INTER-AMERICAN TROPICAL TUNA COMMISSION

1ST EXTERNAL REVIEW OF DATA USED OF STOCK ASSESSMENTS OF TROPICAL TUNA IN THE EASTERN PACIFIC OCEAN

La Jolla, California (USA)

02-06 October 2023
(by videoconference)

REPORT

CONTENTS

1. EXECUTIVE SUMMARY

IATTC called for the first time for an external data review panel to screen the data available for stock assessment. The objective of the external review of data used in stock assessments of tropical tuna in the eastern Pacific Ocean was to provide information that will improve the inputs into the stock assessments and consequently the assessments and management advice.

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

a. identify the best available science for use of data in the EPO tropical tuna stock assessments and associated analyses (e.g. what data to use, how to analyze it, standardize it, or raise it to the totals);

b. provide an independent review of the data used in the stock assessments;

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

The panel met with IATTC staff between 2 and 6 of October 2023 in La Jolla, US, IATTC headquarters. The panel was comprised of Simon Hoyle (Hoyle Consulting Ltd, simon.hoyle@gmail.com), Tom Peatman (Shearwater Analytics Ltd, tom.peatman@gmail.com), Cody Szuwalski (NOAA, cody.szuwalski@noaa.gov) and Ernesto Jardim (MSC, ernesto.jardim@msc.org, chair).

The panel would like to commend IATTC's staff for the support provided to the meeting, both in preparing analyses and engaging in fruitful discussions about the topics.

The panel is of the opinion that the analyses provided are in general appropriate and should be considered by the stock assessment panel. Notwithstanding this, the panel is of the opinion that taking into account the recommendations below will improve the information.

1.1. Global recommendations

- Any estimate that is provided for stock assessment needs to include its uncertainty.
- Estimates of uncertainty need to be clarified in some of the models used.
- Where model estimates are used as input data for other models, it is important to ensure that uncertainty is propagated across the analysis to avoid false precision in the final estimates.
- Sampling programs and biological parameters need to be updated.
- Model-based approaches to deal with missing information are preferred over design-based approaches, so that spatial temporal effects are appropriately taken into account. Wherever possible, include all data and allow the model to deal with uncertainty due to low effective sample sizes.
- Where important information is missing, model outcomes should, if possible, be externally validated.
- Develop a standard set of plots that describe the available data in detail, including factors such as spatial and temporal distributions of effort and size data. These can be updated annually using automated procedures.
- Provide a series of figures that show the stages of development from raw data to stock assessment inputs.

2. CONCEPTUAL MODEL FOR YELLOWFIN TUNA

2.1. Background

Spatial structure has been thought to be one of the main issues in the assessment of yellowfin tuna in the eastern Pacific Ocean (EPO). There is some evidence of at least two groups of fish with different dynamics in the EPO, roughly occupying areas to the north and the south of the equator, but not only are the divisions not clear, mixing may occur with variable magnitude and periodicity. The 2020 benchmark assessment considered that spatial structure was the main uncertainty, and a conceptual model for the stock which encompassed a set of overarching hypotheses to address this issue was proposed (SAC-11-INF-J). The two fish groups may have "High mixing", "Episodic/high variability mixing", or "Negligible mixing". The "High mixing" overarching hypothesis may be represented by single-stock models for the whole EPO. The "Episodic/high variability mixing" overarching hypothesis may be represented by single-stock models that are driven by data from each of the groups. This means that the model is fitted to data for one group while the selectivity for the other group is fixed. The “Negligible mixing” hypothesis may be represented by two independent assessments, one for each group. The 2020 assessment (SAC-11-07) focused on modeling a hybrid “High mixing” and “Episodic/High variability mixing” by fitting the assessment models to an index derived from the standardized CPUE of the purse-seine fishery with sets on dolphins (roughly representing the “north” area and the majority of the catches), and not fitting to the longline index (roughly representing the “south” area) or longline length composition data.

In 2023, the conceptual model for yellowfin tuna in the eastern Pacific Ocean (EPO) was refined based on a review of all available information, including information relevant to stock structure. The main component of the updated conceptual model is the idea that there are at least two groups of yellowfin tuna in the EPO associated with different biogeochemical provinces, the epipelagic and mesopelagic. This idea is supported by a suite of information. The distribution of these two groups varies seasonally and interannually following the expansion and contraction of the biogeochemical provinces. The main challenges that this pattern poses to stock assessment is to determine where to place the boundary between stocks and how to estimate mixing rates. The dynamic shape of the biogeochemical provinces can be summarized using oceanographic variables, which allows a path forward to delimit the preferred habitat of each stock.

A new methodology was proposed to split the groups dynamically based on biogeochemical characteristics of the ocean as explanatory variables and length frequency data as a multivariate response variable. Several biogeochemical variables for each purse-seine set from 2000-2017 were condensed using principal component analysis. The first component (PC1) summarized the vertical structuring of the water column, while the second component (PC2) mainly represented the sea surface temperature (SST). Tree analyses were used to split the length composition of purse-seine fleet, where the two principal components (a proxy for the location within the environmental gradients) and seasons (quarters and cyclical quarters) were used as explanatory variables. The length frequency data from the purse-seine sets on floating objects was used because it has the largest proportion of smaller and younger fish, which are more likely to inhabit the shallower parts of the water column, thus may be a better indicator of fish originating in each location. The tree analysis showed the first split on the PC2, separating the areas with lower SST (areas of influence of the Humboldt and California currents) from the warmer tropical areas. Independent tree analyses for the warm and the cold areas both supported splits along similar locations of the PC1 axis, into two distinct areas, an area with lower sea surface height, shallow thermocline, and shallow upper layer of the mesopelagic zone roughly located in the northeastern (NE) region of the EPO, and an area with higher sea surface height, deeper thermocline, and deeper upper layer of the mesopelagic zones, roughly located in the southwestern (SW) area of the EPO. The locations of these areas vary seasonally and interannually. The catches were split between the two groups according to the membership defined by the tree analysis. The NE group encompasses almost all the purse-seine catches in weight taken in dolphin sets, 96% on unassociated sets and 83% of floating objects, while only 17% of the longline catch in numbers (average for 1995-2017). The SW group, in contrast, encompasses 79% of the longline catches and 17% of the floating object catches, with an increase in that proportion in recent years.

2.2. Comments

The information and analysis provided show some support for the hypothesis of two sub-populations, one in the North-East and one in the South-West, associated with biogeochemical provinces. Genomics, data about larvae presence, length frequencies of the catches, PCA analysis on mean length, and tagging data all partially support the hypothesis. However, all these analyses have deficiencies and other hypotheses, including a single stock hypothesis with spatial structure following a north/south divide, cannot be discarded.

An alternative view is that yellowfin tuna are very broadly distributed across the world's oceans, and it would be surprising if there were separate substocks within the EPO adapted to, or at least specializing in, different biogeochemical provinces. Instead, genetic subgroups, spatial size variation, and limited tag displacements may all be caused by high viscosity of the population, i.e., movements of adult fish that are on average more limited than has traditionally been assumed. Could the observed covariation of size structure and oceanography / biogeochemical provinces be because tuna adapt their behavior / foraging to the conditions, and which affects their catchability at size?

Due to the considerable impacts of El Nino/La Nina events in the environmental conditions in the EPO, which change the areas where the sub-populations can exist, the spatial domain of the sub-populations can vary over time. When one spatial subdomain was smaller it would be expected to compress the subpopulation and increase its density, resulting in higher CPUE, with inverse CPUE patterns in the other subpopulation, and both correlated with the SOI (positively & negatively). These relationships are not apparent in the dolphin and longline CPUE indices, which conflicts with the proposed hypothesis of two subpopulations divided at the fluid boundary of two biogeochemical provinces. This should be further investigated.

Developing a dynamic spatial domain stock assessment is not trivial. For example, since the index of abundance available has a fixed spatial domain, the relationship between the survey area and the sub-population spatial domain is not constant, and the representativeness of the index becomes unknown. The same applies for any biological parameter collected/modeled under the assumption of a fixed domain.

2.3. Recommendations

- Do not conclude these groups are two different stocks, most likely there is spatial structure within the same stock with unknown levels of mixing.
- Consider dealing with the 2 sub-populations within a single model, which allows analysts to share parameters.
- If two stock assessments are performed, one for each subpopulation, need to check that the scale of the combined stock size is not very different from a single stock assessment outcome.

3. BIGEYE TUNA STOCK HYPOTHESES

3.1. Background

Following the recommendation provided by the review panel of the last benchmark assessment for bigeye tuna, the current stock assessment of bigeye tuna in the EPO uses the "areas-as-fleets" approach. This approach implicitly assumes that the bigeye population is well-mixed within the EPO. Genetic studies so far have found no evidence against panmixia across the Pacific Ocean. However, the main reason for using the "areas-as-fleets" approach is that there is not enough tagging data to support a spatially explicit stock assessment model, rather than because there is no spatial pattern found in the population of bigeye tuna in the EPO. Both archival and conventional tagging studies suggest limited mixing of bigeye tuna between the far east and far west of the Pacific Ocean, and significant mixing in the central Pacific with more bigeye tuna moving towards the eastern Pacific than the western Pacific.

In addition to tagging studies, other studies also support three main bigeye populations in the Pacific Ocean. Japanese larval samples in the Pacific Ocean suggest that there are three tropical spawning grounds for bigeye tuna in the tropical Pacific: eastern, central, and western Pacific. Biological studies conducted for bigeye in the Pacific Ocean show notably different growth curves and maturity curves between the eastern and western Pacific Ocean. In summary, there are likely three bigeye populations in the Pacific Ocean separately by distance with the central Pacific population residing on both sides of the management boundary (150W). However, a spatially structured stock assessment model for bigeye in the EPO is not practical due mainly to spatially varying biological parameters and a lack of tagging data for adult bigeye.

3.2. Comments

Differences in the depletion rates between areas in the EPO for bigeye tuna could be reason to consider finer population structure, however issues in the construction of the indices of abundance that show this differential depletion may also provide rationale for maintaining a single stock definition in the EPO. Until definitive evidence for stock separation has been identified, the EPO bigeye population should be considered a single stock. Although stock separation is plausible, it currently appears to be impractical to develop useful models based on this scenario.

It is not necessarily appropriate to define, say, three stocks. For a start – what is the working definition of a stock? Clearly there is spatial structure (size at maturity, growth curve, depletion, sizes caught by longline, sizes caught by PS). However, you divide up the Pacific into groups there is likely to be a) spatial structure within each group, and b) connectivity between groups. The rationale for spatial subdivision should be the feasibility of developing useful models. When fitting an age structured model to size data, one of the most important issues is spatial variation in the growth curve.

Splitting of the fishery data with regression trees appears to identify somewhat unintuitive regions to use as fisheries. The poor deviance explained in size from the regression trees likely comes from the use of latitude and longitude as splitting parameters, which results in rather rigid assignment of areas. The panel asked about other available data to identify fishery splits, like differences in gear or vessel characteristics. Using GAMs to model the fishery data could allow seasonality to be represented and control for other covariates like fleets and year effects, after which spatial effects could be more accurately represented. Those spatial effects could then be used to identify spatial areas. Seasonality in the size frequencies could suggest that you should do the fishery definition on a seasonal basis. Quarter was in the regression tree model but was not important, but that does not mean it might not be important using other methodologies.

The regression tree method may be appropriate for designing spatial structure for a fishery used in a stock assessment, under a particular set of assumptions. However, additional analyses using alternative methods are also needed to provide understanding of the factors that may drive spatial size structure. These may include growth variation, ontogenetic movement, environmental effects on fish behavior, and selectivity effects of gear and set time.

3.3. Requests and responses

The review panel requested consideration of alternative methods for splitting the EPO into fisheries. Clustering analyses could provide more flexible boundaries for fisheries and could provide better identification of areas with similar fisheries data. GAMs could be useful for identifying and controlling for other covariates. The authors presented a hierarchical clustering algorithm to cluster the fishery data, but it did not improve the variance explained appreciably.

3.4. Recommendations

- Continue to try other methods for splitting the fisheries. Providing a comparison with an assessment that does not split fisheries, but has a single fishery with time-varying selectivity would be useful for understanding the relative merits of each approach.
- As with the general recommendations, finding corroboration of trends from spatiotemporal models with external data sets would be useful to boost confidence in the signal provided by data used in the assessment. The tagging data and the buoy data, in particular, seemed to be potentially useful. When splitting fisheries and developing data inputs with spatiotemporal models, do not trim data sets–every datum counts!

4. PURSE-SEINE OBSERVER PROGRAM, PORT SAMPLING PROGRAM, LOGBOOK, ESTIMATION OF CATCHES, LENGTH COMPOSITION FROM PORT SAMPLING AND SIZE CATEGORIES FROM OBSERVER DATA

4.1. Background

(See section 5.1 on purse-seine catch)

4.2. Comments

Trips are selected opportunistically for sampling through the port sampling program. This may result in sampling bias, particularly for class 1-5 vessels which have a lower coverage rate of trips. Additionally, samples are only taken from wells that contain catches from a specific stratum. This results in a tendency for sampling of wells with catches from sets with relatively high catch volumes. This may introduce bias in species and size compositions if the compositions vary with the catch volume of a set (e.g., Peatman *et al*., 2019).

The current sampling protocol requires alternating between taking species counts for a predetermined number of fish, and taking measurements for lengths. As such, individual fish are either counted for species or measured, but not both. This approach will introduce additional variance in estimates of species compositions, as the length measurements are used to convert catch compositions from numbers to weight.

The Enhanced Monitoring Program pilot study has provided a valuable high-resolution port sampling dataset that allows investigation of within-well variation in catch compositions. Within-well variation in species compositions has been identified, likely reflecting differences in species composition between sets. The current port sampling protocol results in sampling during a short period of time relative to the duration of the unloading process for a sampled well. This likely introduces additional variance in estimated catch compositions, due to the within-well variation in species composition.

The high-resolution EMP dataset provides a means with which to review and update the port sampling protocol, as it provides information on the relative strengths of within-well and among-well variation in species compositions. This should allow simulations to explore the performance of alternative sampling protocols with varying coverage rates of trips, wells, and samples per well, building on the simulations in Lennert-Cody *et al*. (2023). Consideration should also be given to testing approaches that sample for species counts and length measurements separately, versus sampling the same individuals for both, in order to most efficiently achieve desired levels of precision in estimated species and size compositions.

Species and size compositions of purse seine catch are currently estimated using a design-based approach, using stratified ratio estimators. Species compositions for strata with no available port samples are based on estimates from strata with available port-sampling data, using substitution rules. The resolution of the strata used to estimate catch compositions depends on the specific usage case. Moving to a model-based estimation framework would allow for spatio-temporal variation to be explicitly accounted for when estimating species compositions. A model-based approach would also allow for direct estimation of species compositions for strata with no available port samples.

The species-composition models would ideally estimate proportions of catches by weight directly, e.g., using beta or Dirichlet regression models. Zero and/or one-adjusted models may be required if there are frequent observations of 0s and 1s, or ad hoc data substitutions applied. Alternatively, the species counts could be modeled, allowing estimation of proportions of catches by number, e.g., using multinomial or Dirichlet-multinomial likelihoods. Modeling of the counts allows 0s and 1s, in terms of proportions by numbers, but then requires estimation of average weights by species to convert compositions from numbers to weights. Alternative modeling approaches include random forest classification approaches, which have been fitted to port-sampling data for EU purse seiners in IOTC (Duparc *et al*., 2020), and modeling log-ratio transformed compositions. Ideally, the modeling framework should account for the fact that the observations are compositions, i.e., that the species-specific catch proportions should sum to one, or that the species-specific counts should sum to the number of sampled fish. However, species-specific models may represent a pragmatic starting point. It may also be necessary to compromise on certain features of the modeling framework, e.g., the use of single-species models to allow incorporation of spatio-temporal variation.

The panel noted that length-weight relationships are used when converting species compositions from numbers to weight. These length-weight relationships were estimated in the 1960s to 1980s, when the tuna fisheries in the region were developing and more coastal in nature. As such, there are concerns as to whether the length-weight parameters are still appropriate for current use (see Section 16 Biology).

Species counts were first collected through the port sampling program in 2000, with port sampling prior to 2000 focusing exclusively on collection of length composition data. In recent assessments, the species compositions for purse seine fisheries prior to 2000 were set at strata-specific averages for the period 2000-2004. However, the estimates of species compositions for the period pre-2000 are likely less precise than the estimates for 2000 onwards and may be systematically biased. The panel requested model runs to assess the sensitivity of the assessments to errors in species composition estimates for the purse seine fishery pre-2000 (see Section 4.3). This was explored for bigeye tuna as the species likely to be most impacted by errors in purse seine species compositions, given that bigeye accounts for a relatively low proportion of overall purse seine catch from the floating object fishery, but the floating object fishery accounts for a relatively large proportion of total bigeye catch in the EPO.

The panel also noted the importance of calculating the variance of species composition estimates, which would allow uncertainty in catch estimates to be propagated to the assessment models. However, the estimated coefficients of variation are unlikely to reflect all major sources of uncertainty in catch compositions, and so may need to be considered as relative levels rather than absolutes.

4.3. Requests and responses

The bigeye assessment model was rerun assuming that the catches of bigeye pre-2000 in the floating object fishery were either underestimated by 50% for each year-quarter or overestimated by 50%. Over-estimation of catches pre-2000 resulted in a large increase in estimated SSB, whereas SSB was remarkably insensitive to under-estimation of catches pre-2000. It is not clear why the model was insensitive to underestimation of purse seine catches pre-2000, though this may be explained by compensatory changes in other estimated quantities, e.g., recruitment. Alternatively, there may have been issues with model fit, e.g., convergence on a local minimum or parameter estimates hitting bounds. Regardless, the runs demonstrated that the bigeye assessment model can be sensitive to errors in catches for the floating object purse seine fishery pre-2000.

4.4. Recommendations

- Review the port sampling protocol in light of high-resolution samples collected through the enhanced monitoring program pilot.
- Move to model-based estimates of purse seine species compositions, that will better reflect spatio-temporal variation, and facilitate estimation of species compositions for strata with no available samples.
- Explore the sensitivity of assessments to uncertainty in catch estimates during periods with no, or more limited, species composition information from port samples. This could be achieved by including time-varying coefficients of variation for catch data in assessment models, or constructing alternative plausible catch histories.

5. PURSE-SEINE CATCH

5.1. Background

Estimates of total catch by species for the purse-seine fishery are provided by strata (area x month x set type x vessel size class category) (Tomlinson 2002; 2004; Suter 2010). For the assessments, stratum estimates are summed over months, vessel categories, and possibly areas, to obtain estimates by fishery area x quarter x set type. The estimation methodology for total catch by species is different for years 1975 - 1999, compared to 2000 and onwards. For 1975 – 1999, the species catch amounts from several data sources (canner/processor, observer, logbook) are combined and adjusted based on correction factors computed from the 2000 – 2004 species composition estimates.

For 2000 and onwards, port-sampling data for estimation of species composition, in addition to length composition, have been collected, and are used to estimate the total fleet catch by species. The total purse-seine fleet catch of tropical tunas (yellowfin + bigeye + skipjack) from several data sources (canner/processor, observer, logbook) are combined and prorated to catch strata. The port-sampling data are then used to estimate the species composition by stratum, considering different average weights among species and that sample species proportions are from numbers not weight. In each year, there are a number of strata with catch but not port-sampling data, which happens because sampling does not occur in all ports where catch is unloaded, mixed-stratum wells are not sampled, and additional logistical constraints (the sampling is largely opportunistic; active strata for a year are not known in advance; coverage of wells is relatively low, even though about 50 - 60% of trips are sampled). In the catch estimation methodology, port-sampling data from 'neighboring' strata are used to estimate the species composition of catch of strata with no port-sampling data. A set of hierarchical rules define the choices of 'neighbor' strata. The length data of the neighboring stratum are 'grown'/'shrunk' to the month of the stratum with no port-sampling data, if necessary. Even though the rules for selecting the 'best' neighbors are fixed, neighbors used as 'substitutes' for the same stratum may be different in different years because of differences in fishing activity and sampling by year. Substitution can effectively change the catch estimation 'model' from one year to the next, which may introduce added variability into the catch time series.

5.2. Comments

Considerable discussion was had around the uncertainty in catch estimates and the need to more appropriately represent this uncertainty. The currently specified CVs in the assessment model are 0.01 and this does not seem to match the uncertainty described in the catch estimation process. For example, the way in which data were borrowed via the 'substitution' method to calculate catch estimates seemed antiquated and ripe for improvement.

5.3. Requests and responses

(See Section 4.3)

5.4. Recommendations

- Spatiotemporal models should be used for the 'substitution' methods used to determine species composition of catches. Propagating the uncertainty through each step of the catch estimation should also be paid close attention. Enumerating the places in the process of catch estimation where uncertainty enters could be a useful starting point in trying to understand how best to propagate uncertainty.

6. PURSE-SEINE DISCARDS

6.1. Background

Discards are defined as fish that are caught, but not landed, and are assumed to die. Two types of discards were considered in previous assessment models: i) independent of size/age - those resulting from inefficiencies in the fishing process (e.g. catch from a set exceeding the remaining storage capacity of the fishing vessel), those are considered to have composition the same as the retained catches, ii) dependent of size/age - those related to the sorting of catches (when fishers discard tuna that are under a certain size), those are considered as independent fisheries. For purse-seine fisheries in unassociated sets and sets on dolphins, the two types of discards are simply added to the retained catch and considered as the total removals of those fisheries. For purse-seine on floating objects, the discards resulting from inefficiencies in the fishing process are added to the retained catch, but the sorting discards are treated as a separate fishery in the assessment because the port sampling does not provide information about the size distribution of those fish, which are assumed to be small and young. The selectivity is assumed to be at age uniform over 1-3 quarters old. The discards are estimated from the information obtained by observers using the method described in BET-02-06. The observer program for Class-6 vessels attained 100% coverage in 1993 with the AIDC. There is no information to estimate discards prior to 1993, and discards in that period are considered zero. The discards due to sorting in the floating-objects fisheries show a reduction beginning around 2001, and ceased almost completely following resolutions adopted by the IATTC which prohibited discarding of small tunas (e.g. C-04-05). Inefficiency discards have decreased over time and are around 0.45% for yellowfin tuna and 0.25% for bigeye tuna in floating objects sets in recent years.

6.2. Comments

The assumption of no discards before 1993 seems to be somewhat arbitrary, however, given a lack of data, it is not clear how to best estimate discards before this period. It was further unclear what impact the discards have on assessment outcomes.

6.3. Requests and responses

The panel requested an assessment run be performed in which the discard data were removed and the output of this model compared to one in which the discard data were included. The authors did this. There was little difference between the outcomes including and excluding the discards. This is perhaps unsurprising given the small magnitudes of discards compared to other catch sources.

6.4. Recommendations

- Discards do not appear to drive the assessment results and therefore improving estimates of discards should consequently likely be a lower priority than, say, revising fishery definitions of refining indices of abundances. A lack of impact in the assessment does not, however, imply that continued monitoring of discards is not important. It is possible that discards could reach a level that could impact the stock in the future, so monitoring and periodic re-analysis in the assessment should continue. Further, whatever approach is used for the discards (e.g., including or excluding them from the assessment), this should be consistent across data types and assessments by species.

7. HOW TO TREAT THE PANDEMIC YEARS FOR PURSE SEINE AND LONGLINE

7.1. Background

A simulation study was conducted to try to quantify the effect of pandemic-related data loss on catch estimation (Lennert-Cody *et al*. 2022). This study concluded that systematic loss of port-sampling data in 2020 may have led to a bias in the OBJ fishery catch estimates, particularly for bigeye tuna (BET), because while the COVID-19 pandemic generally limited the ability of IATTC port-samplers to collect data in 2020 – 2021, the disruption to data collection was greater in some ports than in others. This may have resulted in bias in the estimated catch composition for 2020 – 2021 for the OBJ fishery because some fleet segments preferentially unload in specific ports. While results from the simulation study indicated that the 2020 BET OBJ catch estimates could be biased, by as much as +/- 20%, the exact magnitude and direction of that bias could not be determined from this simulation study alone. Thus, to further investigate the issue of bias and to adjust the species catch estimates for the OBJ fishery for bias caused by the pandemic, a Conditional Auto Regressive (CAR) spatio-temporal model was developed and used to predict OBJ fishery total species catches for 2020 – 2021 (Majumdar *et al*. 2022). This approach used observer and logbook data, which had much higher coverage during the pandemic, to 'supplement' port-sampling data. The CAR spatio-temporal model was used to predict the port-sampling species proportion in the catch, using the observer (logbook) species proportion and spatial and temporal information. These estimated species proportions were then multiplied by the total tropical tuna catch, by stratum, and summed over strata to obtain estimates of the total OBJ fishery species catch of BET and SKJ. The total OBJ fleet catch of YFT was computed as the difference between the OBJ fleet total catch of tropical tunas and the estimates of BET and SKJ from the CAR models.

7.2. Comments

The effect of the COVID pandemic affected port sampling. This data is used to compute species compositions, which is afterwards used to compute catches by species.

The total volume (aggregated for the 3 species of tunas) is split by species using the port sampling programme. Historically there ~ 25% of the total volume that doesn't have [related] samples, while in 2020 and 2021 that volume was ~50%. The port of Manta in the second half of 2020 and first quarter of 2021 was the main affected.

The algorithm used to estimate catches by species includes a set of substitution rules which allocated data to strata without data. The substitution rules were not designed to deal with such a large amount of missing data and as such seem to be creating a large bias when compared with estimates derived from observers.

An alternative spatial temporal model to compute catches was derived and presented. The model showed good behavior to recover historical data.

The two algorithms show similar impacts on the outcome of the stock assessment model.

7.3. Requests and responses

- Run models with both algorithms to check if it has an impact in management quantities.

Model runs show there is an effect that needs to be dealt with and that the two algorithms do differ from each other with relation to estimates of F/Fmsy.

- CAR model estimates to be added to the plot with both design-based estimates and observers estimates.

7.4. Recommendations

- Try to deal with the missing data through the assessment model settings, e.g., increase CV of relevant fleets and quarters. Compare the outcomes with the 2 algorithms. If similar, keep the stock assessment model approach to maintain consistency within the model and avoid adding another source of uncertainty through external estimation of missing values.
- Considering the two algorithms provided to predict the missing values of catches, the CAR model seems to be the best option.
- Explore the use of a multidimensional response variable spatial temporal model to estimate the species compositions for the three species together instead of individually estimating two and deriving the proportion of the third by difference.
- Deal with length frequencies missing data through model specifications.

8. LONGLINE CATCH

8.1. Background

Five types of data from the longline fleets are submitted routinely by the CPCs to the IATTC in fulfillment of the obligations from several resolutions (e.g. resolution on provision of data C-03-05, resolution on observers on longliners C-19-08): 1) the list of the vessels predicted to operate in the EPO, 2) catches without spatial information (aggregated by year; for the most recent year, the monthly catches of bigeye tuna need to be submitted as mandated by the tuna conservation resolutions, e.g. C-21-04), 3) catch and effort data with spatial information aggregated at the 5° latitude by 5° longitude –month by species, 4) size (length and/or weight) data with spatial and temporal information aggregated at different scales, 5) observer data with spatial and information at fine scale resolution ('operational level data'). Other data sets are available for special projects with national scientists using other instruments such as Memorandums of Understanding.

The catches are reported in numbers, in weight or both. JPN and USA only report the catches in numbers, PYF and TWN always report both weight and numbers. Other flags report either both quantities or only weight in some years. When the catches are reported in weight it is not clear if the reported weight is whole weight or processed weight and what type of processing was done. The comparisons of the catches among the flags are complicated by the mixture of units, as assumptions are needed. For the stock assessments, whole weight is assumed. In the assessment longline fisheries are defined for the two types of units. When a flag reports both, numbers are used in the assessment.

The procedure to obtain the total catch by fishery in the assessment, described in WSBET-02-03, includes the aggregation by the area (and quarter) definition of the fisheries. Longline fisheries are also defined by unit. Thus, catch in numbers and catch in weight are treated as separate fisheries with the same selectivity, so that the transformations between units are dealt with inside the assessment models.

In recent years the industrial longline fleets effort is contracting closer to the western and central Pacific Ocean, with the exception of China, which is expanding. The catches of swordfish have increased in recent years off Peru and in equatorial areas east of 120°W. The catches of albacore in the southern hemisphere had also increased in recent years. The longline catches of tropical tunas in the EPO since the year 2000 are on average 73% bigeye tuna, 25% yellowfin tuna and 1.5% skipjack tuna.

8.2. Comments

Changes in spatial patterns in catches is likely an issue for catch estimation. Swordfish targeting could have implications for the access to operational data if it is not possible to distinguish swordfish targeting without the operational data.

8.3. Recommendations

- Comparing time series of average weight (calculated from total catch in weight divided by total catch in numbers) might be a better indicator for stock trajectories than the size composition data. This was the case for Taiwan because of inconsistent sampling protocols and looking at average weight for countries that have difficulties with size sampling and have both numbers and weights.

9. LONGLINE DISCARDS

9.1. Background

The current stock assessments of tropical tunas in the EPO assume all longline fleets operated in the EPO did not discard tropical tunas in the EPO. This strong assumption is due mainly to the fact that there is very limited research estimating the degree of discard in EPO distant-water longline tuna fisheries. The tunas caught by distant-water longline fisheries are consumed mainly in the sashimi market where there is a strong discrimination against small fish. Preliminary work conducted for longline tuna fisheries in other oceans suggests a possibility of high-grading, a behavior that longline fishers tend to discard small fish because of its relatively low economic value. Ignoring high-grading could cause several issues in the stock assessment including underestimated longline catch, biased longline index of abundance, and biased longline selectivity. The only longline discard study for EPO longline fisheries was conducted by Korean scientists. Using data provided by Korean observers, the mean discard rate for bigeye and yellowfin tuna in the Korean longline fishery in the EPO was estimated to be 3.4% and 3.6%, respectively. It is important to note that these two mean discard rates are estimated based on only three years (2016-2018) of data. No length information is currently available for longline discards. A difference is found between the length composition collected by fishers and observers, but it is unclear whether it is caused by measurement, discard, or both.

9.2. Comments

The panel inquired about the decision to include discards for purse seines but not for the longline fleet. Even though the discards are relatively small (~3%), understanding their impact should still be examined.

9.3. Requests and responses

The panel requested an assessment model run that includes the longline discards and compare the assessment outcomes to a model run that excludes the longline discards. The differences in the outcomes were minimal.

9.4. Recommendations

- Discard work for SBT might be useful to explore for ideas relevant to the EPO given their relatively in-depth consideration of the topic.
- More broadly, a consensus with a defensible rationale on whether to include highly uncertain discards estimates in the assessment vs. excluding them for lack of impact on assessment outcomes should be reached and then applied to each stock consistently. Also consider consistency over time—if you include discard data in one time period, they should probably be in another time period.

10. LONGLINE LENGTH COMPOSITION

10.1. Background

Since the last benchmark assessment, survey fleets are disconnected from the fisheries structure, total catch, and catch composition. In the EPO, there were no fishery-independent surveys of tuna abundance and size composition, with the term "survey" in this context referring to a fleet that has data (e.g., abundance index and size composition) but takes no catch.

The exploratory assessment models for bigeye tuna in the EPO include twelve longline fishery fleets and one longline survey fleet. The longline length composition data provided by Japan and Korea are used to compute longline length compositions for this assessment because they are the two most important longline fleets for bigeye in the EPO. The comparison of length frequency data collected by longline observers shows that Japan and Korea have similar selectivity within the same spatiotemporal windows. As such, they are treated as a joint fishery fleet and the longline length compositions for bigeye tuna in the EPO are based on both Japanese and Korean size data. For Korean size data, only those from observers were used because preliminary work suggests the length frequency data collected by Korean fishers may be biased towards larger bigeye.

The Japanese longline composition data used to compute the length compositions for the survey fleet are from commercial longline vessels only. The data includes both length and weight compositions and is reported at various spatial resolutions and bin sizes. Survey length compositions should be spatially weighted by CPUE so the spatiotemporal fields of both length frequency and fish abundance are needed. The spatiotemporal field of fish abundance can be extracted from the spatiotemporal model that has been developed to provide the index of relative abundance. Another spatiotemporal model is used, which should be length-specific, to predict the spatiotemporal field of length frequency.

VAST is used as the platform to develop the length-specific spatiotemporal for standardizing longline length frequencies due mainly to its ability to account for multiple categories (length bins in this case) simultaneously. VAST models encounter probability and positive catch rate separately to account for zero-inflated length frequency observations. Specifically, VAST was specified to use the logit and log link functions for the linear predictors of encounter probability and positive catch rate, respectively, for each length bin. Both linear predictors include an intercept (year-quarter) term, a time-invariant spatial term, and a time-varying spatiotemporal term. Of these three terms, the intercept term is estimated as fixed effects and the other two terms are estimated as random effects. Neither the catchability covariate (HBF) term nor the vessel effects term is included in this model because they are not available in the Japanese longline length composition data. This VAST model treats the four quarters equally (no seasonal component) to be consistent with the "quarters-as-years" approach used in the stock assessment model.

Due to the high dimensions of the length-specific spatiotemporal model, several simplifications are made to make the model computationally more feasible: 1) only 40 spatial knots are used to estimate the spatial and spatiotemporal random effects in the EPO; 2) length bins are regrouped from the original resolution to 10 cm; 3) length frequencies for < 60 cm are negligible and are assumed 0 (length bins in the model: 60-70 cm, 70-80 cm, …, 190+ cm); and 4) all hyperparameters are assumed to be shared among length bins. It should be noted that the predicted length frequencies for each knot and time do not necessarily sum to 1 across length bins, as the spatiotemporal field of length frequency is predicted for each 10 cm length bin without a multinomial constraint. To solve this problem, the predicted length frequencies are scaled to have a sum of 1 for each knot and time.

The longline length compositions for the survey fleet and fisheries fleet have a bin size of 10 cm and are derived from the same predicted length frequency from the length-specific spatiotemporal model. The main difference lies in how the predicted length frequency is raised: to fish density across the EPO for the survey fleet and to catch within a fishery region for the fishery fleets. In the exploratory assessment model for bigeye tuna in the EPO, all longline length compositions are formatted with a bin size of 10 cm from 60 cm to 190+ cm and are reported in the size composition section of the Stock Synthesis. Overall, the sample size of the Japanese longline length composition data decreased rapidly since 2011, especially in the inshore tropical region.

10.2. Comments

Weight composition data should be used cautiously and with awareness of their potential biases. Length-weight and weight type conversion factors are outdated and mostly based on small samples close to the coast. They are therefore likely to be biased and inaccurate for most of the population, and until updated will compromise the value of weight frequency data – though the size of the bias is also unknown. Weight composition data should whenever possible be converted to length compositions outside the model because this is the most accurate approach given the various sources of variability (e.g., spatial) associated with length-weight and weight-weight relationships.

Training vessel data have been omitted from size data standardizations for both fisheries and surveys because they are not representative of the commercial catch, since they are often sourced from areas with no commercial catch. However, predictions for surveys require estimates of fish sizes across the population, not just the catch. Similarly, predictions for fisheries may be improved by having size information for areas without catch, if the standardization accounts for spatial variation in size. Moreover, training vessel data are measured as lengths and therefore unaffected by errors in weight-length conversion that affect some commercial size data. Training vessel size data may therefore be useful to include in the size data standardization. The key remaining issue is whether the gear selectivity is the same for training and commercial vessels, but we are not aware of evidence for substantial variation in selectivity among longline fishing methods.

Similarly, data cleaning to set a minimum sample size of measured fish per 1x1 stratum is unnecessary and discards potentially useful information when using a spatio-temporal model. The model can both smooth across space and account for uncertainty if effective sample sizes are assigned appropriately.

Noting that size distributions for Korean observer data are similar to Japanese observer and commercial size data, and that Korean commercial size data appear biased high, it makes sense to drop the Korean commercial size data. The panel agreed that a single fishery could be used for all longline fleets, rather than separating them by fleet. Korean commercial size data may not always have been biased, so it would be useful to continue to liaise with Korean scientists to identify if and when the sampling quality changes.

Uncertainty in predicted size distributions can be estimated with the spatiotemporal model and used to assign relative effective sample sizes among quarters, to be subsequently adjusted using Francis weighting. Uncertainties should be presented as part of the data analysis. Input sample sizes are based on fish sampled / 100. This is subsequently adjusted by VAST to provide a scaled estimated sample size. Whether this adjustment adequately accounts for uncertainty associated with imputation into areas without sampling should be evaluated.

Observer data are available by sex and size. Including observer size data by sex in the model should be considered.

At the end of the time series the assessment model residuals show conflict between the ‘observed’ length frequencies in the survey – predicted by the VAST model to be representative of the population - and the length frequencies predicted by the assessment, which decline given the increasing fishing mortality. It would be informative to look at spatial variation in the predictions of the time series and evaluate whether the predictions in eastern Area 2 are affected by imputation based on trends from sampled areas in the west.

The stock assessment must assume that growth is spatially unvarying. However, the growth curves and asymptotic lengths are very different in the WCPO and EPO. Therefore, unless one or both of these growth curves are wildly wrong, there is very likely to be spatial growth variation within the EPO. In a model that fits to catch at length, size data provides much of the information about total mortality, but this info will only be reliable for locations where growth is consistent with the model growth curve. If spatial size variation is seen as plausible, size data ESS should consider consistency between the local growth curve and the model growth curve – based on the sampling locations of the data used to develop the model growth curve.

10.3. Recommendations

- Include as much data as possible, don't filter data to limit extrapolation and imputation of data.
- Consider including training vessel size composition data in standardization analyses, since it provides information about fish sizes in part of the population that receive little sampling effort.
- When fitting to size data, allow the model to deal with uncertainty due to sparse data rather than setting a minimum sample size per spatial cell. This avoids removing potentially informative data.
- If spatial size variation is seen as plausible, size data ESS should consider consistency between the local growth curve and the model growth curve – based on the sampling locations of the data used to develop the model growth curve.
- Check the number of trips "sampled", if the same vessel is repeated sample size may need to be adjusted due to correlation across samples.
- Effective sample size for the stock assessment model should take into account the increased extrapolation over time.

11. LONGLINE STANDARDIZED AVERAGE BODY WEIGHT

11.1. Background

Logbook data collected by longline fleets may contain information on both catch in numbers and weight. Weight data is important for commercial reasons, numbers are required reporting. From this information the average weight by set could be estimated. The coverage of catch in number and catch weight is larger than length composition, which is based on samples taken by fishers or observers. Average weight could be used as an alternative data in the assessment model. Patterns could be compared with large scale patterns in size from composition data.

11.2. Comments

Trends in average weight through time are apparent – should these be included in the assessment?

The estimates will be affected by the length-weight relationship and other morphometric relationships, which need to be accurate for this data type to provide reliable information. The possibility that they are affected by changes through time in measurement and/or preparation methods should be evaluated.

The index does provide information but seems to be noisy.

If doubts about morphometric relationships and other factors make it difficult to include average weight in the assessment, there is still potential to use the index for monitoring outside the assessment.

As an initial step it would be useful to put the index in the model without fitting it, to see if it's consistent.

Spatial coverage varies through time. The spatial component of the model should adjust for these changes in coverage, but this needs to be checked.

11.3. Recommendations

- Standardized average body weight has potential as an information source, and should be further developed and tested.

12. ECHOSOUNDER BUOY INDEX

12.1. Background

Tropical tuna stock assessments traditionally relied on catch and effort data from fishing records, forming the basis for evaluating fish stocks and their trends. Metrics like Catch-Per-Unit-Effort (CPUE) offered insights into abundance, but factors like changing fishing methods and environmental shifts complicated the accuracy of these estimations. Standardizing CPUE became crucial to filter out these influences and pinpoint changes solely related to population abundance.

The introduction of Fish Aggregating Devices (FADs) in tuna fishing significantly boosted efficiency but brought challenges in incorporating their CPUE into stock assessment models due to the complexity of standardization. Innovative collaborations between scientists and industry pioneers facilitated the integration of new technology, like satellite-linked echosounder buoys attached to FADs. These buoys offered a non-invasive means to monitor fish biomass beneath FADs, providing valuable data for understanding tuna behavior and estimating abundance indices, eventually leading to the creation of the Buoy-derived Abundance Index (BAI).

This BAI has now been integrated into stock assessments, marking a milestone in collaboration efforts. A partnership involving tuna fishing companies, research bodies, and buoy providers aims to produce reliable BAI for tropical tuna in the Eastern Pacific. Continuously updating the methodology and sharing recent and future progress in meetings appears to be key factors in establishing this indicator as crucial for upcoming population assessments.

12.2. Comments

The echosounder buoy indices were generated using data from ‘virgin segments’, i.e., the period from 20 to 35 days after the inferred deployment of a drifting FAD. The 90th percentile of estimated biomass over the ‘virgin segment’ was used as a proxy for the abundance of tropical tunas. However, there is potentially additional information in the time series of the echogram to inform the local abundance of tropical tunas, for example in the rate of accumulation of biomass around the FAD. As such, future work should consider analyzing the time series of estimated biomass, rather than the asymptote. This is an area of ongoing research by IATTC staff.

The biomass estimated to be associated with the FAD during the ‘virgin segment’ was apportioned by species using available catch compositions from the area and time of the virgin segment, taken from observer data (Class-6 vessels) and logbook data (Class 1-5). This approach assumes that the catch compositions are representative of the species compositions during the virgin segment. However, sets on

FADs are generally not made during the virgin segment. As such species compositions during the virgin segment may differ from catch compositions from sets on drifting FADs. Sonic tagging experiments could help improve the understanding of the behavior of the different tropical tuna species when associated with FADs, and potentially identify drivers for differences in the species compositions of biomass associated with FADs compared with catch compositions (e.g., Scutt Phillips *et al*., 2019). There have also been a relatively large number of recent archival tag releases. The expanded dataset may increase understanding of associative behavior of tuna around drifting FADs, particularly for yellowfin and possibly skipjack.

Additionally, work is ongoing to use machine-learning algorithms to predict species compositions from echosounder buoy data, trained using echograms with a corresponding set for which catch compositions are available. Predictive models of species compositions could then be used to estimate species compositions directly from echogram data, removing the need to infer species compositions from an external data source. Multi-frequency echosounder buoys may be able to better differentiate between species and size ranges of tropical tunas associated with FADs.

The models used to analyze the echosounder buoy data assumed linear effects for continuous variables. Non-linear effects, for example using cubic splines, should be considered in future analyses. For example, the effect of local FAD densities may not be linear, for example in the case of potential FAD saturation effects.

There are a number of different data filtering rules applied in the current analysis of the echosounder buoy index. It is not clear how sensitive the estimated indices are to the different assumptions made in the data selection process, for example the timing of the virgin segment. Sensitivity analyses may help identify the key assumptions that are made and their influence on assessment models, as well as prioritize future research. This is an area of ongoing research by IATTC staff.

The species-specific echosounder buoy indices of abundance were highly correlated. Potential drivers for the similarity between indices were not clear, though it may be related to the use of catch compositions to split the echosounder estimates of total biomass between species.

Inclusion of the echosounder buoy indices in the assessment models as a survey fleet requires assumptions regarding the selectivity of the survey, or information on the size composition of fish associated with FADs. The echosounder buoy indices currently assume that the species compositions of catches are reflective of the species composition of biomass associated with FADs. As such, it would be consistent to assume that the size composition of catches is also reflective of the size composition of the fish associated with the FAD.

The current echosounder dataset in the Eastern Pacific Ocean consists of data from 2010 onwards. The coverage of the dataset is highest in the western region of the EPO, and more limited in the eastern region where FAD densities are thought to be highest. Additional historic data from other fishing companies operating in the EPO would be very informative in increasing the spatial coverage of the dataset, noting that area has been identified as an important source of variation in abundance in the current analysis. This would allow the spatial representation of the echosounder buoy index to better match the spatial extent of the purse seine fishery. Additional historic data would also provide a more powerful dataset for testing for variance in abundance related to environmental conditions and local estimates of FAD densities. Additional data would provide a more comprehensive dataset for fitting predictive models of species compositions.

Echosounder buoy indices have the potential to inform assessments of all three tropical tunas. This is particularly true for skipjack, which does not currently have an alternative index of relative abundance that is considered reliable. However, it is difficult to assess how robust the current echosounder buoy indices are, and whether they should be included in the assessment models. The panel requested additional analyses for the three species, to explore the sensitivity of the models to the inclusion of the echosounder buoy indices, and to assess the extent to which the indices were consistent with other data sources fitted to by the models.

12.3. Requests and responses

The bigeye assessment was reasonably insensitive to the inclusion of the echosounder buoy index in terms of the estimated levels and temporal trends of spawning stock biomass. However, model fits to the echosounder buoy index were relatively poor for some years, particularly in the early 2010s. Additionally, recent estimates of recruitment and F/Fmsy were more sensitive to the inclusion of the echosounder buoy index, and the object-associated purse seine fishery selectivity that was assumed to apply to the echosounder buoy index. This is problematic as the recent estimates of F/Fmsy are used as the basis for management advice of EPO tropical tuna fisheries.

The yellowfin assessment was similarly influenced by inclusion of the echosounder buoy index, though recent estimates of spawning stock biomass were more strongly impacted compared to bigeye. Recent estimates of depletion in spawning stock biomass were higher when the echosounder buoy index was included, along with higher estimates of recent F/Fmsy and lower estimates of recent recruitments. Similarly to yellowfin, the model fits to the buoy index were relatively poor.

The skipjack assessment model was the most sensitive to the inclusion of the echosounder buoy index, with large increases in SSB and large corresponding decreases in F/Fmsy since 2018. Model fits to the echosounder buoy index were reasonable, though in some time periods the temporal variation in the buoy index was not reflected in model fits.

The requested runs for all three species identified sensitivity in recent F/Fmsy estimates with inclusion of the echosounder buoy index. These estimates at the end of the time-series are most influenced by the echosounder buoy index, as the buoy index tracks relatively small fish compared with the other survey fleets included in the models, and so the recent cohorts will not be fully selected by the other survey types. Additionally, there was apparent conflict between the echosounder buoy index and other data inputs to the assessments, with either poor fits to the buoy index (bigeye and yellowfin), or high sensitivity of assessment model outputs to the inclusion of the buoy index (skipjack).

12.4. Recommendations

- There are a number of recommended avenues of exploration related to the analysis of the echosounder buoy index covered in the panel comments, some of which are ongoing areas of research by IATTC.
- The echosounder buoy index has great potential to inform assessments of tropical tuna in the EPO as a fisheries-independent survey. However, the assessment models are sensitive to the inclusion of the echosounder buoy index, particularly recent estimates of F/Fmsy that are used as the basis for management advice. It is not currently clear to what extent the species-specific estimates of biomass associated with FADs are representative of the relative abundance of the underlying

population. In this context, the panel recommends that the echosounder buoy indices are not included in assessments of the tropical tuna species at this time.

13. LONGLINE INDICES

13.1. Background

Longline indices of relative abundance are crucial input data for the stock assessments of bigeye tuna and skipjack tuna in the eastern Pacific Ocean. The stock assessment of bigeye tuna in the EPO is fit to only one index of relative abundance. It is the standardized longline index computed based on Japanese longline catch and effort data aggregated at the vessel-year-month-1° grid level. In comparison, the stock assessment of skipjack tuna in the EPO is fit to both longline and echo-sonar buoy index of abundance while assuming the longline index is more precise than the echo-sonar buoy index. This longline index is also computed based on Japanese longline catch and effort data. Among all distant-water longline vessels operated in the EPO, Japanese longline vessels have the highest spatial coverage within the EPO and the longest history of high-quality logbook data, providing the information needed for the standardization of a reliable abundance index with a large contrast across time.

A delta-generalized linear mixed spatiotemporal model VAST is used to standardize the Japanese longline CPUE data for bigeye tuna in the EPO. VAST separately models encounter probability and positive catch rate to account for zero-inflated catch rate observations. Specifically, VAST is specified to use the logit and gamma link functions for the linear predictors of encounter probability and positive catch rate, respectively. Both linear predictors include an intercept (year-quarter) term, a time-invariant spatial term, a time-varying spatiotemporal term, a catchability covariate (using HBF as a 2-knot spline) term, and a vessel effects term. Of these five terms, the intercept term and the catchability covariate term are estimated as fixed effects and the other three terms are estimated as random effects.

The VAST model used to standardize the longline index of abundance for bigeye tuna in the EPO covers 1979-2022 because the ID of Japanese vessels was not available before 1979. This VAST model treats the four quarters equally (no seasonal component), consistent with the “quarters-as-years” approach used in the stock assessment model. The coefficient of variation (CV) of the longline index of relative abundance is also estimated by VAST and is scaled to have a mean of 0.15 between 1979-2014.

13.2. Comments

The focus was on analysis of the bigeye indices, but some of the issues would apply equally to yellowfin if those indices were used in the assessment.

The analyses show considerably more decline in CPUE in the east (Area 2) than in the west (Area 1), suggesting some viscosity (limited mixing at the sizes caught in longline fisheries) in the stock structure. There is uncertainty about this decline, and about the methods used to estimate it. Given the spatial variation in CPUE trend it may be appropriate to model the two areas separately in the assessment, because the different declines imply that total mortality and age structure would differ between the areas. But the role of the data meeting is to consider the trend and its reliability.

There may be a component of target switching to focus on other species such as swordfish and albacore. The presentation notes that the Japanese fleet persistently targets bigeye 1979-2022 (Document RVDTT-01-02). However, the cluster analysis figures (slides 47-51) indicate a strongly increasing proportion of effort in the non-bigeye clusters, particularly albacore. These transitions in species composition may be linked to species spatial and temporal abundance patterns or to genuine changes in targeting behavior by the fleet. It is very important to better understand these changes.

To address this issue, we suggest working with Japanese scientists to explore targeting indicators other than HBF, such as hooks per set, bait use, and the movement behavior of individual vessels. For example, both high numbers of hooks per set and the use of squid bait have been shown to be associated with swordfish targeting by Japanese longliners (Hoyle & Okamoto 2011). It may also be useful to explore different clustering approaches. Ward's hierarchical clustering method is commonly used. Clustering without restricting to 5x5 squares may more effectively associate the cluster with targeting by the vessel, rather than with the species available at a particular location, which is a risk with the current method. Working with operational data would permit clustering over a 10 day or 1-month period rather than a full quarter – depending on how often vessels may change their targeting strategy. Including additional shark and billfish species would provide more information about fishing strategies, because when the targeting information comes from the species mix, more species provide more information. However, note that when including non-target species, it is important that reporting quality is consistent during the analysis period, so analysts should identify a suitable period to focus on.

There is a need for a more thorough characterization process, documenting as many aspects of the fishery as possible. For example, maps of HBF and hooks per set through time at 1 x 1 scale, aggregating across periods such as every 5 years, or by season, would help to identify changes that occur seasonally and through time. Plots to show patterns of vessel turnover would also be useful.

The analysis started in 1979, due to lack of vessel call-signs before 1979, and limited HBF 1975-1978. Japanese scientists may be able to recover earlier vessel ids for the EPO, as they have for other oceans (e.g., available from 1975 in the Atlantic: Matsumoto 2023). This is worth exploring if it would be useful to extend the time series.

In the last benchmark assessment, the indices were split into early and late periods, but the split has been removed in current analyses. The panel supported the decision not to split the indices.

The IATTC have also expanded the spatial domain to include all 1x1 cells with at least 20 quarters (rather than 80 quarters) of CPUE data between 1979 and 2022, allowing the eastern EPO to be included. It is important to index as much as possible of the spatial domain, but also risky to impute too far beyond the available data. Analysts need to be careful about imputation into large data gaps over long periods.

The panel suggested exploring the influence of environmental density covariates, such as those related to season. They also emphasized the need for additional sources of information to corroborate the imputed depletion, such as alternative indices, or evidence from changes in tag return rates through time.

There is also a need for auxiliary information about reasons for changes in fleet behavior to stop fishing in the eastern EPO, which may be obtained by interviewing fishers.

Supporting evidence for the Area 2 decline was provided by nominal indices of associated purse seine catch rates, which declined more in the eastern than the western EPO.

The panel supported the decision to correlate the ST term in both time and space, rather than in space alone, because this increases the information available to the model. Applying the temporal correlations via a random walk was considered appropriate. Simulation shows that the approach works in principle. The panel noted the increasing effort concentration, as seen in other oceans.

As noted for the longline size data, the panel suggested that, when smoothing across space, it is not necessary to clean data at fine (1x1) spatial scale, because individual cells are not fitting independently. Retaining all the data increases the information available for the analysis. The statistical model can allow for the uncertainty associated with low sample sizes.

The panel commented that the CPUE CV used in the assessment appears low compared with intuitions about uncertainty given the data gaps. To some extent this may be due to VAST's assumptions and approach for sharing uncertainty across space.

Vessel effects changed the index in the expected direction, but with only a small impact. The change in the trend associated with vessel effects represents effort creep associated with vessel turnover. The small change is likely due to long-term stability of the fishing fleet in the analysis – the data filter removes vessels with fewer than 40 quarters of fishing effort. If the filter was set at a lower level, such as 5 quarters, more vessel turnover would occur and there would likely be more difference between the trends with and without vessel effects. The very strong vessel filter may be necessary to speed up the VAST analyses but including more vessels (particularly in Area 2) may improve data coverage which would be helpful given the large data gaps.

Effort creep estimated in this way does not include changes in fishing power associated with changes in vessel behavior or the technology available to the vessels. These effects are not included because data are unavailable, so CPUE is likely to be overly optimistic.

Adjusting HBF after the introduction of monofilament mainlines, by subtracting 3 from HBF in sets using monofilament, appears to be appropriate, if ad hoc. It may be useful to explore alternative adjustments, but we agree that this does not seem a high priority given that the change did not make much difference to the indices.

Although confidentiality requires data to be aggregated, it is useful to have data at the relatively detailed level provided (vessel x 1° cell x quarter x HBF). However, there remains a need for operational data for analyses of targeting. There is also a need for operational data to explore issues such as vessel behavior.

The observed change in gear configurations through time by the Korean fleet suggest they may have changed fishing strategies to target BET.

The panel agreed that it would be useful to explore joint analyses incorporating data from multiple fleets, and that there is a need for further work with operational data from other fleets such as Korea to develop this index. It was noted that the highest priority is to increase the limited information about the eastern EPO, given the estimated decline, and that there is no significant Korean fishing effort in area 2 - effort here is mostly Japan and 'Other' fleets. The availability of logbooks from 'Other' fleets fishing in this area should be explored.

The correlation between CPUE and ONI (SOI) varies spatially, so it is difficult to include it as a catchability or density covariate in the analysis. A local oceanographic variable that is associated with the ONI, which may act as a density covariate, should be explored. If such an association occurs, it would be better to use this variable in the VAST model.

13.3. Recommendations

- Explore the availability of logbooks that can provide information about recent catch rates in the eastern Area 2.

- More detailed documentation is needed of the data characterization and CPUE standardization.
- Investigate the factors that cause the observed changes in species composition through time.
- Investigate the potential to include more vessels in the analysis, particularly in the eastern area.
- Continue to collaborate with data holders in conducting analyses with operational data, to explore hypotheses and to validate the results using aggregated data.

14. PURSE INDEX ON DOLPHIN SETS FOR YELLOWFIN TUNA

14.1. Background

The 2020 assessment was fit to an index of abundance derived from the purse-seine CPUE of sets associated with dolphins. Until 2019 the main index for yellowfin was derived from the CPUE of the Japanese longline fleet, and purse-seine indices were secondary. The purse-seine and longline indices were not compatible and the Japanese fleet continued to retract farther from the main purse-seine yellowfin fishing grounds, closer to the western and central Pacific Ocean. The incompatibility of the indices from the two gears may be due to spatial heterogeneity. The current methodology to obtain the purse-seine index includes the standardization of the catch and effort data using spatio-temporal models. The data is obtained by the observers. A strict data selection is performed (only vessels with 75% or more of sets on dolphins), so that vessels that have their main fishing strategy as fishing associated with dolphins are retained. This allowed for the use of days fished as a unit of effort. The spatial domain was also restricted to areas north of 5N and 1 by 1 cells with more than 30 years of data. The effect of different data selection criteria was shown. The spatiotemporal model estimates the encounter probability (logit link) and positive catch rate (log link). It is implemented in the VAST R package (https://github.com/James-Thorson-NOAA/VAST). Spatial correlation is assumed in different directions (anisotropy). Vessel effects are also included as random effects. Size compositions are also standardized to represent the index. The standardized size compositions show cohorts moving through the population.

In the assessment model (Stock Synthesis platform) the abundance index is entered in the model as a "survey", a fishery without catches but with associated size compositions. An extra component of variation is added to the CVs estimated by the spatiotemporal model so that it averages about 0.15. Several hypotheses regarding the relationship between abundance and the index are considered in the assessment.

14.2. Comments

The association between tuna and dolphins is not well understood and exploration of this relationship would be useful for interpreting indices based on dolphin sets. Potential topics for exploration include the relationship between dolphin pod size and tuna school size, the fraction of tuna that are dolphin associated in the EPO, and time trends in the size and number of dolphin pods. As with other indices of abundance presented during the panel, the uncertainty does not seem to be well represented and a more in-depth description of how the variances are calculated could be useful. A specific example pointed out was that the data in the first quarter is frequently much more sparse, but the associated uncertainty does not appear to reflect that.

Some discussion was had about the relative coverage of the indices for each stock: the spatial distribution of the bigeye data encompass a larger amount of the perceived extent of the population than for yellowfin and this should likely be a factor in the way these data are analyzed and included in the assessment.

14.3. Recommendations

- A standardized comparison of indices across assessments and presentations would be useful for consistency and comparability in the review process.

15. TAGGING DATA AND THE SPATIO-TEMPORAL TAGGING MODEL

15.1. Background

Tagging data used to infer movements, spatial structure, growth, and natural mortality of bigeye (BET) and yellowfin tuna (YFT) stocks in the eastern Pacific Ocean (EPO), as well as providing the foundation for the EPO skipjack tuna (SKJ) TP was primarily derived from two independent large-scale tagging campaigns conducted during 2000-2006, and 2019-2022 (regional tuna tagging project, RTTP). Data was also included from several collaborative and opportunistic tagging campaigns conducted during 2002-2015 where significant numbers of BET and YFT were released with plastic dart (PDT) and archival tags (AT).

Focused on primarily tagging bigeye tuna, the 2000-2006 efforts released 19,174 BET, 2,234 YFT, and 3,425 SKJ with PDTs, and another 323 BET, 53 YFT, and 135 SKJ with ATs. To date, 8293 BET (43.3%), 419 YFT (18.8%), and 579 SKJ (16.9%) PDTs and 168 BET (52.0%), 10 YFT (18.8%), and 7 SKJ (5.2%) ATs have been returned. Considering the elapsed time since these tagging efforts took place, it's unlikely additional tags will be returned. Con-current to these tagging efforts, tag shedding experiments were conducted, where a portion of tunas released with PDTs were double tagged, as well as all AT fish receiving two PDTs. Seeding experiments were also conducted to evaluate tag reporting dynamics, and absolute reporting rates.

Effort during the RTTP was focused on tagging SKJ and a total of 265 BET, 1,679 YFT, and 6,181 SKJ were tagged and released with PDTs. An additional 57, 472, and 250 were released in BET, YFT, and SKJ, respectively. To date, 109 (41.1%) BET, 306 YFT (18.2%), and 1705 SKJ (27.6%) PDTs, and 23 BET (40.4%), 94 YFT (19.9%), 60 SKJ (24.0%) ATs have been returned. To facilitate the collection of the best possible recovery data, three tag recovery specialists (TRS) were employed for the duration of the tagging efforts, one in each of the following IATTC field offices: Mazatlán, Mexico, Manta, Ecuador, and Playas, Ecuador. Each TRS, in addition to facilitating the collection of high confidence tag return data, was responsible for advertising the tag recovery program throughout the region. No shedding rate experiments were conducted, but extensive tag seeding was, providing insights into reporting rate, reporting accuracy, and absolute reporting rates. While there is some bias in the dissemination of distribution of tag seeding kits to observers likely to return to ports where TRS were operating, nearly 15% of the total trips returned to ports with no TRS. Results from preliminary analysis indicate that reporting accuracy, and absolute reporting rates were considerably low in ports without TRS.

Collaborative tagging conducted with the Secretariat of the Pacific Communities (SPC) Oceanic Fisheries Program, and opportunistic tagging aboard sportfishing vessels accounted for the release of 30,793 BET and 5,722 YFT with PDTs, and an additional 449 BET, 1,292 YFT, and 53 SKJ with ATs between 2009 and 2015, and 2002 to 2011, respectively. To date, there have been 10,033 (32.5%) 981 (17.2%) PDTs recovered in BET and YFT, respectively. In addition, there have been 75 (16.7%) and 443 (34.3%) ATs recovered from BET and YFT, respectively. There were no SKJ ATs recovered from these releases, which is likely a result of tagging mortality and poor tag attachment.

A length-structured spatiotemporal population model allows estimation of movement as an advection-taxis-diffusion process and length-based mortality rates utilizing available tagging and effort data and might ultimately allow estimation of population size, distribution and sustainable harvest levels. While advection might be informed by ocean currents, taxis can be based on smooth habitat preference functions of environmental covariates such as sea surface temperature or the mixed layer depth. Movement rates and recapture probabilities can be estimated by means of the matrix exponential of instantaneous rates or the classic Kalman filter. Results indicate that the movement of SKJ in the EPO is inversely related to the velocity of ocean currents and depends on sea surface temperature. SKJ prefers intermediate sea surface temperatures around 25-26°C and exhibits stronger undirected movement at low and high temperatures. Further, the model estimates length-based fishing mortality rates in space and time for each fleet and a length-based natural mortality rate in line with previously reported rates.

15.2. Comments

The reporting rate of recovered tags is an essential parameter in many analyses of mark-recapture datasets. It is rare for all recoveries of tagged fish to be detected and reported, and incomplete tag reporting will lead to downwards bias in fishing mortality estimates if it is not accounted for. Tag seeding experiments have recently been undertaken during the Regional Tuna Tagging Project (RTTP), which provide information on reporting rates. These suggest high reporting of tag returns, with tag recoveries reported for over 85% of seeded tags. This likely reflects the extensive efforts to advertise the RTTP throughout the region. However, the majority of observer trips returned to Manta, Playas, and Mazatlán, the three ports with a Tag Recovery Specialist. This may introduce upwards bias in reporting rates if reporting of tag returns is easier at ports with Tag Recovery Specialists and therefore more likely to occur. Bias in estimates of reporting rates could be mitigated by attempting to undertake tag seeding experiments to achieve a representative sample of landing ports in the region and taking a weighted average of port-specific reporting rates, weighted by the volume of tropical tuna unloaded at the respective ports. IATTC are currently undertaking additional tag seeding experiments, aiming to improve the coverage of ports in the region.

Tag seeding experiments also provide a means to assess errors in reported dates and locations of recaptures, by comparing reported dates and locations to the equivalent information for the set where seeding took place. These comparative analyses can provide information on errors in reported information, as well as the relative reliability of tag returns that were not reported directly to Tag Recovery Specialists. This information can then be used to ensure appropriate treatment of tagging data in analyses.

Tagging-induced mortality and tag shedding rates are also critical variables with respect to analysis of mark-recapture data. Both act to reduce the numbers of tagged fish that are available for recapture, and lead to downward bias in recapture rates if they are not accounted for. Experience from other tuna tagging programs suggests that tagging-induced mortality and tag shedding can substantially reduce the numbers of tagged fish available for recapture. IATTC have been collecting information associated with tag releases that may be informative with respect to estimation of tagging-induced mortality and tag shedding rates. Additionally, IATTC are currently exploring options for at-sea experiments to estimate tagging-induced mortality rates.

Recapture rates of bigeye tuna were comparatively high (39% of 265 releases with plastic dart tags) relative to recapture rates for skipjack (27% of 6181 releases) and yellowfin (17% of 1679 releases). This difference may be explained by relatively strong associative behavior of small (<90cm) bigeye to drifting FADs in the region, though it may also reflect the limited number of bigeye releases.

There was a low proportion of ‘high confidence’ recoveries of skipjack, i.e. recoveries for which vessel, well position and number were validated by Tag Recovery Specialists, relative to bigeye and yellowfin. This appears to reflect the high proportions of skipjack releases in 2020, and difficulties in placing observers and accessing vessels in port at this time due to the impacts of COVID-19.

Tag releases during the RTTP demonstrated a bimodal size distribution, particularly for skipjack and bigeye. The causes of this pattern in size at release have not been identified.

Tagging cruises provide a potential dataset for constructing indices of relative abundance, particularly for skipjack given that the tagging cruises were skipjack focused. However, experience from the tagging cruises, e.g., difficulties in getting fish to ‘bite’ after schools are found, suggests that there may not be a clear relationship between the numbers of tag releases and local abundance. Regardless, the tagging dataset may have utility as a point of comparison for other potential data inputs to the assessments, for example the echosounder buoy index.

Mark-recapture data can be used to make inferences regarding the underlying population if the tagged fish are representative of the population, i.e., the tags are ‘mixed’ at the resolution used in the analysis. As such, to incorporate tagging data into assessment models, it is necessary to assume that tagged fish are mixed with the underlying population within each model region after an assumed period of time (the ‘mixing period’). This may lead to difficulty in incorporating tagging data directly inside the assessment models, as the majority of tagged fish may die before mixing has been achieved within large areas, given the constraints around where tuna can be tagged and released. However, the spatio-temporal model fitted to IATTC’s tagging dataset uses a relatively fine-scale spatio-temporal structure, within which tagged fish will mix more rapidly.

The use of a fine-scale spatio-temporal structure may make the tagging model more sensitive to observation errors in reported times and positions of recovery. As described above, the tag seeding dataset can provide estimates of observation error in the tagging dataset, and may be helpful in identifying appropriate spatio-temporal resolutions for models of tagging data. Excluding tag return data that is considered less reliable, e.g., that has not been validated by Tag Recovery Specialists, may introduce bias in estimated movements of tagged fish if there is spatial structure to reliability of the tagging data.

The relatively limited spatial coverage of tag releases is typical of tuna tagging programs elsewhere, due to the inherent logistical challenges involved, e.g., access to bait resources. Parameterizing movement in the spatio-temporal tagging model as a function of environmental conditions allows movement rates to be estimated for areas with limited available tagging data. However, the relatively limited spatial coverage of tag releases by IATTC, coupled with the relatively short displacements for most tagged fish, raises concerns regarding the robustness of estimated relationships between movement rates and environmental conditions.

The spatio-temporal tagging models have clear potential to provide external estimates that can be used as inputs to assessment models, or as a point of comparison for estimates from the assessments. These estimated quantities could include natural and fishing mortality rates, as well as a relative index of abundance. As such, the tagging dataset and the spatio-temporal tagging model represent an important area of research with respect to strengthening the scientific basis for management advice of tuna fisheries in the region, particularly for skipjack.

Tagging cruises in the region are impacted by difficulties in finding schools of tuna for tagging. This situation could be improved through collaboration with the fishing industry, given the extensive network of drifting FADs deployed in the region. Releasing tagged tuna more broadly throughout the EPO should provide a more statistically powerful dataset with which to inform the spatio-temporal movement model.

15.3. Requests and responses

The Panel requested plots of linear displacement against time-at-liberty from the conventional tagging dataset. The plots of linear displacement against time-at-liberty for the three species appeared to be fairly typical for tropical tuna, indicating fairly rapid dispersal with increasing time-at-liberty though with considerable variability. For yellowfin, there was an apparent tendency for some individuals to remain close to their point of release, i.e., with displacements < 100 nm. This feature was not observed for the other species. For bigeye, there appeared to be (at least) two groups of individuals, characterized by either relatively short or long displacements for a given time at liberty. This may reflect tag releases at different locations relative to concurrent fishing effort. However, it is difficult to interpret plots of linear displacements, as the observed displacements are influenced by the spatio-temporal distributions of both tag releases and fishing effort, as well as the underlying movement of fish.

15.4. Recommendations

- Continue to explore collaboration with the fishing industry to increase efficiencies and spatial coverage of tagging cruises. This would also be required to run sonic tagging experiments on tuna associated with drifting FADs.
- Explore indices of relative abundance using data collected on tagging cruises.
- Continue to develop the spatio-temporal tagging model, given its potential to inform assessments of skipjack in the EPO.

16. BIOLOGY

16.1. Background

Morphometric relationships (e.g., length-weight relationships) for tropical tunas are outdated and are likely no longer representative of the spatial extent of industrial fisheries in the EPO. Considering the evidence supporting stock structure, these antiquated relationships for BET (Nakamura and Uchiyama 1966), YFT (Wild 1986), and SKJ (Hennemuth 1959), are woefully inadequate for many statutory requirements of the organization. These data are a critical component to several research and reporting activities required to fulfill the objectives of the Antigua Convention and IATTC's Strategic Science Plan (SSP). IATTC staff have drafted a proposal to address these deficiencies in these relationships, however, there is little enthusiasm for such projects.

Methods utilized to estimate the age and growth are derived from direct counts of validated daily increments for BET (38 - 135 cm) and YFT (40 - 148 cm: Schaefer and Fuller 2006; Wild *et al.*, 1995; and Wild and Foreman 1980). While daily increment deposition is well described for these species in the EPO, validation of daily increment deposition does not encompass the entirety of the harvested length range for BET and YFT, which may lead to bias in the age estimates of larger fish. However, corroborating evidence from tagging studies where fish tagged at smaller sizes (<90cm) and recovered at larger sizes (>150cm), over long periods at liberty (years), supports the belief that BET at lengths greater than those included in the validation studies still produce readable daily increments up to at least 160 cm. A model to improve age estimates for BET was developed by Aires-da-Silva *et al*. (2015), where high confidence tagging data was integrated into a modified Richards model to better estimate asymptotic lengths, and provide more robust age estimates of larger fish. The growth model for YFT developed by Wild (1986), while outdated due to expansion of the fishery encompassing a greater spatial area and the adoption of different fishing strategies, provides the most reliable means to estimate length-at-age.

Estimating the age of SKJ using hard parts (otoliths, spines, vertebrae) has proven quite difficult as increment deposition rates are variable and not 1:1 as they are in BET and YFT. Wild and Foreman (1980) evaluated tetra-cycline marked otoliths from 26 SKJ ranging in times at liberty from 17-249d and 42-64 cm in fork length. Results from the study indicated SKJ produce increments at a rate significantly less than 1 per day. Because of the reliability issues, a paper (SAC-13 INF-J) presented to the scientific advisory committee of the IATTC in 2022 by Maunder *et al*. uses tagging data to estimate growth.

A sampling program executed between 2009 and 2016 collected ovaries and otoliths from female YFT to evaluate the spatial variability in growth and maturation, and to evaluate possible temporal changes in the growth rates and sizes at maturity. The work on reproductive biology by Schaefer and Fuller (2022) is completed and is described below. The daily aging of the otoliths from this sampling is in progress and is expected to be completed by the end of 2024. However, preliminary analysis indicates that growth rates are not significantly different from those estimated by Wild (1986), but there is evidence of spatial variability whereby YFT in the north of 20°N grow slower than those south of 15°N. There also appears to be a difference in growth rates of YFT in nearshore versus offshore areas between 5N and 15°N. No fish south of the Equator had yet been evaluated.

Reproductive characteristics for BET were described by Schaefer *et al*. 2005 from sampling conducted aboard purse-seine and longline vessels operating throughout the EPO and portions of the central Pacific ocean. Sex ratios, an ogive, and estimates of batch fecundity were derived from the analyses of nearly 2000 samples collected between 2000 and 2003. The ogive for female BET was derived from the classification of 683 histological samples of ovarian tissue, resulting in an estimated size at 50% maturity of 135 cm. Relative batch fecundity was derived from seven ovary samples and estimated to be about 24 oocytes/gram body weight (range: 14-43 oocytes/g). The mean spawning interval was 1.3 days for 102 females classified as reproductively active and spawning occurred entirely at night.

Reproductive characteristics of YFT in the EPO were first described by Schaefer (1998) consisting of the analyses of nearly 20,000 fish (male and female) collected from 1987-1989. Between 0° and 20°N spawning took place throughout the year with no pronounced seasonal pattern, however, samples collected from north of 20°N indicated spawning occurred primarily between July and November. The size at 50% maturity was estimated to be 92 cm based on the histological classification of over 7000 ovarian tissue samples. Relative batch fecundity was derived from the assessment of 345 ovaries and was estimated to be about 67 oocytes/ gram body weight. The mean spawning interval was 1.19 days for 565 females classified as reproductively active and spawning appears to occur entirely at night.

A more recent study of the reproductive biology of YFT by Schaefer and Fuller (2022) derived from 1,728 samples collected between 2009 and 2016 over a length range of 40-160 cm indicated that there is both spatial and temporal variation when considering the earlier work of Schaefer (1998). Spawning was widespread, occurring from 16°S and 25°N, between 78°W and 148 °W, in sea-surface temperatures mostly >25°C and throughout the year. The size at 50% maturity was estimated to be 81.6 cm, substantially smaller than those reported in Schaefer, 1998. Relative fecundity was reported to be about

61.0 oocytes/g of body weight from the evaluation of 146 ovaries. Spawning frequency for reproductively active females was estimated to be 1.23 d. Based on the presences of hydrated oocytes and post-ovulatory follicles, spawning appears to take place almost entirely at night.

The reproductive biology of SKJ in the EPO was evaluated by Schaefer and Fuller (2019) and is based on the sampling of nearly 8,200 fish. Histological classification of ovaries sampled from 3,732 female SKJ provides the foundation for estimates of length-specific reproductive characteristics. Spawning was widespread from about 19 °N to 12 °S and from 79 °W to 136 °W, and continuous throughout the year between about 15 °N and 10 °S. The size at 50% maturity was estimated to be 52.9 cm across the aggregate area, and showed a decreasing trend from N to S (56 cm in the N, 52.7 in the Central, and 47.0 in the S areas). Spawning frequency for reproductively active females was estimated to be 1.18 d, and based on the presences of hydrated oocytes and post-ovulatory follicles, spawning appears to take place almost entirely at night. Relative batch fecundity was estimated to be 55 oocytes/g body weight from the evaluation of 129 ovary samples.

16.2. Comments

Morphometric relationships are important for stock assessment outcomes, but the available length-weight and weight-weight relationships are very outdated. Estimating conversion factors requires large datasets with stratification across sources of variation, which may include fishery, location, season, year, sex, observer, sampling method, and fish state (frozen/fresh/etc.). The panel supported the proposed project F.3.a to estimate morphometric relationships for tuna species. To maximize data availability and to ensure that consistent methods are used, studies of morphometric relationships should collaborate with members and with other RFMOs, particularly WCPFC / SPC. It could be useful to develop a large joint database for IATTC with contributions from multiple members so that analyses can include data from all relevant sources and across long periods of time.

Where possible, analysts should convert data to a single measurement type (FL) outside the assessment model, to avoid using a single (and usually inaccurate) conversion factor inside the assessment model.

Studies of reproductive biology have identified spatial and temporal variation in reproductive biology for yellowfin tuna. From 1998 to the mid-2000s a substantial reduction was observed in mean length at maturity in multiple locations. This may reflect stock depletion.

Both spawning fraction and batch fecundity vary spatially, with length, and through time for YFT in the EPO (Schaefer, 1998; Schaefer and Fuller, 2022). Given the reduction in length at maturity between 2006 and 2022, similar changes in spawning frequency and batch fecundity should be investigated.

Analysts should use smoothers to jointly model location, size, and other covariates when analyzing data, rather than grouping data into subareas and analyzing them separately or using subarea as a categorical variable. Environmental covariates such as SST should also be considered. Comparisons between sampling programs in different years may be affected by natural (interannual) variability. There is a need for more sampling.

Spatial and temporal variability in reproductive parameters is also likely for bigeye tuna but has not been explored. There is a need for similar sampling work for bigeye.

It is useful to evaluate whether there is sufficient information to base stock status on spawning potential (SP) rather than spawning stock biomass (SSB). IATTC assessments for BET (two-sex model) and SKJ use SSB in stock status indicators, based on the proportion of female biomass at age that is sexually mature for BET. WCPO assessments (single-sex models) and the EPO YFT (two sex model) use SP, which is based on egg production and is the product of biomass, proportion female, proportion mature, spawning fraction, and fecundity at age. SP may be a more relevant metric than SSB. This has significant implications for tuna stock assessments (Hoyle and Nicol, 2008; Minte-Vera *et al*., 2019). These changes can be significant even when steepness is set to 1, but they increase with lower values of steepness. Uncertainty and possible spatial and temporal variation in the age distribution of reproductive output (SSB and SP) can also be important for assessment outcomes.

Growth and natural mortality differences between sexes also affect the ogives of relative SP at age. There is a need to better understand why sex ratios change with length for both bigeye and yellowfin tuna. This further highlights the importance of being able to age fish that are too large to age from daily rings.

Age and growth are associated with some of the most important issues affecting stock status estimates. IATTC aging of using daily rings has been validated for bigeye up to 135 cm (Schaefer and Fuller, 2006) and for yellowfin up to 148 cm (Wild and Foreman, 1980; Wild *et al*., 1995). However, the lack of validation for larger fish and the difficulty of ageing them makes it difficult to estimate the shape of the upper part of the growth curve, which is highly influential for tuna stock assessments.

In other RFMOs (WCPFC, IOTC, ICCAT) tuna of all ages are aged using otolith annuli, but IATTC have not identified consistent annuli in EPO tuna otoliths. Other tuna ageing scientists have criticized the use of daily rings for tunas at some of the larger sizes. It is a high priority to resolve uncertainty about ageing.

There is a need for continued injection of OTC into tagged fish to permit ageing validation via mark-recapture of chemically tagged wild fish.

Annual ageing has been validated using bomb radiocarbon for Atlantic and WCPO tunas (Andrews *et al*., 2022; Andrews *et al*., 2020). Bomb radiocarbon dating is generally considered to be a reliable validation method (e.g., https://uni.hi.is/scampana/otoliths/scientists/age-validation-methods/). Initially it was restricted to fish hatched in the 1960s but with new methods has been extended to recent periods (Andrews *et al*., 2016). The method has been peer-reviewed by ageing experts and tuna ageing validation has been published in reputable journals. However, IATTC scientists expressed doubt about these validations and the applicability of the approach to the EPO. We are not sufficiently expert in this area to assess these concerns, but it seems important to resolve this issue to help develop ageing methods for large tunas.

Otolith images presented to the panel showed very clear daily rings. Ageing validation indicated that daily ageing was remarkably accurate, with no apparent increase in uncertainty with the number of days at liberty. However, counting daily rings is very time consuming with a rate of approximately 4-6 otoliths per day. There are also few experienced readers. Automation with AI appears very feasible given the high quality of the images presented to the panel. If successful this approach could speed up ageing enormously, providing very large efficiency gains. It may make production ageing feasible, at least for the smaller purse-seine caught bigeye and yellowfin tuna that can be aged using this method. Efficient ageing would permit analysts to explore biological hypotheses such as spatial-temporal patterns and sex differences in growth.

There is a need to resolve the current disagreements about ageing methods between tuna ageing scientists. Differences between groups and alternative hypotheses need to be clearly identified so that they can be tested.

Ongoing work could include exchange of otoliths and otolith images with tuna ageing groups working in other oceans, along with blind ageing of otoliths marked with OTC or strontium chloride.

Given the very different growth rates and asymptotic lengths estimated for the EPO and the WCPO, there is a need for information on spatial growth variation within the EPO. Analysts should consider space and time in analyses of historical data and new samples. It is usually better to use spatial models (e.g., GAMs with 2D splines) than to compare growth between groups, because pre-selected groups may not reflect the underlying spatial patterns of growth variation. If spatial patterns of size variation are linked to growth variation, stratification of sampling across observed spatial size variation may help to identify growth patterns. Selectivity can affect mean length at age, and total mortality can affect mean age at length, so be aware of these when exploring spatial factors affecting growth rates.

16.3. Recommendations

- We strongly recommend work to develop methods for ageing bigeye and yellowfin tuna of all sizes. Developing methods to age mature fish in the EPO is a top priority for improving the assessments and resolving critical uncertainties.
- We recommend an independent expert review of ageing methods.
- Develop sufficient information about reproductive parameters and their variation in space and time to permit the use of spawning potential in stock assessments.
- Update estimates of morphometric relationships using datasets that are sufficiently large to identify sources of variation (e.g., spatial / annual / seasonal / fishery / sampling method).

17. CONCLUSIONS

We thank the IATTC stock assessment scientists for their constructive and positive approach to the review. They undertook a considerable amount of work and fully engaged with the panel, provided detailed information in response to questions and engaged in scientific discussions.

This is the first time that the IATTC has undertaken paired data and assessment reviews. A draft version of the data review was provided to the assessment review panel which convened the following month, and one panelist participated in both reviews.

For the data reviewers, the narrower focus was useful and generally preferable to the standard TORs that review data inputs and assessment together. The focus on data inputs permitted deeper exploration of the complex issues involved in developing these inputs. Data issues require at least as much discussion time as modeling decisions. Given the complexity of some issues and their unique characteristics, IATTC may wish to consider appointing the same reviewers repeatedly in future reviews so as to save time and build experience and specialist knowledge.

The panel involved in the assessment review noted that the separation was liberating. It allowed them to focus on the modeling and application issues, with some peripheral consideration of data quality issues. They noted that having one of the panelists at both reviews was crucial, since the report of the data review was still in draft form during the assessment review, and the volume of documentation made it difficult to absorb an additional report.

18. REFERENCES

Aires-da-Silva, A. M.; Maunder, M. N.; Schaefer, K. M.; Fuller, D. W. Improved growth estimates from integrated analysis of direct aging and tag–recapture data: an illustration with bigeye tuna (*Thunnus obesus*) of the eastern Pacific Ocean with implications for management. Fisheries Research. 163: 119-126; 2015.

Andrews, A.; Okamoto, K.; Satoh, K.; Welte, K.; Eveson, P.; Roupsard, F.; Macdonald, J.; Lockheed, B.; Farley, J. Final report on bomb radiocarbon age validation for yellowfin and bigeye tunas in the WCPO (Project 105) - 2022. WCPFC-SC18-2022/SA-IP-14a. Western and Central Pacific Fisheries Commission, 18th Scientific Committee. Online; 2022

Andrews, A.H.; Pacicco, A.; Allman, R.; Falterman, B.J.; Lang, E.T.; Golet, W. Age validation of yellowfin (*Thunnus albacares*) and bigeye (*Thunnus obesus*) tuna of the northwestern Atlantic Ocean. Canadian Journal of Fisheries and Aquatic Sciences. 77:637-643; 2020.

Andrews, A.H.; Siciliano, D.; Potts, D.C.; DeMartini, E.E.; Covarrubias, S. Bomb radiocarbon and the Hawaiian Archipelago: coral, otoliths, and seawater. Radiocarbon. 58:531-548; 2016.

Duparc, A.; Depetris, M.; Cauquil, P.; Floch, L.; Lebranchu, J. Improved version of the Tropical Tuna Treatment process: new perspectives for catch estimates of tropical purse seine fishery. IOTC-2020-WPTT22(AS)-13_Rev1; 2020.

Hennemuth, R.C. Additional information on the length-weight relationship of skipjack tuna from the Eastern Tropical Pacific Ocean. Inter-American Tropical Tuna Commission Bulletin. 4: 23-37; 1959.

Hoyle, S.D.; Nicol, S. Sensitivity of bigeye stock assessment to alternative biological and reproductive assumptions. WCPFC Scientific Committee. Nouméa, New Caledonia; 2008.

Hoyle, S.D.; Okamoto, H. Analyses of Japanese longline operational catch and effort for bigeye and yellowfin tuna in the WCPO, WCPFC-SC7-SA-IP-01. Western and Central Pacific Fisheries Commission, 7th Scientific Committee. Pohnpei, Federated States of Micronesia; 2011

Lennert-Cody, C.E., Maunder, M.N., Majumdar, A. 2022. The effect of pandemic-related port-sampling data loss on the 2020 purse-seine catch estimate of bigeye tuna in floating-object sets. Inter-American Tropical Tuna Commission Document SAC-13-INF-L. https://www.iattc.org/GetAttachment/cce0a9d8-7aff-4be2-b9dd-b258a4262ff8/SAC-13-INF-L_The-effect-of-pandemic-related-port-sampling-data-loss-on-the-2020.pdf

Lennert-Cody, C.E.; De La Cadena, C.; Chompoy, L.; Nieto, E.A.; Vogel, N.W.; Wiley, B.A.; Maunder, M.N.; Aires-da-Silva, A. Enhanced monitoring program for bigeye tuna catches: preliminary results of pilot study and workplan for 2023. SAC-14-10. Inter-American Tropical Tuna Commission, 14th Scientific Advisory Committee; 2023.

Majumdar, A., Lennert-Cody, C.E., Maunder, M.N., Aires-da-Silva, A. 2022. Identifying and correcting the purse-seine fleet catch for bias caused by the COVID-19 pandemic in 2020-2021. Inter-American Tropical Tuna Commission Document SAC-13-05 CORR. https://www.iattc.org/GetAttachment/5c79d113-8720-46cf-aa3f-df90f7a58d4b/SAC-13-05_Investigation-on-potential-bias-on-the-tropical-tuna-catch-estimates-caused-by-the-COVID-19-pandemic.pdf

Matsumoto, T. Standardization of yellowfin tuna cpue in the Atlantic Ocean by the Japanese longline fishery. SCRS/2023/019. Collect. Vol. Sci. Pap. ICCAT, 80(2): 133-149 (2023).

Maunder, M.N.; Fuller, D.W.; Schaefer, K.M. 2022. Growth estimates for skipjack tuna in the eastern Pacific Ocean. Inter-American Tropical Tuna Commission Document SAC-13-INF-J. https://www.iattc.org/GetAttachment/69ccc445-7df6-4baf-bde8-079f2033c814/SAC-13-INF-J_Growth-estimates-for-SKJ-in-the-eastern-Pacific-Ocean.pdf

Minte-Vera, C.V.; Maunder, M.N.; Schaefer, K.M.; Aires-da-Silva, A.M. The influence of metrics for spawning output on stock assessment results and evaluation of reference points: an illustration with yellowfin tuna in the eastern Pacific Ocean. Fisheries Research. 217:35-45; 2019.

Peatman, T.; Fukofuka, S.; Park, T.; Williams, P.; Hampton, J.; Smith, N. Better purse seine catch composition estimates: progress on the Project 60 work plan. WCPFC-SC15-2019/ST-WP-02; 2019.

Schaefer, K.M. Reproductive biology of yellowfin tuna (*Thunnus albacares*) in the eastern Pacific Ocean. Inter-American Tropical Tuna Commission Bulletin 21; 1998.

Schaefer, K.M. ; Fuller, D. W.; Miyabe, N. Reproductive biology of bigeye tuna *Thunnus obesus* in the eastern and central Pacific Ocean. Inter-American Tropical Tuna Commission Bulletin. 23: 1-32; 2005.

Schaefer, K.M.; Fuller Daniel, W. Estimates of age and growth of bigeye tuna (*Thunnus obesus*) in the eastern Pacific Ocean based on otolith increments and tagging data. IATTC Bulletin. 23: 35-76; 2006.

Schaefer, K.M.; Fuller, D.W. Spatiotemporal variability in the reproductive dynamics of skipjack tuna (Katsuwonus pelamis) in the eastern Pacific Ocean. Fisheries Research. 209: 1-13; 2019.

Schaefer, K.M.; Fuller, D.W. Spatiotemporal variability in the reproductive biology of yellowfin tuna (*Thunnus albacares*) in the eastern Pacific Ocean. Fisheries Research. 248:106225; 2022.

Scutt Phillips, J.; Leroy, B.; Peatman, T.; Escalle, L.; Smith, N. Electronic tagging for the mitigation of bigeye and yellowfin tuna juveniles by purse seine fisheries. WCPFC-SC15-2019/EB-WP-08, 2019.

Suter, J.M. 2010. An evaluation of the area stratification used for sampling tunas in the Eastern Pacific Ocean and implications for estimating total annual catches. Inter-American Tropical Tuna Commission Special Report 18. https://www.iattc.org/GetAttachment/e1648324-575c-4aaa-b5c1-629288715843/No-18-2010-SUTER,-JENNY-M_An-evaluation-of-the-area-stratification-used-for-sampling-tunas-in-the-eastern-Pacific-Ocean-and-implications-for-estimating-total-annual-catches.pdf

Tomlinson, P.K. 2002. Progress on sampling the Eastern Pacific Ocean tuna catch for species composition and length-frequency distributions. In: Inter-American Tropical Tuna Commission Stock Assessment Report 2: 339 – 356. https://www.iattc.org/GetAttachment/f0672474-ea40-46f6-8edc-708a1fa5d8f4/No-2-2002_Status-of-the-tuna-and-billfish-stocks-in-2000.pdf

Tomlinson, P.K. 2004. Sampling the tuna catch of the Eastern Pacific Ocean for species composition and length-frequency distributions. In: Inter-American Tropical Tuna Commission Stock Assessment Report 4:311-324. https://www.iattc.org/GetAttachment/5bc36019-eec7-4685-a1fc-c8c35d930ff6/No-4-2004_Status-of-the-tuna-and-billfish-stocks-in-2002.pdf

Nakamura, E.L., and J.H. Uchiyama. Proceedings of the Governor's Conference on Central Pacific Fishery Resources. Hawaii. 197-201; 1966.

Wild, A.; Foreman, T.J. The relationship between otolith increments and time for yellowfin and skipjack tuna marked with tetracycline. Inter-American Tropical Tuna Commission Bulletin. 17:509-560; 1980.

Wild, A. Growth of yellowfin tuna, *Thunnus albacares*, in the eastern Pacific Ocean based on otolith increments. Inter-American Tropical Tuna Commission, Bulletin 18: 421-482; 1986.

Wild, A.; Wexler, J.B.; Foreman, T.J. Extended studies of increment deposition rates in otoliths of yellowfin and skipjack tunas. Bulletin of Marine Science. 57:555-562; 1995.

ANNEX 1. List of participants

Name	Organization	Email
Cody Szuwalski	Panelist	cody.szuwalski@noaa.gov
Ernesto Jardim	Panelist (Chair)	ernesto.jardim@msc.org
Simon Hoyle	Panelist	simon.hoyle@gmail.com
Tom Peatman	Panelist	tom.peatman@gmail.com
Andrés Ceballes	CeDePesca	andres.ceballes@cedepesca.net
Mayra Palacios	CeDePesca	mayra.palacios@cedepesca.net
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
Haikun Xu	Inter-American Tropical Tuna Commission	hkxu@iattc.org
Jeff Morgan	Inter-American Tropical Tuna Commission	jmorgan@iattc.org
Mark Maunder	Inter-American Tropical Tuna Commission	mmaunder@iattc.org
Marisol Aguilar	Inter-American Tropical Tuna Commission	maguilar@iattc.org
Monica Galvan	Inter-American Tropical Tuna Commission	mgalvan@iattc.org
Robert Sarazen	Inter-American Tropical Tuna Commission	rsarazen@iattc.org
Heewon Park	National Institute of Fisheries Science	heewon81@gmail.com
Youjung Kwon	National Institute of Fisheries Science	kwonuj@korea.kr
Leonel Caicedo	Transmarina	leonelcaicedolc@hotmail.com