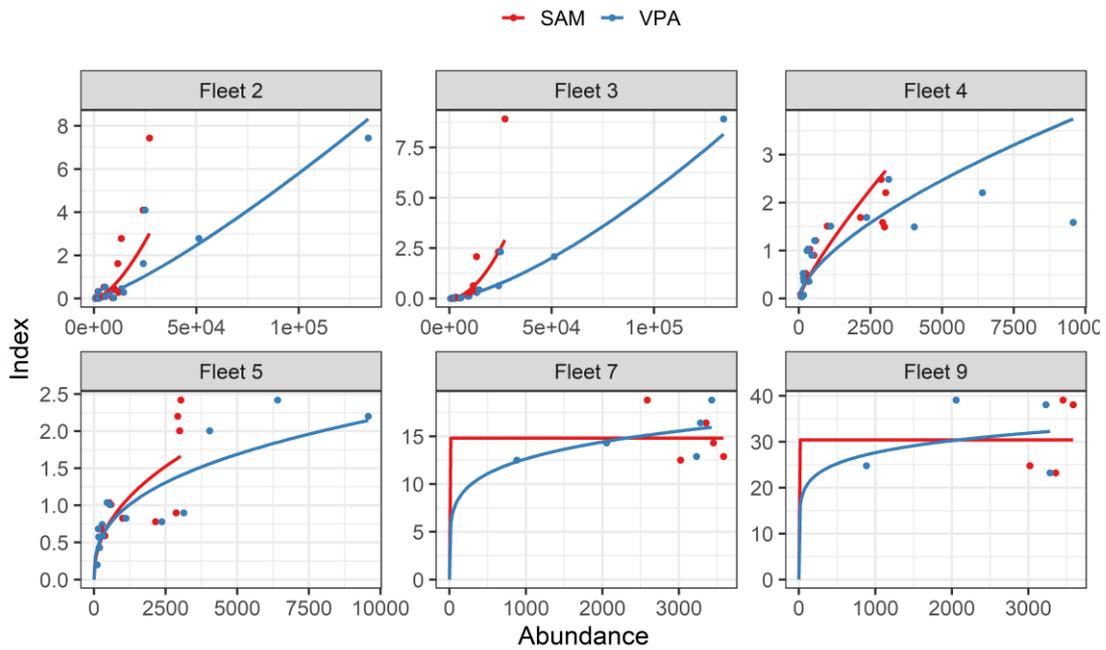
Figure 11: Relationships between abundance indices and their corresponding abundance estimates in SAM (red) and VPA (blue) under the scenario C.
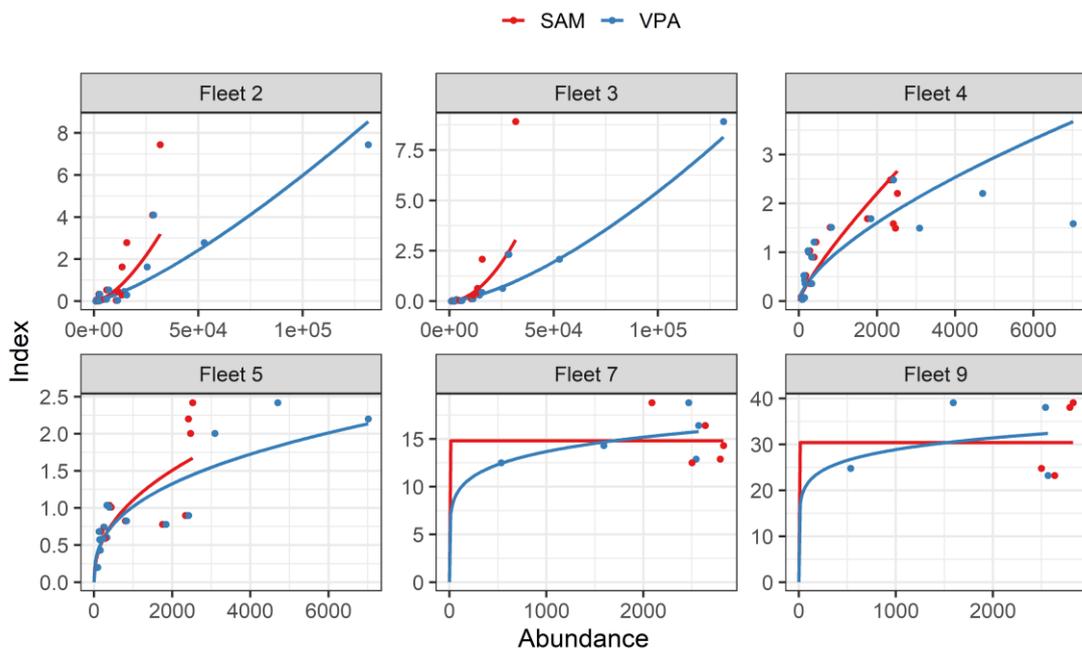
Figure 12: Relationships between abundance indices and their corresponding abundance estimates in SAM (red) and VPA (blue) under the scenario D.
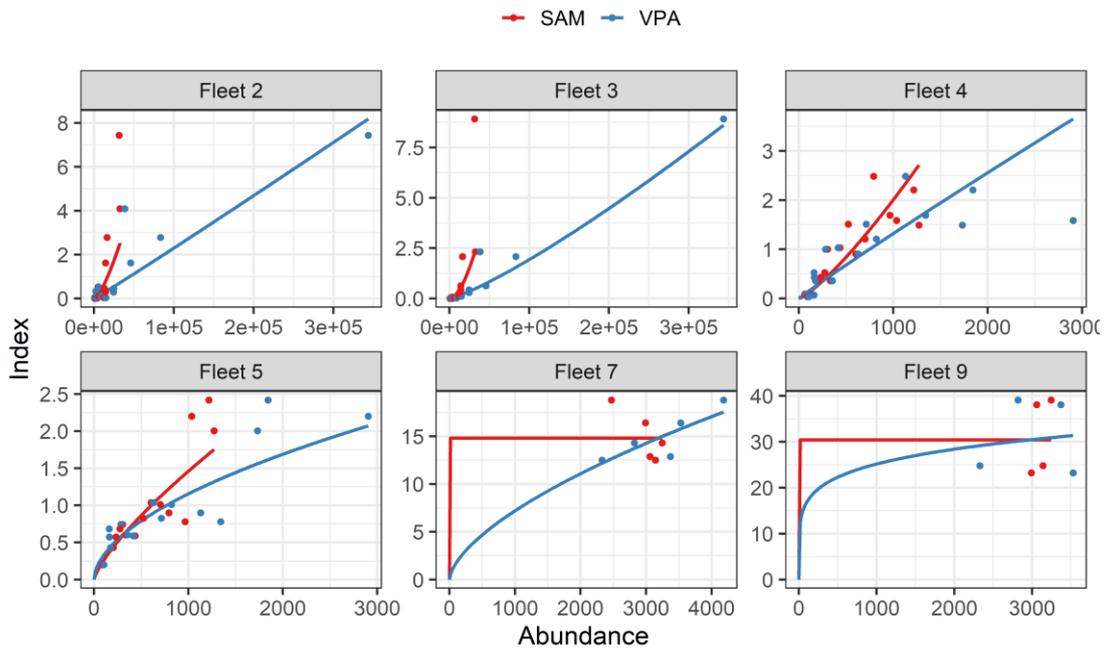
Figure 13: Relationships between abundance indices and their corresponding abundance estimates in SAM (red) and VPA (blue) under the scenario E.
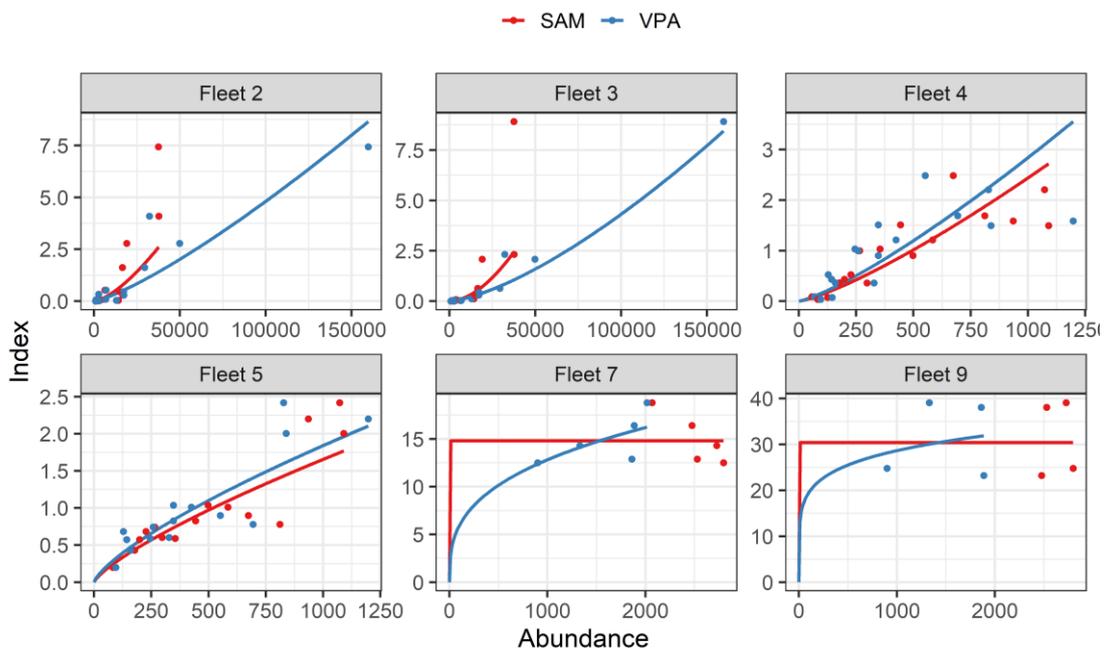
Figure 14: Relationships between abundance indices and their corresponding abundance estimates in SAM (red) and VPA (blue) under the scenario F.

Figure 15: Retrospective pattens SAM under the scenario A.

Figure 16: Retrospective pattens of SAM under the scenario B.
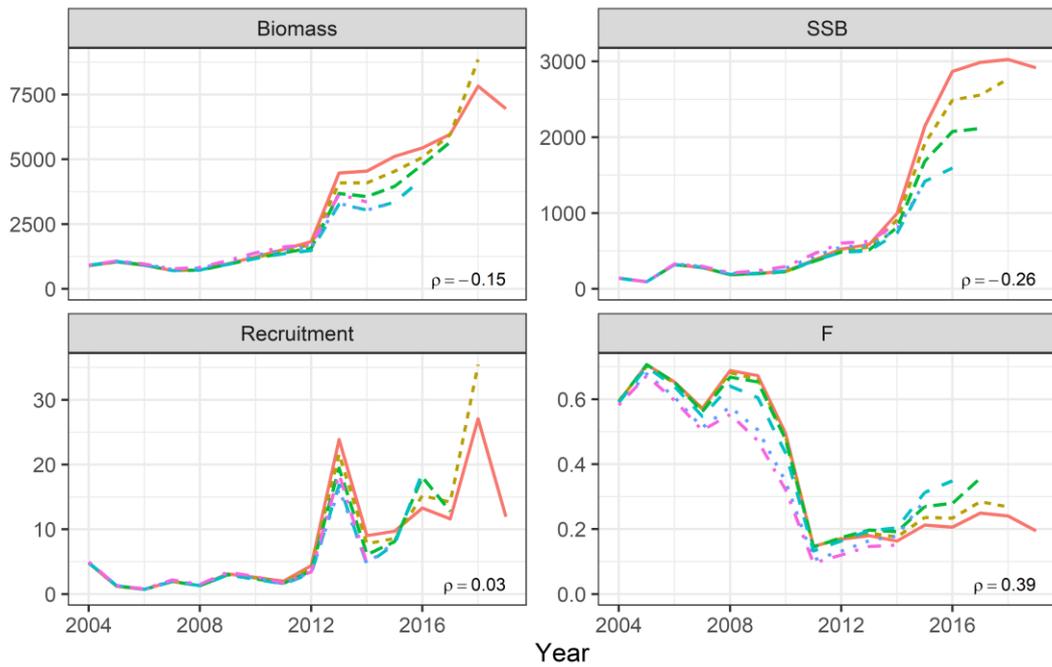
Figure 17: Retrospective pattens of SAM under the scenario C.

Figure 18: Retrospective pattens of SAM under the scenario D.
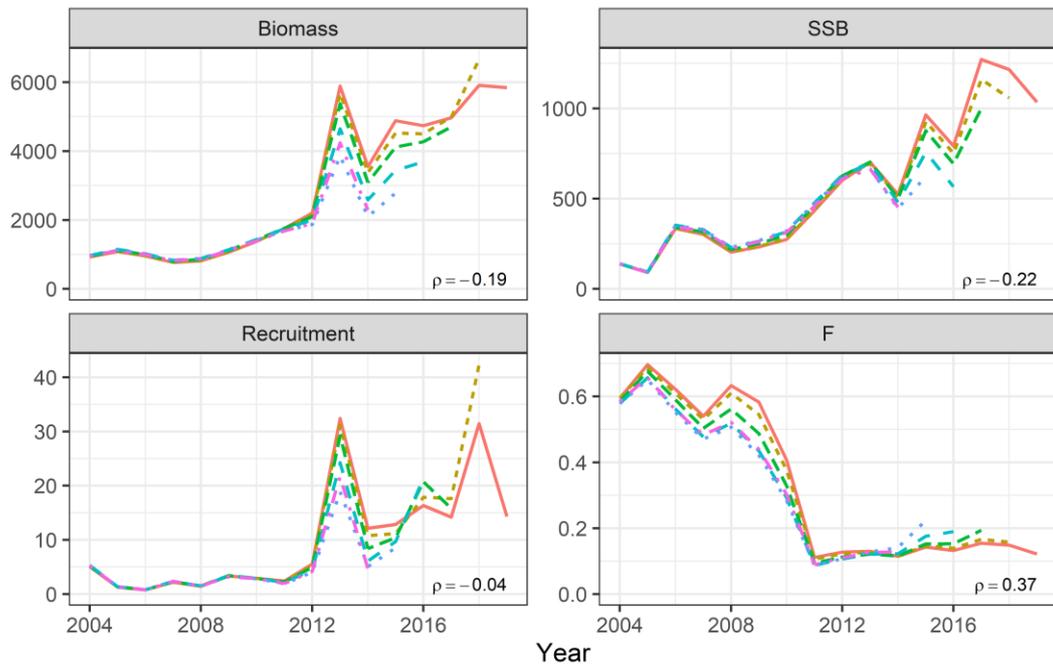
Figure 19: Retrospective pattens of SAM under the scenario E.
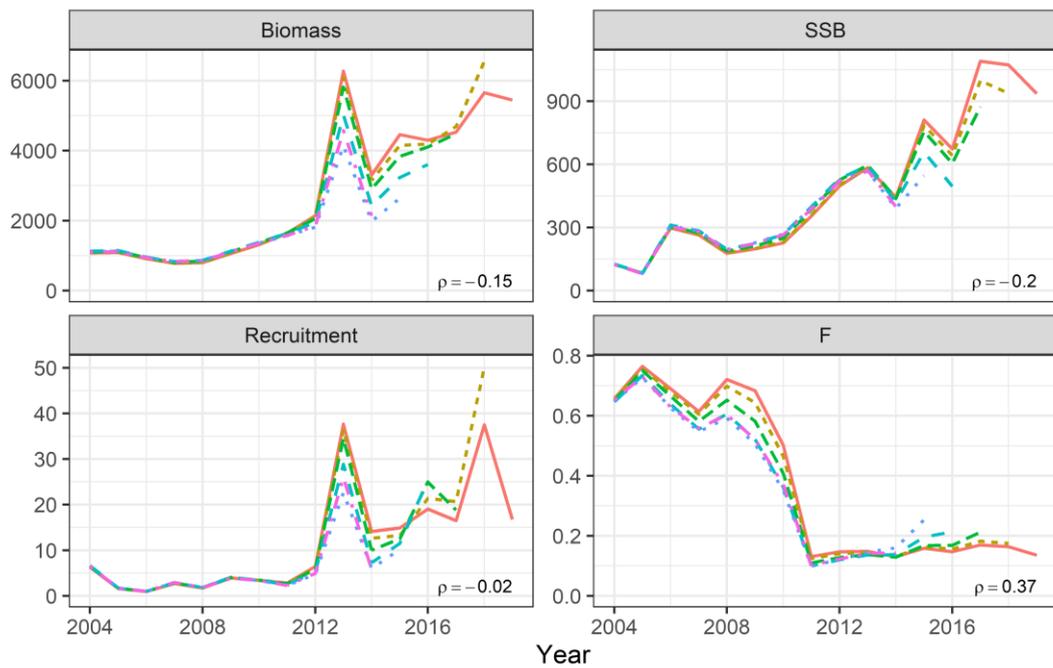
Figure 20: Retrospective pattens of SAM under the scenario F.
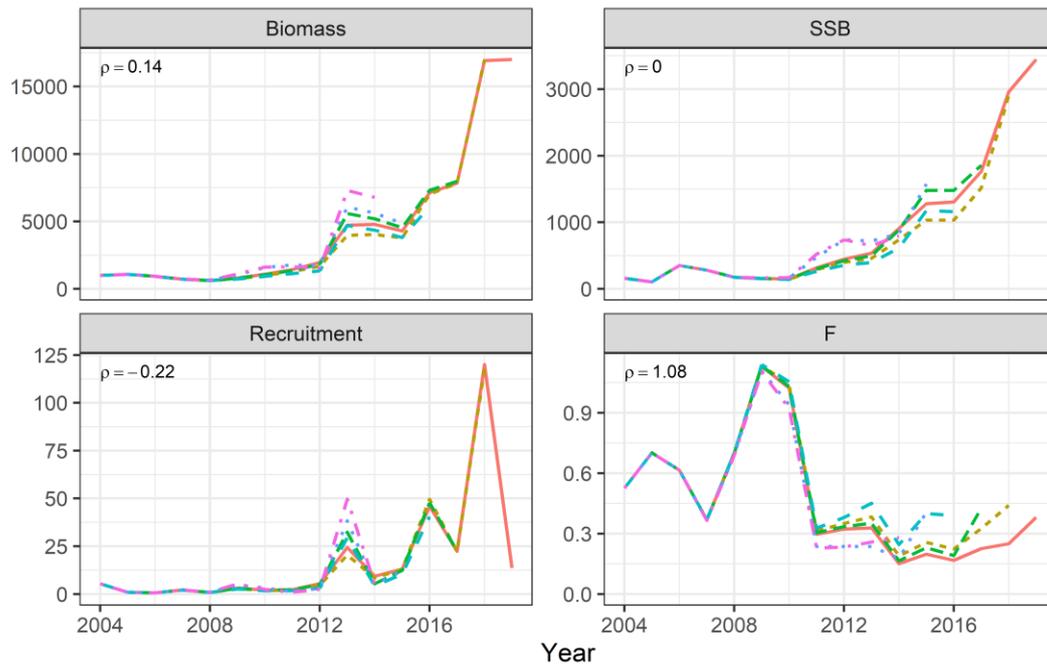
Figure 21: Retrospective pattens of VPA under the scenario A.
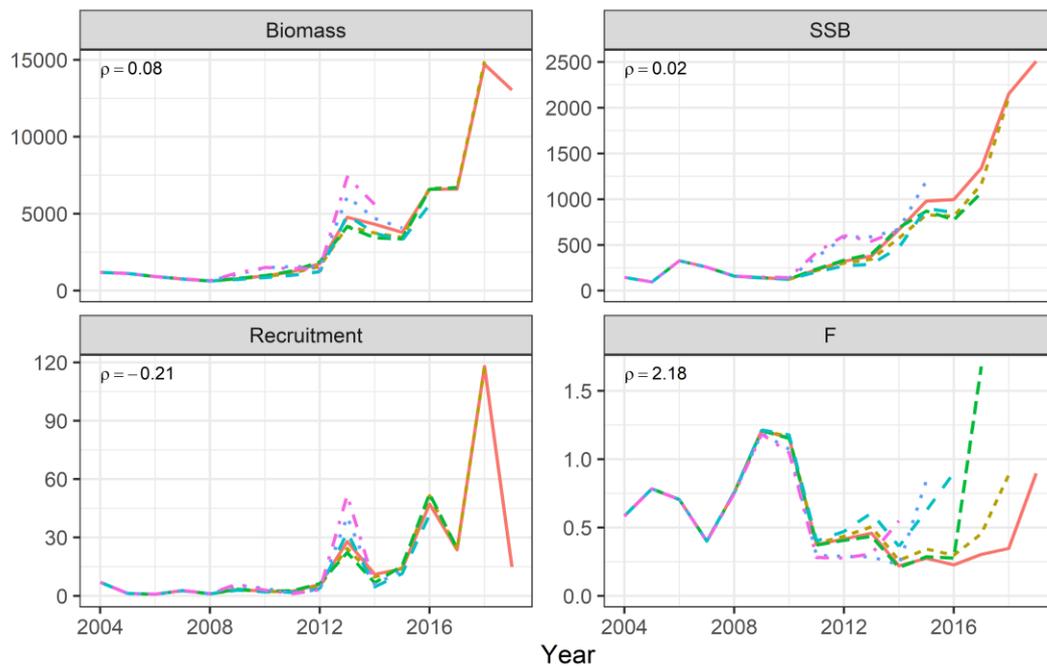
Figure 22: Retrospective pattens of VPA under the scenario B.
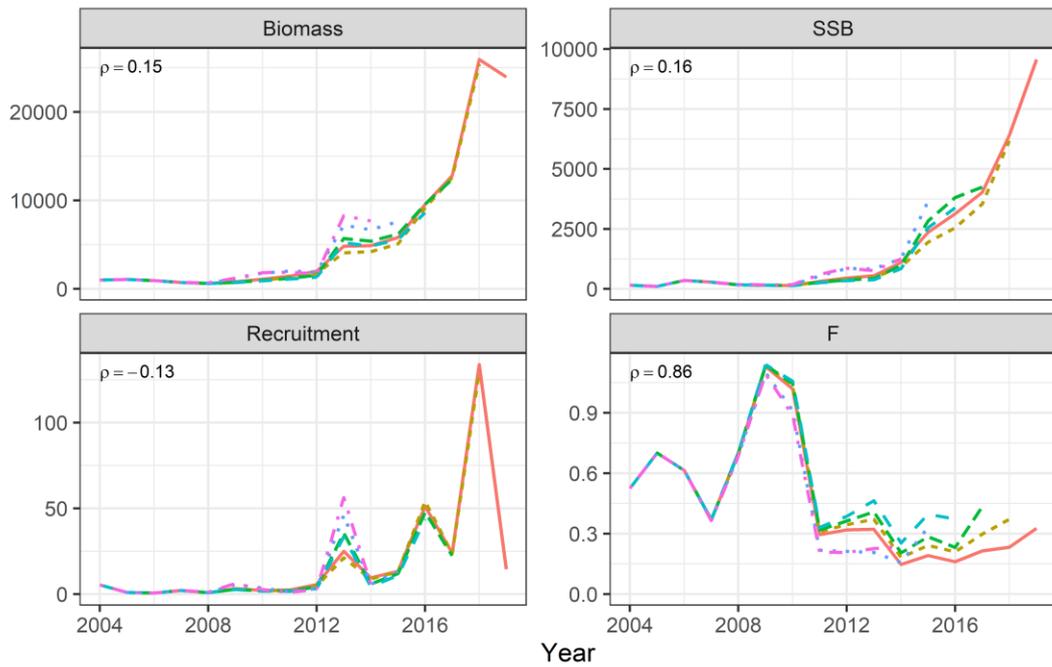
Figure 23: Retrospective pattens of VPA under the scenario C.
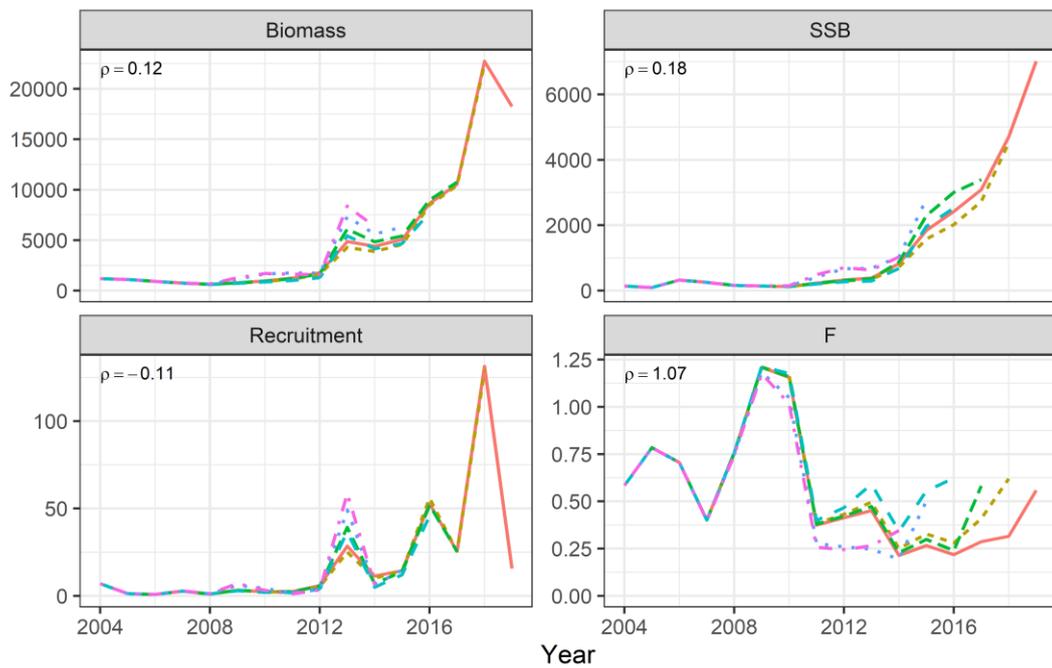
Figure 24: Retrospective pattens of VPA under the scenario D.
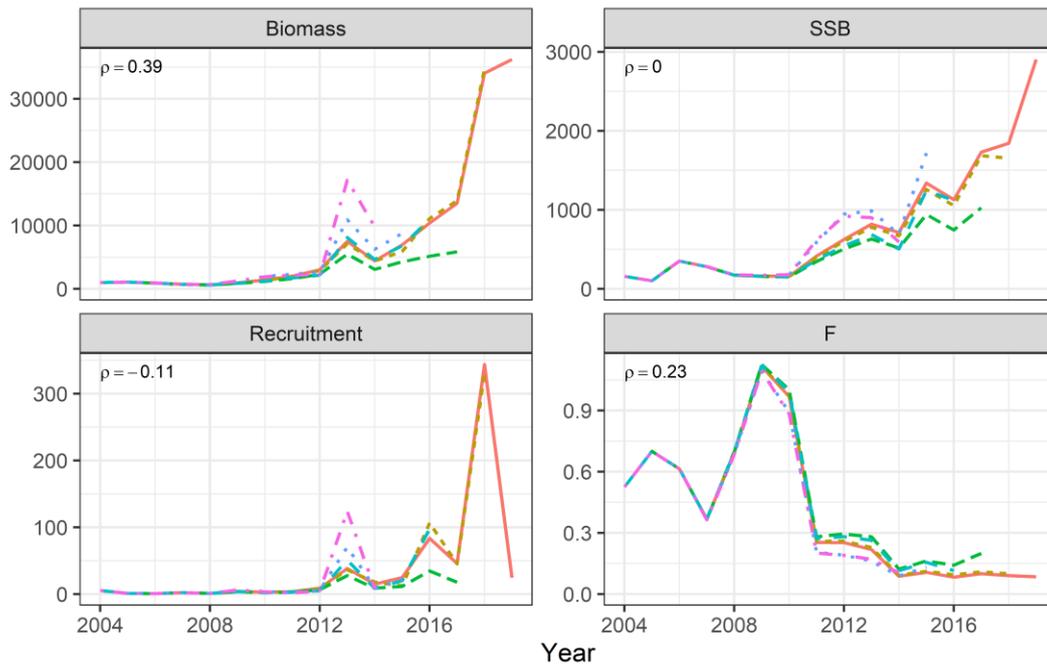
Figure 25: Retrospective pattens of VPA under the scenario E.
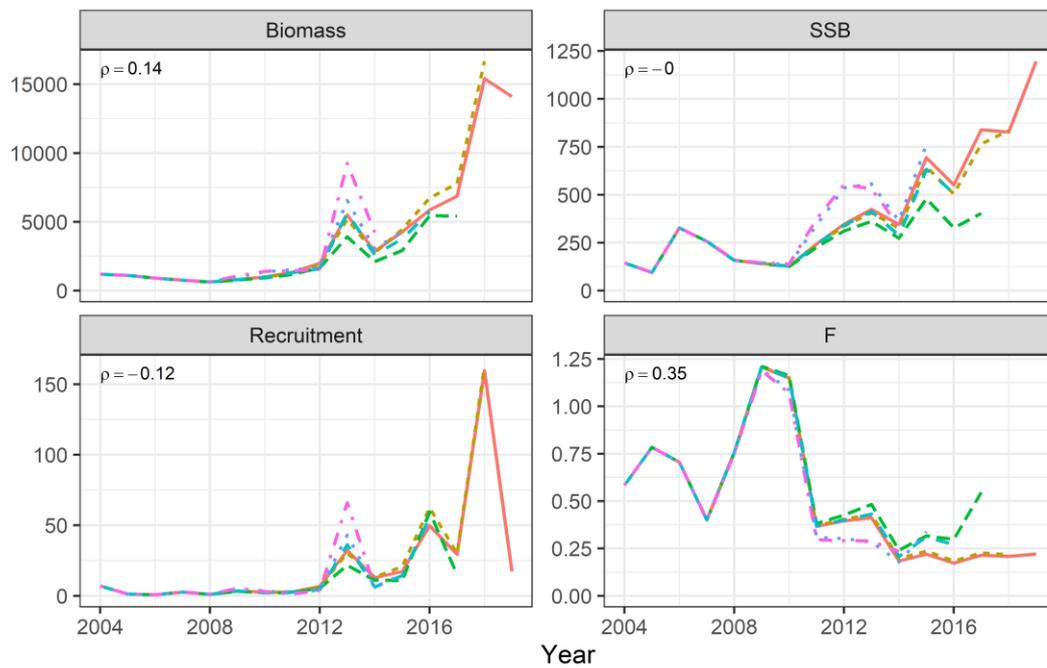
Figure 26: Retrospective pattens of VPA under the scenario F.
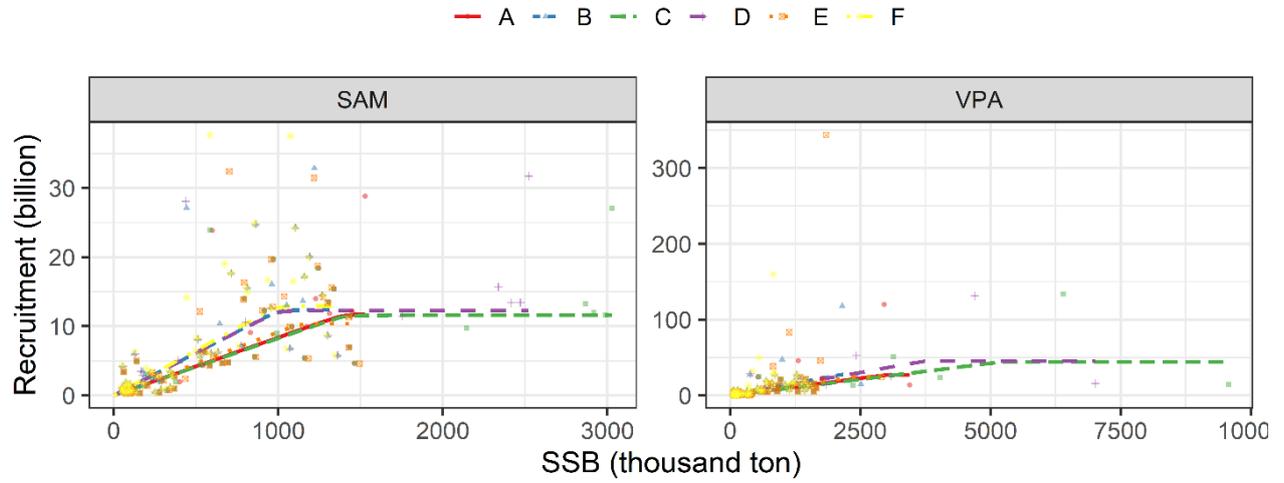
Figure 27: Continuous hockey-stick stock-recruit relationships in SAM and VPA with different scenarios.
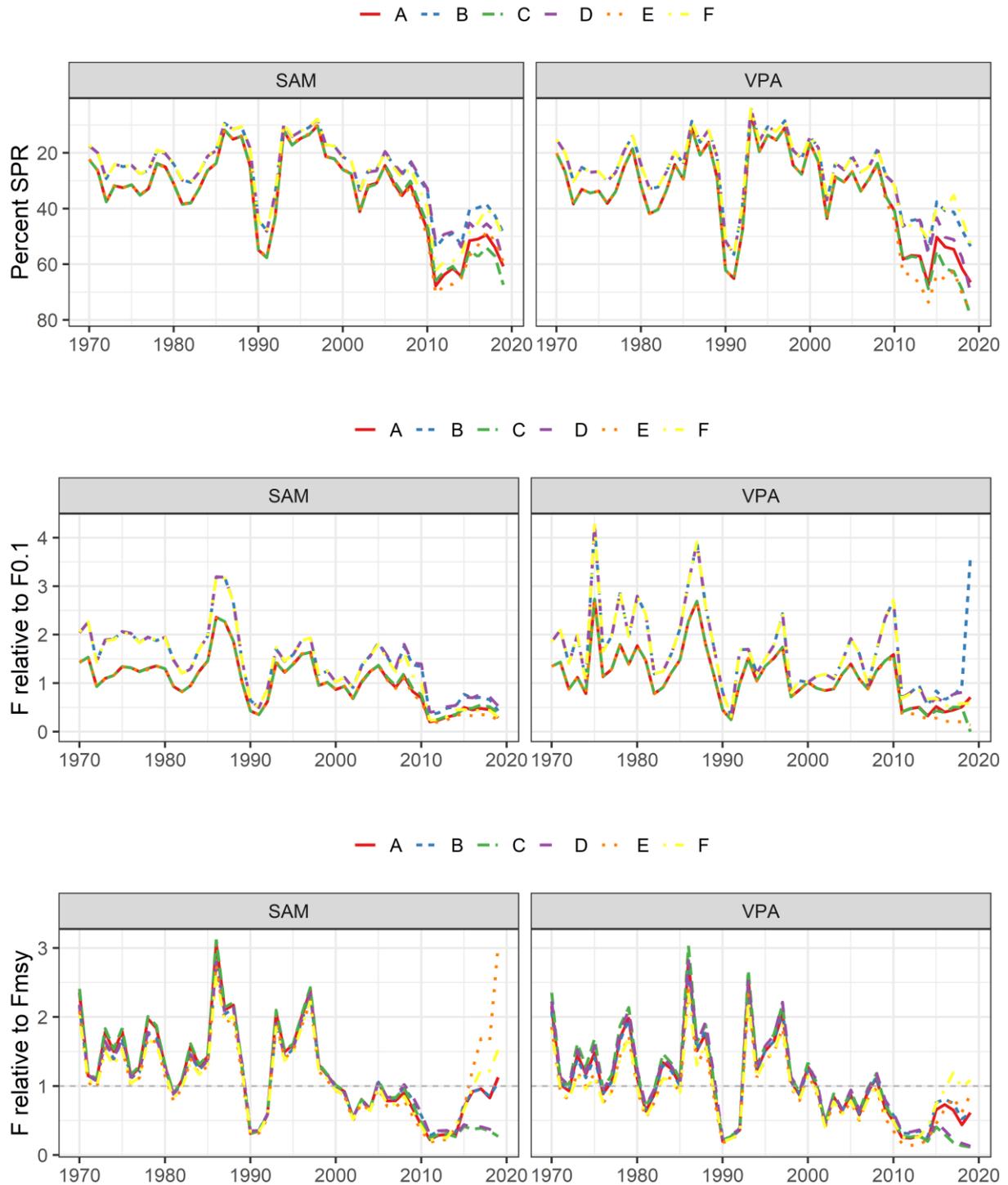
Figure 28: Temporal trends of percent SPR (top), F relative to F0.1 (middle), and F relative to Fmsy (bottom) in SAM (left) and VPA (right) under the scenarios A to F when using per-year biological parameters and F-at-age estimates. The values of Fmsy here is based on the time-varying estimates shown in Fig. 29.
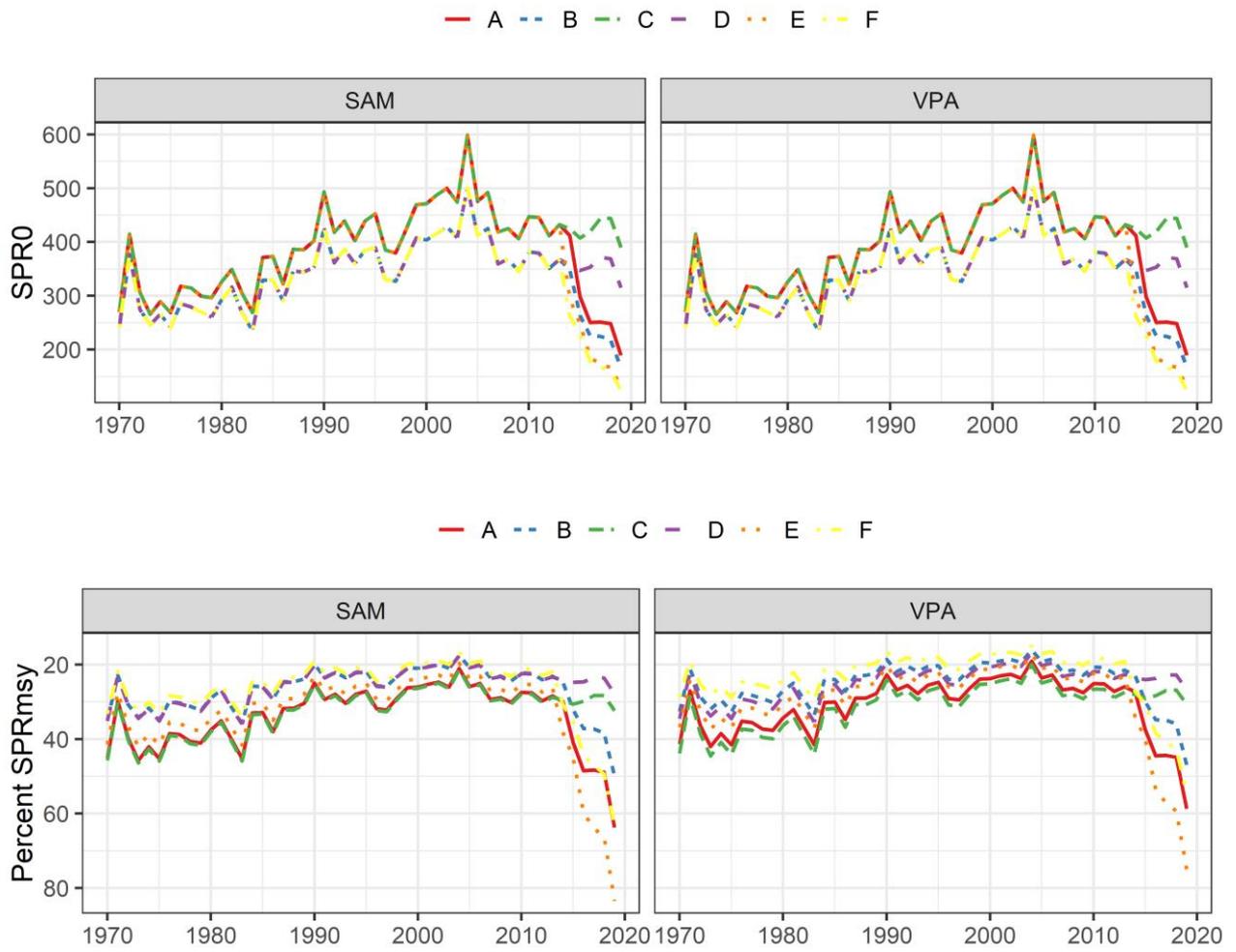
Figure 29: Temporal trends of SPR0 (top) and %SPRmsy (bottom) in SAM (left) and VPA (right) under the scenarios A to F.
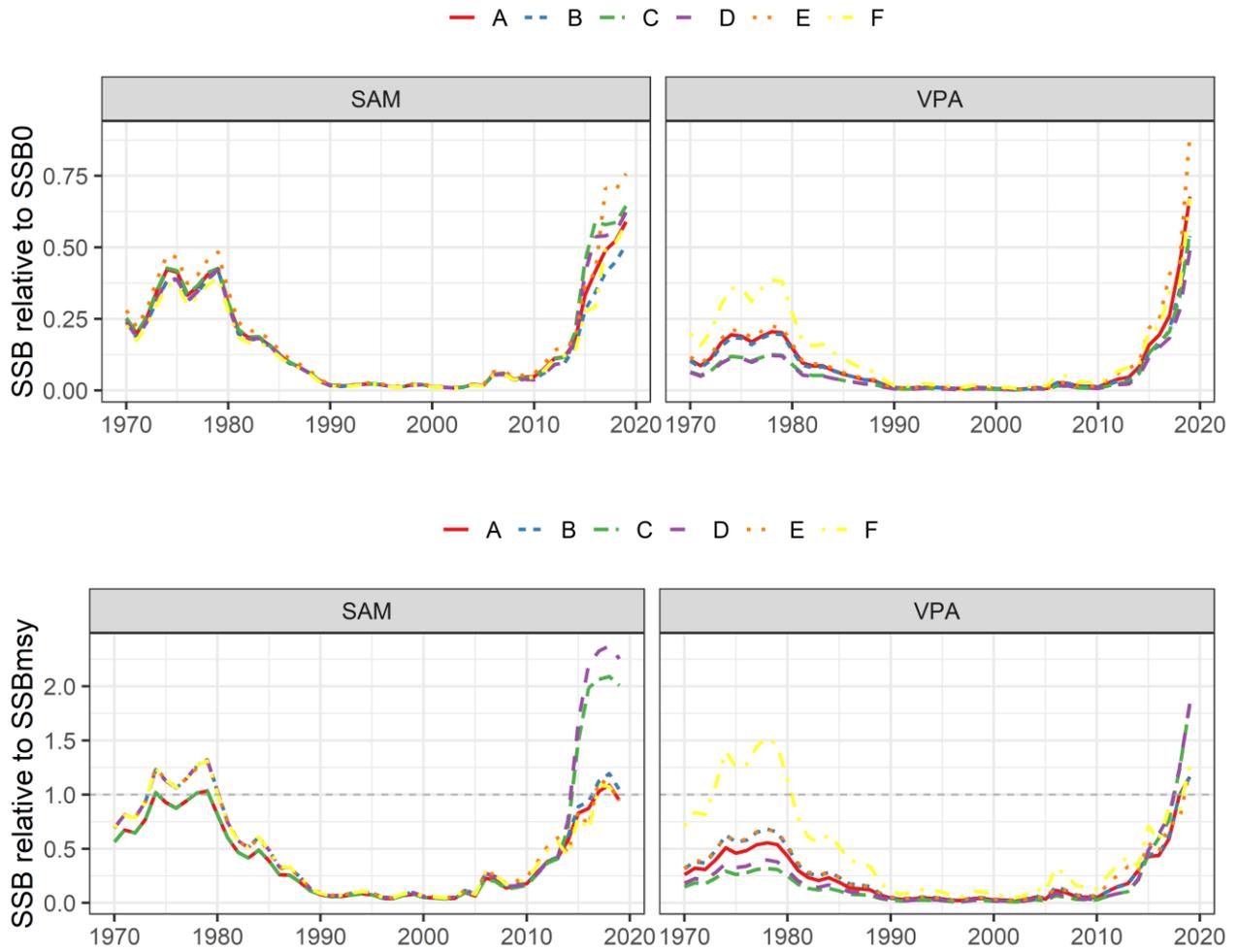
Figure 30: Temporal trends of SSB relative to SSB0 (top) and SSB relative to SSBmsy (bottom) in SAM (left) and VPA (right) under the scenario A to F.
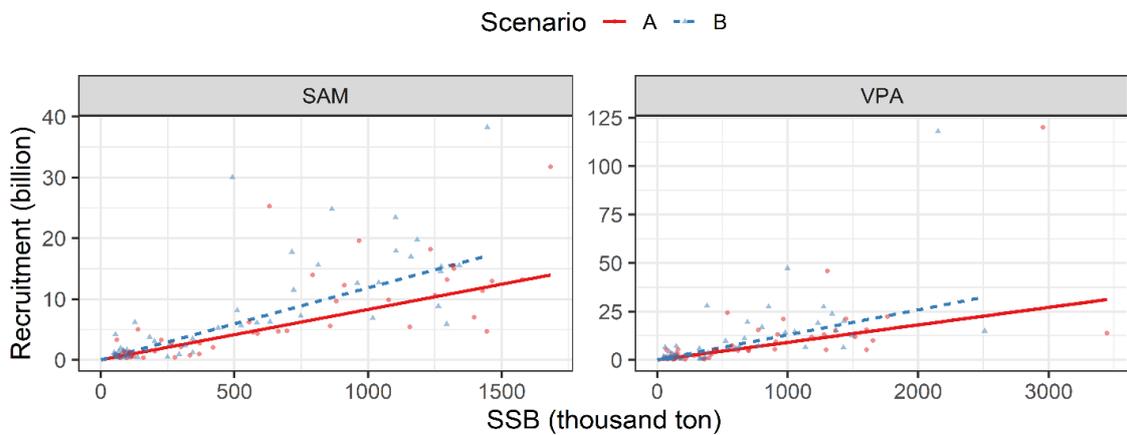
Figure S1: The Beverton-Holt stock-recruit relationships applied to SAM (left) and VPA (right).
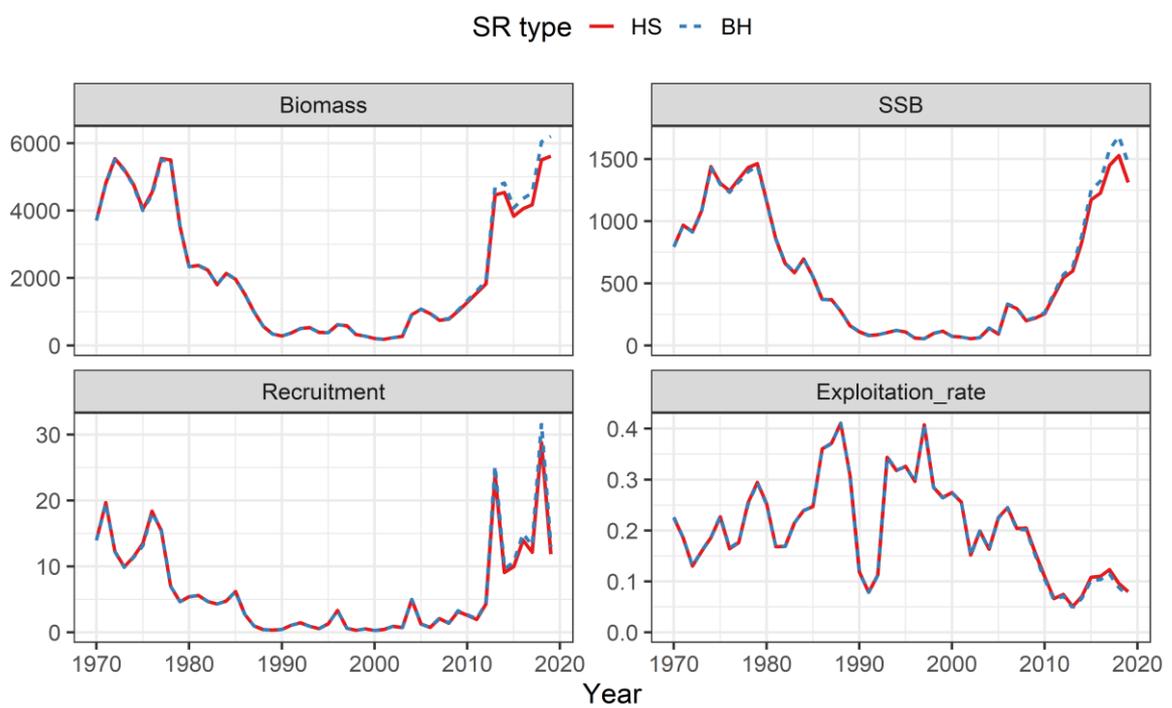
Figure S2: Comparison of estimates in biomass, SSB, recruitment, and exploitation rate between continuous hockey-stick (HS) and Beverton-Holt (BH) relationship under the scenario A with SAM.
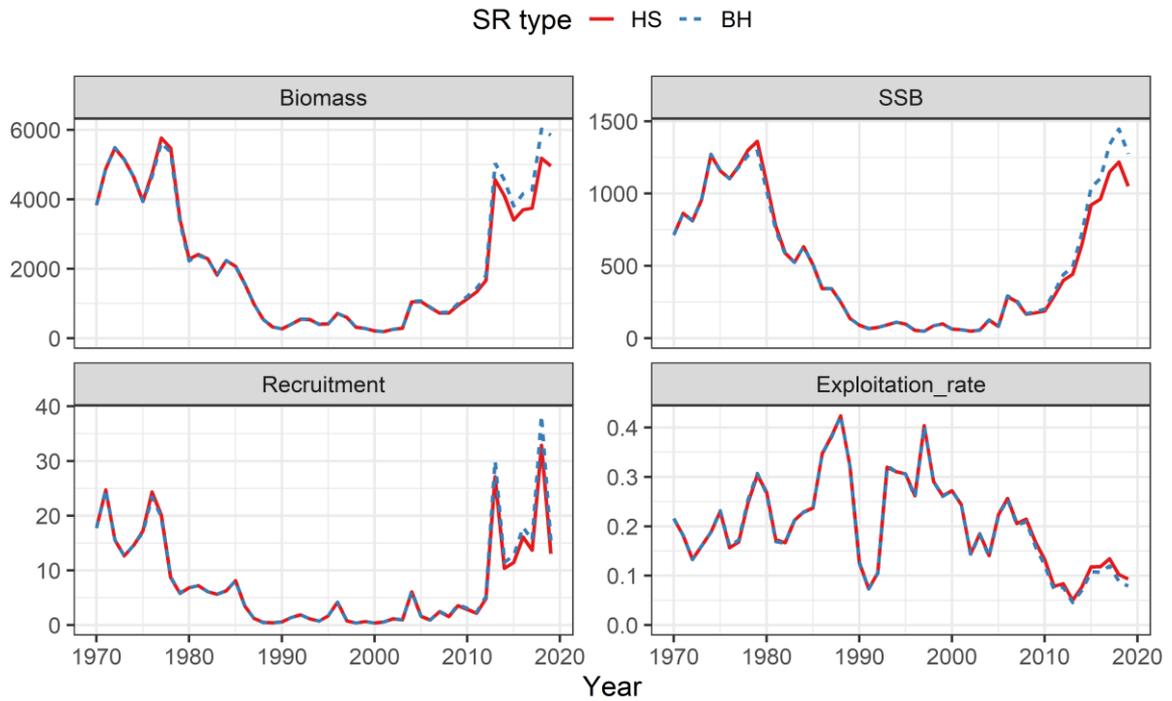
Figure S3: Comparison of estimates in biomass, SSB, recruitment, and exploitation rate between continuous hockey-stick (HS) and Beverton-Holt (BH) relationship under the scenario B with SAM.
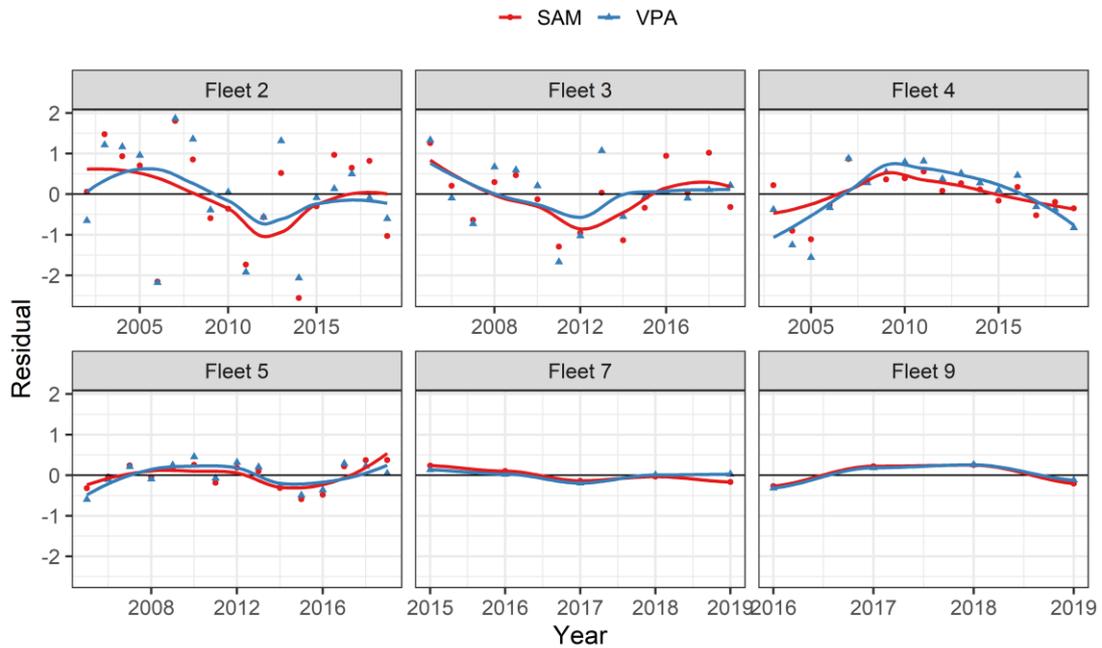
Figure S4: Residuals of abundance indices under the scenario A in SAM and VPA. The curves are the prediction by the LOESS (locally estimated scatterplot smoothing) regression.
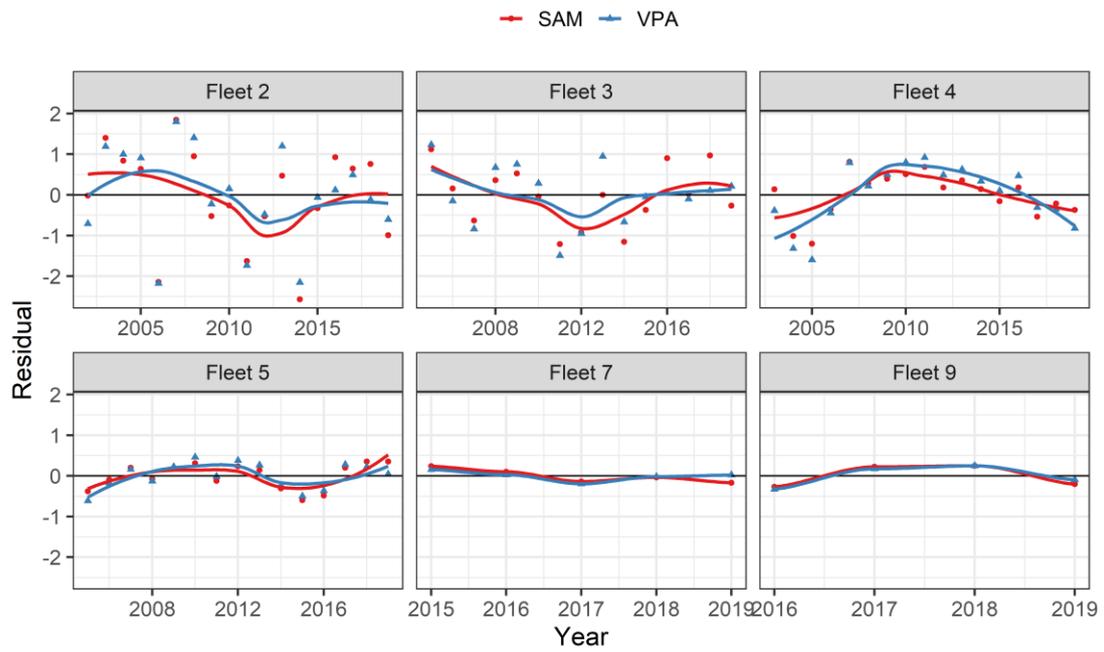
Figure S5: Residuals of abundance indices under the scenario B in SAM and VPA.
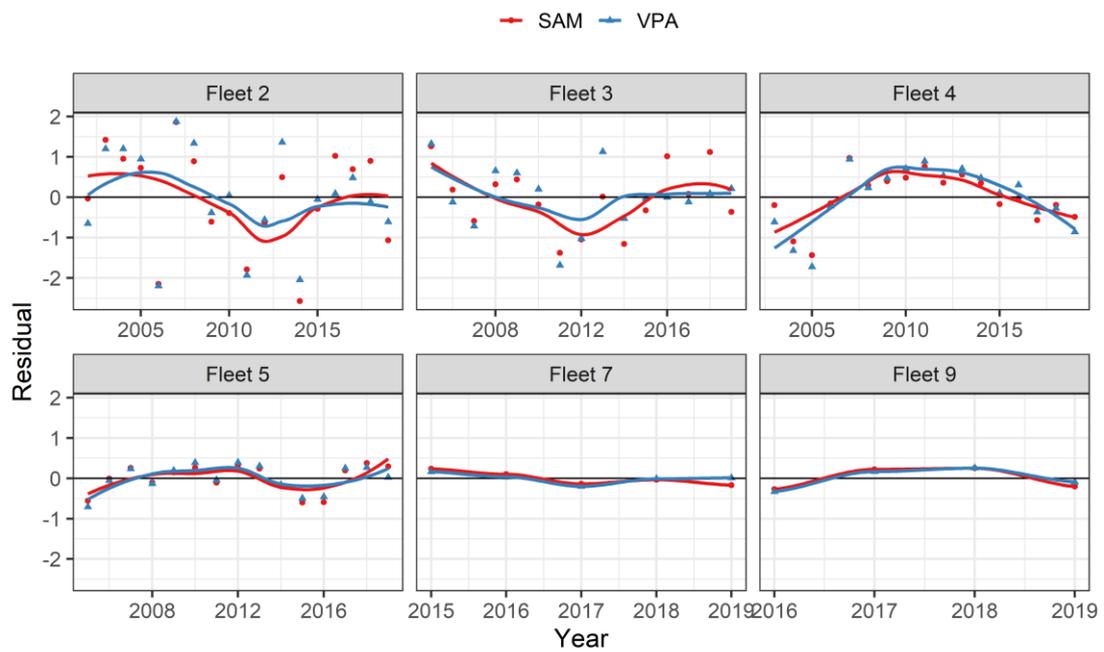
Figure S6: Residuals of abundance indices under the scenario C in SAM and VPA.
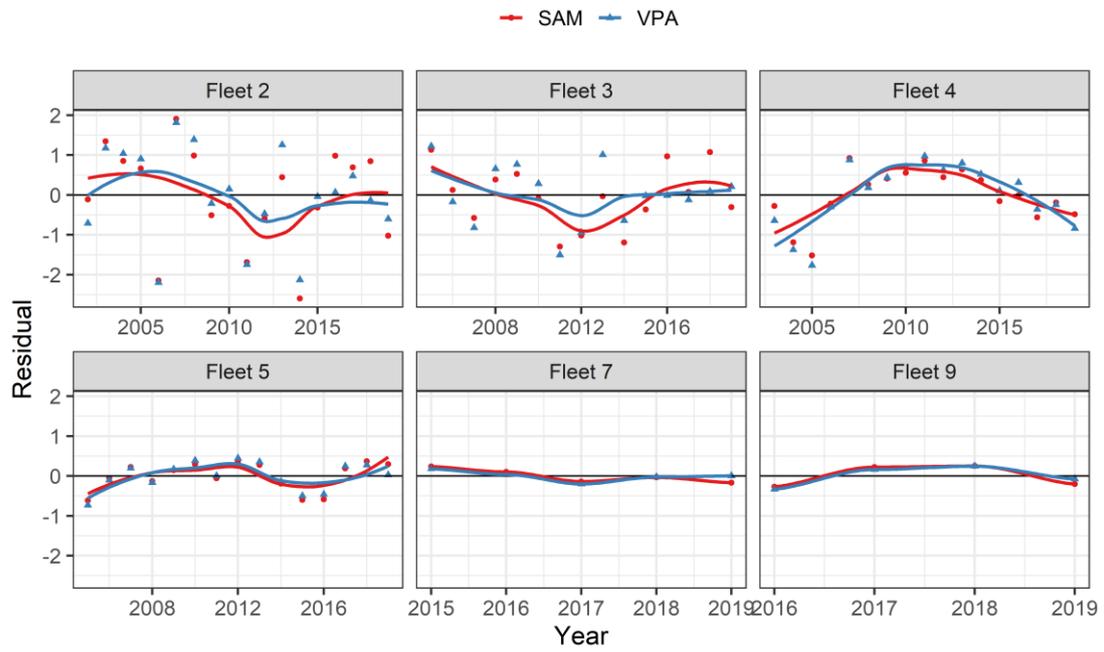
Figure S7: Residuals of abundance indices under the scenario D in SAM and VPA.
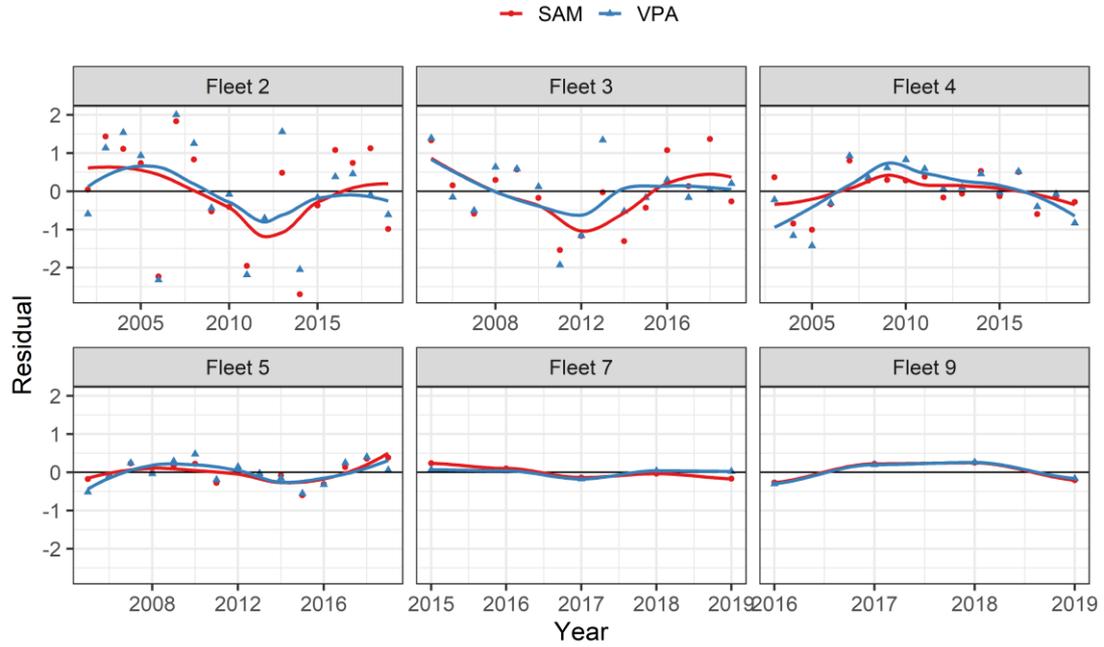
Figure S8: Residuals of abundance indices under the scenario E in SAM and VPA.
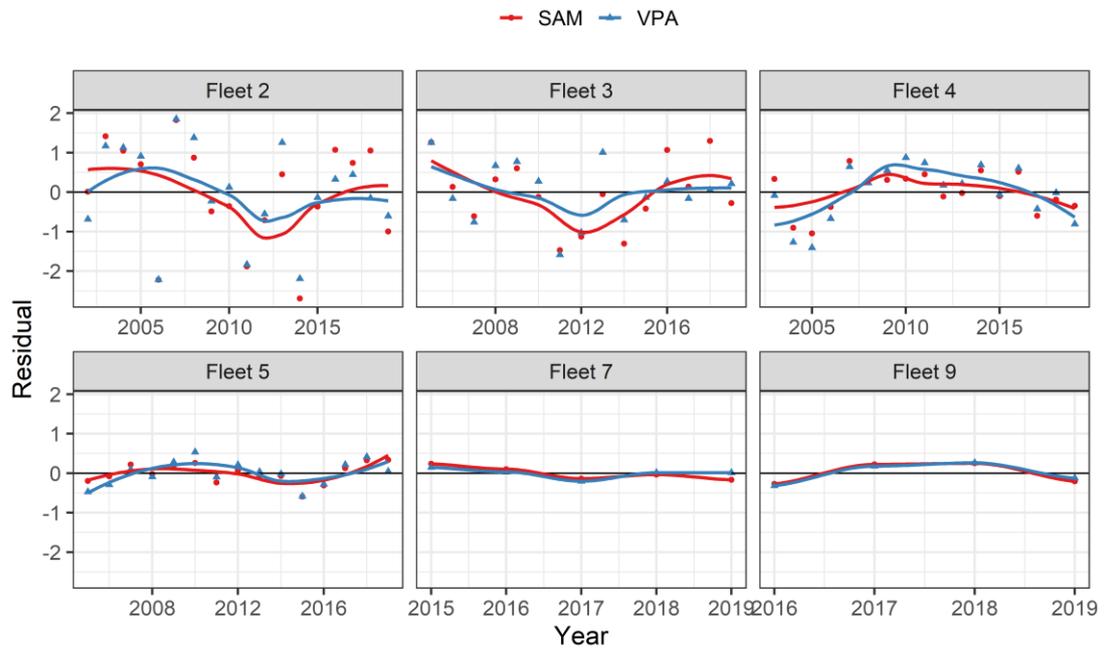
Figure S9: Residuals of abundance indices under the scenario F in SAM and VPA.
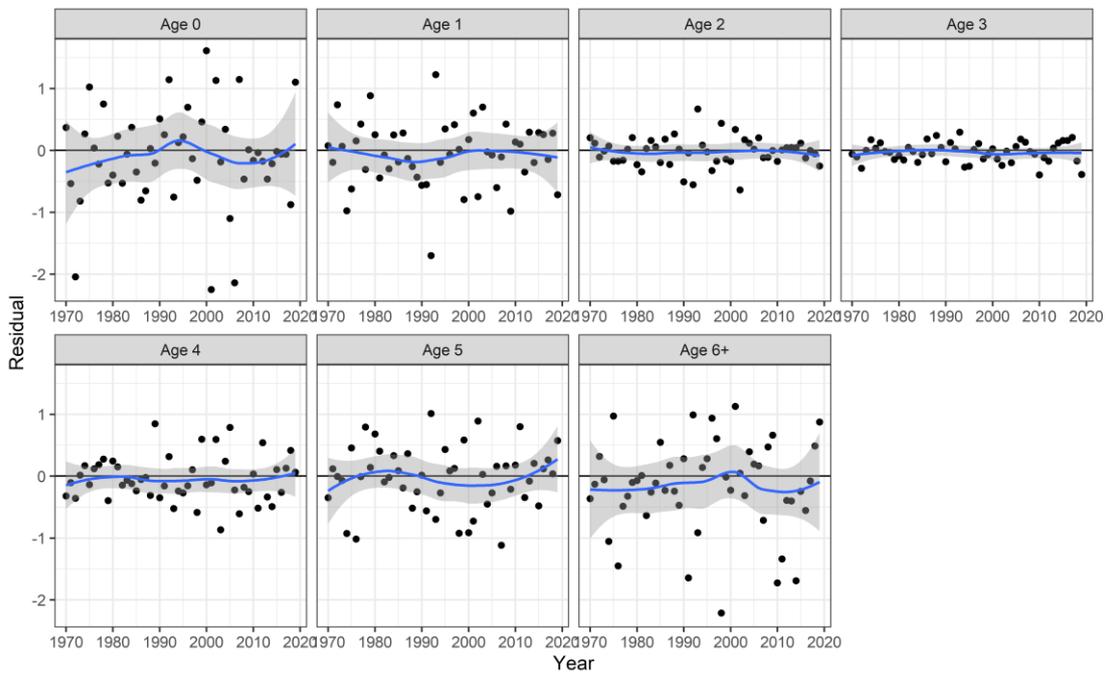
Figure S10: Residuals of catch-at-age under the scenario A in SAM. The blue curves are the prediction by the LOESS regression with 95% confidence intervals.
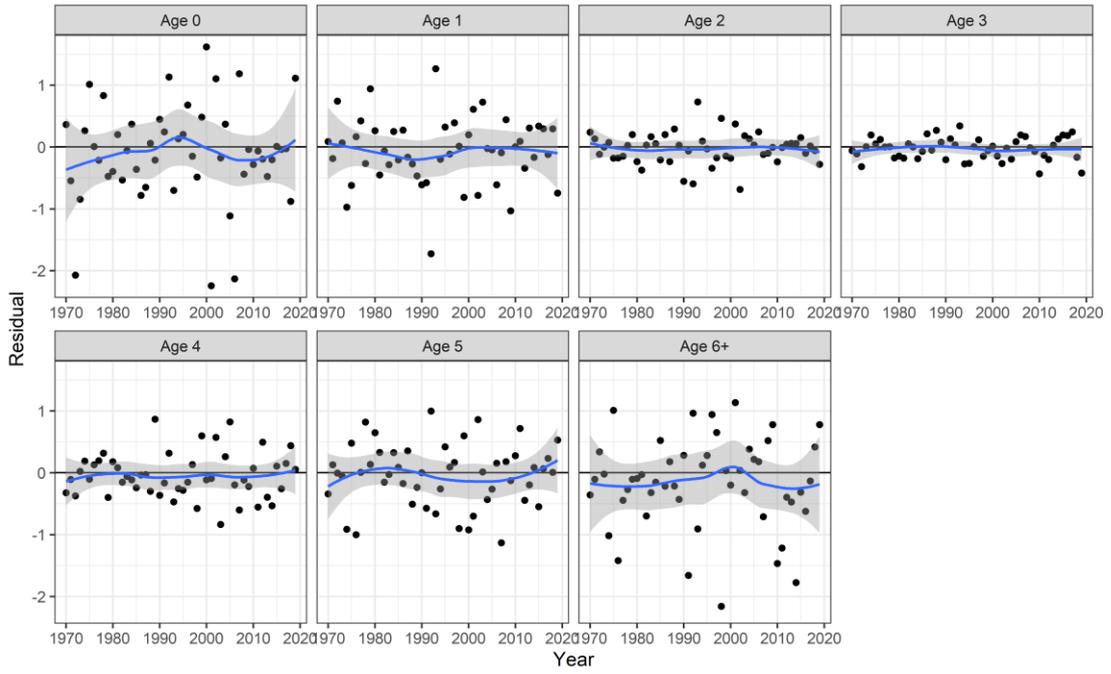
Figure S11: Residuals of catch-at-age under the scenario B in SAM.
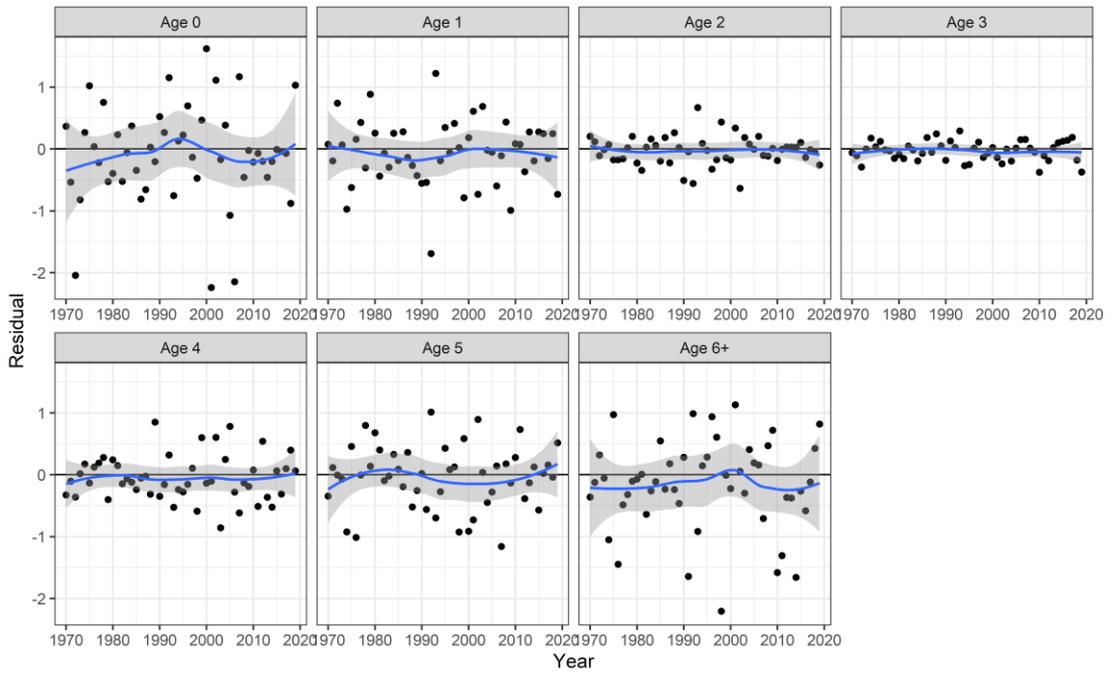
Figure S12: Residuals of catch-at-age under the scenario C in SAM.
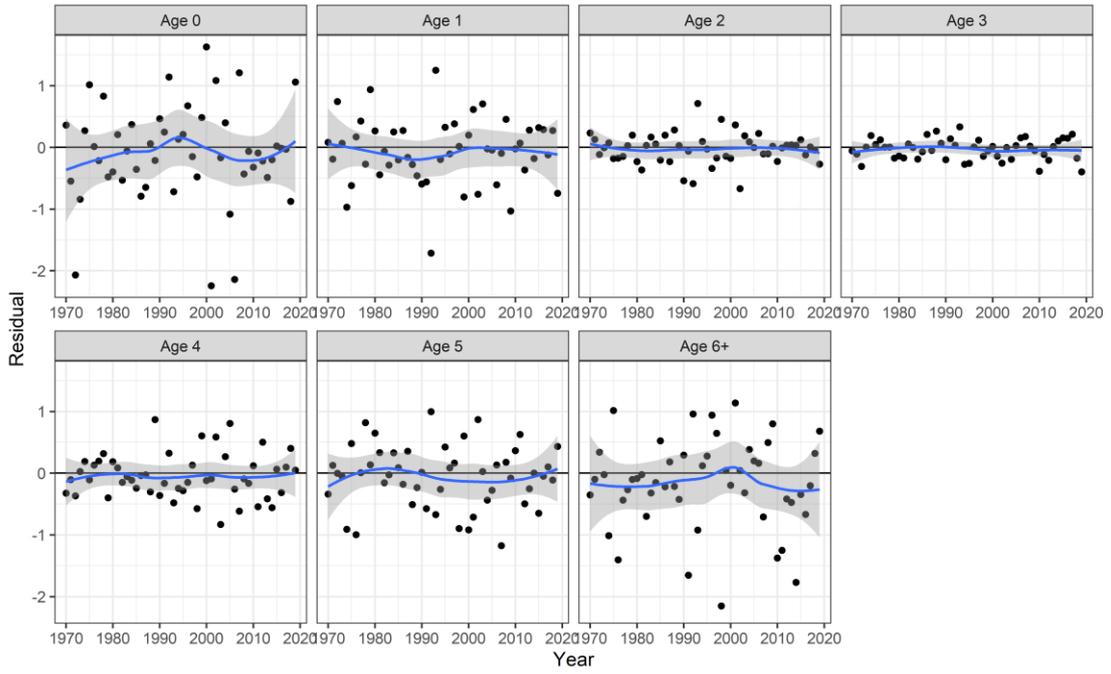
Figure S13: Residuals of catch-at-age under the scenario D in SAM.
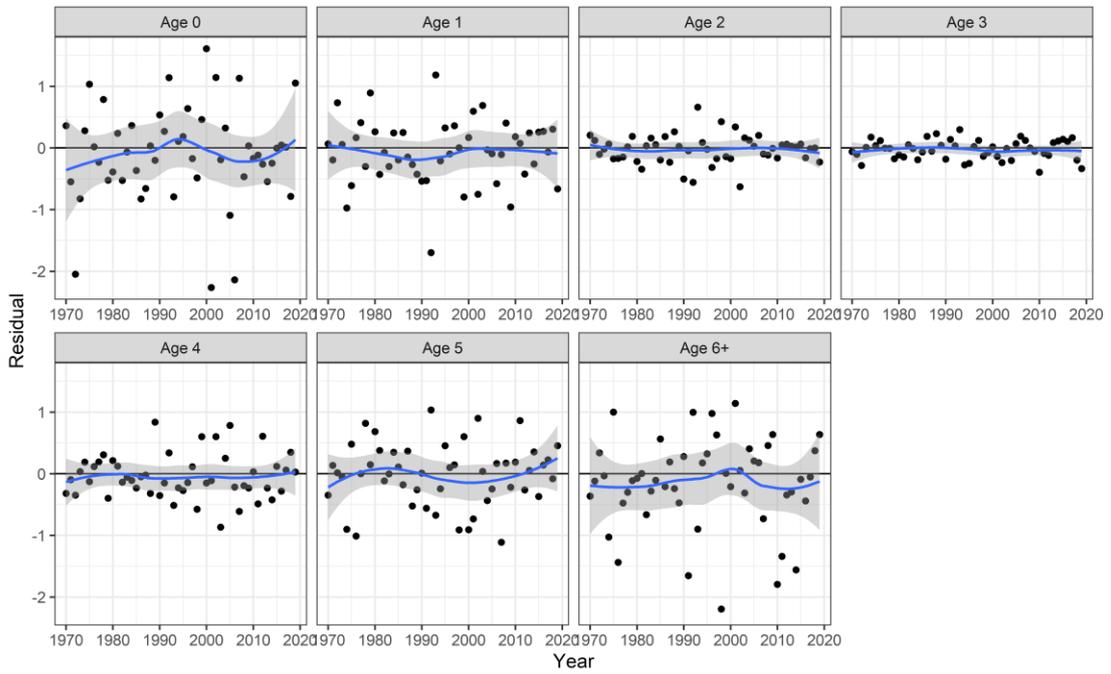
Figure S14: Residuals of catch-at-age under the scenario E in SAM.
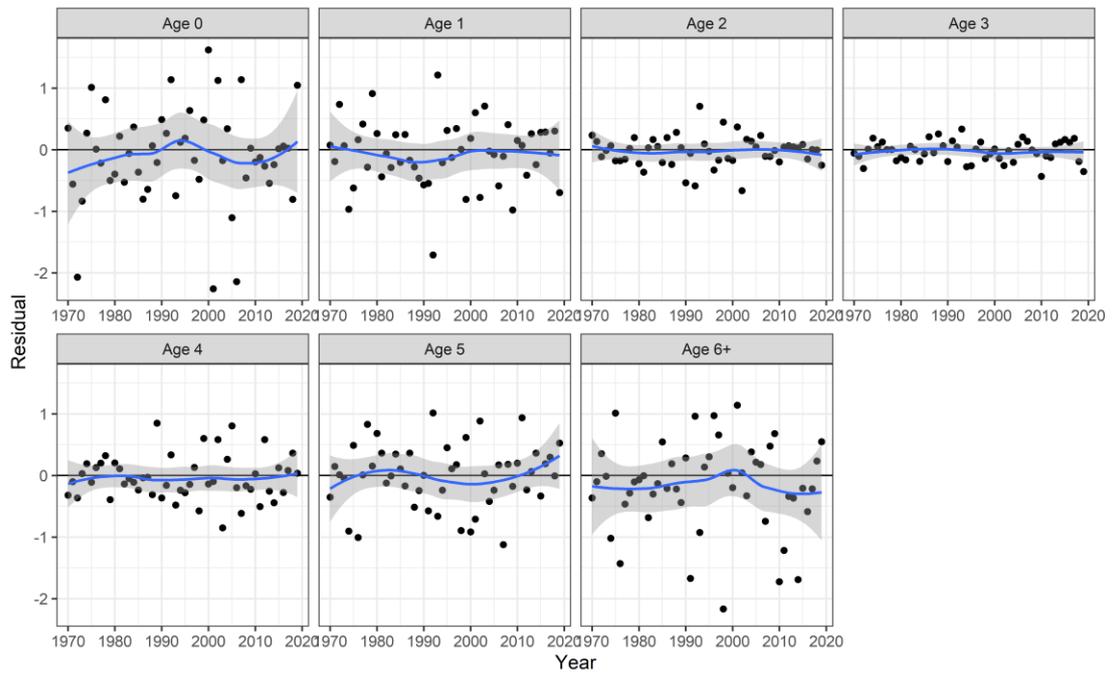
Figure S15: Residuals of catch-at-age under the scenario F in SAM.