North Pacific Fisheries Commission

NPFC-2022-TWG CMSA06-WP08 (Rev. 1)

Calculation of performance measures by fitting VPA and SAM to pseudo data for chub mackerel in Northwestern Pacific

Shota Nishijima, Momoko Ichinokawa, and Kazuhiro Oshima

Fisheries Resources Institute,
Japan Fisheries Research and Education Agency (FRA)
(Corresponding author: Shota Nishijima, nishijima_shota02@fra.go.jp)

Background

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 CMSA04, the TWG members have determined six scenarios or input data to fit each stock assessment model to (NPFC 2021). The members have also determined performance measures for comparing stock assessment models, which include state variables, depletion statistics, biological reference points (BRPs), and relative fishing impacts, and Mohn's rho calculated from retrospective analysis (NPFC-2022-TWG CMSA05-WP01). As intersessional works, the members have then submitted estimated results for the use of input data for the operating model PopSim-A to the invited expert. The invited expert in turn distributed pseudo datasets produced by PopSim-A under the determined six scenarios with several modifications. We fitted VPA and SAM to the pseudo datasets (PseudoData_SCN_Ver3_05July.zip, available from here) and calculated performance measures. In this document, we report the results of performance measures obtained by fitting VPA and SAM to the pseudo datasets.

Calculating performance measures

We developed the R package 'OMutility' that easily computes the performance measures (NPFC-2022-TWG CMSA05-WP01) from output results of stock assessment models. Although we found several bugs in the previous version of OMutility (Nishijima et al. 2022a), we fixed these bugs and a new version (Version 1.0.1) was used in this document.

Fitting VPA to pseudo data

There are zero data of catch number at age in some pseudo data. Since VPA cannot directly treat zero catch data, we added a small constant value to zero catch samples as an ad-hoc approach. We used the half of the minimum value of positive catch-at-age in a single iteration to zero catch samples in the same way as the previous analysis (Nishijima et al. 2022a). We fitted VPA to the pseudo data in the same way as in fitting VPA to the actual data (Nishijima et al. 2022b).

The median values of biomass and recruitment estimates among iterations were generally lower than the corresponding estimates of VPA with the real data in recent years, although the median values of estimates in the past years were close to the estimates of VPA with the real data (Fig. 1). This suggest that VPA with the real data estimated high biomass and recruitment in recent years. Variations in recruitment estimates among iterations were low in past years but high in recent years (Fig. 2). However, variations in total biomass among iterations were high in not only recent years but also past years especially scenario E, because fishing mortalities or exploitation rates were also varied among iterations (Fig. 2).

Regarding BRPs related to fishing mortality (F), some iterations in some pseudo datasets generated F_{ref} = 10 for F0.1, Fmax, Fmsy (Fig. 3). This suggests that the optimization reached the upper boundary of F value. By contrast, Fmsy_HS sometimes exhibited extremely low values (≈0). MSY-based BRPs related to total biomass and spawning stock biomass (Bmsy and SBmsy) were sometimes estimated as zero or negative values when Fmsy was estimated at the maximum value (10). In addition, some extraordinarily high and low estimates in Bmsy and SBmsy been found in some pseudo datasets (Fig. 4). Relative fishing impacts (ratios of current F to F reference points) show very low values in some iterations probably because of large values of F reference points (Fig. 5). Regarding depletion statistics, we found large variations when stock abundances were high (1970s, 1980s, and 2010s) (Fig. 6).

Fitting SAM to pseudo data

SAM also cannot treat zero catch data. We therefore employed the same ad-hoc approach, adding a small value, as done in VPA. In addition, SAM often caused estimation errors (not finish optimization), failures to converge (detected by nlminb function), and unrealistic estimates ($F \approx 0$ or $N < 0$) when we fitted SAM to pseudo datasets in the same model configurations in fitting SAM to the real data (Nishijima et al. 2022b). Therefore, when we found these signs of errors and failures in estimation, we gradually simplified model configurations from model #1 to #7 as follows:

1. The same model configurations in fitting SAM to the real data were used.
2. The same configurations as #1 except that the same standard deviation in observation errors for all the six abundance indices were assumed.
3. The same configurations as #2 except that the same standard deviation in process errors of F random walk among age classes were assumed.
4. The same configurations as #3 except that the same standard deviation in observation errors of catch at age among age classes were assumed.
5. The same configurations as #4 except that the nonlinear coefficients of abundance indices from Chinese and Russian fisheries were assumed to be 1.
6. The same configurations as #5 except that we estimated F random walk from 2010 to 2011 (removed in the original analysis).
7. The same configurations as #6 except that correlation coefficients of multivariate normal distribution of F random walk were assumed to be 1 (selectivity became constant over years).

There was no iteration in which even model # 7 did not converge. We could converge SAM for more than 90% (90.8%) of iterations with the original model configurations (model #1, Fig. 7). Interestingly, the scenarios with age-specific M (B, D, and F) were higher likelihoods of convergence with model #1 (94.0% on average) than the scenarios with age-common M (A, D, and E) (87.7%). In this document, we show figures including the iterations in which model #1 did not converge, but another model did, although SWG OM has agreed to preferentially use only iterations in which model #1 achieved convergence.

The median values of state variables among iterations were close to the estimate of SAM with the real data (Fig. 8). Comparing VPA results (Fig. 1), SAM estimated lower total biomass and recruitment and higher fishing mortality and exploitation rate in recent years (Fig. 8). SAM had larger between-iteration variations of total biomass, weighted average F, and exploitation rate in past years under the scenarios with age-constant M (A, C, E) than under the scenarios with age-specific M (B, D, F) (Fig. 9).

F-related BRPs sometimes reached the maximum value (10) especially for Fmax under scenarios A, B, E and F (Fig. 10). Extremely small values in Fmsy were less likely to be observed in SAM than in VPA (Fig. 11). Bmsy and SBmsy estimated by the Beverton-Holt stock-recruitment relationship showed extremely large values in some iteration under scenarios E, whereas Bmsy and SBmsy estimated by hockey-stick stock-recruitment relationship sometimes exhibited extremely low values (≈0) under scenarios E and F (Fig. 11). As in VPA, relative fishing impacts (ratios of current F to F reference points) show very low values in some iterations probably because of large values of F reference points (Fig. 12). Regarding the depletion statistics of ratios to median, we found large variations when stock abundances were high (1970s, 1980s, and 2010s) in a similar vein to VPA (Fig. 13).

Retrospective analysis

In this document, we firstly report the results of Mohn's rho obtained by conducting 7-year retrospective analysis for pseudo data. The retrospective analysis was performed for 20% of iterations (i.e., 80 iterations per scenario) and the list of iterations for the retrospective analysis was provided by the external expert. Since we have two short time-series abundance indices (5 years for Chinese fishery and 4 years for Russian fishery), we removed these abundance indices in the process of retrospective analysis so that estimation achieved convergence; we removed the abundance indices when the available years were less than three in VPA and four in SAM.

Mohn's rho showed distinct patterns between VPA and SAM (Fig. 14). The Mohn's rho calculated from VPA was small in weighted average F, but positive values were shown in total biomass and spawning stock biomass. By contrast, Mohn's rho from

SAM was small in total biomass and spawning stock biomass, but positive values were shown in weighted average F. The tendency is similar with the true value of Mohn's rho calculated from the real data.

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.

Nishijima, S., K. Oshima, M. Ichinokawa, N. 2022a. Fitting VPA and SAM to pseudo data generated from POPSIM: A first analysis for chub mackerel in the Northwestern Pacific. NPFC-2021-TWG CMSA05-WP08.

Nishijima, S., K. Oshima, M. Ichinokawa, N. Shinohara. 2022b. Update of Virtual Population Analysis and State-Space Assessment Model for Operating Models of Chub Mackerel Stock Assessment in NPFC. NPFC-2021-TWG CMSA05-WP06.

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 2021. 4th Meeting of the Technical Working Group on Chub Mackerel Stock Assessment. NPFC-2020-TWG CMSA03-Final Report. (available at www.npfc.int)

NPFC-2022-TWG CMSA05-WP01. Summary of 2nd Meeting of the Small Working Group on Operating Model for Chub Mackerel Stock Assessment.

Fig. 1: Time series of state variables (TBy: total biomass, Ry: the number of recruits, AFy: F at age weighted by catch weight at age, and Ey: exploitation rate) estimated by VPA from scnario A to F. The coloured lines indicate the median values among iterations per pseudo dataset. The points indicate estimates when we fitted VPA to the real data under scenarios A to F. Note that the points do not indicate true values because pseudo data were generated from not only VPA but also other models.

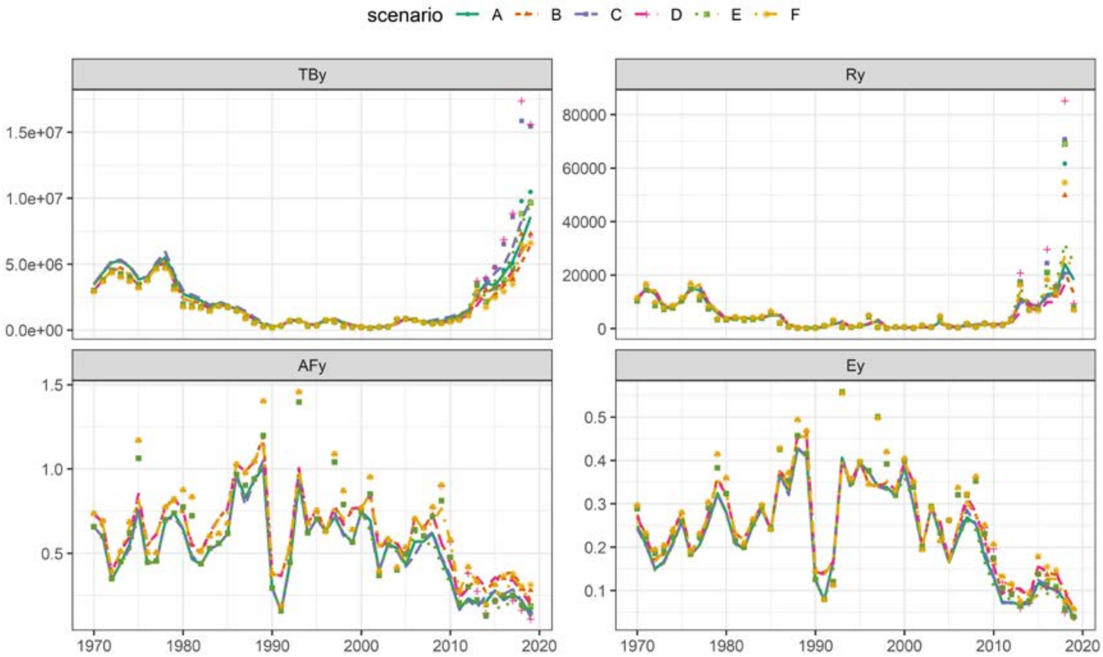

Fig. 2: Time series of state variables (TBy: total biomass, Ry: the number of recruits, AFy: F at age weighted by catch weight at age, and Ey: exploitation rate) with VPA under scnarios A to F. The coloured lines indicate the median values among iterations per dataset while the shadowed areas represent 80% estimation intervals among iterations. Note that the *y*-axes are log-transformed for visualization of large intervals.

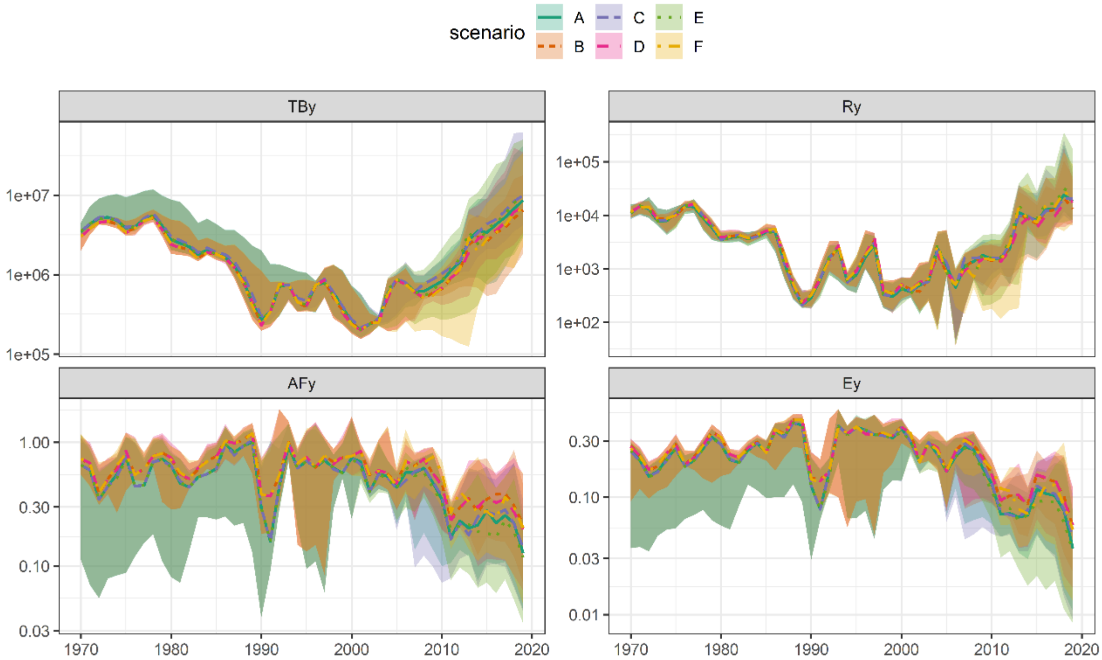

Fig 3: BRPs related to fishing mortality with VPA from scnario A to F. The points of X indicate the values in fitting VPA to the real data under scenarios A to F. Note that the points do not indicate true values because pseudo data were generated from not only VPA but also other models. The *y*-axes are log-transformed for visualization of large variations.

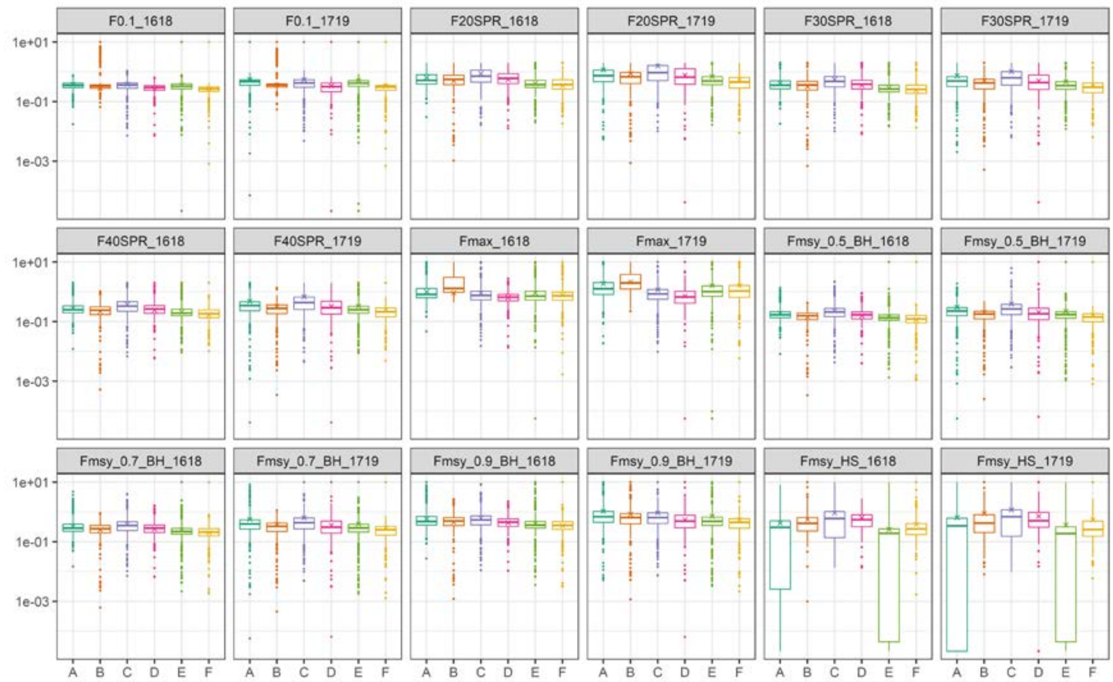

Fig 4: BRPs related to total biomass and spawning stock biomass with VPA from scnario A to F. The points of X indicate the values in fitting VPA to the real data under scenarios A to F. Note that the points do not indicate true values because pseudo data were generated from not only VPA but also other models.The *y*-axes are log-transformed for visualization of large variations. We excluded the cases in which Bmsy or SBmsy was estimated as zero or negative values from this figure.

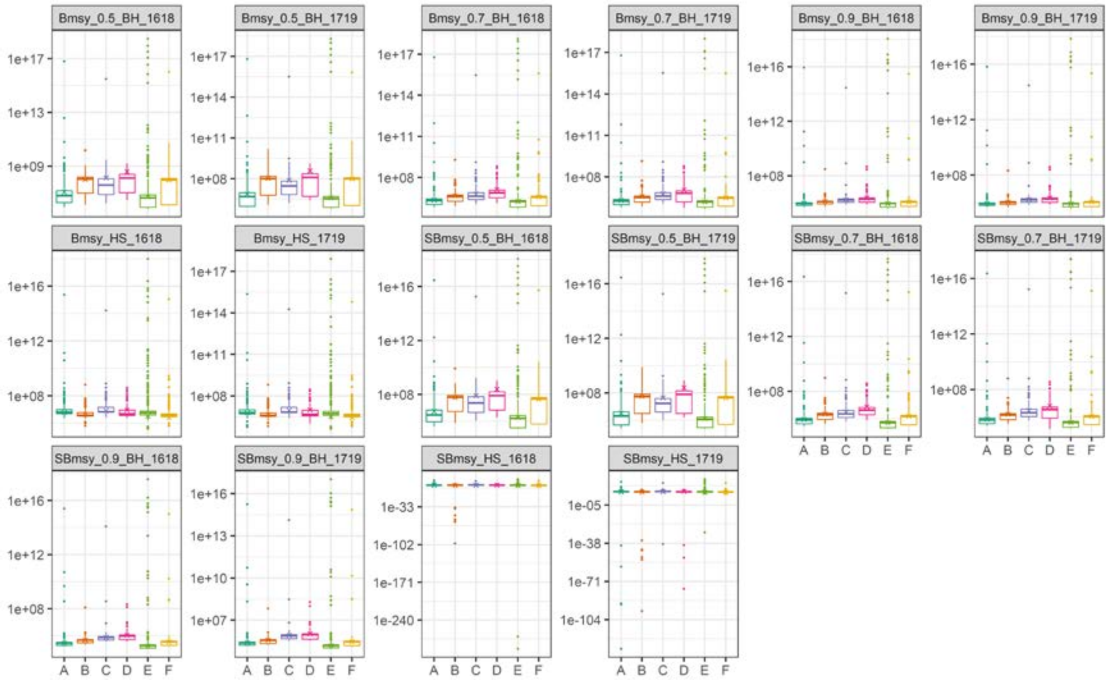

Fig. 5: Relative fishing impacts (ratio of current F to F reference points shown in Fig. 3) with VPA from scnario A to F. The points of X indicate the values in fitting VPA to the real data under scenarios A to F. Note that the points do not indicate true values because pseudo data were generated from not only VPA but also other models. Note that the *y*-axes are log-transformed for visualization of large variations. The horizontal dashed lines indicate 1 (i.e., current F is equal to F reference point).

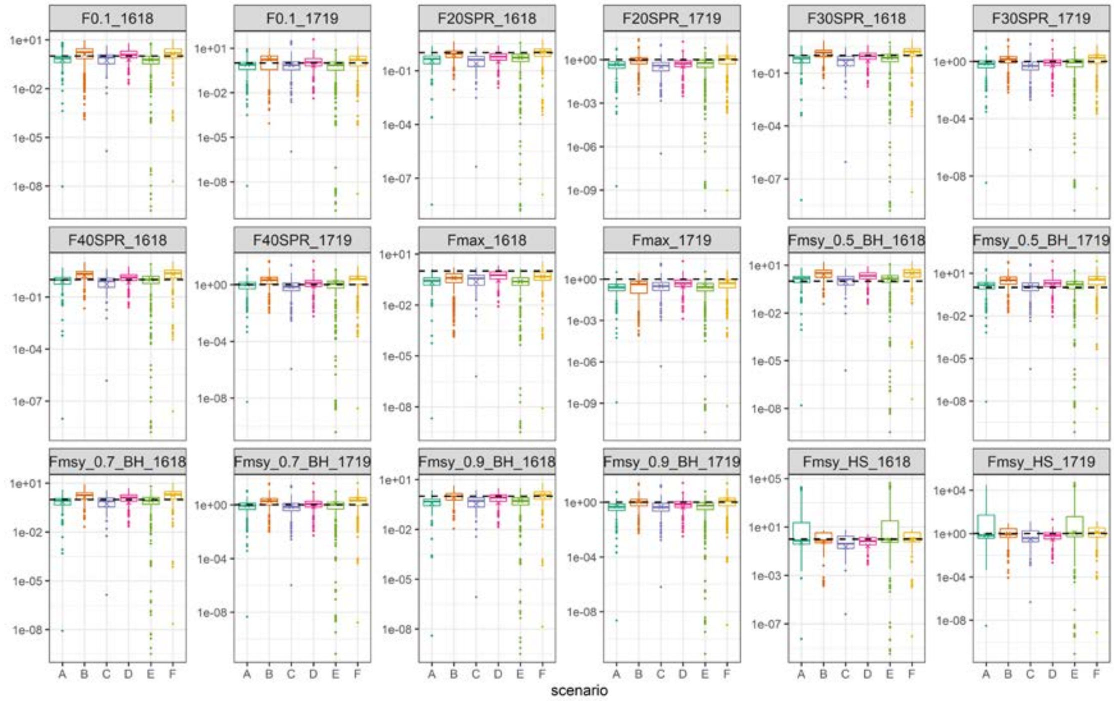

Fig. 6: Depletion statistics with VPA from scenario A to F. The colors indicate each scenario. The points of ‘X’ indicate the values in fitting VPA to the real data under scenarios A to F. Note that the points do not indicate true values because pseudo data were generated from not only VPA but also other models. The relative values to median are log-transformed for visualization of large ranges.

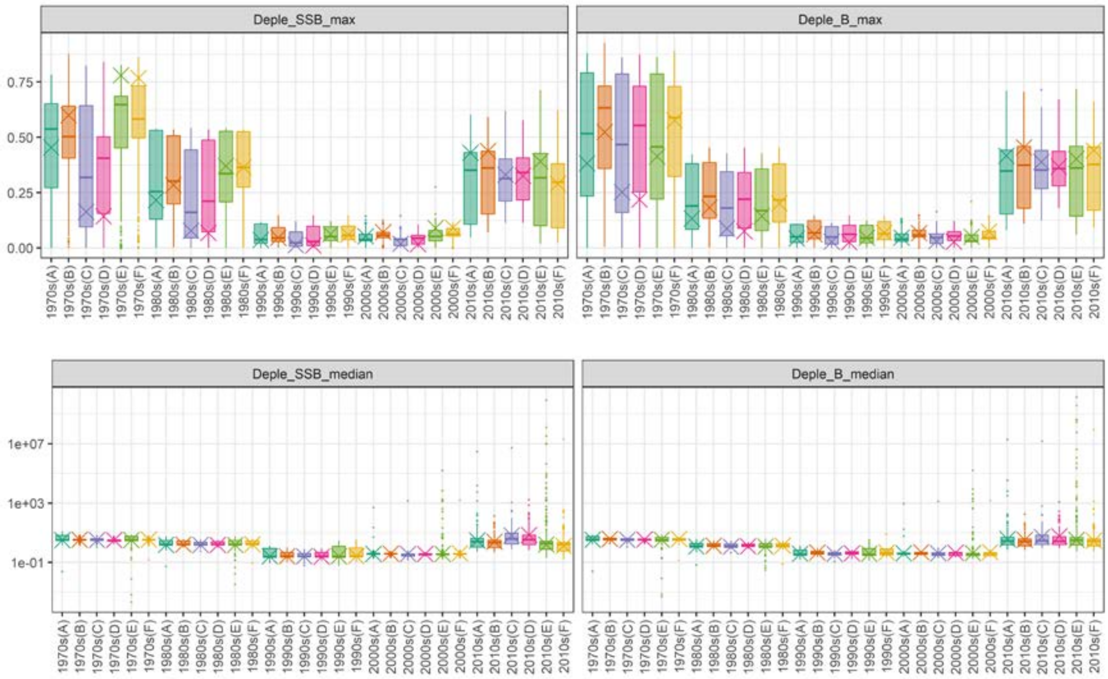

Fig. 7: Frequencies of models that were used for SAM fitting for each pseudo dataset.

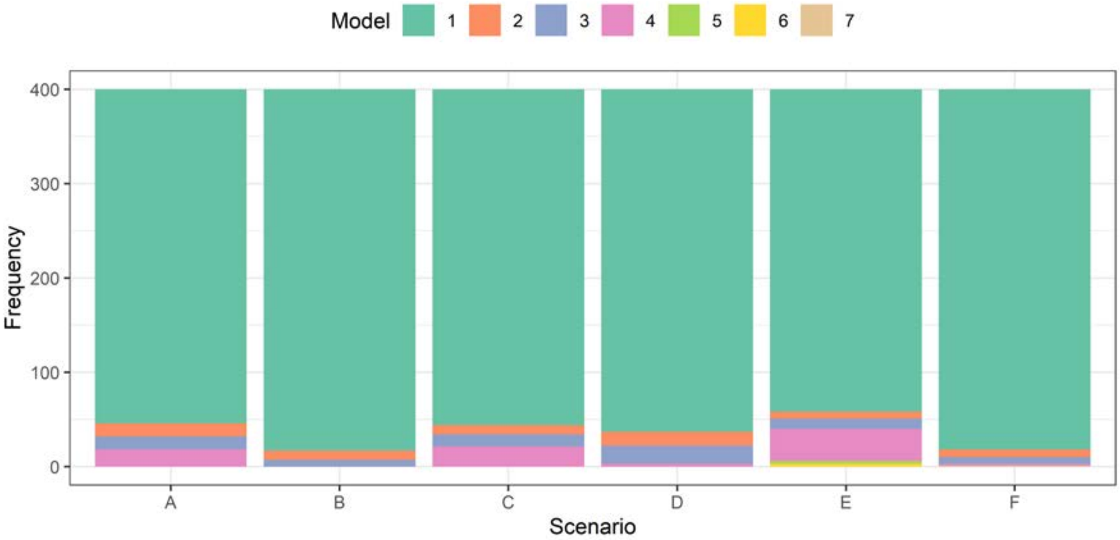

Fig. 8: Time series of state variables (TBy: total biomass, Ry: the number of recruits, AFy: F at age weighted by catch weight at age, and Ey: exploitation rate) with SAM under scnarios A to F. The coloured lines indicate the median values among iterations per dataset while estimation variations among iterations are shown in the next figure. The points indicate estimates when we fitted SAM to the real data under scenarios A to F. Note that the points do not indicate true values because pseudo data were generated from not only SAM but also other models.

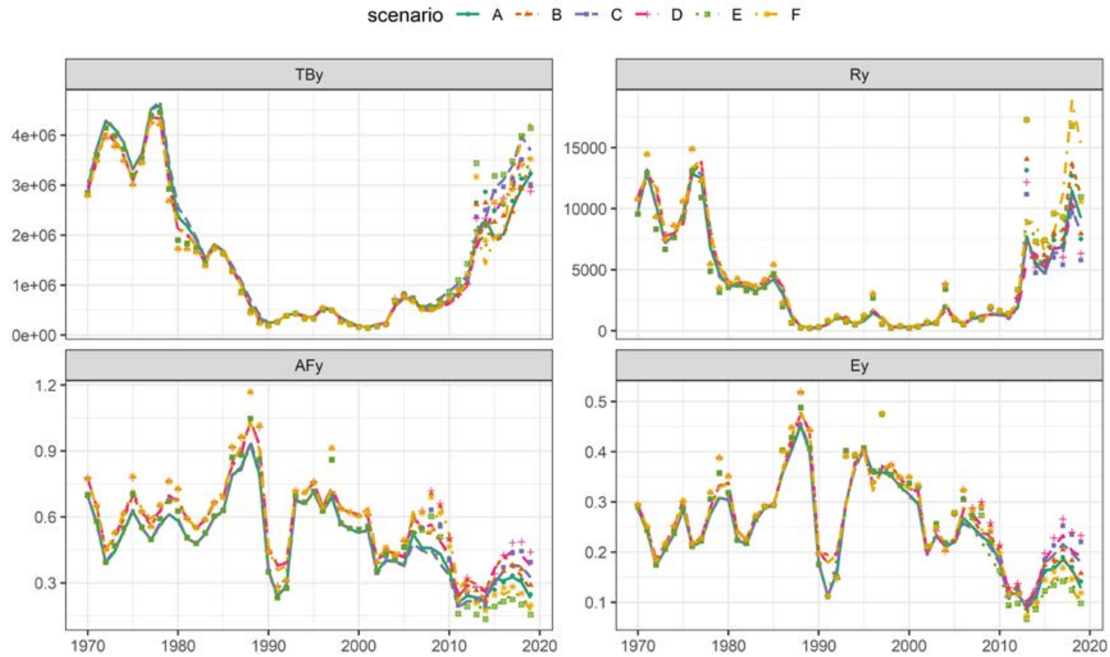

Fig 9: Time series of state variables (TBy: total biomass, Ry: the number of recruits, AFy: F at age weighted by catch weight at age, and Ey: exploitation rate) with SAM under scnarios A to F. The coloured lines indicate the median values among iterations per dataset while the shadowed areas represent 80% estimation intervals among iterations. Note that the *y*-axes are log-transformed for visualization of large intervals.

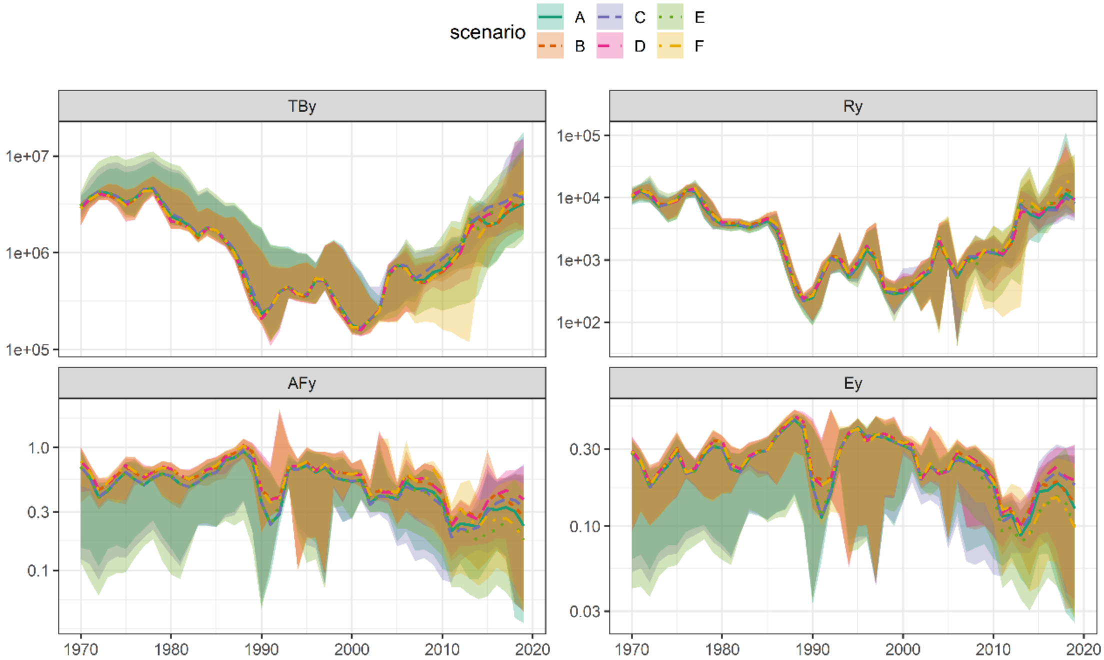

Fig 10: BRPs related to fishing mortality with SAM from scnario A to F. The points of X indicate the values in fitting SAM to the real data under scenarios A to F. Note that the points do not indicate true values because pseudo data were generated from not only SAM but also other models. The *y*-axes are log-transformed for visualization of large variations.

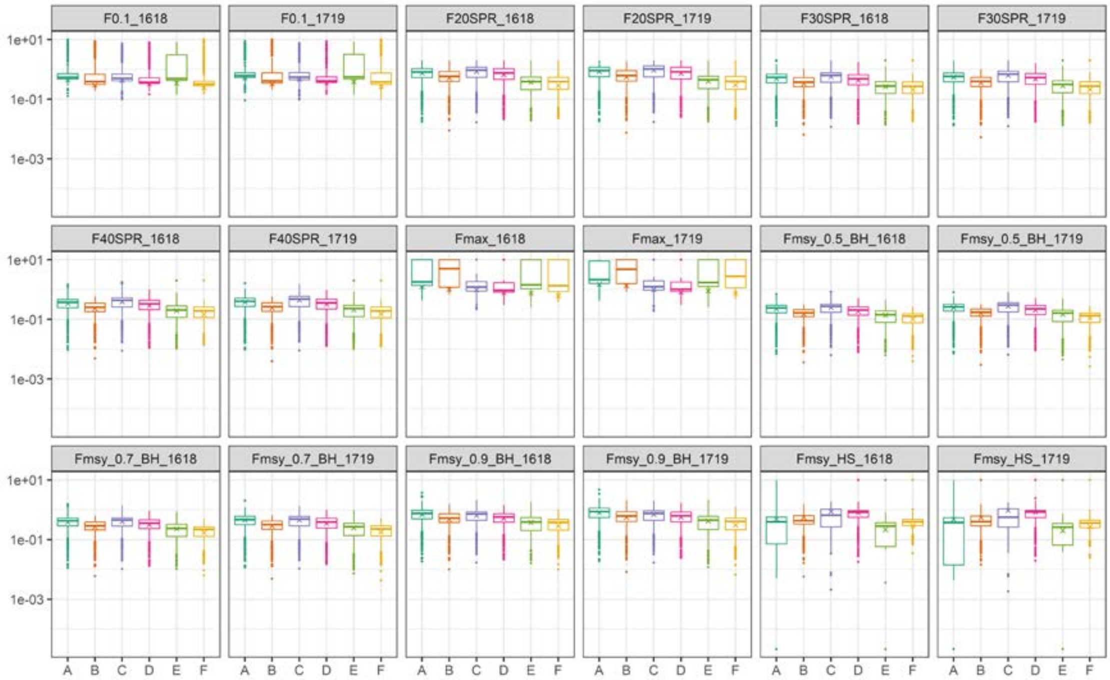

Fig 11: BRPs related to total biomass and spawning stock biomass with SAM from scnario A to F. The points of X indicate the values in fitting VPA to the real data under scenarios A to F. Note that the points do not indicate true values because pseudo data were generated from not only SAM but also other models. The *y*-axes are log-transformed for visualization of large variations.

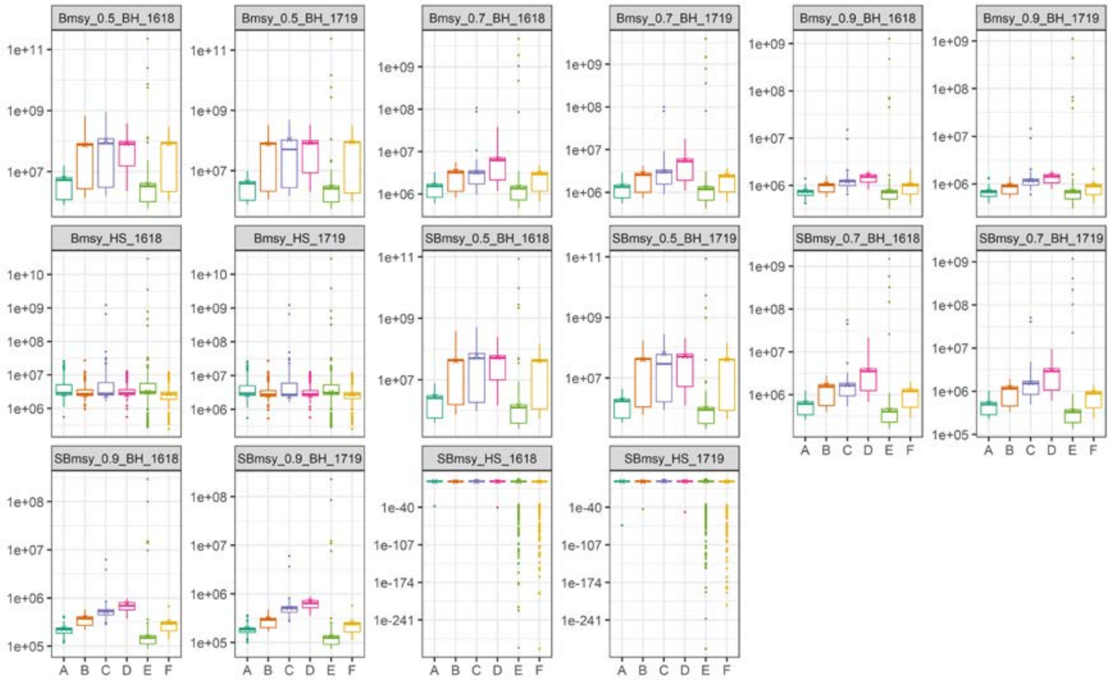

Fig. 12: Relative fishing impacts (ratio of current F to F reference points shown in Fig. 3) with SAM from scnario A to F. The points of X indicate the values in fitting SAM to the real data under scenarios A to F. Note that the points do not indicate true values because pseudo data were generated from not only SAM but also other models. The *y*-axes are log-transformed for visualization of large variations. The horizontal dashed lines indicate 1 (i.e., current F is equal to F reference point).

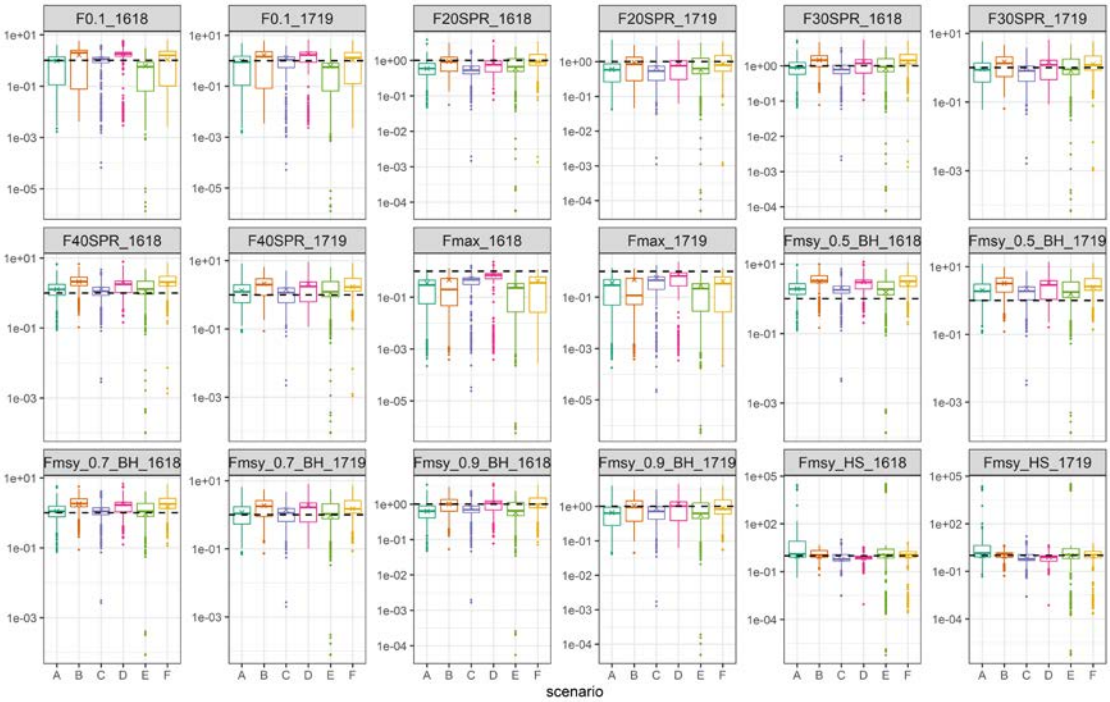

Fig. 13: Depletion statistics with SAM from scenario A to F. The colors indicate each scenario. The points of ‘X’ indicate the values in fitting SAM to the real data under scenarios A to F. Note that the points do not indicate true values because pseudo data were generated from not only SAM but also other models. The relative values to median are log-transformed for visualization of large ranges.

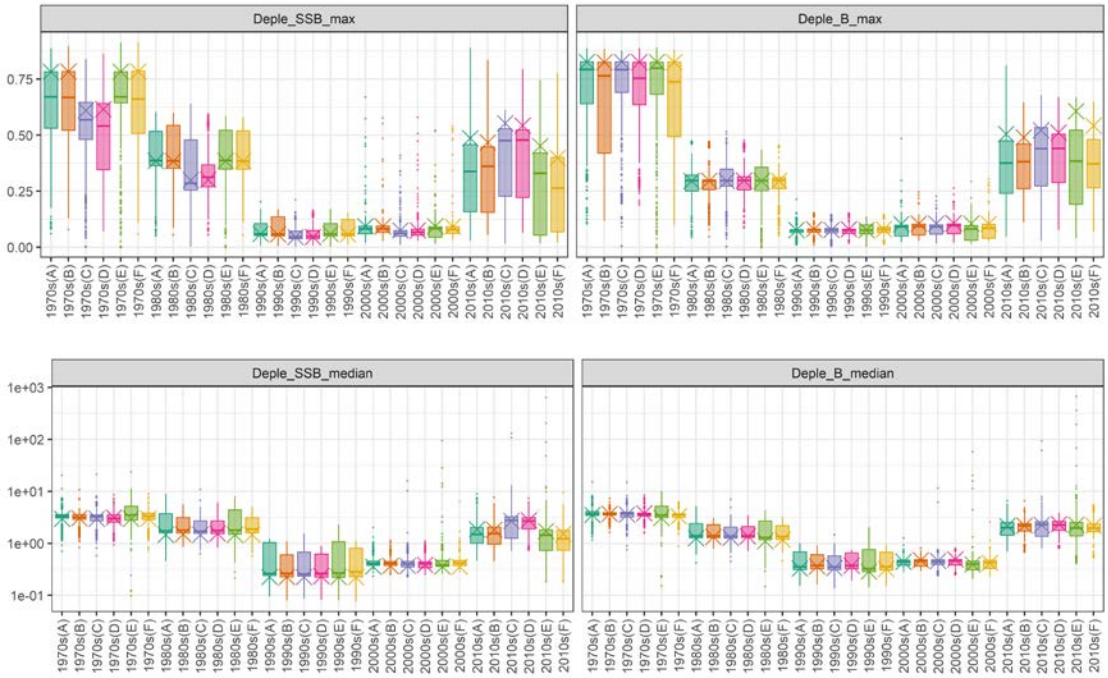

Fig. 14: Frequencies of Mohn's rho of weighted average F (AF_y), spawning stock biomass (SB_y), and total biomass (TB_y) in the 7-year retrospective analysis from scenario A to F. Blue and red indicates SAM and VPA, respectively. The solid horizontal lines and the upper values show the median of Mohn's rho among iterations. The dotted vertical lines and the lower values show Mohn's rho scores with true data. Note that we plotted Mohn's rho values more than 5 as 5 for visualization purpose.