North Pacific Fisheries Commission

NPFC-2022-TWG CMSA06-WP06

3rd Meeting of the Small Working Group on Operating Model for Chub Mackerel Stock Assessment
30 June 2022 (10 a.m. – 2 p.m. Tokyo time)
Webex

Summary

Agenda Item 1. Opening of the Meeting

The 3rd intersessional meeting of the Small Working Group on Operating Model for Chub Mackerel Stock Assessment (SWG OM) commenced at 10 AM on 30 June 2022, Tokyo time in the format of video conferencing via Webex. The meeting was attended by China (Qiuyun Ma, Libin Dai), Japan (Shota Nishijima, Kazuhiro Oshima, Momoko Ichinokawa) and Russia (Oleg Katugin, Vladimir Kulik, Igor Chernienko) as well as the Secretariat (Alex Zavolokin, Sungkuk Kang, Mervin Ogawa). Dr. Joel Rice attended the meeting as an invited expert. The meeting was opened by Dr. Shota Nishijima (Japan) who served as the SWG OM Lead.

Agenda Item 2. Adoption of Agenda

There were no amendments to the agenda.

Agenda Item 3. Review of revisions of pseudo data generated from PopSim

The invited expert (Dr. Joel Rice) reported on the progress since TWG CMSA05 regarding the generation of pseudo data by POPSIM-A as input to the candidate stock assessment models. POPSIM-A data were re-estimated using a maximum age of 15, with previous age range of 1 to 7 being revised to 1 to 15 and summarized to seven age classes: from 0 to 6+.

The invited expert also reported on the re-estimation of selectivity for KAFKA, the inclusion of observation error for SAM non-linear coefficients and the provision of scenarios for the models and explained the random observation errors in nonlinear abundance indices.

The invited expert informed participants that he did not exclude any data and asked for any recommendation on what can be construed as unrealistic data, concluding that it is difficult to determine if data is unrealistic or otherwise as what is realistic for one model can be unrealistic for another.

Japan pointed out that the 15 age groups were not summarized into the seven age classes from 0 to

2nd Floor Hakuyo Hall
Tokyo University of Marine Science and Technology
4-5-7 Konan, Minato-ku, Tokyo
108-8477 JAPAN

TEL +81-3-5479-8717
FAX +81-3-5479-8718
Email secretariat@npfc.int
Web www.npfc.int

6+ in the provided pseudo data. The invited expert answered that he would check this issue later.

Agenda Item 4. Review of revisions of OMutility for calculating performance measures

Japan reported the release of the new version of the OMutility package (v1.0.1) that was redistributed on 2 June 2022. The new version of OMutility fixed the following bugs in the previous version: (1) error when installing the downloaded package from the NPFC Collaboration site; (2) missing identifier of pseudo dataset (A_3, A_15, …and F20); (3) missing depletion statistics output in the 2010s; (4) incorrect output of depletion statistics of total biomass where it shows values similar to those of the spawning stock biomass; (5) failure to find the biological reference point of F0.1 in some cases; (6) incorrect output of parameters of stock-recruitment relationship that are not included in the performance measures.

China commented that there were no problems encountered after using the new version and that they will submit their results. Russia also used the package and assured their submission of results.

Agenda Item 5. Discussion of performance measures and prioritization

Participants discussed what metrics and methods to be used for comparing performances among candidate stock assessment models. Participants also discussed which pseudo datasets to include in the subset of models for retrospective analysis.

Participants recognized the need to hold further discussions on priority performance measures for evaluating the stock assessment models, including consideration of the following:

(a) Reducing their dimension – which performance measures are correlated or uncorrelated?
(b) How do different performance measures conflict among the models?
(c) How should the relative weight for self-test and cross-test be allocated? (Good performance in self-test is usually of particular importance.)
(d) What situations, in particular, should be avoided?
High values in Mohn's rho
Overestimation of recruitment in most recent years
Overestimation of Fmsy and underestimation of Bmsy
Underestimation of relative fishing impact (Fcurrent / Fref)

Which metrics?

- Measure of deviation to true values
  - ✓ Root mean squared error (RMSE): $\sqrt{\frac{1}{100}\sum_{i=1}^{100}\left(X_i^{estimated} - X^{true}\right)^2}$

✓ Median absolute relative bias: $median\left(\left|\frac{X_i^{estimated}-X^{true}}{X^{true}}\right|\right)$

- Measure of bias
  - ✓ (Median) Relative percent bias (%bias): $median\left(\frac{X_i^{estimated}-X^{true}}{X^{true}}\right)$
- Measure of variance
  - ✓ CV

Other matters about metrics

- Should X be log-scaled and/or standard scaled?
- Convergence check: What is the proportion of model convergence?
  - ✓ What is the treatment for iterations with no convergence?
- Should we collect per-iteration results rather than summary statistics like mean and median to investigate the attributes of metrics?

The Chair, with the agreement of all participants, instructed the invited expert to collect twenty percent (20%) of the total iterations, proportionally by model for each scenario and provide the participants a list of those random iteration numbers for retrospective analysis.

Agenda Item 6. Report on progress of candidate stock assessment models

The Chair reported that Japan has no specific progress for the improvement of stock assessment models but has received from Russia answers to questions regarding the model configuration of KAFKA.

Russia explained that KAFKA (cohort analysis with Kalman filter) is a statistical cohort model developed for age-structured stocks which takes the following inputs from the research survey: catch-at-age, CPUE and averaged catch per haul (or absolute abundance estimation). KAFKA includes three consecutive calculating stages with common input and independent outputs for each stage. Russia concluded the presentation by showing the result of running the model.

The Chair requested Russia to post the KAFKA specifications on the NPFC Collaboration to share with participants. Russia replied that the document needs further revision but promised to post the updated document approximately in two weeks' time.

The Chair noticed the big difference between SSB filter and SSB cohorts in recent years for the KAFKA model Scenario A and asked if the difference is caused by the filtering after cohort analysis, which Russia answered in the affirmative. When asked by the invited expert, Russia confirmed that SSB filter is the final estimate. The Chair then clarified if no difference in past years is because of the unavailability of the abundance stock indices. Russia responded that prior to 2000, the calculation used only models, but the recent years used the combination of models and observations.

Agenda Item 7. Timeline up to the SWG OM04 meeting and the TWG CMSA06 meeting

Participants reported that they have already submitted the performance measures based on true data for VPA, SAM and ASAP to the invited expert. The deadline for KAFKA and BSSPM was extended to 8 July 2022.

Japan agreed to provide the aggregate plus age group data by 5 July 2022.

The invited expert will send the list of 20 iterations per model per scenario for retrospective analysis by 5 July 2022.

Participants agreed to submit performance measures for scenarios A and B by 9 August 2022. Participants will also submit of Mohn's rho of retrospective analysis for scenarios A and B by the same date, 9 August 2022, if possible.

Participants also agreed to submit performance measures for all other scenarios and for Mohn's rho by 22 August 2022.

The timeline of tasks is attached (Annex).

Agenda Item 8. Other matters

No other matters were raised.

Agenda Item 9. Closing of the Meeting

The meeting closed at 2:11 p.m. Tokyo time.

Annex

Timeline of Tasks

Category	May 2022	June 2022	July 2022	August 2022	5-8 September 2022 (TWG CMSA06)
ASAP		Submit performance measures based on true data by 30 June. (Done for ASAP, VPA, and SAM, deadline extended for KAFKA and BSSPM)	Done	・Fit to pseudo data and calculate performance measures ・Submit performance measures based on pseudo data for scenarios A and B by 9 Aug ・Submit Mohn's rho of retrospective analysis under scenarios A and B by 9 Aug if possible ・Submit performance measures for all other scenarios and for Mohn's rho by 22 Aug	
VPA			Done		
SAM			Done		
KAFKA	Answers to the questions from Japan by 31 May. (Done)		Submit performance measures based on true data by 8 July		
BSSPM	Improve by 15 June if needed.				
Generation of pseudo data (PopSim)		By 15 June (Almost done, see the next column)	・Aggregate plus age group data by 5 July ・Provide the information of 20 iterations per model per scenario for retrospective analysis by 5 July		
Performance measures (OMutility)		By 15 June (Done)			
Scoring and Ranking	How to use the performance measures for scoring and ranking of the candidate stock assessment models.			How to use the performance measures for scoring and ranking of the candidate stock assessment models.	Rank the candidate stock assessment models.
Intersessional SWG OM meetings		•30 June Agenda: progress in generating pseudo data, initial check if there are any issues with fitting models and calculating performance measures.(Done)		•12 Aug Agenda: TBD	