INTER-AMERICAN TROPICAL TUNA COMMISSION

EXTERNAL REVIEW OF MODELLING ASPECTS FOR STOCK ASSESSMENTS OF TROPICAL TUNA IN THE EASTERN PACIFIC OCEAN

La Jolla, California (USA)
6-10 November 2023

REPORT OF THE MEETING

CONTENTS

Executive summary 2
1. Main recommendations of the Panel 2
2. Exploring risk analysis hierarchies........ 3
3. Introduction 5
4. Description of the base/ancestral models for bigeye and yellowfin tuna in the EPO 5
4.1. Bigeye tuna in the EPO 5
4.2. Yellowfin tuna in the EPO 6
5. Topics addressed during the review 7
5.1. Comparison of the methods for assessments for the 3 species 7
5.2. Risk analysis approach and hypotheses for the risk analysis for bigeye tuna 7
5.3. Hypotheses for the risk analysis for yellowfin tuna 9
5.4. Fisheries definitions, length compositions, and fishery selectivity assumptions 10
5.5. Time step of the model, number of age classes, initial conditions and start year 11
5.6. Recruitment 12
5.7. Growth 13
5.8. Natural mortality 15
5.9. Indices of abundance 15
5.10. Data weighting 17
5.11. Model diagnostics 18
5.12. Model weighting 19
5.13. Estimation of management quantities and associated uncertainty 20
6. List of requests 21
7. Figures 26
Annex 1. Terms of Reference for the Panel review 32
Annex 2. Tentative agenda of the Panel review 35
Annex 3. Participants in the review 37

EXECUTIVE SUMMARY

The first external review of IATTC modelling aspects for stock assessments of tropical tuna met with the aim of improving the stock assessments and consequent management advice for bigeye and yellowfin tuna in the EPO. Eight high level recommendations were made, including the need for additional research to resolve uncertainties about growth, natural mortality, and absolute abundance, employing close-kin mark-recapture, increasing computational resources, further use of ensemble approaches, and maintaining the hierarchical risk analysis. Fifteen topics were discussed, and the Panel made further specific recommendations on growth, natural mortality, stock structure and fishery definition, model consistency, and indices of abundance. These were incorporated into recommendations on potential future hypotheses for risk analysis. The Panel appreciated the recent developments in stock assessment on these stocks, and thanked participants for the clarity of the presentations and openness of the dialogue.

1. MAIN RECOMMENDATIONS OF THE PANEL

1. Conduct biological research in order to resolve important unknowns; especially age and growth, natural mortality, movement, and spatial variation in biological processes.
2. Evaluate the feasibility of employing close-kin mark-recapture to help resolve uncertainties with regards to absolute abundance, natural mortality, and stock structure.
3. Suggest investigating options for increasing computational resources available for stock assessment analysts via high-performance/high-throughput computing, either on premises or in the cloud.
4. Encourage further development of the ensemble approach with respect to increasing the number of scenarios investigated in order to capture the full range of uncertainty.
5. Encourage approaches that are robust to spatially varying processes (e.g., growth).
6. Use analytically derived data weights within the stock assessment model where possible.
7. Support inclusion of models in the risk analysis that have resolved or mitigated data conflicts.

Topic specific recommendations are provided in the body of the report.

2. EXPLORING RISK ANALYSIS HIERARCHIES

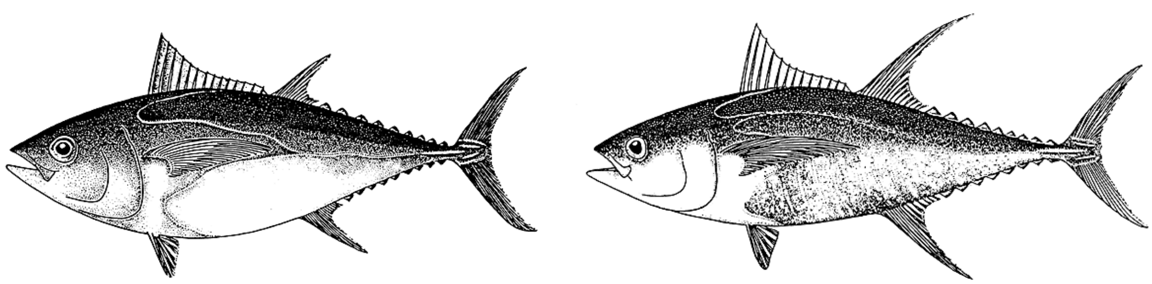
In light of the documents, presentations, the risk analysis hierarchies from the 2020 stock assessments, and discussions with IATTC staff, the Panel explored potential future approaches for the hierarchies of future risk analysis of bigeye and yellowfin tuna in the EPO. The Panel recommends that the following structure be considered for the next iteration of the risk analyses.

Bigeye tuna

- (Level 1a) Alternative CPUE assumptions (how to fill un-sampled areas)
- (Level 1b) Spatial variation in growth:
  - Down-weight non-core comps
  - Areas-as-fleets (movement)
- (Level 1c) Regime shift
  - True (low priority), not true (short/medium models)
- (Level 2) Growth estimation
  - use tagging data to inform prior on L2
  - Consider sensitivity to sigma R
  - Functional form
  - Estimate without the otolith data (no prior on L2 if asymptotic selectivity assumed)
- (Level 2) 1% effort creep
- (Level 2) Sex-ratio change (females)
  - Growth
  - Natural mortality
    - Fixed vs. estimated
    - Functional form: Lorenzen
  - Selectivity
- (Level 2) Asymptotic selectivity
  - Time-varying selectivity on fleet that catches the largest fish (increase the frequency of time variation)
- (Level 3) Steepness of the stock-recruitment curve

Yellowfin tuna

- (Level 1) Stock structure (static fishery definitions; dynamic is research issue)
    - Whole EPO model (fit to core data NE, include catch from SW)
    - 2 models: NE model, SW model
        - 2 models: NE model, SW model (fix parameters at some NE model values)
        - 2 box model (no movement, but parameter sharing)
- (Level 2) Sex-ratio change (females)
    - Growth
    - Natural mortality
        - Fixed vs. Estimated
        - Functional form: Lorenzen
    - Selectivity
- (Level 2) Growth estimation
    - use tagging data to inform prior on L2
    - Consider sensitivity to sigma R
    - Functional form
    - Estimate without the otolith data (no prior on L2 if asymptotic selectivity assumed)
- (Level 2) Asymptotic selectivity
- (Level 2) Catchability for DEL index
- (Level 2) 1% effort creep
- (Level 3) Steepness of the stock-recruitment curve



Images: FAO

3. INTRODUCTION

The first external review of IATTC modelling aspects for stock assessments of tropical tuna in the eastern Pacific Ocean met in La Jolla, California from 6th to 10th November 2023. This report reflects the consensus views of the review Panel. The purpose of the external review was to provide information that will improve the stock assessments and consequently the management advice.

The goals and objectives of the review were to:

a. identify the best available science for the EPO tropical tuna stock assessments;

b. provide an independent review of the stock assessment approach;

c. provide advice on future research and data collection that will improve the assessments and the provision of management advice.

The Terms of Reference of the review are shown in Annex 1, the draft agenda in Annex 2 and the list of the participants in Annex 3 of this report.

This report highlights the main findings and recommendations in its opening sections (executive summary, main recommendations, and exploration risk analysis hierarchies). The methods used for the base/ancestral stock assessment methods are described. Further specific details on the 15 topics listed in the Terms of Reference for bigeye and yellowfin tuna assessments are provided in the subsequent sections. These include summaries of the presentations by IATTC staff, and the Panel's views on the technical merits and/or deficiencies in the assessment approach and recommendations for remedies, and any unresolved problems and major uncertainties. List of analyses requested by the Panel, rationales for each request, and brief summaries of the responses are shown in tabulated form at the end of the report.

No public comment was received during the review. There were no unresolved differences of opinion between Panel members.

4. DESCRIPTION OF THE BASE/ANCESTRAL MODELS FOR BIGEYE AND YELLOWFIN TUNA IN THE EPO

The base/ancestral models were presented to the Panel. These were based on the 2020 assessments with further modifications.

4.1. Bigeye tuna in the EPO

- Six changes have been made to the bigeye assessment model since the last benchmark assessment in 2020.
- Those changes lead to a faster decreasing longline index of relative abundance and a better fit to the longline length compositions associated with an asymptotic selectivity.
- The recruitment regime shift observed in the last benchmark assessment is not obvious in the new assessment model, likely indicating fewer model misspecifications.

Two issues remain in the new model:

- No longline fishery selectivity is predicted to have an asymptotic selectivity.
- A pronounced temporal trend in the residual of longline survey length comps

Model assumptions

- Quarter-as-year: quarterly time step from 1979
- 40 age groups (0-39 quarters)

- Population length bins: 2 cm – 220 cm
- One initial R regime parameter and two initial Fs (one for OBJ and one for LL)
- Twenty-eight early recruitment devs
- Areas-as-fleets approach: one survey fleet and twenty fishery fleets
- B-H stock-recruit relationship with a steepness of 1
- Fixed natural mortality (sex-specific) and growth that are estimated outside the assessment model

Data used

- Catch time series for each fishery (1979-2020)
- One standardized longline index of relative abundance based on Japanese CPUE (1979-2020)
- Standardized catch-weighted length compositions for longline fisheries (1986-2020)
- Standardized CPUE-weighted length compositions for longline survey (1986-2020)
- Unstandardized catch-weighted length compositions for purse-seine fisheries (mostly since 2000)
- Conditional age-at-length (used in some reference models where growth is internally estimated)

4.2. Yellowfin tuna in the EPO

- The yellowfin assessment uses the “areas-as-fleets” approach.
- Areas defined using tree analysis on length frequency data.
- Multimodal patterns in length composition: selectivities are splines for most fleets.
- The model implements the “high mixing” hypothesis focus on the core of the catches: assuming the longline index does not represent the main group of fish.
- The model is fit to an index derived from the purse-seine on dolphin fishery north of 5N.
- Data that may be from another group is not fit (longline data and purse-seine data on dolphins in the “south”), those fisheries have fixed selectivity.

Issues in the model:

- Spatial dynamics may not be fully captured.
- The spline selectivities at length imply in selectivity at age with no fully selected age for some fisheries.
- The index of abundance does not have asymptotic selectivity.

Model assumptions

- One area model
- Quarter-as-year (recruitment estimated in each quarter): quarterly time step from 1984
- 29 age groups (max = 7.25 years)
- Population length bins: 2 cm – 220 cm

- One initial R regime parameter, one initial Fs (F16-DEL_NE), 16 initial recruitment deviations
- Areas-as-fleets approach: one survey fleet (PS-DEL) and 38 fishery fleets
- B-H stock-recruit relationship with a steepness of 1
- Fixed natural mortality M (sex-specific) and growth

Data used

- Catch time series for each fishery (1984-2022)
- One standardized purse-seine index of abundance based on catch per set on dolphins for area north pf 5N
- Standardized length composition associated with the index
- Unstandardized catch-weighted length compositions for purse-seine fisheries

5. TOPICS ADDRESSED DURING THE REVIEW

5.1. Comparison of the methods for assessments for the three species

Presentation by staff

Introductory presentations on IATTC, an overview of the three species of tropical tuna (bigeye, yellowfin and skipjack) stock assessments and the risk analysis approach were given.

Link to presentation1. Link to presentation2. Link to presentation 3.

Panel comment

The Panel asked general questions about IATTC and the three species.

Panel recommendations

There are no recommendations on this topic.

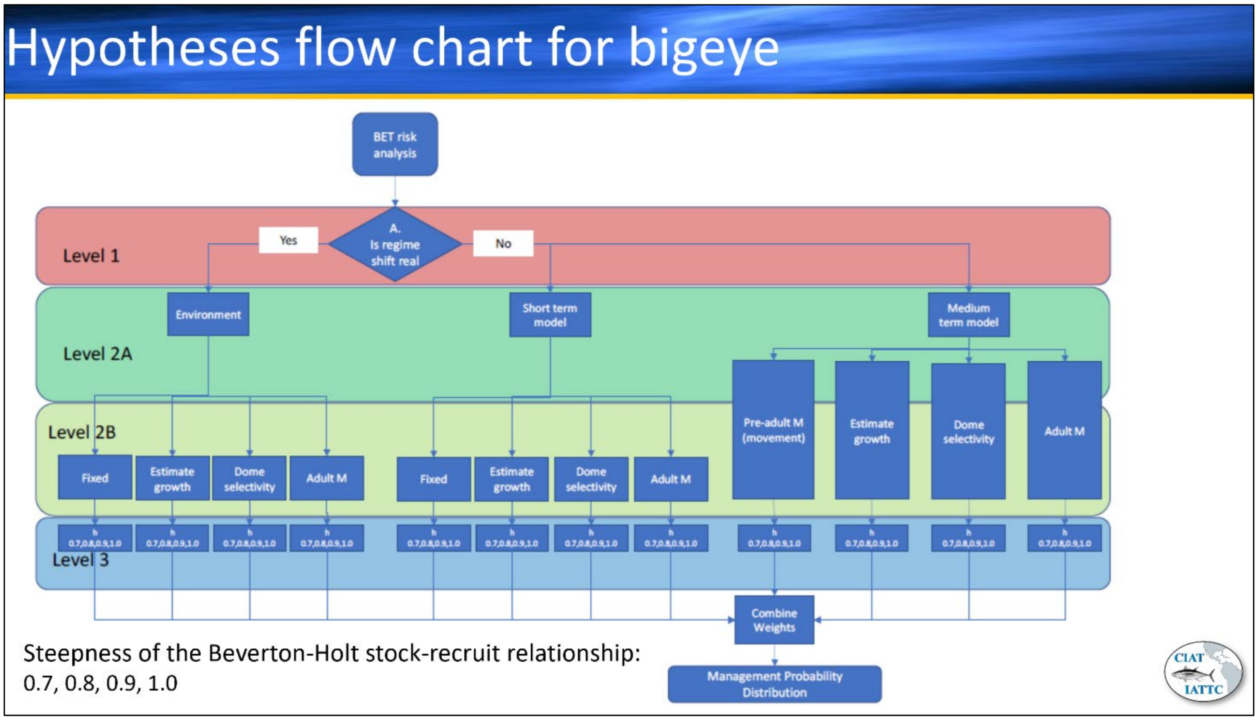
5.2. Risk analysis approach and hypotheses for the risk analysis for bigeye tuna

Presentation by staff

- Tagging data show pronounced mixing of bigeye juveniles between the EPO and WCPO.
- Very different Linf and length at 50% maturity between the EPO and WCPO.
- Spatial heterogeneity of population depletion rate has been observed in both longline and purse-seine fisheries in the EPO.
- The EPO bigeye assessment model is not spatially-structured mainly because of 1) a lack of tagging data for adult bigeye and 2) its inability to resolve the regime shift in R.

Hierarchical hypotheses for the risk analysis

- 1^{st} level hypothesis: the regime shift in R is real or due to model misspecification.
- 2^{nd} level hypothesis: why no longline fishery is estimated to have an asymptotic selectivity.
- 3^{rd} level hypothesis: what is the steepness of the B-H stock-recruitment relationship.



Panel comment

The Panel discussed the mixing of juveniles between the EPO and CPO (WCPO) and the differences in Linf and L50 (maturity) between and within the two areas. There is spatial heterogeneity in depletion of longline (LL) and purse seine (PS-OBJ) CPUE in the EPO. The stock is more depleted in the inshore EPO area. The Panel considered how to deal with spatially varying biological parameters and the lack of tagging data for adult bigeye. The current assessment framework (areas-as-fleets) assumes a single well-mixed stock. Differences in age/length comps are assumed to be due to spatial differences in availability.

The Panel had extensive discussion related to the existence of the recruitment regime shift in the mid-1990s. Model developments since the previous benchmark assessment appear to mediate the magnitude of this recruitment shift, and it was discussed that it is more probable that the recruitment shift is indicative of a model misspecification rather than a real phenomenon. However, it is recommended that the possibility that the recruitment shift is real be retained as a hypothesis in the risk analysis, albeit one with lower probability.

The Panel also discussed how the empirical selectivity diagnostic indicates a lack of asymptotic selectivity for the longline fishery, which could be indicative of potential model misspecification in growth, natural mortality, or selectivity. However, as noted in [diagnostic section] the empirical selectivity diagnostic is potentially overly-sensitive to the largest individuals.

Panel recommendations

The Panel focused on how to deal with spatial patterns in the growth/longline selectivity issue and after discussion with IATTC staff some alternative approaches for dealing with these issues are suggested:

i. Focusing on modelling the data from the core area (this is further discussed in the section on growth below);

ii. Developing a spatiotemporal age-length key to convert length observations to age external to the model, and assuming a single growth curve within the model;

iii. Approximate intraregional movement (availability) using areas-as-fleets.

Additional recommendations related to alternative index assumptions, growth estimation, effort-creep, and accounting for the sex-ratio change in females at length are described in greater detail in subsequent sections.

5.3. Hypotheses for the risk analysis for yellowfin tuna

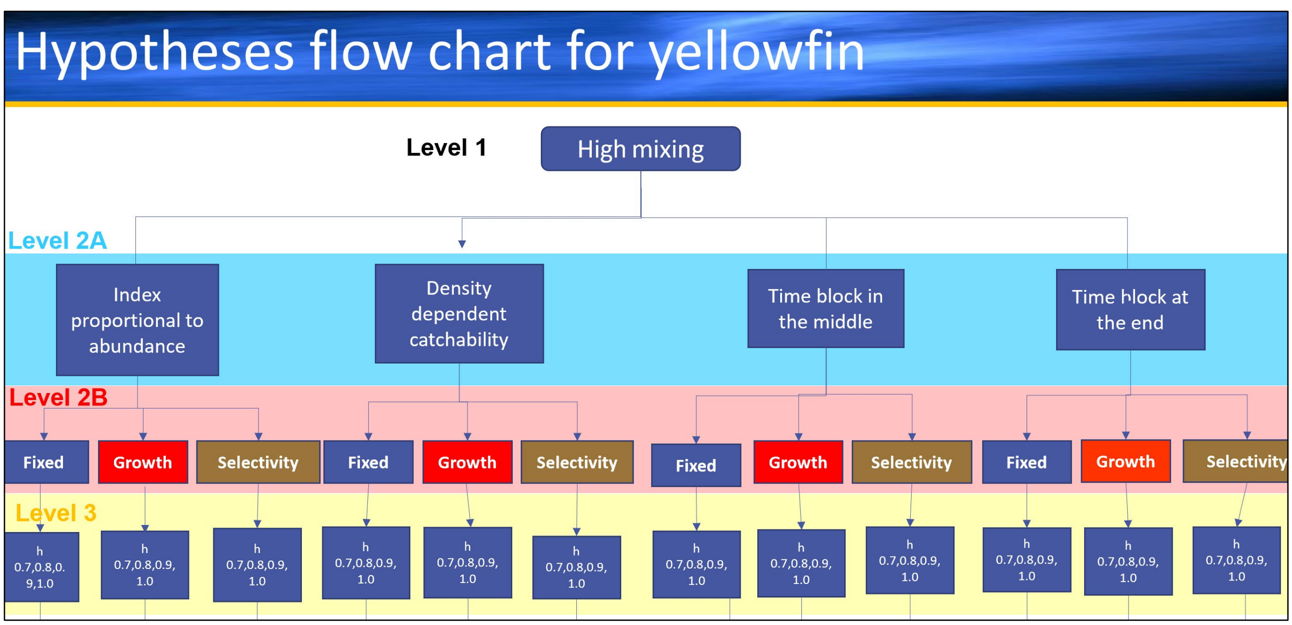
Presentation by staff

- The yellowfin tuna stock in the EPO may be composed of at least two groups of fish with different dynamics.
- The groups may be associated with major habitats and roughly located towards the NE and SW regions in the EPO, and have dynamic boundaries.
- No index comprises both groups, each of the two main indices of abundance may be associated with only one group.
- In 2020, assessment done with focus on the core of the catches and high mixing assumptions.
- Episodic Mixing and No Mixing between groups to be modelled by splitting them based on environment.

Hierarchical hypotheses for the risk analysis

- Level 1 Hypothesis: Mixing: assumed High mixing between groups, other hypothesis not modelled.
- Level 2 hypotheses (data may be informative): related to index of abundance, growth and selectivity.
- Level 3 hypotheses (no information in the data): steepness of the B-H stock-recruitment relationship.

Link to presentation



Panel comment

The key issues to be considered here are the alternative biological hypotheses about stock and/or spatial structure. The evidence presented supported the idea that individual movements of adult fish are likely to be relatively constrained, which suggests that the high mixing scenario is unlikely. There appears to be a degree of viscosity in the population, with spatial variation in mean sizes and lengths at maturity, and potentially in growth, if local productivity affects growth rate and/or asymptotic length.

The degree to which yellowfin tuna sub-stocks are associated with biogeographical provinces / sets of environmental conditions is less clear. The Panel suggested that analysts consider an alternative hypothesis in which fish behaviour and catchability at size are affected by environmental conditions and change when the biogeographic boundaries move.

It is also difficult to model a population that moves with time-varying boundaries based on environmental conditions.

Panel recommendations

Although the high-mixing scenario is unlikely, the areas-as-fleets model is a potentially useful approach.

Important to identify data conflicts and resolve them in each scenario, or further split.

Use fixed boundaries guided by environmental conditions. Boundaries that change inter-annually are a research project.

As a research project, consider the hypothesis in which fish behaviour and catchability at size are affected by environmental conditions and change when the biogeographic boundaries move.

[Consider modeling two separate NE and SW sub-stocks both independently and in a single model with some shared parameters.] Consider using early Japanese CPUE to estimate regional scaling between sub-stocks.

5.4. Fisheries definitions, length compositions, and fishery selectivity assumptions

Presentation by staff

Bigeye tuna

- Double-normal (asymptotic for one longline fleet and dome-shaped for rest fleets) selectivity is used for all fisheries.
- Purse-seine fishery selectivity curves are time-invariant.
- Longline survey selectivity curve is time-invariant.
- Longline fishery selectivity curves have three blocks (before 1994, 1994-2010, and after 2010).

Issues:

- An obvious temporal trend in the residuals of longline survey length comps.
- Empirical selectivity analyses indicate that all longline fleets have dome-shaped selectivity.

Yellowfin tuna

- Fishery definitions in the base model are based on tree analysis of length composition data and lat, lon, cyclic quarter and quarter as explanatory variables.
- Assumptions about the fisheries that catch larger fish are influential and included in the risk analysis.

- Bimodal pattern in PS-NOA may be due to fish behaviour, which justify splitting the fishery.
- New conceptual model based on habitat used to redefine fisheries definitions for PS-OBJ and PS-NOA.
- New fishery definitions allowed to implement negligible mixing hypothesis.
- New runs highlighted the impact of the longline fishery length composition data.
- Most purse seine selectivities are estimated using splines, while longline selectivities are not estimated but fixed.

Issues:

- Longline data is influential.
- Convergence of the "SW" model is problematic.

Panel comment

The Panel supports the analytical tree-based approach that has been used to define fisheries, as well as ongoing refinements to the method using habitat information. The selectivity time blocks in the bigeye assessment are justified by changes in the fishery operations and data collection.

There are outstanding questions about long-term trends in bigeye recruitment and the longline survey mean length residuals, and uncertainty about what the underlying causes might be is incorporated into the risk analysis model ensemble.

The regional structure of yellowfin tuna poses a challenge. In the 2020 assessment, the selectivity of SW longline fisheries was fixed because those length compositions were causing large changes in the overall stock size, even though those fisheries are small in magnitude. Approaches are currently being explored to find the best way to model the southwestern (SW, mainly longline, small part of the total catches) and northeastern (NE, mainly purse seine, large majority of the total catches) parts of the stock.

Panel recommendations

The Panel recommends exploring three general directions for developing the yellowfin assessment: (1) combined NE+SW model, (2) two separate NE and SW models, and (3) two separate NE and SW models sharing parameter values or priors. The results of this exploration will guide the modelling choices that will form the basis of the assessment and management advice.

Further analysis using GAMs to better understand causes of size data variation.

Explore increasing the frequency of time-variation in fishery selectivity, particularly for the large-fish fisheries that are influential for population scale.

5.5. Time step of the model, number of age classes, initial conditions and start year

Presentation by staff

- Models use quarters as time step: allows for recruitment to be estimated in each quarter.
- Starting year chosen to be when reliable indices start.
- Models are not sensitive to the start year, but sensitive to the data in earlier period.

Panel comment

It was noted that the maximum age is 7.5 years in the yellowfin assessment and 10 years in the bigeye assessment; whereas the maximum ages estimated for the Indian, Western Pacific, and Atlantic Oceans are older (e.g., 17 for yellowfin and 18 for bigeye in the Atlantic). If there are still some fish alive at the maximum age in any part of the time series, the assumed maximum age can be influential if the growth curve is not flat at the maximum age. A run was requested for bigeye with the maximum age increased to 60 quarters (15 years). This had no impact on assessment results, although results may differ with different assumptions for natural mortality or growth.

The short-term bigeye model started in the year 2000 in order to avoid including information from the early period before the full expansion of the floating object purse seine fishery. It was suggested that some of the conditions during the early period may be sufficiently different to affect parameter estimates that are applied across all time periods. Including early period data may result in parameter estimates that are inappropriate for the later period, or data conflict that affects assessment outcomes.

Panel recommendations

Retain the scenario that starts the model in 2000 within the bigeye tuna risk analysis.

5.6. Recruitment

Presentation by staff

- Steepness = 1 for estimation
- Lognormal recruitment deviates with fixed Rsd
- No autocorrelation in estimation
- Use bias correction ramp
- Use dynamic B0
- Include precautionary assumptions about steepness in reference points and management goals

Summary of base case

Parameter	YFT	BET
Stock-Recruitment	Beverton-Holt fixed steepness = (0.7, 0.8, 0.9, and 1.0)	Beverton-Holt fixed steepness = (0.7, 0.8, 0.9, and 1.0)
Recruitment variation	Quarterly, lognormal, sd fixed at 1.0, penalized likelihood, bias adjustment ramp	Quarterly, lognormal, sd fixed at 0.6, penalized likelihood, bias adjustment ramp, (recruitment regime parameter)
Spawning biomass	Proportion of mature females, batch fecundity, fraction of females spawning per day, by age (from length)	Proportion mature at length converted into age-at-maturity

Link to presentation

Panel comment

The Panel agreed with the proposed approaches for both BET and YFT. The recruitment regime shift in BET was interpreted as a model misspecification. Viewing how this changed in the stepwise model progression can help improve understanding of what model configurations helped to alleviate this issue. Several factors contributed to alleviating the recruitment regime. However, adding another time block in selectivity to the longline fishery and imposing asymptotic selectivity on the last time block appeared to have the greatest impact. Doing so appears to impact estimates of model scale. By increasing the model scale, a recruitment regime shift is no longer as necessary to explain the rapid increase in catches from the floating object fishery in the mid-1990s.

The recruitment penalty was shown to have a large influence on the likelihood profiles. It is important to understand the sensitivity of the model to the assumed value for this parameter. Model results appeared to be insensitive to the choice of sigmaR when growth was fixed.

Panel recommendations

During estimation, use values of Sigma R that are high enough to avoid affecting recruitment estimates. When estimating management quantities, adjust the recruitment used as appropriate.

5.7. Growth

Presentation by staff

- Use a flexible growth curve.
- Estimate growth inside the stock assessment model.
- Integrate age-length into the stock assessment as data conditional age-at-length.
- Length-increment data should be analyzed outside the model and the results used to create priors to include in the assessment model (until possible to include it in the model)
- Use sex specific growth.
- Need to consider spatial variation in growth and how it might be affecting the results.
- Estimate Lsd inside the model.

Summary of base case

Parameter	YFT	BET
Growth	Richards fixed parameters from previous assessment, (estimated)	Richards, fixed based on otolith and tagging, (estimated)
Variation of length-at-age	Normal, coefficient of variation of 7.5%	Normal, sd assumed proportional to mean length, sd0 estimated and sd40 fixed (estimated)
Length-weight	Allometric Fixed (Wild 1986)	Allometric Fixed, Nakamura and Uchiyama (1966)

Link to presentation

Panel comment

There is a need for more biological sampling and validation to better understand potential spatial and temporal patterns in growth. Strong need to develop methods for ageing older fish.

There is considerable uncertainty about growth for both species. In both cases, estimates inside the assessment tend to have smaller asymptotic length than estimates based on otolith and tagging data, but estimates inside the model are very sensitive to other assumptions such as time blocks, asymptotic selectivity, and even SigmaR.

There was a discussion of approaches for dealing with spatial patterns in growth / availability at size / longline selectivity.

If growth varies spatially and/or temporally, it can substantially affect outcomes in a model that fits to size data and is constrained to have the same growth curve everywhere and throughout the time series. There is a need to develop modelling strategies that are robust to these sources of variation. There is strong evidence for spatial variation in bigeye and yellowfin growth across the Pacific, with considerably smaller asymptotic length (or lengths at older ages) in the western equatorial WCPO compared to the equatorial EPO. There is likely to be spatial variation within the EPO. Similarly, temporal variation is not unlikely and could be associated with density dependence.

The primary issues in a stock assessment are estimating a) the abundance trend, and b) the population scaling. The abundance trend is relatively well-determined by the index of abundance. Observations of the largest fish and assumptions or estimates of asymptotic length (i.e., the growth curve) are influential in determining population scaling.

The Panel proposed an approach that gives priority to information from size data sampled in locations (and period) where growth rates that are consistent with the growth curve. The area (and potentially the time period) that primarily inform the growth curve are identified, and size data outside this area and time period are given lower statistical weight.

An alternative approach is to develop a spatiotemporal age length key outside the model, use this to translate lengths into ages, and fit the model by fitting to these age composition data.

For bigeye, the estimate of L2 based on integrated modelling of tagging data (Aires-da-Silva *et al.* 2015) was higher than when L2 was estimated inside the model while fitting to age at length data and this substantially changed the population scaling. In addition, changing model setup while estimating L2 (e.g., alternative assumptions about SigmaR: 0.6, 1.2, 2.0) gave slightly different L2 estimates which had quite large effects on population scaling.

Uncertainty remained about the best way to model bigeye growth within the assessment model, particularly with respect to including outside data sources and/or priors. There were also concerns about the feasibility of simultaneously estimating growth, natural mortality, and selectivity.

Sex-specific growth appears likely for both species at larger sizes, based on evidence from other oceans and other Thunnus species. However, sex-specific growth is not evident at the sizes that can be read using daily rings. There may be an effect on growth estimates if fish that have stopped growing are less readable and so are not selected for ageing.

Panel recommendations

This issue is very important and influential for the stock assessment. There is a high priority for more biological research. In the short term, explore existing data such as otolith weights and daily age estimates to help evaluate hypotheses about growth variation in space, in time, and between sexes.

As a major source of unresolved uncertainty for both species, scenarios addressing more than one issue may need to be included in the model ensemble.

Explore assessment methods that are robust to spatially varying growth.

Estimation of the growth curve inside the model should include scenarios that apply a prior to L2 based on information in the tagging data.

Separate growth curves for adults by sex should be strongly considered, which may require alternative functional forms such as the growth cessation model. This can be considered in conjunction with natural mortality and availability/selectivity.

Explore the potential to estimate growth using spatial subsets of size composition data.

5.8. Natural mortality

Presentation by staff

- Natural mortality assumed to vary with age and sex.
- Support from theory and sex-ratio at size.
- Adult natural mortality linked to longevity.
- Longevity is difficult to assess because ageing is possible to only to 4-5 years with daily rings – what value to use for longevity?

Panel comment

There was discussion about potential changes to the approach to model M in YFT from that used in 2020. This discussion included adapting to the Lorenzen approach, how to estimate M in juveniles, and in older females and males and the interaction of M, growth and availability. The Panel asked for further exploration of alternative ogives for M in YFT as there is uncertainty about natural mortality ogives, including the base level.

There was little discussion of changing approaches to M in BET. The Panel suggested that the approach to M should be consistent between the two species.

Panel recommendations

Give priority to the Lorenzen approach for both sexes, with the potential for female M to increase associated with reproductive stress.

Estimate the functional form by fitting to the sex ratio outside the model.

Consider 3 hypotheses in the risk analysis to account for the change in sex ratio in older fish; namely that it is due to M, to growth or a combination of M, growth, and availability/selectivity.

Give high priority to the development of close kin mark recapture methods. This is the best option for obtaining empirical data to support estimation of natural mortality.

5.9. Indices of abundance

Presentation by staff

Key issues BET:

- Longline indices of abundance have decreased coverage over time.
- Assumptions on local depletion change the index.

Key issues YFT:

- Longline indices of abundance may not represent the fish targeted by most of the purse-seine fleet.
- Purse-seine dolphin set index may not be directly proportional to abundance of yellowfin.

Panel comment

IATTC staff presented the approaches taken for developing indices of abundance for the bigeye and yellowfin stock assessments. For bigeye tuna, a spatiotemporal model was fit to aggregate (by vessel, month, and hooks between floats) 1x1 Japanese longline data in order to develop a standardized index for the time period 1979-2022. The standardized index was input into the assessment as a single survey fishery (no associated catch and separate selectivity from the extraction longline fisheries) and did not break the index into two time periods as was done in the previous benchmark assessment. A time-varying CV for the index was estimated within the spatiotemporal model, though prior to being input into the assessment model this was re-scaled using an additive constant such that the mean CV for the period 1979-2015 equaled 0.15.

The Panel supported the spatiotemporal modelling approach used to develop an index for use in the assessment. The Panel also found the justification used to support not splitting the index to be reasonable. Selectivity was found to be similar across both time periods and the change in estimates of catchability for the two time periods was counter-intuitive under the hypothesis that effort creep likely occurred within the fishery. The Panel agreed with IATTC staff conclusions that splitting the index can reduce the ability to inform long term trend, and that splitting the index can introduce bias if population scaling across the split is inappropriate.

The Panel noted the concerning systematic decrease in spatial coverage of the Japanese longline fishery in the EPO and discussed alternatives with IATTC in the event that this data stream is no longer a viable option in the future. Data from other longline fleets operating in the region could be used to help fill data gaps. However, these fleets may have different fishing strategies (e.g., selectivity and/or catchability) and more investigation is needed before choosing a replacement data source or modelling different longline fleets together in a joint analysis. In the upcoming benchmark assessment an assumption has to be made for how to deal with the un-sampled areas when constructing the index. The Panel supported using a random-walk in the spatiotemporal random effect term of the standardization model and noted that this assumption was influential on model outcomes. Sensitivity to this assumption should be explored in the upcoming benchmark assessment. It was suggested that a 'preferential sampling' type model could be approximated by assuming that un-sampled areas be filled with the minimum predicted density for cells that had previously supported fishing. Another alternative discussed included truncating the index at the point where the Japanese longline effort distribution was no longer spatially representative of the distribution of the stock.

Similar to bigeye tuna, a spatiotemporal model was used to develop indices of abundance and time-varying CVs for yellowfin tuna. In the northeast EPO a dolphin associated purse seine index was developed since longline fishing operations had primarily shifted to the southwest and may no longer be representative of the population group targeted by most of the purse seine fishing effort. The Panel noted that the dolphin associated purse seine index may not be directly proportional to abundance of yellowfin tuna, and supported the continued inclusion of alternative scenarios for catchability (e.g., density-dependence, and time block to address the peak in catch-rates or residual patterns in the fit to the survey composition data) in the risk analysis.

With regards to the index, the Panel questioned the impact that the temporal closures may have on the index and suggested further characterizations be undertaken to evaluate the potential for the temporal closures to introduce systematic bias into the index. It was noted that the fit to the index appeared to show a strong seasonal pattern in the residuals and it was suggested to explore either a seasonal catchability or splitting the index into quarterly indices. This was investigated by IATTC staff and shown to have no effect on model estimates of stock status in the current model configuration. Further work is being undertaken to increase seasonal resolution of the index, which may increase the impact of seasonality. Lastly, the Panel noted that fisheries that actively engage in searching behaviour may be prone to effort creep as technology, communication networks, and fisher knowledge improve. Collaboration and communication with the industry to identify changes in the fishery is encouraged. Additionally, developing an understanding of encounter rates of dolphin pods and/or dolphin pods with associated schools of tuna could provide auxiliary information on yellowfin abundance.

Panel recommendations

For bigeye tuna, the Panel recommends:

- Consider alternative assumptions within the risk analysis for dealing with un-sampled areas when developing the index;
- Use an analytical approach to set the mean level of CV for the index in the assessment (see data-weighting section);
- Consider an effort creep scenario within the risk analysis (e.g., 1% per year); and
- Investigate solutions, relative to the index, to address the decreased spatial coverage of the Japanese longline fishery in the EPO.

For yellowfin tuna, the Panel recommends:

- As for bigeye tuna, consider an effort creep scenario within the risk analysis and use an analytical approach to set the mean level of CV for the index;
- Continue collaboration with industry to identify changes within the dolphin associated purse seine fishery so that the catch-effort relationship can be more appropriately modelled.

5.10. Data weighting

Presentation by staff

- Main data are length composition and indices of abundance.
- Length compositions that represent the index are standardized.
- Length compositions for fisheries may be sparse: standardization may be a better to impute where data is missing.
- Length composition weights are obtained using Francis method.
- Inputs sample size: number of wells sampled (PS) or number of 1,000 measured fish (LL).

Panel comment

The time-varying data weights used in the 2020 bigeye and yellowfin assessments were initially derived from VAST analysis (abundance index CV), Francis weighting of sample sizes (length compositions), and McAllister-Ianelli weighting of the number of otoliths. However, some of the data weights were then adjusted based on subjective choices: the abundance index CV was scaled using an additive constant to have a mean of 0.15 in the period from 1979 to 2014, while longline effective sample sizes were multiplied by 0.5 to account for potential double use.

Panel recommendations

The Panel notes that scaling the abundance index CV to an arbitrary level may not be necessary. Instead, the maximum-likelihood estimate of sigma could be calculated and used, based on the log differences between the observed and predicted abundance index values. The predictions for this purpose can come from an age-structured production model.

Where available, use analytical estimates of uncertainty from spatiotemporal models to determine initial sample sizes for size composition data.

5.11. Model diagnostics

Presentation by staff

- Diagnostics can be used to learn about the models or to improve/reject models (in 2020 Benchmark also used to contribute to the weight the models in the risk analysis).
- Depletion model diagnostic indicates there is information about absolute population size for yellowfin.

Link to presentation

Panel comment

IATTC staff presented their approach for applying diagnostics to stock assessment models. In addition to the key points summarized above, it was also noted that it is currently unclear which diagnostics are useful for evaluating models, whether suitable acceptance thresholds exist for each diagnostic, and whether diagnostics can be used to identify specific mis-specifications within the assessment model. The Panel acknowledged that it is difficult to fully automate the model diagnostic process due to a lack of power in some diagnostics for identifying mis-specified models, and a lack of objective quantitative thresholds for each diagnostic. These are open questions in fisheries science and more research is needed by the broader stock assessment community to resolve these questions. In the interim, it was suggested that some diagnostics that are currently evaluated by IATTC staff by visual inspection could be quantified (e.g., conflict in a likelihood profile could be quantified as the standard deviation of the MLE values across components) in order to facilitate summarizing and evaluating multiple model runs (e.g., in the risk analysis).

IATTC staff also presented an overview of diagnostics applied to the ancestral model, a subset of which included: examination of model residuals, RMSE, empirical selectivity, ASPM (and variants), in-season depletion model, catch-curve analysis, retrospectives, and likelihood profiles. The Panel questioned the sensitivity of the empirical selectivity diagnostic, particularly as it related to the empirical selectivity of the largest individuals under asymptotic selectivity. The Panel suggested further refinement to the presentation of this diagnostic (optionally in combination with simulation testing) as the current presentation could exaggerate misfits for uncommon (large) size classes. The Panel also discussed that as models become more parametrized and complex (e.g., simultaneous estimation of growth, and/or natural mortality; or estimation of time-varying selectivities) it will be important to consider diagnostics related to convergence and model stability such as jittering, self-tests, and parameter characterization (e.g., highly correlated or on bounds).

Panel recommendations

The Panel supported the approach taken, and recommended:

- Where possible, develop quantitative metrics for diagnostics that are visually inspected;
- Investigate the sensitivity of the empirical selectivity diagnostic and alternative ways of presenting this diagnostic; and
- Consider, where computing resources allow, diagnostics related to model stability and convergence.

5.12. Model weighting

Presentation by staff

- Use diagnostics to fix models.
- Only keep good models.
- Use equal weight until a better approach is developed.

Panel comment

IATTC staff presented their framework for stock assessment model development, ultimately resulting in the risk analysis. At the core of the framework was an iterative cycle where diagnostics are used to determine the quality of assessment models and refine as needed. At this stage, alternative hypotheses could be developed to address model misspecifications identified in the stock assessment models. A trade-off was identified between the number of models/hypotheses considered in model development and the ensuing risk analysis. Having fewer models allows for a more hands-on approach to diagnosing and improving the fit of each individual model. When the ensemble set is small, risk in management quantities may be inappropriately characterized; but when the ensemble set is large, less attention can be devoted to any individual model and poorly behaving models may simply be dropped from the analysis rather than improved. Additionally, if the large ensemble set is not constructed carefully, implausible models may end up in the risk analysis. A large model set requires more computational resources and automated methods for triaging models and results.

The Panel supports the hierarchical approach used by the IATTC staff when developing their risk analysis. However, the Panel also discussed alternative approaches, such as the Monte Carlo Bootstrap or fractional factorial design, that could be combined with the hierarchical approach to more appropriately reflect the number of true uncertainties considered in the risk analysis while keeping the number of total runs manageable.

The Panel discussed the ‘equal weighting of reasonable models’ approach proposed by IATTC staff to combine models in the risk analysis and found it to be a sound approach. The Panel agreed that selection of which models to be included in the risk analysis needs to be carefully considered in order to not over-represent certain scenarios or include implausible models. The Panel concurred with IATTC that uncertainty in data inputs can be best accommodated using multiple scenarios (either fitting to alternative CPUE indices or applying a Monte Carlo Bootstrap approach with different catch time series). Lastly, the Panel agreed that model-weighting be agreed to a-priori and documented transparently.

Panel recommendations

The Panel recommends that:

- IATTC staff gain access to high-throughput/high-performance computing to address the increased computational demands of the risk analysis framework including automated calculation of diagnostics as a part of model execution.

- Where appropriate, alternative approaches such as the Monte Carlo Bootstrap or fractional factorial design be combined with the hierarchical approach to efficiently account for known uncertainties.
- Encourage further development of the ensemble approach with respect to increasing the number of scenarios investigated in order to capture the full range of uncertainty.
- Support inclusion of models in the risk analysis that have resolved or mitigated data conflicts.

5.13. Estimation of management quantities and associated uncertainty

Presentation by staff

- Biomass limit reference point is based on equilibrium B0.
- Biomass target reference point is based on dynamic B0.
- Fishing mortality reference points are evaluated using the average estimated F for the last 3 years.
- Uncertainty is approximated by a normal distribution around the estimates.
- Variability taken from the estimated variability in each model, or proxy.

Link to presentation

Panel comment

The uncertainty about management quantities is calculated using the delta method and symmetric normal approximation. IATTC staff presented an analysis that uses MCMC to evaluate the reliability of the normal approximation for calculating the probability of the stock being above or below a reference point. The MCMC approach can be expected to be more reliable than the normal approximation, as MCMC does not enforce the probability distribution to have any predefined shape. The example from the analysis showed that the normal approximation could lead to skewed probability calculations that could affect the management advice. On the other hand, MCMC takes considerable computation time, especially when running long MCMC chains on a single computer.

Panel recommendations

The Panel recommends that three approaches be considered to save time when running MCMC for the uncertainty about management quantities: distributed computing, several short chains using different random seeds (-mcseed), and the No U-Turn Sampling (-nuts) algorithm. Using MCMC rather than the delta method is especially relevant when the fishery status is estimated close to a given reference point.

However, since the main source of uncertainty in management advice is scenarios in the model ensemble, improving the ensemble is the main focus of research, and assuming a normal distribution is a reasonable approach for the parameter uncertainty.

6. LIST OF REQUESTS

Request #	Topics	Request	Rationale	Response
1	11	Plot the distribution of the dolphin associated purse seine fishery, and how this distribution may be impacted by the two temporal closures (particularly the Dec-Jan closure).	Since the dolphin associated purse seine index is the main index for the yellowfin tuna stock assessment it is important to understand if the temporal closures introduce any systemic bias into the index.	Preliminary investigations were conducted, however IATTC staff indicated that this issue would be more fully explored in a document on the YFT PS index which will be prepared for the SAC.
2	11	Investigate the fleet composition of the dolphin associated purse seine fishery to identify if the characteristics of the fishery change during the closure period.	Since the dolphin associated purse seine index is the main index for the yellowfin tuna stock assessment it is important to understand if the temporal closures introduce any systemic bias into the index.	Preliminary investigations were conducted, however IATTC staff indicated that this issue would be more fully explored in a document on the YFT PS index which will be prepared for the SAC.
3	2	Plot the recruitment from the stepwise progression of bigeye tuna model M0 to M6, and calculate the regime shift diagnostic for each model.	The recruitment regime shift was interpreted as a model misspecification. Viewing how this changed in the stepwise model progression can help improve understanding of what model configurations helped to alleviate this issue.	Several factors contributed to alleviating the recruitment regime. However, adding a time block selectivity to the longline fishery and imposing asymptotic selectivity on the last time block appeared to have the greatest impact. Doing so appears to impact estimates of model scale. By increasing the model scale, a recruitment regime shift is no longer as necessary to explain the rapid increase in catches from the floating object fishery in the mid-1990s.
4	2	Explore the feasibility of developing a spawning potential ogive for bigeye tuna.	The yellowfin tuna assessment assumes a spawning potential ogive as a function of female maturity, spawning fraction, and fecundity. It is	Requires biological research.

			recommended that a similar ogive be constructed for bigeye in order to be consistent between the two species, and also since this can have an effect on management reference points.	
5	8	Run models with alternative sigmaR (double the current value) and calculate the variability in estimated recruitments for the period where the bias-correction ramp is fully applied.	The recruitment penalty was shown to have a large influence on the likelihood profiles. It is important to understand the sensitivity of the model to the assumed value for this parameter.	Model results appeared to be insensitive to the choice of sigmaR when growth was fixed.
6	8	Conduct an R0 profile with the recruitment deviates held fixed at the estimated values from the full model run.	This is an alternative formulation of the R0 profile to investigate if the influence of the recruitment penalty on R0 is an artefact of how the profile is conducted.	Completed following the close of the in-person review. Fixing the rec-devs removed the impact of the recruitment penalty on the estimate of R0. However, doing so removed information about data conflict between the different likelihood components. More research (outside the scope of the review) is needed into this topic.
7	5	Plot the tails of the length frequency by fisheries in the yellowfin tuna assessment model over time and with respect to the L2 value(s).	This will help identify which fisheries catch the largest yellowfin tuna and which is the most appropriate choice for an asymptotic selectivity curve.	While some spatiotemporal strata indicate that PS and longline fisheries have similar length frequency of the largest individuals, this is not constant across space and time. Some strata show larger individuals caught in the longline while others show larger individuals are caught in the PS.
8	5	Plot the length frequency distribution of yellowfin tuna for the longline and dolphin associated purse seine operating in similar areas.	Identify if both fisheries can be assumed to have similar selectivities.	Preliminary figures suggested that selectivity was similar in some areas. However, preliminary

				results also suggested that there could be spatiotemporal patterns in which fishery caught the largest fish. Further investigation is needed but sensitivity to the choice of which fishery is assumed to have asymptotic selectivity may be warranted.
9	11	Run a model with effort creep of 1% for the longline survey.	Indices of abundance are affected by unmodelled catchability change. Effort creep scenarios should be considered in the model ensemble. Need to understand how it affects the model.	Inclusion of 1% effort creep affected the bigeye abundance trend but did not substantially change the stock status. Effort creep affected the yellowfin model by increasing early recruitment and spawning biomass, and reducing early fishing mortality.
10	9	Fit spatial GAM model to length at age data from bigeye otolith daily ages.	Could provide evidence about spatial patterns of growth variation for bigeye.	The preliminary fitted model supported the existence of spatial growth variation, but the model was unstable, with too many parameters. More work will be needed to refine the analysis.
11	11	Split the index into 4 seasonal index fisheries, with the following runs: a. share selectivity and catchability across all four quarters (i.e., equivalent to the unsplit fishery) b. estimate independent q in each fishery (maybe don't bother with this one) c. estimate independent q and selectivity in each fishery.	Index residual means vary by season	Preliminary results showed that splitting the index into quarters had no impact on population scale, fishing mortality estimates, or the temporal pattern in recruitment.

12	9	Estimate L2 inside the model as before, except with R sigma x 2.	Likelihood profile on L2 shows that R sigma is affecting the estimate. R sigma is a modelling convenience parameter that should not be informative about L2.	Increasing R sigma results in slightly smaller L2, and significantly higher biomass.
13	5	Explore what is causing the odd structure in the YFT selectivity ogives. This needs a lot of plots and data exploration. We suggested spatial plots of the residuals by time step, but more is probably needed.	Selectivity ogives have multiple modes. If they can they identify any factors associated with the different sizes, they can be split into different fisheries.	Putting this aside until after the meeting because it requires a lot of exploration.
14	9	Set up a model for each species that is robust to spatially varying growth. To do this, use the normal approach for size data from the core area that provides data for the growth curve, and downweight size data for fleets that catch large fish and are in areas where growth may be different (i.e., outside the core area). For the index fisheries, predict size data only for the core area.	EPO growth curves for both bigeye and yellowfin indicate much larger asymptotic lengths than WCPO growth curves. It is therefore very likely that there is growth variation within the EPO. Pacific stock assessments, which rely heavily on size data, are highly sensitive to asymptotic length.	Putting this aside until after the meeting. The proposed approach was seen as promising. However, it requires development and is not feasible in the time available.
15	10	Run the YFT assessment with alternative ogives for natural mortality.	There is uncertainty about natural mortality ogives, including the base level, and whether the cause of the change in sex ratio at length is caused sex differences in growth, natural mortality, or a combination.	Results are shown in figure R15. Spawning output through time was quite similar with the Lorenzen approach. Depletion was greater than the Base with the Lorenzen and greater again with the two-stage approach.
16	8	Compare BET model runs with three different values of sigmaR (0.6, 1.2, 2.0) and three different growth options (fixed, estimated, estimate all growth parameters except L2).	Understand why the stock size is highly sensitive to changes in sigmaR when estimating growth. Understand whether this is caused only by minor changes in L2. Examine appropriate model options related to growth and recruitment.	The estimated stock size is highly sensitive to the value of L2 and not sensitive to the value of sigmaR. The calculation of relative stock status can be sensitive to sigmaR, depending on how it is calculated.

17	6	Increase the maximum age for bigeye from 40 to 60.	The maximum age can be influential if there are still fish alive and the growth curve is not flat.	This change had no impact on bigeye model estimates.

7. FIGURES

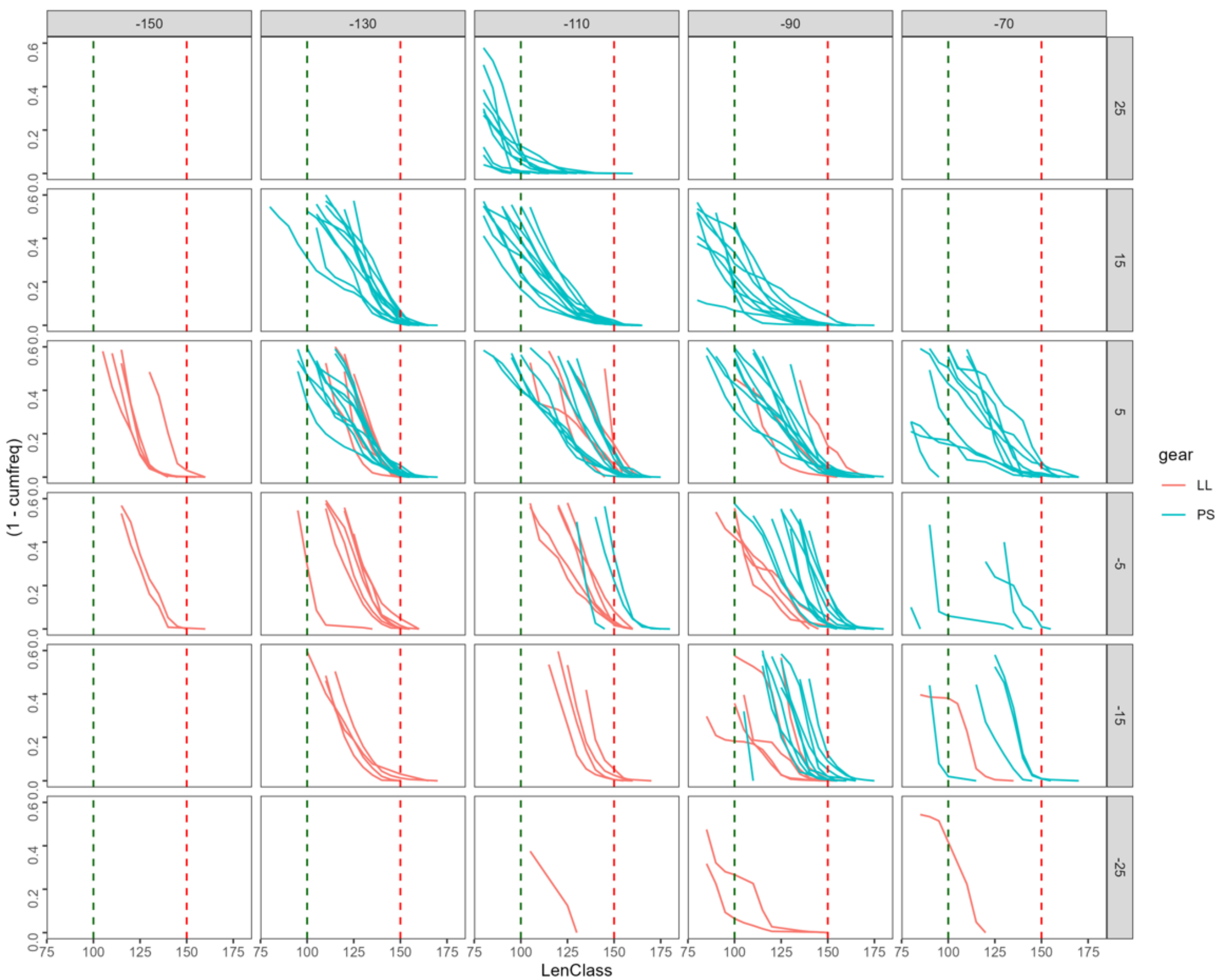


FIGURE R7.1. Upper tails (30%) of logbook length composition data for the years 1975-1985, caught by Japanese longliners, and purse seine lengths from port sampling.

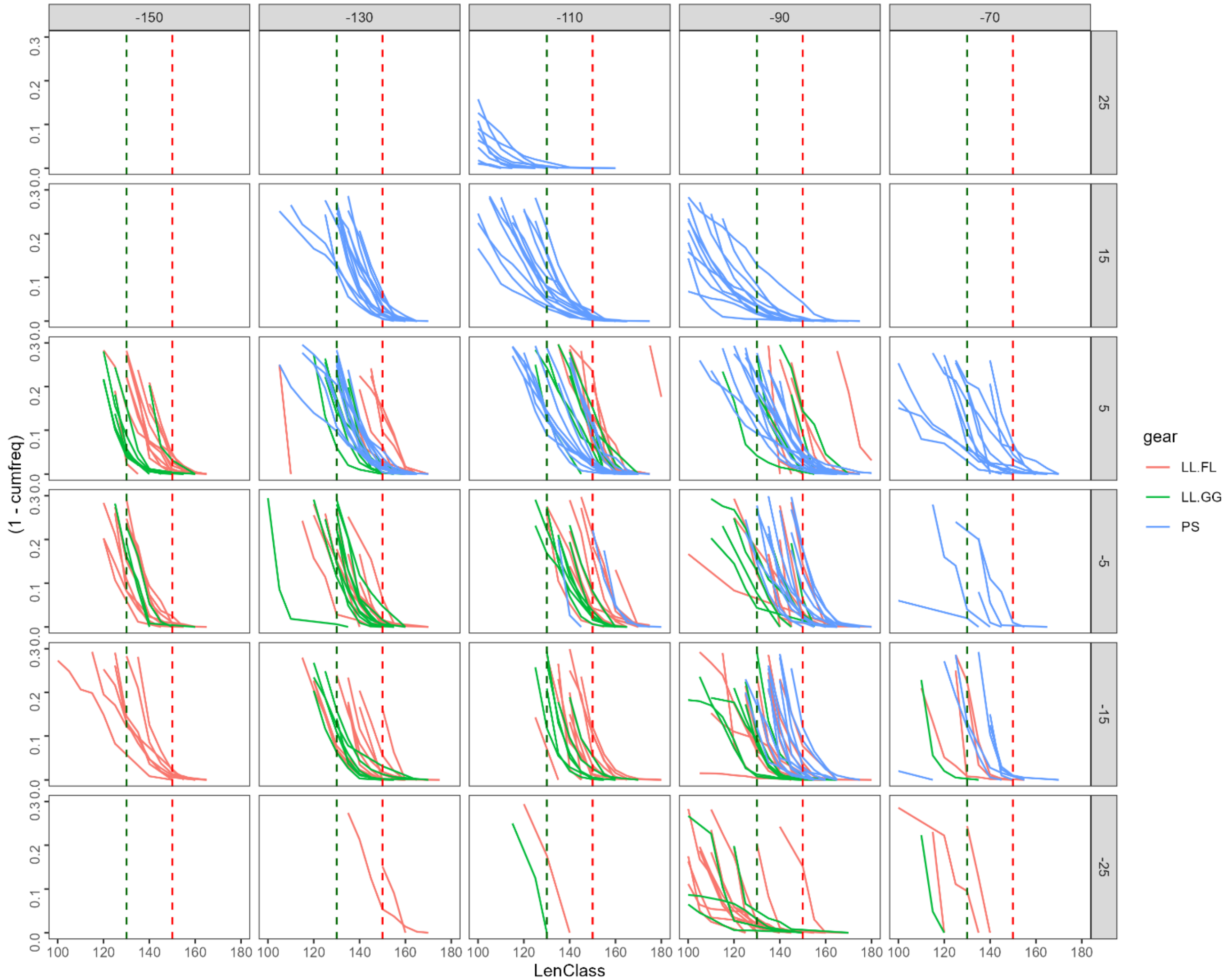


FIGURE R7.2. Upper tails (30%) of size composition data for the years 1975-1987, caught by Japanese longliners (red are logbook lengths and green are logbook weights converted to lengths), and purse seine lengths from port sampling (blue).

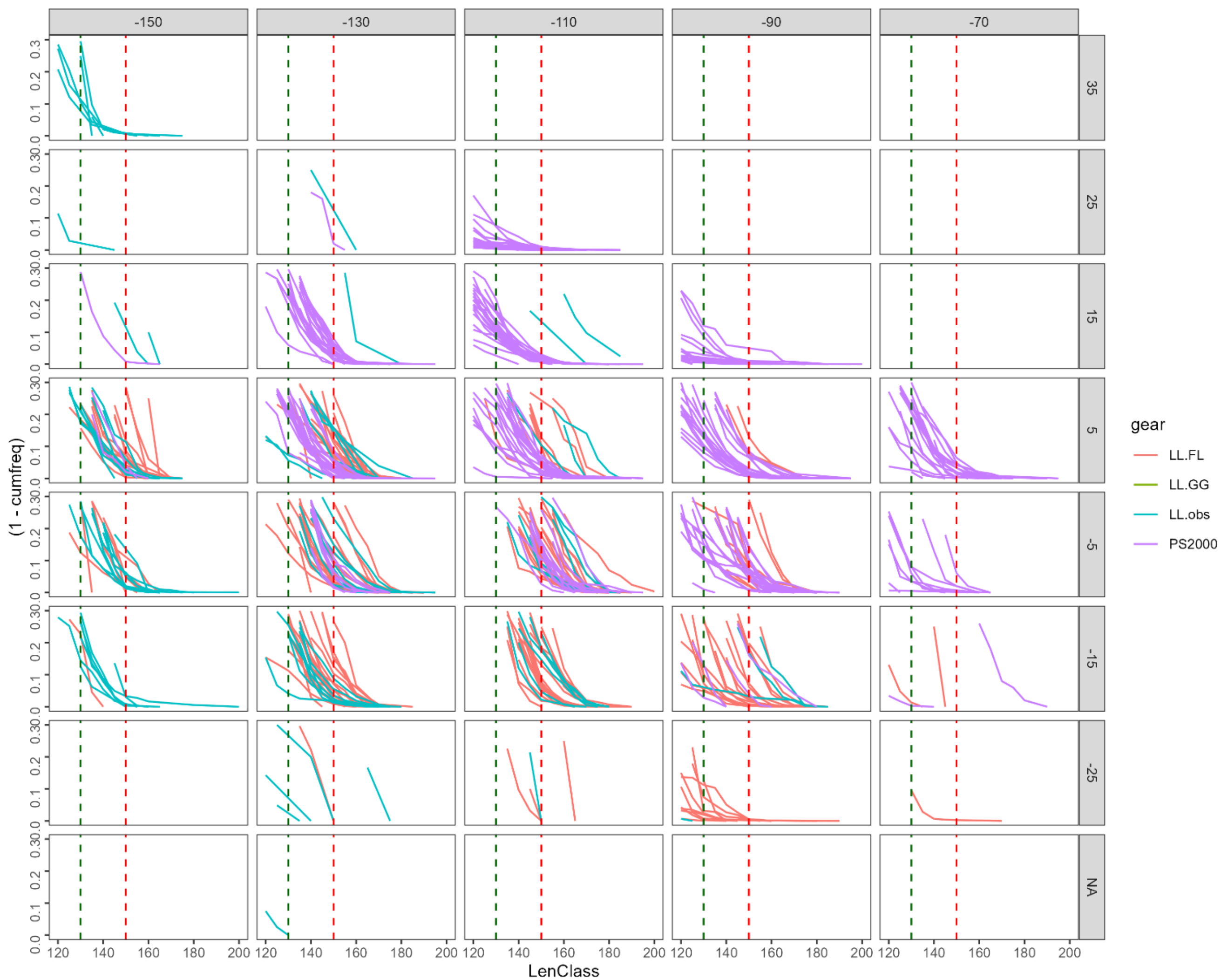


FIGURE R7.3. Upper tails (30%) of size composition data for the years 2001-2021, caught by longliners (red are Japanese logbook lengths and turquoise are Japanese, Korean and Taiwanese observer lengths), and purse seine lengths from port sampling (purple).

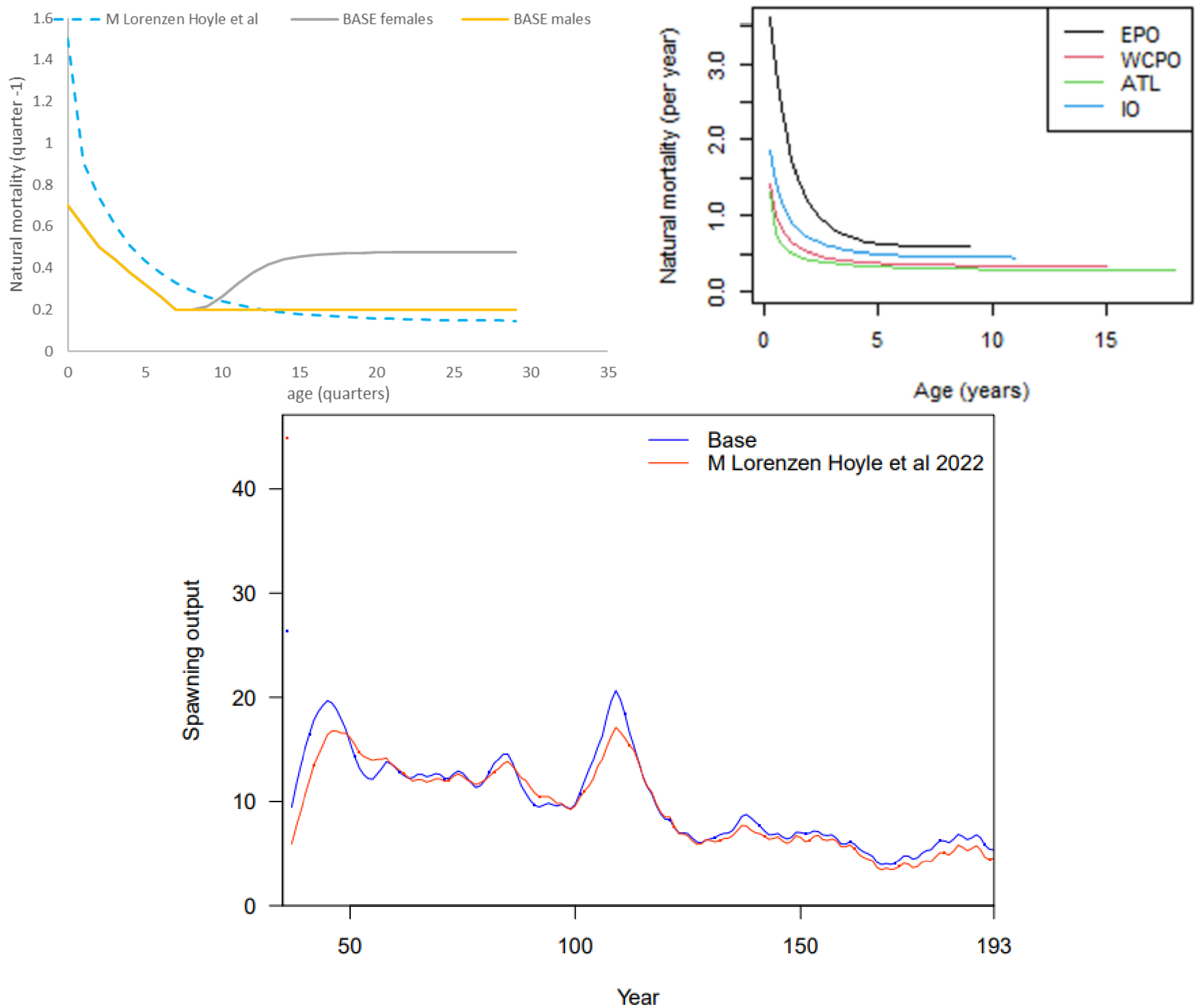


FIGURE R15.1: Comparison of spawning output between the Base assumptions about natural mortality ('Base') and a natural morality ogive defined by the Lorenzen curve and with mean adult natural mortality determined by the oldest known fish in the EPO ('M Lorenzen Hoyle et al 2022').

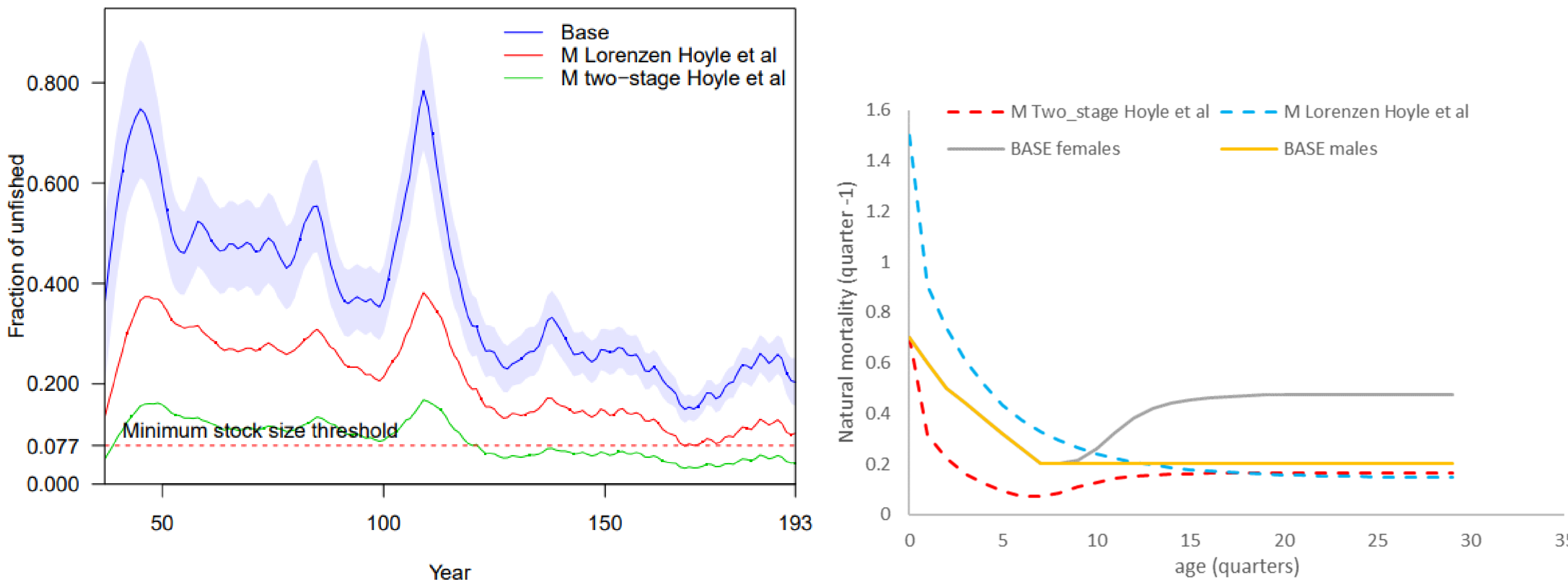


FIGURE R15.2. Depletion of spawning output (left) under the Base model, the Lorenzen model, and the two-stage model. Ogives associated with each model are show in the figure on the right.

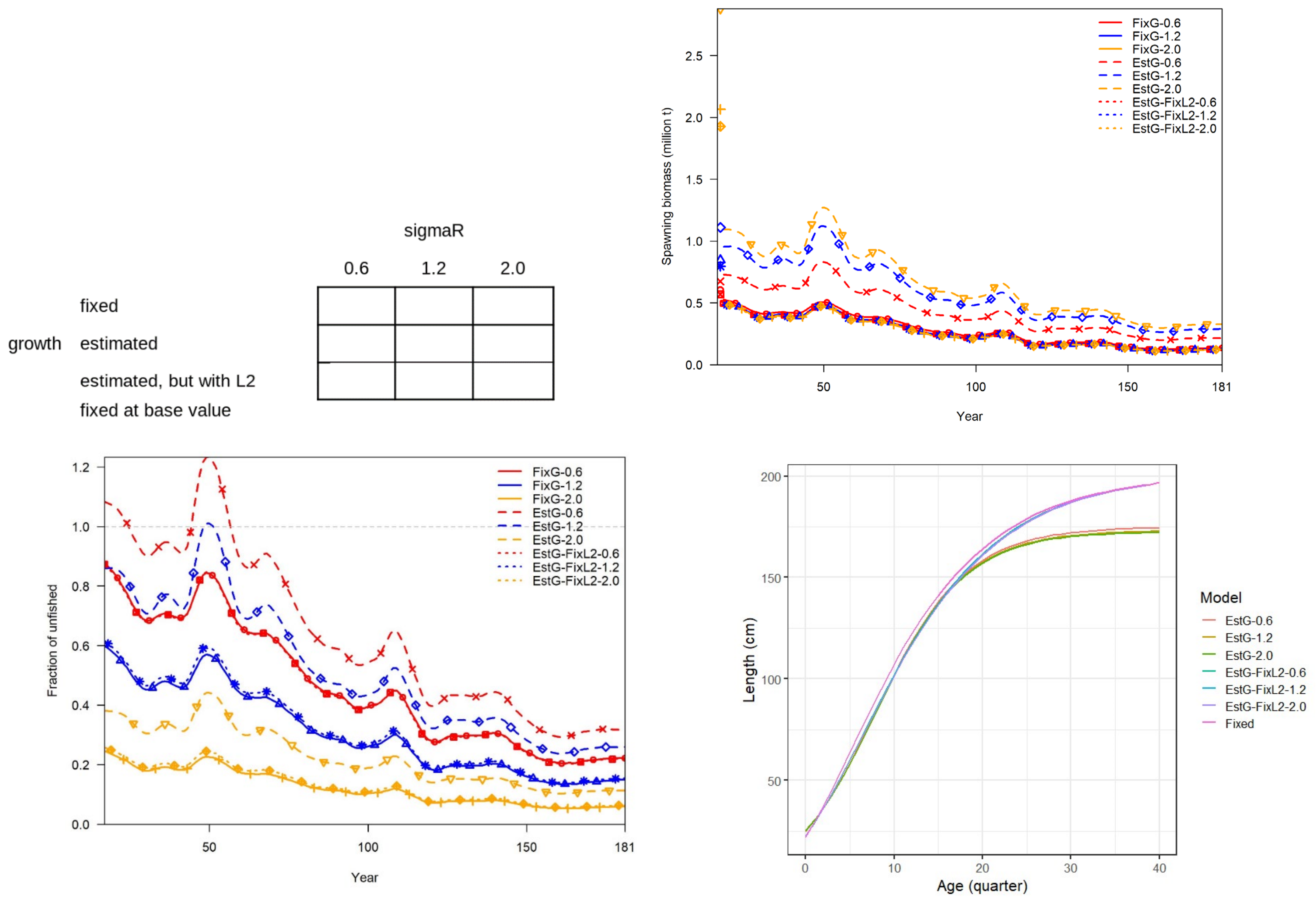


FIGURE R16.1. Top-left: Diagram describing the nine models run for Request #16. Top-right: Spawning biomass estimated by nine models from Request #16. Bottom-left: Fraction of unfished biomass estimated by nine models from Request #16. Bottom-right: Growth curves in the nine models from Request #16.

ANNEX 1. TERMS OF REFERENCE FOR THE PANEL REVIEW

INTER-AMERICAN TROPICAL TUNA COMMISSION

EXTERNAL REVIEW OF MODELLING ASPECTS FOR STOCK ASSESSMENTS OF TROPICAL TUNA IN THE EASTERN PACIFIC OCEAN

La Jolla, California (USA)
6-10 November, 2023
(by videoconference)

TERMS OF REFERENCE

1. GOALS AND OBJECTIVES

The purpose of the external review of modeling aspects for stock assessments of tropical tuna in the eastern Pacific Ocean is to provide information that will improve the stock assessments and consequently the management advice.

To this end, the goals and objectives of this review are to:

a. identify the best available science for the EPO tropical tuna stock assessments;

b. provide an independent review of the stock assessment approach;

c. provide advice on future research and data collection that will improve the assessments and the provision of management advice.

2. REVIEW PANEL RESPONSIBILITIES

The main responsibility of the Review Panel is to perform an adequate technical review of the stock assessment approach. To ensure the greatest objectivity in this exercise, the members of the Panel should disclose any conflicts of interest they may have, including, but not limited to, personal financial interests and investments, employer affiliations, and consulting arrangements, grants, or contracts, etc.

The specific responsibilities of the Panel are to:

a. be familiar with the Terms of Reference;

b. review background documents, data, and analytical methods, along with other pertinent information (*e.g.*, previous documents and Review Panel reports);

c. discuss the technical merits and deficiencies of the stock assessment approach, work with the IATTC staff to correct deficiencies, and, when possible, suggest new tools, analyses, and data collection methods to improve future assessments; and

d. draft a report of the meeting, to document the discussions and recommendations.

It is the Panel Chair’s responsibility to coordinate the discussions so that the review is completed in the time available.

3. PUBLIC COMMENT

During the meeting, time will be set aside for public comment. The Panel will take these comments into consideration, as appropriate, when developing its report.

4. REQUESTS FOR ADDITIONAL ANALYSES

Since the purpose of the meeting is to conduct a technical review of the stock assessment approach, it may be beneficial to conduct analyses during the course of the meeting. The Panel may also request additional details on the data and methods presented, further analyses of alternative assumptions, or additional model runs. However, any such requests must be clear, explicit, and be presented in writing, and be practical in terms of the time available. Such requests should be listed individually in the Panel's report, along with their rationale and the response. To the extent possible, analyses requested by the Panel should be completed during the meeting by the assessment team.

5. PANEL REPORT

The Panel's report should be drafted and approved shortly after the meeting. The report writing process will follow these steps:

a. Panel outlines report at meeting;

b. Panel writes and agrees draft report;

c. Panel provides draft report to IATTC staff for comment on technical accuracy; and

d. Panel reviews staff comments, and modifies report as necessary.

The report will include:

a. Names and affiliations of Panel members;

b. Brief overview of the meeting (location, agenda, main recommendations by Panel, *etc*.);

c. Brief summary of current methods used in the assessment;

d. List of analyses requested by the Panel, rationale for each request, and brief summary of the response;

e. Comments on technical merits and/or deficiencies in the assessment approach and recommendations for remedies;

f. Unresolved problems and major uncertainties, *e.g*., any special issues that complicate the use of the data or the analyses conducted;

g. Data, fishery or analysis related issues raised by the public; and

h. Prioritized recommendations for research and data collection for the subsequent assessments.

The Panel and the IATTC staff will strive to resolve any differences of opinion that may arise regarding the contents of the report. Any unresolved differences of opinion must be documented and reflected in the report, which will be published as an IATTC Special Report.

Appendix: Topics covered

Model structure: Time step, number of age classes, sex structure, start year, initial conditions
Spatial structure (stock structure, areas as fleets – fishery definitions)

Growth
Natural mortality
Selectivity
Recruitment
Data weighting

Diagnostics:
Risk assessment (Model weighting)

ANNEX 2. TENTATIVE AGENDA OF THE PANEL REVIEW

INTER-AMERICAN TROPICAL TUNA COMMISSION

EXTERNAL REVIEW OF MODELLING ASPECTS FOR STOCK ASSESSMENTS OF TROPICAL TUNA IN THE EASTERN PACIFIC OCEAN

La Jolla, California (USA)
6-10 November 2023
(*participation by teleconferencing optional*)

Monday 06-Nov

0900:	Welcome
0905:	Introductions, notifications
0915:	Background (terms of reference, documents, objectives, products)
0945:	Panel discussion on review format and approach
1000:	Summary of the management of tropical tunas
1100:	1. Overview of the assessments for the three species of tropical tunas
1130:	1. Overview of the IATTC risk analysis approach
1200:	Lunch
1300:	2. Conceptual model for bigeye tuna and hypotheses for risk analysis
1400:	Discussion
1430:	3. Conceptual model and spatial structure for yellowfin tuna and hypotheses for risk analysis
1530:	Discussion and further reflection on review format
1600:	Data review recommendations
1630:	Discussion
1700:	Day’s summary by Chair, public comment, discussion, and Panel requests for additional analyses

Tuesday 07-Nov

0900:	Introduction to the bigeye tuna assessment “base/ancestral” model
0930:	Discussion
1000:	Introduction to the yellowfin tuna assessment “base/ancestral” model
1100:	Discussion
1130:	6 & 7 Model dimensioning assumptions: time step of the model, number of age classes, initial conditions and start year
1200:	Lunch
1300:	8. Recruitment
1330:	Discussion
1400:	9. Growth
1430:	Discussion
1530:	10. Natural mortality
1600:	Discussion
1630:	Day’s summary by Chair, public comment, discussion, and Panel requests for additional analyses

Wednesday 08-Nov

0900:	Presentation of requested data, analyses, and model runs

0930: 4. & 5. Fisheries definitions, length compositions, and fishery selectivity assumptions
1000: Discussion
1100: 11. Indices of abundance, its catchability and selectivity
1130: Discussion
1200: Lunch
1300: 12. Overview of data and data weighting
1330: Discussion
1400: 13. Model diagnostics
1430: Discussion
1530: Report writing allocations
1600: Presentation of requested data, analyses, and model runs
1630: Day's summary by Chair, public comment, discussion, and Panel requests for additional analyses

Thursday 09-Nov
0900: Presentation of requested data, analyses, and model runs
0930: 14. Model weighting and multimodel estimates
1000: Discussion
1100: 15. Estimation of management quantities and associated uncertainty
1200: Lunch
1300: Chair summary of state of play
1330: Report writing
1600: Presentation of requested data, analyses, and model runs
1630: Day's summary by Chair, public comment, discussion, and Panel requests for additional analyses

Friday 10-Nov
0900: Presentation of requested data, analyses, and model runs
1000: Reviewers' recommendations
1100: Report writing
1200: Lunch
1300: Report writing
1400: Summary by Chair of reviewers' recommendations, public comment
1500: Adjournment

ANNEX 3. PARTICIPANTS IN THE REVIEW

Name	Organization	Email
Arni Magnusson	Panelist	arnim@spc.int
Mark Dickey-Collas	Panelist (Chair)	Mark.dickey-collas@ices.dk
Nicholas Ducharme-Barth	Panelist	nicholas.ducharme-barth@noaa.gov
Simon Hoyle	Panelist	simon.hoyle@gmail.com
Gorka Merino	Azti	gmerino@azti.es
Thaiza Barreto	BioImpact Lab	barreto.thaiza@gmail.com
Martha Betancourt	Fidemar	martha.betancourt@uabc.edu.mx
Michel Dreyfus	Fidemar	dreyfus@cicese.mx
Manuel Correia	IATTC Bycatch WG Chairman	manuelcorreia.a@gmail.com
Alex Da Silva	Inter-American Tropical Tuna Commission	alexdasilva@iattc.org
Barbara Cullingford	Inter-American Tropical Tuna Commission	bcullingford@iattc.org
Carolina Minte-Vera	Inter-American Tropical Tuna Commission	cminte@iattc.org
Cleridy Lennert	Inter-American Tropical Tuna Commission	clennert@iattc.org
Dan Fuller	Inter-American Tropical Tuna Commission	dfuller@iattc.org
Haikun Xu	Inter-American Tropical Tuna Commission	hkxu@iattc.org
Jeff Morgan	Inter-American Tropical Tuna Commission	jmorgan@iattc.org
Jon Lopez	Inter-American Tropical Tuna Commission	jlopez@iattc.org
Juan Valero	Inter-American Tropical Tuna Commission	jvalero@iattc.org
Marisol Aguilar	Inter-American Tropical Tuna Commission	maguilar@iattc.org
Mark Maunder	Inter-American Tropical Tuna Commission	mmaunder@iattc.org
Monica Galvan	Inter-American Tropical Tuna Commission	mgalvan@iattc.org
Robert Sarazen	Inter-American Tropical Tuna Commission	rsarazen@iattc.org
Fukuda Hiromu	Japan Fisheries Research and Education Agency	fukuda_hiromu57@fra.go.jp
Hirotaka Ijima	Japan Fisheries Research and Education Agency	ijima_hirotaka69@fra.go.jp
Makoto Nishimoto	Japan Fisheries Research and Education Agency	m.nishimoto103@gmail.com
José Vélez	Ministerio de Producción, Comercio Exterior,Inversiones y Pesca	jvelezt@produccion.gob.ec
Qinqin Lin	Shangai Ocean University	qqlin@shou.edu.cn
Diego Alves	State University of Maringa	dcalves@uem.br
Daisuke Goto	Swedish University of Agricultural Sciences	daisuke.goto2@gmail.com
Toshihide Kitakado	Tokyo University of Marine Science and Technology	kitakado@kaiyodai.ac.jp
Juan Quiroz	Tunacons	jcquiroz@facilevisual.com
Pedro Santistevan	Tunacons	psantistevan@tunacons.org
Rujia Bi	UW-Madison	rbi24@wisc.edu