

**INTER-AMERICAN TROPICAL TUNA COMMISSION**

**SCIENTIFIC ADVISORY COMMITTEE**

**16<sup>TH</sup> MEETING**

La Jolla, California (USA)

02-06 June 2025

**DOCUMENT SAC-16-03**

**STOCK ASSESSMENT OF YELLOWFIN TUNA IN THE EASTERN PACIFIC  
OCEAN: 2025 BENCHMARK ASSESSMENT**

Carolina Minte-Vera, Mark N. Maunder, Haikun Xu and Rujia Bi

**SUMMARY**

**Full report to be posted shortly**

A benchmark stock assessment and risk analysis were conducted for yellowfin tuna in the eastern Pacific Ocean covering the period from 1984 through to the start of 2024. The main uncertainty addressed in this benchmark assessment was spatial structure. Advances were made in determining the regions and spatial definitions of fisheries (areas) based on a new cluster analysis method using length composition data. Seventy-two models based on three levels of hypotheses were used in the risk analysis. The hypotheses addressed (level 1) the spatial structure; (level 2) effort creep (changes in catchability over time), uncertainty in growth and natural mortality; and (level 3) the steepness of the stock-recruitment relationship. A model starting in 2006 was also conducted to account for the possibility of change in population or fishery dynamics before and after this period to explain differences in information content between the index of relative abundance and length composition data.

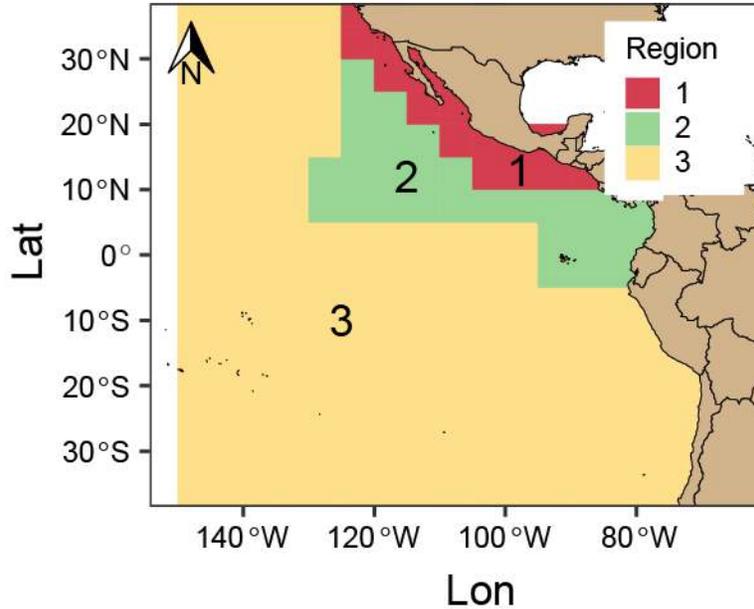
**STOCK AND FISHERY STRUCTURE**

The data suggests that there are either two or more stocks in the EPO or there is spatial structure in the population. A large recruitment in the 1990s enters the core dolphin associated purse seine fishery north of the equator in a different year than the large recruitment that enters the longline fishery south of the equator. Length composition data from the purse seine fishery associated with dolphins has smaller fish in the northeast and larger fish in the south and west, with intermediate sized yellowfin in the core area.

A three-region spatially-structured model was developed to evaluate the stock structure and movement. The regions (see Figure S-1) were delineated based on clustering length composition data. However, the model estimated limited movement among the regions. Therefore, this approach was abandoned until more information (e.g., improved tagging data) becomes available and when the assessment platform includes more flexibility for modelling movement.

Two approaches were used to incorporate spatial structure: 1) a single model for the whole EPO using areas as fleets to allow flexibility in the representation of spatial structure (EPO model) and 2) separate assessments for a) the northeast region where the core of the catches is taken (NE) and b) for the south and west region (SW). See figure S-1 for spatial definitions. A main difference among the assessments is that the indices of abundance for the EPO-wide and the NE assessments are based on dolphin associated purse seine CPUE and the index of abundance for the SW assessment is based on the longline CPUE.

Fisheries were defined in the model based on gear type (purse seine, longline, pole-and-line), purse seine set type (floating object, unassociated, dolphin associated), and area of operation to represent the different sizes of yellowfin caught. The areas were developed based on clustering length composition data. Some of the fisheries were split into small and large fish fisheries to better represent the size of the fish removed from the stock. Fisheries representing discarded small fish were also defined.



**FIGURE S-1.** Regional divisions obtained using cluster analysis of length composition data from the purse-seine fishery associated with dolphins. The EPO models consider all regions, the NE and NE\_short models include regions 1 and 2, the SW model comprised only region 3. Further subdivisions in areas were made within each region based on the cluster analysis results, which were treated as different fisheries.

## DATA

Catch, length composition, and indices of relative abundance were the main data types used in the assessment. The purse-seine catch, CPUE and length composition of yellowfin tuna are estimated by the staff using several data sources including observer reports, logbooks, and port sampling. The longline catch, CPUE and length composition data are obtained by the CPCs and submitted annually to the IATTC. Longline operational-level CPUE data by Japan, Korea and China and high-resolution vessel-specific CPUE data by Chinese Taipei were made available for this assessment.

The dolphin associated purse seine and longline CPUE-based indices of relative abundance and associated length composition data were developed based on spatio-temporal models. The longline fishery length composition data were also developed using a spatio-temporal model, but weighted by catch rather than relative abundance.

Additional data sets of reproductive biology, daily increments in otoliths and tagging obtained by the staff were used to estimate reproductive output (to define spawning biomass), growth, and natural mortality externally from the assessment models.

## MODEL ASSUMPTIONS

The stock assessment was conducted using Stock Synthesis, an integrated statistical age-structured stock assessment modeling platform. The models started from a fished state in 1984 (or 2006 for NE\_short) and were modelled through to the start of 2024 on a quarterly time step. Thirty age classes were defined from 0 quarters to 29 (7.25 years), with the oldest age used as a plus group. The models are sex-structured, but only natural mortality differs between females and males. The models are conditioned on catches and fit to relative abundance indices and length composition data.

The initial conditions include estimating the initial recruitment, the initial fishing mortality, and 16 recruitments deviations to represent the initial age structure. No penalty associated with initial equilibrium catches is used. The fishery used to create the initial conditions depends on the spatial structure assumption, but in general it was chosen as a fishery with a wide range of sizes and large catches.

Growth was updated by fitting the growth cessation model to a combination of new otolith daily increment age and length data and tagging data. Information for younger fish (up to 10 quarters of age) came from the otolith data and the information for older fish came from tagged fish with expected age at recovery of 10 or more quarters and with reliable length information.

Natural mortality ( $M$ ) was updated using a cohort analysis fit to recently collected tagging data and to sex ratio data.  $M$  was assumed to vary by age and sex, using the Lorenzen function to model the decline in  $M$  with age and assuming an increase in female natural mortality related to maturity. The sex ratio data came from both purse-seine and longline fisheries.

Maturity and fecundity were updated based on new data for maturity, batch fecundity, and frequency of spawning at length.

Recruitment was assumed to occur quarterly and follow a Beverton-Holt stock-recruitment relationship. The recruitment variability was implemented using a penalty function. An iterative process was used to set the standard deviation of the logarithm of the recruitment deviations and the lognormal bias correction factor. Three levels of steepness of the stock-recruitment relationship were used in the risk analysis: 1.0, 0.9, and 0.8.

Selectivity was specified using a decision tree based on the magnitude of the catch, reliability of the length composition data, and ability of a double normal selectivity curve to represent the length composition. Fleets with high catch volumes, reliable composition data, and a good fit to composition data, had time blocks in selectivity, the parameters of a double-normal selectivity curve estimated, and Francis weighting for the fit. Other fleets had no time blocks, fixed selectivity, lower data weighting, and/or not fit to the composition data. Asymptotic selectivity was used for the fisheries and surveys that catch the largest individuals (longline fisheries and/or purse-seine fisheries associated with dolphins depending on the model).

A risk analysis approach is used in this benchmark assessment. The approach starts by identifying alternative “states of nature” (i.e. hypotheses) that are considered plausible for describing the population dynamics of yellowfin tuna. The identification of those hypotheses is done in a hierarchical way, with the higher-level hypotheses representing the most important uncertainty (level 1) and lower-level hypotheses nested within the higher level to represent other uncertainties (level 2), and are crossed with the level 3 hypotheses, which encompass parameters for which there is little or no information in the data. The three levels of hypotheses in the risk analysis for yellowfin tuna are: level 1 - the spatial structure (EPO, NE, NE\_short, SW), level 2 - scenarios constructed to represent the uncertainty in biological parameters (growth, natural mortality) and effort creep (1% increase per year in

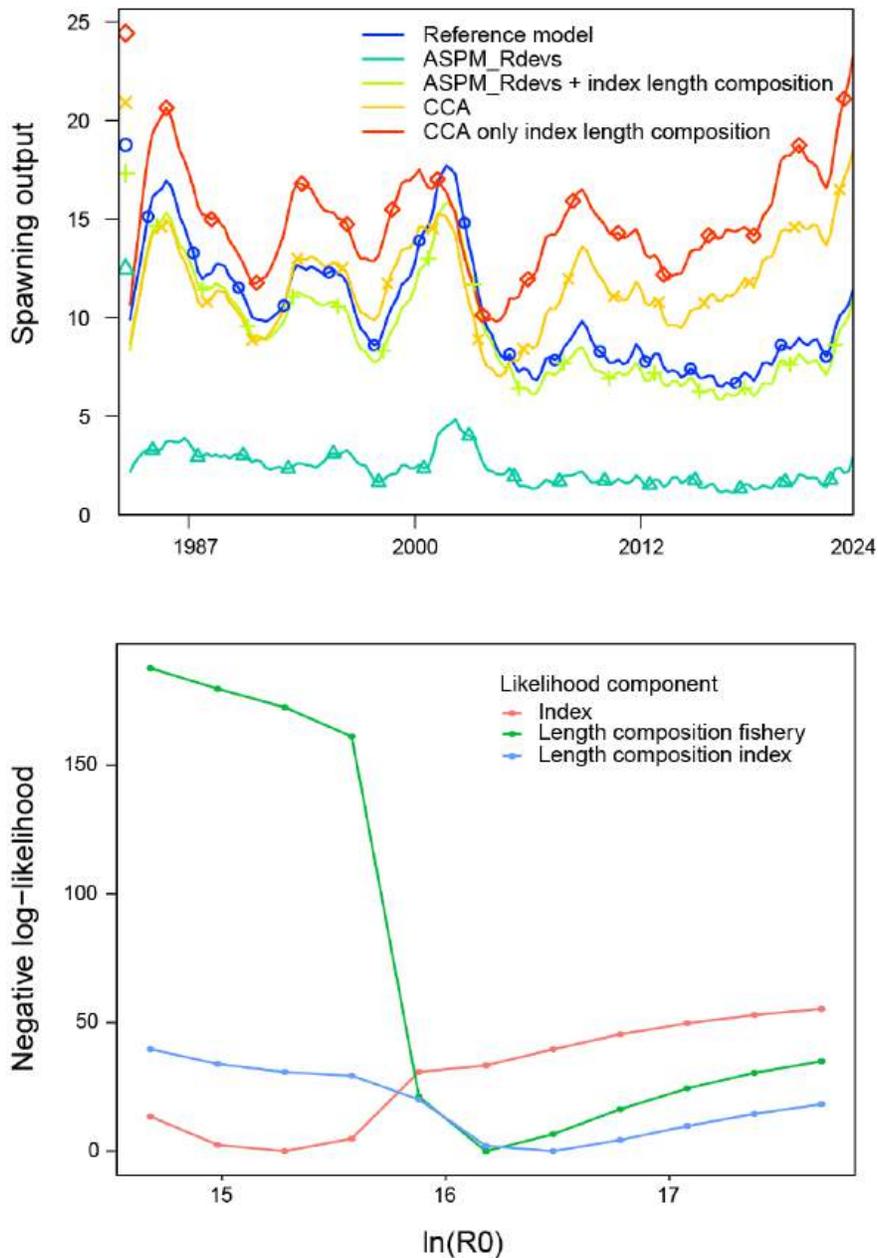
the catchability of the indices of abundance); and level 3 - the steepness of the stock-recruitment relationship. The level 2 hypotheses were implemented by changing one assumption at a time in the base reference model of each spatial structure. The low and high scenarios for growth and natural mortality were based on the uncertainty of the external estimates (values that have approximately half the likelihood as the maximum likelihood estimate). Three values of steepness of the Beverton-Holt stock recruitment relationship ( $h=1.0$ ,  $h=0.9$ ,  $h=0.8$ ) are considered for the third level hypothesis. The combination of the three levels of hypotheses results in  $4*6*3 = 72$  reference models.

The models were fit by minimizing a penalized negative log-likelihood function (NLL). To ensure that the models obtained the global minima, a series of jitter analyses, which randomly change the initial parameter values to test convergence, were performed until the model passed the jitter test (the initial parameters estimates were ones that produced the lowest NLL among all jittered models). The fits were also evaluated using residual analysis. Integrated model diagnostics were used to understand the information content of the data. The diagnostics used were the age-structure production model with estimated recruitment deviations (ASPM\_dev), ASPM\_dev also fit to the index length composition, catch curve analysis (CCA), CCA only fit to the index length composition, likelihood profile on the scale parameter ( $\log_{R0}$ ), and retrospective analysis.

One risk analysis was done for each of the four level 1 hypotheses by combining 18 reference models. Equal weight was used for all level 2 hypotheses. The weights for three values of steepness (level 3 hypotheses) were based on expert judgement from the risk analysis done for the last benchmark assessment:  $P(h=1.0) = 0.46$ ,  $P(h=0.9) = 0.32$ ,  $P(h=0.8) = 0.22$ .

## **ASSESSMENT RESULTS**

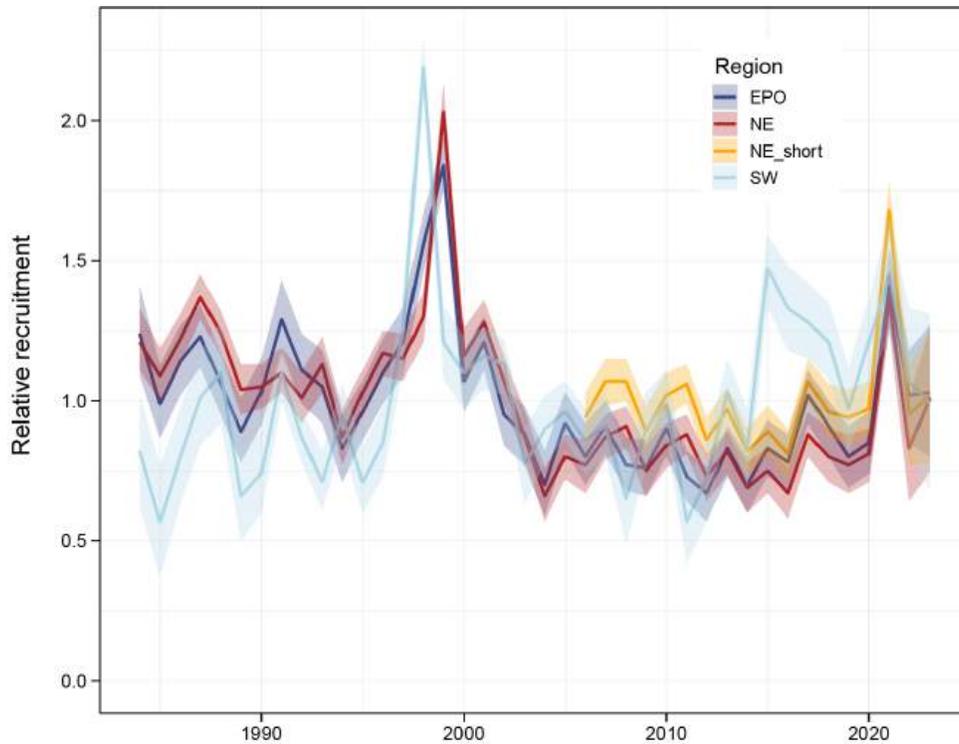
All 72 models converged and are used in the stock assessment and risk analysis. The integrated model diagnostics indicated that the indices and the catches alone are not enough to estimate the absolute scale of the models, and the length composition data provides the information on scale (Figure S-2). The indices provide information on relative trends. The length composition data for the NE and EPO models support higher absolute biomass levels in the second half of the time series (Figure S-2). For this reason, the NE\_short model that starts in 2006 was developed to represent possible changes in the dynamics of the stock or the fishery that are not understood or accurately modelled in the NE and EPO models. In addition, the tagging, otoliths and reproductive biology data come from recent years and may best represent the period in the NE\_short models.



**FIGURE S-2.** Integrated model diagnostics (Top panel: catch curve analysis (CCA) and age structured production model diagnostic (ASPM); Bottom panel: R0 likelihood component profile) for EPO base model with steepness 1 to illustrate the main information content of the data.

### Recruitment

The recruitment trends show patterns of similarities and differences among spatial structure hypotheses (Figure S-3). All models that compose the ensemble for each spatial structure hypothesis have similar trends in recruitment. Models for all four spatial structure assumptions estimate two peaks in recruitment, but for the SW model the largest peak occurs in 1998, while in the others it occurs in 1999. The second peak is in 2021. The SW models also estimated high recruitment in 2015-2017. The EPO and NE models estimate a regime shift in recruitment to a lower level after this peak, while the SW model does not.



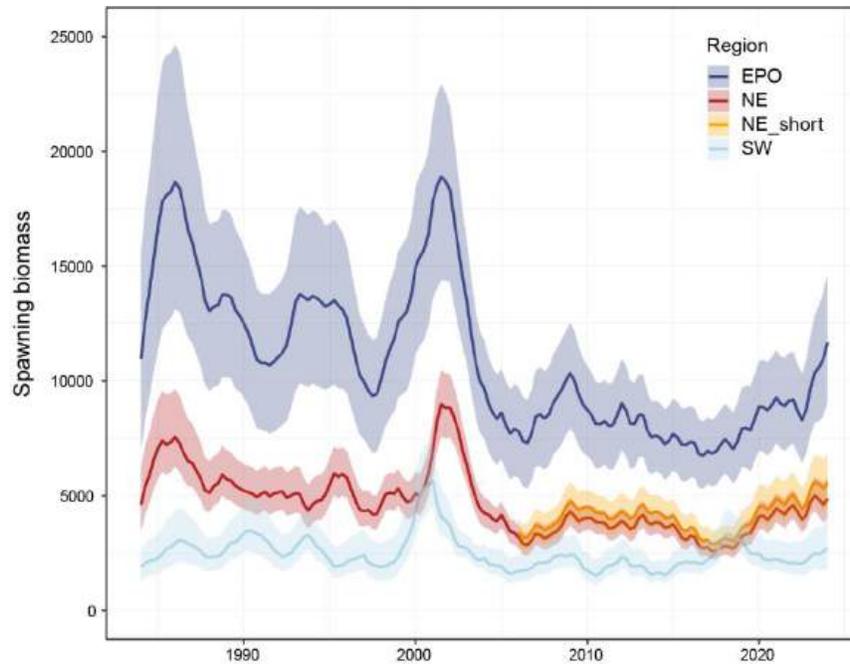
**FIGURE S-3.** Comparison of multi-model estimates of median relative annual recruitment and 80% confidence interval of yellowfin tuna for each hypothesis of spatial structure. The multi-model estimates include all level 2 and level 3 uncertainty scenarios.

### Biomass

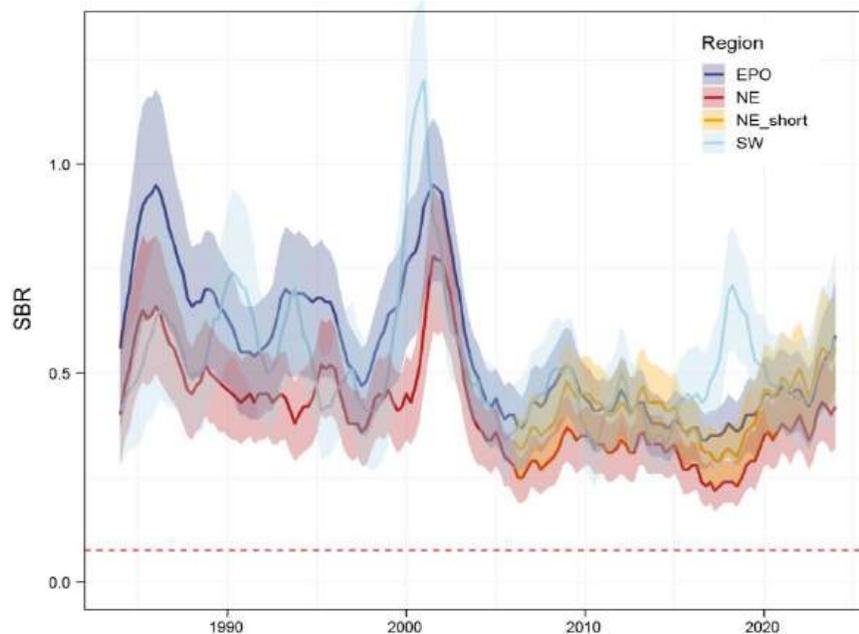
The spawning biomass in the NE is estimated to be about twice the level of that estimated for the SW. The estimate for the EPO is larger than the sum of the estimates for the two component stocks. The biomass trends (Figure S-4) generally follow the recruitment trends. Large spawning biomasses are a result of strong recruitment 2 or 3 years prior. The strong cohorts of 1998 and 1999 in the NE and SW regions show up as large spawning biomasses in 2001 and 2002 in the two regions, respectively. The trends in biomass since 2010 are diametric for the NE and SW regions.

The EPO-wide models, which use the areas as fleets approach to model spatial structure, estimates larger and more uncertain spawning biomass levels than the NE and SW combined, indicating that the EPO-wide models have difficulty fitting data with incompatible signals.

The NE and NE\_short models estimate very similar spawning biomasses.



**FIGURE S-4A.** Comparison of multi-model estimated spawning biomass of yellowfin tuna for each hypothesis of spatial structure with 80% confidence intervals.



**FIGURE S-4B.** Comparison of multi-model estimated spawning biomass ratio (spawning biomass over equilibrium virgin spawning biomass) of yellowfin tuna for each hypothesis of spatial structure with 80% confidence intervals. The red dashed line (at 0.077) indicates the SBR at the limit reference point  $S_{LIMIT}$ .

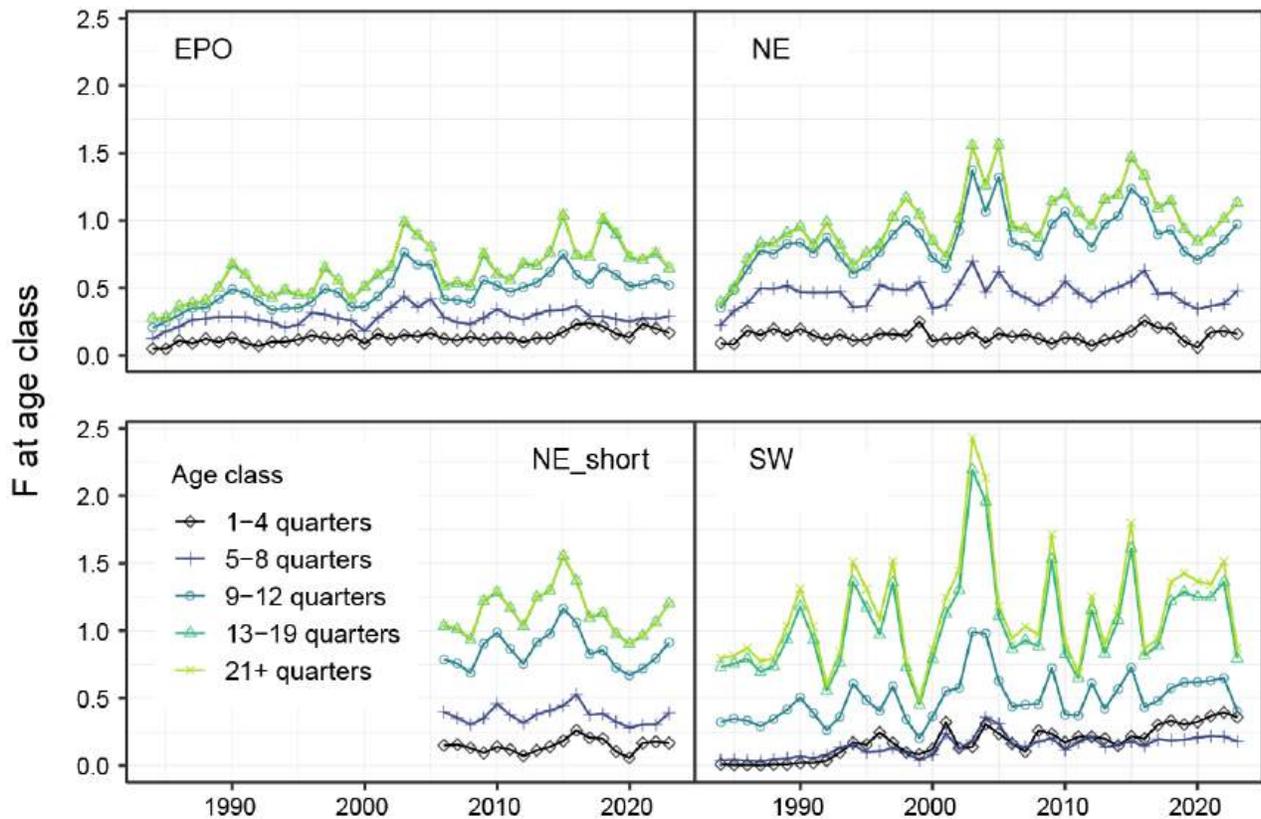
### Fishing mortality

The relative distribution of fishing mortality at age is similar for the EPO, NE and NE\_short models: the fishing mortality is much higher for the older age classes. The magnitude of the fishing mortality, however, is lower for the EPO model, which is a consequence of its biomass being estimated higher than

the sum of the biomasses for the NE and SW regions. The relative distribution of fishing mortality at age of the SW region follows a different pattern. The fishing mortality on the intermediate aged yellowfin (9-12 quarters of age) is lower since the unassociated catches are lower and the purse-seine fishery associated with dolphins generally catches larger yellowfin. The fishing mortality on the youngest yellowfin (1-4 quarters of age) has steadily increased following the expansion of the FAD fishery in the mid 1990's. After 2015 the fishing mortality of this age group surpasses the 5-8 age class.

The trends in fishing mortality are similar between the NE and the NE\_short models, indicating that starting the model later does not change the perception of the effects of fishing in recent years. For those two hypotheses, there is a general increase in fishing mortality in all age classes after the year 2006, decline after 2015, with the lowest at the start of the covid19 pandemic, in 2020. After that, the fishing mortality increases, particularly for older yellowfin.

The increase in fishing mortality noticed in the last five years in the NE area is not shared by the EPO model. This may be due to the influence of the SW area, which has stable fishing mortality followed by a sharp decline in 2023. This indicates that using an EPO-wide model may underestimate and mask regional trends in fishing mortality.

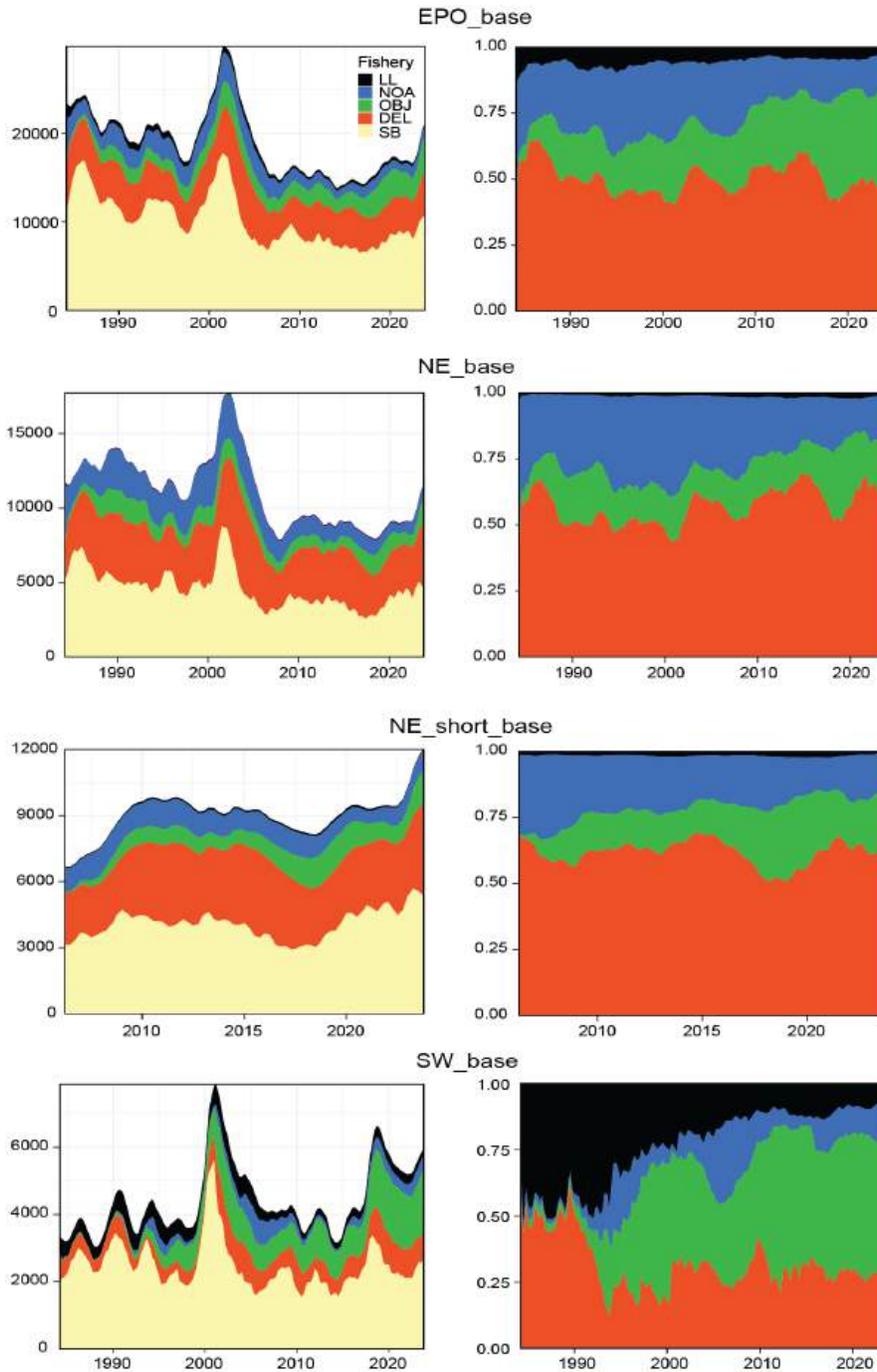


**FIGURE S-5.** Annual fishing mortality at age (sum of the four quarterly estimates within a year) of yellowfin by age group for each hypothesis of spatial structure (level 1). The values for each age group are weighted across level 2 and level 3 hypotheses.

## **FISHERIES IMPACT**

The EPO, NE and NE\_short models estimate similar impacts of the different types of fisheries (Figure S-6). The longline fisheries have the smallest impact, while the purse-seine fisheries associated with dolphins have the greatest impact during most of the modelled period. The unassociated fisheries had the second largest impact in the early years, but in the 1990s the impact of the floating-object fisheries started to increase and surpassed that of the unassociated fisheries around 2008.

For the SW models, the impact of the different purse seine set type has changed considerably over time. The longline fishery and the purse-seine associated with dolphins had the largest impact until mid-1990's, when there was an expansion of the floating object fishery, which steadily increased its impact and became the fishery with the largest impact in this region, larger than all other fisheries combined. The longline fishery has decreased both its effort and its impact on yellowfin in that area. The fishery associated with dolphins has slowly increased its absolute impact in this region, but in proportion it has stayed stable since the year 2000.



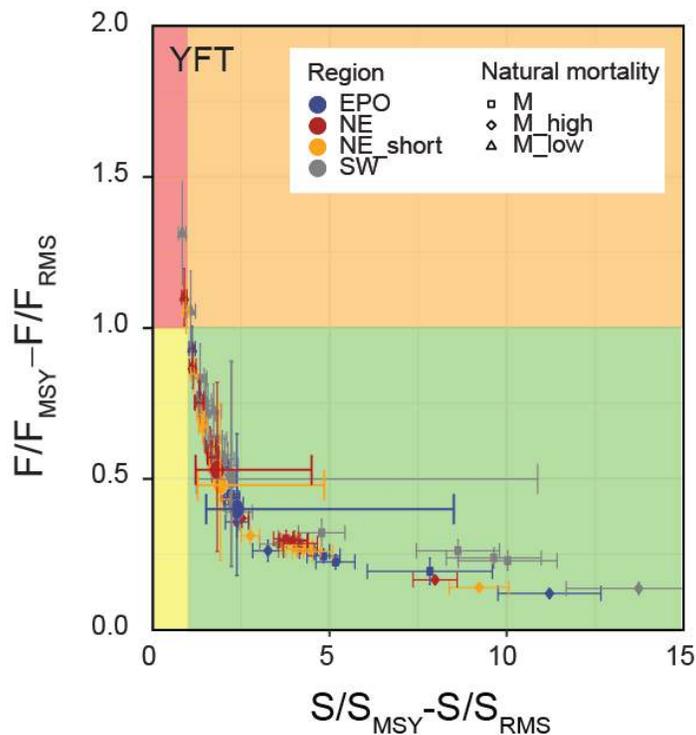
**FIGURE S-6.** Impact of the different fishing methods on the spawning biomass. Left panels: comparison of spawning biomass trajectory of a simulated population of yellowfin tuna that was never exploited (colored area) and that predicted by the stock assessment model (SB, yellow shaded area), and the impact of each fishing method (purse-seine on floating objects OBJ, also includes sorting discards and pole and line, purse-seine associated with dolphins DEL, purse-seine unassociated NOA and longline LL fisheries) for each stock structure hypothesis calculated from the base reference models with steepness of 1. Right panels: Proportional impacts.

## STOCK STATUS

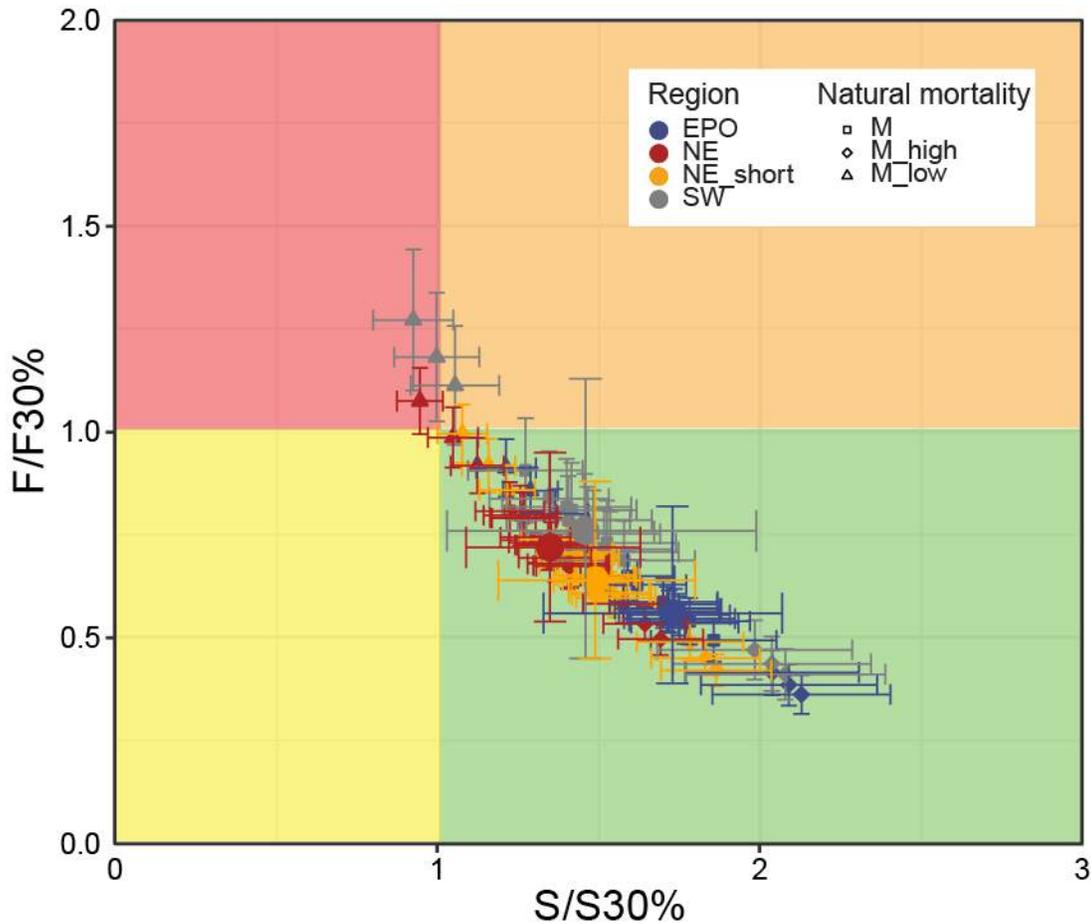
With respect to the IATTC interim target and limit reference points, all four spatial-structure hypotheses estimate the same stock status (Table S-1). The stock(s) is estimated to be well above the spawning biomass correspond to MSY ( $S_{MSY}$ ) and the Staff proposed MSY proxy  $S_{30\%}$  (SAC-15-05) with low probability of being below these. The fishing mortality is estimated to be well below the level corresponding to MSY and the MSY proxy  $F_{30\%}$  with low probability of being above these. The assessment estimates zero probability that the spawning biomass or fishing mortality limit reference points have been breached. The EPO model is the most optimistic.

The most pessimistic models are those with low natural mortality (Figures S-7A, S-7B and S-8). Some of these models estimate that the spawning biomass is below the  $S_{30\%}$  level and the fishing mortality is above the  $F_{30\%}$  level. The high natural mortality levels are generally the most optimistic.

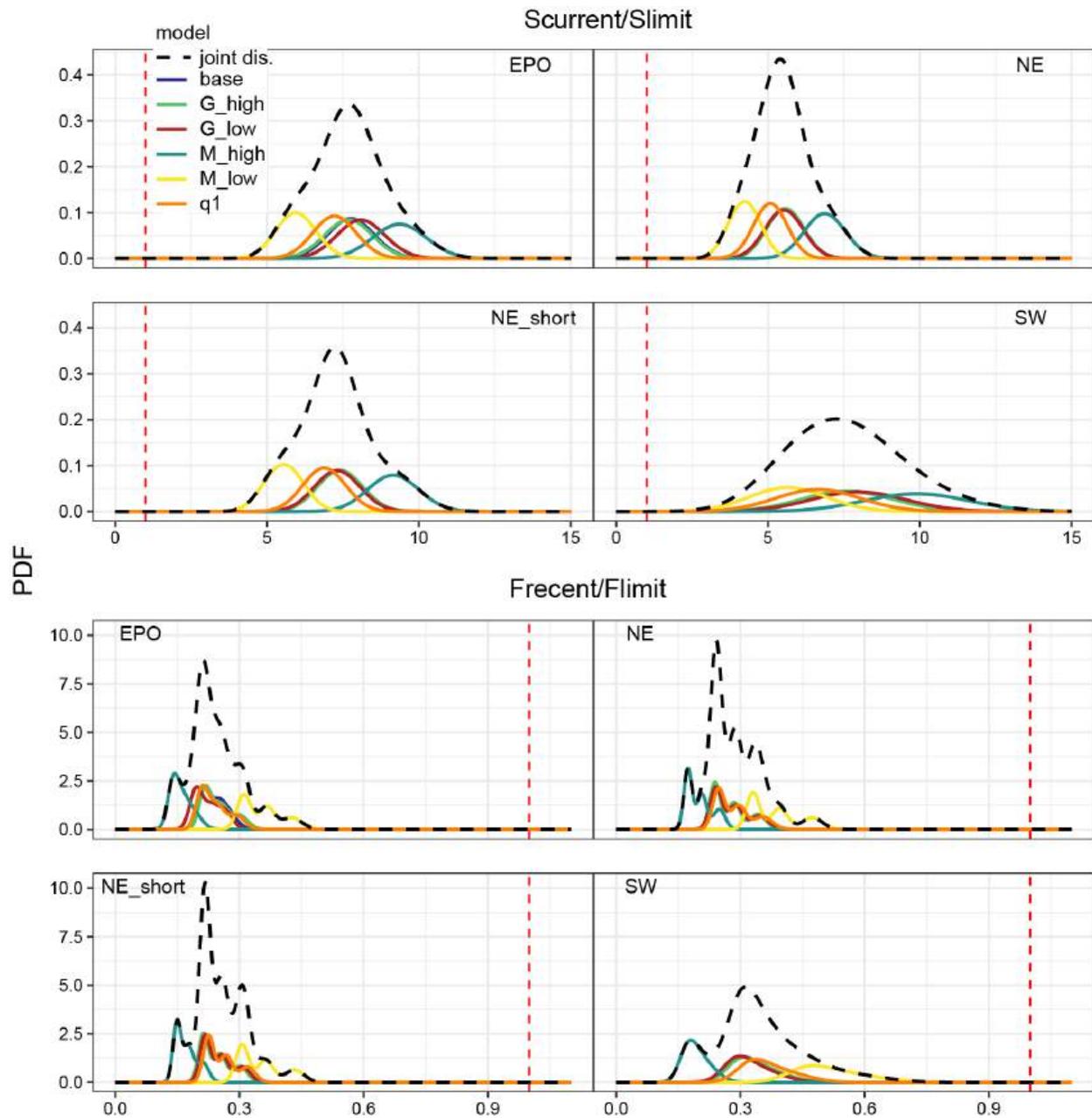
The estimates of the SBR (the ratio of the spawning biomass to the virgin spawning biomass) corresponding to MSY are low (generally below 20%, Table S-1) even though the highest fishing mortality is on older yellowfin. The value is higher with lower steepness of the stock-recruitment relationship and lower natural mortality. For example, the SW model with no relationship between stock size and recruitment (steepness equals 1) and high natural mortality has a value of 5% while the NE\_short model with steepness equal to 0.8 and low natural mortality has a value of 32%). The low level of SBR corresponding to MSY might be due to the assumptions about natural mortality declining with age (i.e., high M for juveniles).



**FIGURE S-7A.** Kobe plot of the most recent estimates of spawning biomass ( $S$ ) and fishing mortality ( $F$ ) relative to their target reference points ( $S_{MSY,d}$  and  $F_{MSY}$ ) for each hypothesis of spatial structure. Each dot is based on the average  $F$  over the most recent three years, 2021-2023, and the  $S$  for the first quarter of 2024 and the error bars represent the 80% confidence interval of model estimates. The larger dots represent the combined result for each spatial structure hypothesis.



**FIGURE S-7B.** Kobe plot of the most recent estimates of spawning biomass ( $S$ ) and fishing mortality ( $F$ ) relative to their proxy target reference points ( $30\%S_d$  and  $F_{30\%S_d}$ ) for each hypothesis of spatial structure. Each dot is based on the average  $F$  over the most recent three years, 2021-2023, and the  $S$  for the first quarter of 2024 and the error bars represent the 80% confidence interval of model estimates. The larger dots represent the combined result for each spatial structure hypothesis.



**FIGURE S-8.** The joint probability distributions for spawning biomass ( $S$ ) in the first quarter of 2024 and average fishing mortality ( $F$ ) in 2021-2023 relative to their limit reference points ( $S_{Limit}$  and  $F_{Limit}$ ). The distributions are provided for each of the four spatial structure hypotheses separated into different components (level 2 hypotheses). The level 3 hypotheses (steepness values) were integrated out.

**TABLE S-1.** Management quantities for yellowfin tuna in the EPO for each spatial structure hypothesis. The medians (or expected values \*) and probabilities were obtained from the joint probability distributions across models.

	<b>EPO</b>	<b>NE</b>	<b>NE_short</b>	<b>SW</b>
$SMSY/SO$ *	0.180	0.189	0.194	0.162
$SMSY\_d/SO\_d$ *	0.190	0.192	0.201	0.170
$F_{current}/F_{30\%SO\_d}$	0.559	0.718	0.643	0.757
$p(F_{current} > F_{30\%SO\_d})$	0.002	0.059	0.020	0.161
$F_{current}/F_{MSY}$	0.397	0.532	0.484	0.502
$p(F_{current} > F_{MSY})$	0.004	0.034	0.031	0.075
$F_{current}/F_{LIMIT}$	0.232	0.272	0.243	0.330
$p(F_{current} > F_{LIMIT})$	0.000	0.000	0.000	0.000
$S_{current}/30\%SO\_d$	1.73	1.35	1.49	1.46
$p(S_{current} < 30\%SO\_d)$	0.0000588	0.044	0.004	0.081
$S_{current}/SMSY\_d$	2.38	1.82	1.91	2.22
$p(S_{current} < SMSY\_d)$	0.000	0.000	0.000	0.000
$S_{current}/S_{LIMIT}$	7.67	5.43	7.23	7.48
$p(S_{current} < S_{LIMIT})$	0.000	0.000	0.000	0.000