

INTER-AMERICAN TROPICAL TUNA COMMISSION

SCIENTIFIC ADVISORY COMMITTEE

11TH MEETING

La Jolla, California (USA)

11-15 May 2020¹

DOCUMENT SAC-11-07 REV

**YELLOWFIN TUNA IN THE EASTERN PACIFIC OCEAN, 2019:
BENCHMARK ASSESSMENT**

Carolina Minte-Vera, Mark N. Maunder, Haikun Xu, Juan L. Valero, Cleridy E. Lennert-Cody, and
Alexandre Aires-da-Silva

CONTENTS

Executive summary	2
1. INTRODUCTION.....	3
1.1. Background	3
1.2. The new approach	4
2. DATA	5
2.1. Fisheries and ‘surveys’	5
2.2. Catch	8
2.3. Indices of abundance	10
2.4. Size-composition data	11
2.5. Conditional age-at-length	13
3. Assumptions and parameters.....	13
3.1. Biological and demographic information	13
3.2. Fisheries dynamics.....	16
3.3. Data weighting.....	17
3.4. Model diagnostics.....	17
4. REFERENCE MODELS.....	18
5. RESULTS	19
5.1. Model diagnostics.....	19
5.2. Stock assessment results	22
6. STOCK STATUS	25
6.1. Definition of reference points	25
7. FUTURE DIRECTIONS.....	26
ACKNOWLEDGEMENTS	27
REFERENCES	27
Appendix 1	61
Appendix 2	66
Appendix 3	79

¹ Postponed until a later date to be determined

EXECUTIVE SUMMARY

1. This year's benchmark assessment of yellowfin tuna in the eastern Pacific Ocean (EPO) is the basis for a risk analysis used to provide management advice (SAC-11-08, SAC-11-INF-J). The risk analysis encompasses alternative hypotheses on the states of nature. The hypotheses were developed in a hierarchical framework that addressed uncertainties and issues with previous assessments
2. Unlike previous assessments that relied on one base-case model with an assumed steepness of 1.0 for the stock-recruitment relationship, this benchmark assessment integrates 12 reference models, each with four steepness assumptions (0.7, 0.8, 0.9, and 1.0), for a total of 48 models.
3. The 12 reference models are developed within a hierarchical framework, and combine components that address three major uncertainties in the previous assessment: a) oversensitivity to the inclusion of new data, mainly from the longline index of abundance and inconsistencies between that index and the purse-seine ones, b) misfit to length-composition data for the fishery that is assumed to have asymptotic selectivity, and c) the steepness of the stock-recruitment relationship. In addition, new fishery definitions were implemented, and spline selectivity functions were adopted for most fisheries.
4. The 48 models for this assessment indicate that:
 - a. At the beginning of 2020, the spawning biomass (S) of yellowfin ranged from 49% to 219% of the level at dynamic MSY ($S_{MSY,d}$); 12 models suggest that it was below that level.
 - b. During 2017-2019 the fishing mortality (F) of yellowfin ranged from 40% to 168% of the level at MSY (F_{MSY}); 14 models suggest that it was above that level.
 - c. At the beginning of 2020, the spawning biomass (S) of yellowfin ranged from 145% to 345% of the limit reference level (S_{LIMIT}); no models suggest that it was below that limit.
 - d. During 2017-2019, the fishing mortality of yellowfin ranged 22% to 65% of the limit reference level (F_{LIMIT}); no models suggest that it was above that limit.
5. Every reference model suggests that lower steepness values correspond to more pessimistic estimates of stock status: lower S and higher F relative to the reference points. However, models that assume fixed growth, a linear relationship between the index of abundance and population abundance, no changes in selectivity through time, and asymptotic selectivity for the purse-seine fishery that catches the largest fish (BASE) estimated that, regardless of steepness, the stock was below the MSY level ($S < S_{MSY,d}$) and the fishing mortality was above that level ($F > F_{MSY}$) at the beginning of 2020. Conversely, models that assume dome-shape selectivity for the purse-seine fishery that catches the largest fish sizes (DS, TBM.DS, TBE.DS, DDQ.DS) estimate the opposite. The stock status at the beginning of 2020 estimated by the remaining models depends on the value assumed for steepness.
6. A key uncertainty not addressed in this assessment is the spatial structure of the stock of yellowfin tuna in the EPO. Future work to further improve the assessment will focus on it.
7. The results from the reference models are combined in a risk analysis to provide management advice (SAC-11-08).

1. INTRODUCTION

This report presents the results of a benchmark stock assessment² of yellowfin tuna (*Thunnus albacares*) in the eastern Pacific Ocean (EPO), conducted using Stock Synthesis (version 3.30.15), an integrated statistical age-structured stock assessment modeling platform. It is the first assessment of the species undertaken by the Commission's scientific staff under the 2018 [Work plan to improve stock assessments of tropical tunas](#) and, although it uses the same modeling platform, the methodology is quite different. The assessment now forms the foundation of a risk analysis, which takes uncertainty into account explicitly when determining stock status and formulating management advice. All model input files and output results for this benchmark assessment are available in [html and pdf formats](#)

1.1 Background

In recent years, some problems and sources of uncertainty had arisen in the staff's assessment of yellowfin, leading to the staff not considering it reliable enough for management advice, and ultimately its inclusion in the [work plan to improve stock assessments of tropical tunas](#) in 2019. The main problem was that the assessment results became overly sensitive to the inclusion of new data, in particular recent observations for the indices of relative abundance from the longline fishery ([SAC-10 INF-F](#)). As part of the [work plan](#), in 2019 a [workshop](#) and collaborative work with the main longline CPCs³ (China, Japan, Korea, Chinese Taipei) were conducted to better understand the longline data used in the assessments of both yellowfin and bigeye tunas. As a result, the over-sensitivity to the inclusion of the new data was found to be partially due to the contraction of both the spatial extent and the fishing effort of the Japanese longline fishery (whose data was used to estimate the longline index of relative abundance), resulting in less accurate and precise indices of relative abundance for recent years.

Other issues were identified which also related to the longline information, such as a change in length composition data towards larger fish ([SAC-10 INF-F](#)) while the longline index showed a decline in recent years. The collaborative work with longline CPCs suggested that this may be due to changes in the fishery (*e.g.* gear or operation), but the recent increase in the mean size of yellowfin is also seen in the catches of the dolphin-associated purse-seine fisheries and some unassociated purse-seine fisheries. This gives plausibility to the hypothesis that changes in some processes (*e.g.* selectivity) or model misspecification (*e.g.* growth) may be related to this increase. In some models of this assessment, growth and time blocks in selectivity are estimated to represent these hypotheses.

Another influential issue was inconsistencies between the indices of abundance based on standardized CPUE from the longline fishery and nominal CPUE from the dolphin-associated purse-seine fishery; the stock assessment model did not fit either of them well. A new spatio-temporal modeling framework was developed and applied to both CPUE data sets to create new indices, but the inconsistencies and lack of fit remained.

One major source of uncertainty, and potentially also the explanation for the inconsistencies between the indices, is the possibility of the EPO yellowfin stock having a spatial structure of that is not captured in the model. The staff's work for the 2nd [external review](#) of the yellowfin stock assessment included exploring separate models for hypothetical "southern" and "northern" stocks, but the review panel concluded that the *"evidence supporting a two-stock hypothesis was thought to be suggestive, rather than conclusive"*, and that *"there was further evidence suggesting that [yellowfin] in the EPO was somewhere between a single, well-mixed stock and the two independent stocks"*. The panel suggested various research avenues to explore to better account for stock structure in the assessment (one-stock and two-stock hypotheses).

² "Benchmark" stock assessments are a full analysis of model assumptions, methodologies and/or data sources, whereas in an "update" assessment only the data used in the assessment are updated. .

³ Members and Co-operating Non-Members of the IATTC

This assessment was conducted as if there were a single stock of yellowfin in the EPO, but the staff will continue to evaluate alternative spatial structure hypotheses. To minimize potential biases resulting from ignoring spatial heterogeneity, this assessment is mainly fitted to datasets representative of the core area of the fisheries, which is north of 5°N, where the bulk of the catch is taken. The catches south of 5°N are fully accounted for in the assessment, but the model does not fit to indices of abundance and length-composition data from fisheries in that area. This limits the influence of data that may be representative of another population unit.

There is also uncertainty about the stock-recruitment relationship. In this assessment, as in previous ones, a Beverton-Holt stock-recruitment relationship is assumed, but the uncertainty in steepness is taken into account. Steepness (h) is a parameter that specifies the average effect on recruitment of a reduction of the spawning stock. Previous assessments of EPO yellowfin have consistently presented analyses of the sensitivity of their results to different assumptions about steepness, but only to show the impact of these assumptions on estimated management quantities; the results of these analyses were not explicitly incorporated into the management advice. In this assessment, models assuming h values from 0.7 to 1.0 are included.

Neither external review, of the [bigeye](#) or the [yellowfin](#) assessment, identified a specific replacement for the staff's modelling methodology, but both panels suggested a variety of alternatives for the staff to consider, particularly those that incorporate model uncertainty to derive information for management advice.

1.2 The new approach

This 2020 benchmark assessment of yellowfin tuna in the EPO, and the companion assessment of bigeye tuna ([SAC-11-06](#)), represent a new approach to stock assessments by the staff. Previously, a 'best assessment' approach was used for the evaluation of stock status using a single 'base-case' model. The new assessments are based on 'risk analysis' methodologies, which use several *reference models* to represent various plausible *states of nature* (assumptions) about the biology of the fish, the productivity of the stocks, and/or the operation of the fisheries, and takes into account the different results, thus effectively incorporating uncertainty into the formulation of management advice⁴. This change, which represents a paradigm shift at IATTC, both for the staff's work and for the Commission's decision-making regarding the conservation of tropical tunas, also allows the staff to evaluate explicitly the probability statements specified in the IATTC harvest control rule for tropical tunas established in Resolution [C-16-02](#).

This new approach to formulating advice for the management of tropical tuna fisheries includes the following four components:

1. Two **stock assessment reports**, for yellowfin (this document) and bigeye ([SAC-11-06](#)), presenting the results from all models for each species (model fits, diagnostics, stock status);
2. A **risk analysis** ([SAC-11-08](#)), assessing the consequences of using each model and the models combined as a basis for managing the fishery for tropical tunas by quantifying the probability of meeting the target and limit reference points specified in the IATTC harvest control rule;
3. **Stock status indicators** ([SAC-11-05](#)) for all three tropical tuna species (yellowfin, bigeye, skipjack); and
4. The **staff's recommendations** ([SAC-11-15](#)) for the conservation of tropical tunas, based on the above.

⁴ See SAC-11 INF-F (Maunder *et al.* 2020a) for a description of the technical details of the risk analysis, using bigeye as a case study.

1. DATA

2.1 Fisheries and 'surveys'

The fisheries are defined using several criteria, among them their geographical area of operation. This is consistent with the 'areas-as-fleets' approach, and allows spatial information to be taken into account without explicitly constructing a spatial model.

Fishery-independent surveys are the gold standard to gather data to assess an exploited population. Due to their nature, most tuna fisheries worldwide have no surveys. The EPO is no exception: all data available for assessing the stocks are obtained from the fishery. By a process of statistical standardization of fisheries-dependent data, however, an index of abundance and its associated length composition were estimated. Within the stock assessment models, the index and its length composition are treated as 'survey' data, thus modelled as having no catch and having a different selectivity than the fisheries.

The fisheries and 'surveys' defined for this assessment are illustrated in Figure 1, summarized in Table 1, and described in detail below.

2.1.1 Fisheries

Thirty-eight fisheries are defined for the stock assessment of yellowfin tuna in the EPO, classified by gear (purse-seine, longline, or pole-and-line), purse-seine set type (floating object, unassociated, or dolphin), unit of catch (number or weight), quarter within a year, and geographical area of operation (Figure 1, Table 1).

All the fisheries in this assessment, except the discard and pole-and-line fisheries, are defined by their area of operation delimited based on spatial patterns in the length-frequency data, identified with regression tree analyses (Lennert-Cody *et al.* 2010). The data used in the analyses were, for purse-seine fisheries, from the IATTC's port-sampling program (Section 2.4.1.a) during 2000-2018, with a 5°x5° spatial resolution, and for longline fisheries, from the National Research Institute of Far Seas Fisheries (NRIFSF) of Japan's annual data, aggregated in space and time, for the Japanese fleet during 2000-2009, with a 5°x10° spatial resolution. The predictors for the tree analyses were quarter; cyclic quarter (cyclic combinations of quarters, such as Q1 and Q4 vs. Q2 and Q3, Q1, Q2, and Q4 vs. Q3, *etc.*); 5° latitude; 5° (10°) longitude. The response variable, which was multivariate, was the proportion of individuals in each 10-cm length interval, with plus groups for the smallest and largest fish (≤ 39 cm; 40-49 cm; (...); 150-159 cm; ≥ 160 cm). For both the floating-object and longline fisheries, cyclic quarter (specifically, Q1 and Q4 vs. Q2 and Q3) was found to be an important predictor for explaining variability in the length frequencies, so seasonal fisheries were defined. Longline catches are reported in number by some fleets and in weight by others, so two longline fisheries, one in number and one in weight, are defined for each area-cyclic quarter combination.

The pole-and-line (LP) fishery represents a small portion of the catches, declining over time, so was treated as one homogeneous fishery for the whole EPO (F22). The fisheries used to model discards (F25-F28) have the same structure as in the previous assessment ([SAC-10-07](#)).

2.1.2 'Surveys'

In Stock Synthesis, a 'survey' is modeled as a fleet that has data, such as indices of abundance and age/length compositions, but no catch. One 'survey' used in this assessment was based on data from the dolphin-associated purse-seine fishery (section 2.3); additional 'surveys' based on data from longline fisheries were also defined, and used in [preparation](#) for the benchmark assessment presented at the [external review](#). The longline surveys are not fitted in this assessment, because their datasets are not representative of the 'core' area of the longline fishery north of 5°N.

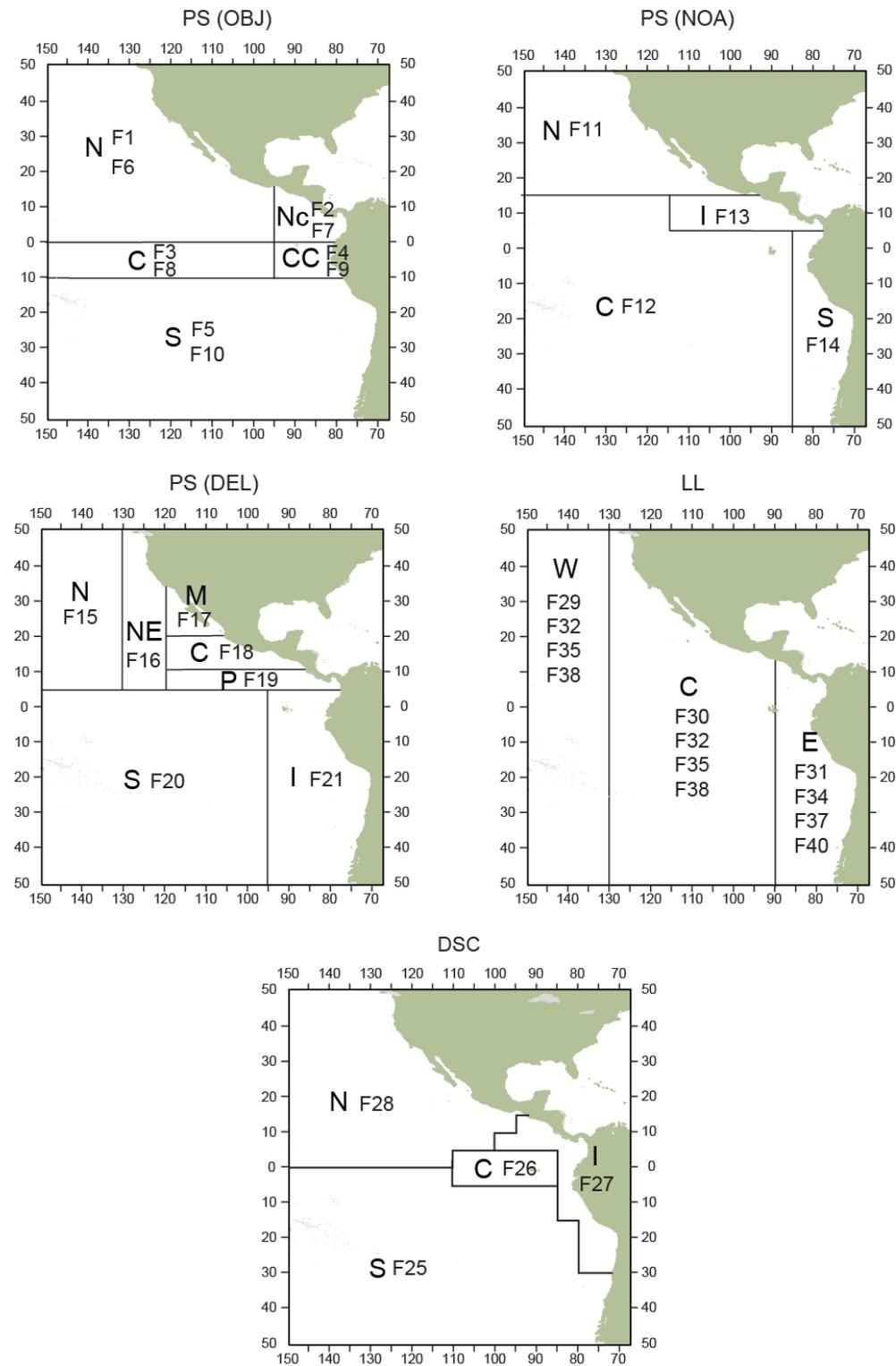


FIGURE 1. Areas corresponding to the fishery definitions (Table 1) used in the stock assessment of yellowfin tuna in the EPO in 2019.

FIGURA 1. Áreas correspondientes a las definiciones de las pesquerías (Tabla 1) usadas en la evaluación de la población de aleta amarilla en el OPO en 2019.

TABLE 1. Fisheries defined for the stock assessment of yellowfin tuna in the EPO in 2019. **Gear:** PS: purse seine; LP: pole and line; LL: longline; **PS set type:** OBJ: floating object; NOA: unassociated; DEL: dolphin; **Area:** see Figure 1; **Discards:** see Section 2.2.3; **Surveys:** see Section 2.1.2.

TABLA 1. Pesquerías definidas para la evaluación de la población de atún aleta amarilla en el OPO en 2019. **Arte:** PS: red de cerco; LP: caña y anzuelo; LL: palangre; **Tipo de lance PS:** OBJ: objeto flotante; NOA: no asociado; DEL: delfín; **Área:** ver Figura 1; **Descartes:** ver sección 2.2.3; **Estudios:** ver sección 2.1.2.

	Fishery	Gear	Set type	Quarters	Area	Catch data	Unit			
1.	F 1	PS	OBJ	1, 4	1- N	Retained catch + discards (inefficiencies)	t			
2.	F 2				2- Nc					
3.	F 3				3- C					
4.	F 4				4 -Cc					
5.	F 5				5- S					
6.	F 6			2, 3	1- N					
7.	F 7				2- Nc					
8.	F 8				3- C					
9.	F 9				4 -Cc					
10.	F 10				5- S					
11.	F 11		NOA	All	1- N	Retained catch + discards (all)				
12.	F 12				2- C					
13.	F 13				3- I					
14.	F 14				4- S					
15.	F 15		DEL	All	1- N					
16.	F 16				2- NE					
17.	F 17				3- M					
18.	F 18				4- C					
19.	F 19				5- P					
20.	F 20				6- S					
21.	F 21				7- I					
22.	F 22	LP	--	All	All	Retained catch	t			
23.	F 25	PS	OBJ	All	S	Discards (size-sorting)	t			
24.	F 26				C					
25.	F 27				I					
26.	F 28				N					
27.	F 29	LL	-	1, 4	W	Retained catch	1,000s			
28.	F 30				C					
29.	F 31				E					
30.	F 32				W					
31.	F 33			C						
32.	F 34			E						
33.	F 35			W						
34.	F 36			C						
35.	F 37			E						
36.	F 38			W						
37.	F 39			C						
38.	F 40			E						
	Survey			Gear	Years		Quarters		Observation	Unit
	S 1			PS	1984-2019		All		Used in assessment	t
	S 2a			LL	1984-1992		1, 4		Not used	1,000
	S 2b						2, 3			
	S 2c	1995-2018	1, 4							
	S 2d		2, 3							

2.2 Catch

The following types of catch data are defined for this assessment:

- **Retained:** catch retained aboard the vessel;
- **Discarded:** catches not retained aboard the vessel;
- **Total:** retained catch + discards;
- **Unloading:** retained catch unloaded from the vessel.

2.2.1 Purse seine

The information used to estimate the total catch by species comes from four main sources: in order of importance, canneries, on-board observers, vessel logbooks, and in-port sampling by IATTC staff. If landing information from canneries is unavailable, catch information in the observer or vessel logbook databases, in that order, is used instead. The observer and logbook databases also contain other information about the catches, such as location and date caught, set type (on dolphin-associated tunas (DEL), on floating objects (OBJ), and on unassociated tunas (NOA) and vessel carrying capacity (<364 t (Classes 1-5) and ≥364 t (Class 6)); 'year' is the only ancillary information available in the unloading database. Additionally, since 2000, the port-sampling program for collecting length-composition data has also provided information on species composition (see section 2.4.1 a.).

For this assessment, EPO total catches by species were estimated by catch strata, and then aggregated across to obtain quarterly estimates for each fishery. The catch strata are defined as the combination of area, month, set type, and vessel fish-carrying capacity. The method used to estimate the species composition of the catch depends on the sources of information available. Estimates prior to 2000 are based on the recorded species totals in the unloading, observer or logbook data, as applicable. To correct for underestimated bigeye catches, a factor that adjusts the catches of all three species, based on the port-sampling data from 2000-2004, is applied. The adjusted species totals are prorated to catch strata using the ancillary information in the observer and logbook databases. Since 2000, the port-sampling data have been used to determine the species composition of the total catch. The total catch of all three species combined (from unloading, observer and logbook data) is prorated to catch strata, using the information in the observer and logbook databases. The port-sampling data on the species and size composition of the catch are then used to estimate the catch of each species by catch stratum. Detailed explanations of the sampling and estimators can be found in the appendix of Suter (2010) and in [WSBET-02-06](#).

2.2.2 Longline

The IATTC staff does not collect data on longline catches directly; they are reported annually to the IATTC by Members and Cooperating non-Members (CPCs), pursuant to Resolution [C-03-05](#) on data provision. Catches are reported by species, but the availability and format of the data vary among fleets: the principal fleets report catch and effort aggregated by 5° cell-month. IATTC databases include data on the spatial and temporal distributions of longline catches of yellowfin in the EPO by the fleets of distant-water CPCs (China, Chinese Taipei, French Polynesia, Japan, Korea, Vanuatu) and coastal CPCs (principally Mexico and the United States). For this assessment, these data are aggregated by area of operation of the fishery (Figure 1) and cyclic quarter (Table 1). Because longline catches are reported in number by some fleets and in weight by others, two longline fisheries, one in number and one in weight, are defined for each combination of area and cyclic quarter (Q4, Q1 vs. Q2, Q3), and the numbers are converted to weights internally in the assessment model.

Updated and new catch data for the longline fisheries (Fisheries 29-40), available to the IATTC staff as of 10 April 2020, were incorporated into the current assessment. New or updated catch data were available for Vanuatu (2018), Chinese Taipei (2016-2018), French Polynesia (2018), China (2018), Japan (2016-2018), Korea (2018) and the United States (2014-2018). For 2019, and for other years when catch may not

be available, catches were set equal, by CPC, to the last year for which catch data were available. For fleets that reported catch aggregated by year and space, the data was disaggregated using the proportion of catches by quarter and area for the closest year for which data on that resolution were available. The catches of a coastal CPC that reported aggregated catches were added to the area which contained that CPC's Exclusive Economic Zone (EEZ). The algorithm to calculate the catches is described in [WSBET-02-03](#).

2.2.3 Discards

Two types of discards are considered, those resulting from inefficiencies in the fishing process and those related to the sorting of catches. Examples of inefficiency are catch from a set exceeding the remaining storage capacity of the fishing vessel or dumping unwanted bycatch species, while catch sorting is assumed to occur when fishers discard tuna that are under a certain size. Both types of discard by purse-seine are estimated, whereas discards by the longline fisheries cannot be estimated with the minimal data available due to the low observer coverage, so it is assumed that the retained catch represents the total catch (Table 1).

For the purse-seine fishery, the amount of yellowfin discarded, regardless of the reason, is estimated with information collected by IATTC or national observers using the methods in Maunder and Watters (2003). No observer data are available to estimate discards prior to 1993, and it is assumed that there were no discards before that time. Also, there are periods for which observer data are not sufficient to estimate the discards, in which case it is assumed that the discard rate (discards/retained catches) is equal to the discard rate for the same quarter in the previous year or, if quarterly data are not available, a proximate year. Removals by Fisheries 1-10 (purse-seine on floating objects) are retained catch plus some discards resulting from inefficiencies in the fishing process. The removals by Fisheries 11-14 (purse-seine unassociated) are retained catch, plus some discards resulting from inefficiencies in the fishing process and from sorting the catch. Discards that result from the process of sorting the catches in the floating-object fisheries are treated separately (Fisheries 25-28), following the rationale of Maunder and Watters (2001). These discards are assumed to be composed only of fish that are 1-3 quarters old. Sorting is infrequent in the other purse-seine fisheries.

2.2.4 Catch and discard trends

Yellowfin has been fished in the EPO since the early 1900s (Estes 1983). Prior to the 1950s, the fishery occurred mostly within 250 miles of the coast or around islands and seamounts, and was done mainly by pole-and-line vessels (Peterson and Bayliff 1985). In the 1950s the longline fisheries started expanding in the EPO from the western and central Pacific, reaching coastal areas around the mid-1960s, mainly targeting bigeye tuna, but catching yellowfin as a secondary species (Shimada and Schaefer 1956; Matsu-moto and Bayliff 2008).

The purse-seine fisheries, mainly associated with dolphins, became the main fishing method in the 1960's, and has continue since (Figure 2). The main dolphin-associated fisheries are close to Central America and southern Mexico (F18, F19; Figure 1). The purse-seine fishery associated with floating objects has been important since the 1970s in areas north of the equator (F1, F2, F6, F7; Figures 1 and 2) and close to the coast of South America, between 10°S and the equator (F4, F9). The fisheries on floating objects had a widespread expansion in the EPO after 1992. In the last 15 years, the number of sets on floating objects has been steadily increasing towards its current record level([SAC-11-05](#)). Catches from floating-object fisheries in areas C (F3, F8) and S (F5, F10) increased after 1992. The main unassociated purse-seine fisheries have been in the north (F11) and, after 2010, in the inshore and south coastal EPO (F13, F14). 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).

Longline catches represent a small proportion of the total catches of yellowfin tuna in the EPO (Figure 2).

The main longline fishing areas have always been in the western EPO (F29, F32, F35, F38), where a decline occurred in the late 2000s, but an increasing trend is apparent since 2010. The longline catches in the eastern EPO (F31, F34, F37, F40) area are characterized by a marked seasonality, and have declined in recent years.

2.3 Indices of abundance

Although both purse-seine and longline indices of abundance are available for yellowfin in the EPO, this assessment includes only the purse-seine index. Inconsistency between the index based on Japanese longline CPUE and the indices based on dolphin associated purse-seine CPUE have been found. Extensive work was done in collaboration with the longline CPCs to better understand the data, incorporate new data, and conduct new analyses. A [workshop](#) was conducted, and scientists from Japan and Korea collaborated with the staff to further address the issues. A new spatio-temporal modeling framework was developed and applied to the CPUE data to create new indices (Xu *et al.* 2019), but the inconsistencies remained unresolved. The stock assessment model cannot adequately fit both types of indices simultaneously. The purse-seine index was selected for this benchmark assessment, because the longline catches represent a small proportion of the yellowfin catches, and the distribution of the Japanese fleet has been contracting towards the western EPO, away from the “core” catch areas for yellowfin (see Figure A1 in [SAC-11-06](#)). It is noted that, given these spatial changes, one potential explanation for the inconsistencies among indices in the model is the possibility of spatial structure in the EPO yellowfin population.

2.3.1 Data selection

The data used to construct the index are the set-by-set catch and effort observations from purse-seine vessels. On-board observers of the Agreement on the International Dolphin Conservation Program (AIDCP) have been collecting these data for large purse-seine vessels (fish-carrying capacity >363 t) since 1992 (Joseph 1994; Scott *et al.* 2016). Logbook data were used for previous trips by such vessels, for which no observer data were available.

Because it is not possible to separate searching effort by set type, and to limit the data used to standardize effort to vessels that fish preferentially for dolphin-associated tuna, the following procedure was used to limit the standardization to the main dolphin-associated fishing grounds and vessels. Only data for 1°x1° sampling cells north of 5°N with at least 30 years of data during 1985-2019 were included, and only vessels that made at least 75% of their sets on dolphin-associated tunas during at least 10 of 18 years of data coverage were selected. The 52 selected vessels were classified as “dolphin-associated vessels”, and their data were used to obtain the index (Appendix 1, Figure A1).

2.3.2 Standardization procedure

The standardization of the catch and effort data y was conducted using the R library VAST (version 3.0.0) (Thorson and Barnett 2017, Xu *et al.* 2019, Maunder *et al.* 2020b). VAST fits a delta-generalized linear mixed-effects spatiotemporal model to data. It models separately the encounter probability and positive catch rate, which are assumed to have a logit and log link, respectively, and combines the results to produce the final estimates. There are several advantages of using mixed-effects spatiotemporal models over the fixed-effects generalized linear models conventionally used in CPUE standardization. First, the estimation of spatiotemporal correlations allows for the prediction of catch rates in unfished locations based on the information from neighboring areas/times. Second, the uncertainty estimates take into consideration the spatial coverage and sample size. Third, the final estimates are naturally weighted by the area related to each knot in the spatial domain, rather than by the sample size. Both the encounter probability and the catch rates are modeled with linear predictors that include an intercept term (year-quarter effect), vessel effects on catchability and spatial effect (Xu *et al.* 2019). The spatial effect is represented by a mesh of 200 knots. The model converged (gradient = 0.0004) with a positive definite Hessian.

The index shows four noticeable periods (Figures 3 and 4):

1. 1984-2000: high abundance, with a spike in 1996
2. 2001-mid-2003: marked increase in abundance
3. Mid-2003-early 2015: decreased abundance
4. Early 2015-2019: lower abundance

The lower spatial coverage and sample size in the early years (Figure 4) resulted in a higher coefficient of variation (CV) for those years.

2.4 Size-composition data

2.4.1 Fisheries

2.4.1.1 Purse-seine

The length-frequency data for the purse-seine fisheries are obtained through the sampling program conducted by IATTC personnel at ports of landing in Ecuador, Mexico, Panama, and Venezuela. The ancillary information available in the port-sampling database is determined by the governing protocol (Tomlinson 2002, Suter 2010), which specifies the strata from which samples are collected: fish-carrying capacity of the vessel, type of set (DEL, NOA, OBJ), month and area of catch (13 areas; see Figure 1 in [WSBET-02-06](#)). Wells are the primary sampling unit within a stratum, with unequal numbers of wells sampled per stratum, and fish within a well are the secondary sampling unit. Sampling at both stages is largely opportunistic, except that a well is sampled only if all the catch within it came from the same stratum. This restriction can result in sets with large catches predominating in the samples (Lennert-Cody and Tomlinson 2010). More than one well may be sampled per vessel if the catch in the other wells comes from different strata, but typically only one or two wells per trip are sampled. For large and small purse-seine vessels, about 50%-60% and 10-20% of trips, respectively, have typically been sampled per year, for a total of over 800 wells sampled in most years (IATTC 2010a; Vogel 2014). The sampling coverage in terms of percentage of the catch is lower (SAC-02-10). The sampling areas were designed for yellowfin prior to the development of the fishery on FADs. Since 2000, both the 5° cell and the sampling area have been recorded for almost all samples (Lennert-Cody *et al.* 2012); the 5° cell has been recovered for many samples prior to 2000. Ideally, 50 fish of each species in the sampled well were measured, and since 2000 samplers alternate between counting fish by species and measuring fish for length. The protocol varies to some extent with the set type associated with the catch in the well and with the species composition of the catch in the well, as recorded by the observer or in the vessel's logbook. More details on the port sampling program can be found in the Appendix of Suter (2010) and in [WSBET-02-06](#).

As with the species composition, the size composition of the catch, in numbers of fish by 1-cm length interval, is estimated by stratum and then aggregated across strata to obtain quarterly estimates for each fishery. The estimated number of fish are then converted to proportion of fish at length for the assessment. The estimated numbers at length are obtained by multiplying the well-level estimates of the proportion at length, combined across sampled wells, by the estimated total catch in numbers for the species in the stratum. Since 2000, the well estimates of proportions at length make use of both the species counts and the length-measurement data. Details of the estimators can be found in [WSBET-02-06](#).

The pole-and-line fishery and purse-seine fisheries on floating objects catch the smallest fish (Figures 5A and 5B), and the longline fisheries and dolphin-associated purse-seine fishery catch the largest fish. The average size of the fish caught by the purse-seine fisheries shows trends over time (Figure 5B). The sample sizes for the dolphin-associated purse-seine fisheries has been slightly decreasing over time, while those for the floating object fisheries have increase steadily since the mid-1990s.

2.4.1.2 Palangre

The length-composition data for longline fisheries were used in preliminary models runs to obtain the best fit for the asymptotic selectivity of the longline fisheries. All models composing the reference set (see section 4. Reference models) were not fitted to these data because of the potential spatial structure of the yellowfin tuna population in the EPO and the focus of the assessment is on the portion of the population exploited by the purse-seine fishery.

The length-composition data for longline fisheries in this assessment are based on 1) new monthly 1°x1° length-frequency data for the Japanese commercial fleet; 2) new monthly 1°x1° catch and effort data for individual Japanese commercial vessels; and 3) 5°x5° quarterly longline catch data reported by CPCs. The length-composition data should be representative of all longline catches, so the monthly 1°x1° length-frequency observations are raised to the fishery catch in a quarterly 5°x5° in the same strata, as follows:

1. Raise monthly Japanese 1°x1° length-frequency data (with 1 or 2 cm resolution) to total Japanese catch in the same strata;
2. Aggregate the raised data from step 1 to quarterly 5°x5° catch;
3. Raise the aggregated length-frequencies from step 2 to the total catch of all CPCs in the same strata;
4. Aggregate the raised length-frequencies from step 3 in the longline fisheries 29-35 in Table 1.

The length compositions, from 20 cm to 198+ cm, are aggregated at 2-cm intervals, and their input sample sizes are computed as the total number of fish sampled divided by 100. The input sample sizes for every longline fishery have decreased continuously since the mid-1990s. and reached very low values in most recent quarters due to the decline in the catches by the Japanese fleet and the switch in sampling strategy from crew samplers to on-board observers, and fewer fish have been sampled..

2.4.2 Survey

The length frequencies of yellowfin associated with the index of abundance (“survey”) were also obtained from the standardization of the data from the dolphin-associated purse-seine fishery using VAST, with the inclusion of a multivariate response variable (Thorson and Haltuch 2018, Maunder *et al.* 2020b). The data used were the length frequencies collected by the port-sampling program. The length frequencies, raised to the well catch, were aggregated by quarter, 5° cell and set type. The aggregated data were raised to the catch in a stratum using data from the observer and logbook databases. Strata were defined as quarter-5° cell combinations. The vessel and spatial cell selection criterion was the same as the CPUE. The multivariate response variable was length-specific catch rate (in ton day⁻¹ fished). The length frequency classes were defined by 10 cm intervals, from 20 to 190 cm.

The standardization model treats the encounter probability and positive catch rate separately, with logit and log links, respectively. The linear predictors are spatial and the temporal (year-quarter) components. The spatial component is represented by 30 spatial knots (that aggregate the 5° cells to improve computational efficiency). The sum of the indices by length class were similar to the overall index (section 2.3), indicating that the standardized length frequencies are a good representation of the length classes represented in the index of abundance. The model converged (gradient = 0.000006) with a positive definite Hessian.

The classes with largest frequencies ranged from 40 to 160 cm (Figure 6) with most lengths between 70 and 120 cm, except in two periods: 1) 2002-2007, when an increase in the proportion of small fish (<70 cm) was maintained for several consecutive quarters; 2) 2015-2019, when the proportion of large fish (>120cm) increased (Figure 6).

2.5 Conditional age-at-length

Age and length data (Wild 1986) were used to provide information when growth is estimated in the stock assessment model. Wild's data consist of ages, based on counts of daily increments in otoliths, and lengths for 196 fish collected between 1977 and 1979. The sampling was conducted by collecting 15 fish in each 10-cm interval in the length range of 30 to 170 cm. For the largest size ranges, Wild was unable to complete the sample size of 15 fish, due to the scarcity of fish of those sizes or unreadability of the otoliths (Wild 1986). The daily periodicity of the rings has been validated for fish from 25 to 146 cm (Wild and Foreman, Yamanaka 1990, Wild *et al* 1995) and in larvae up to 16 days after hatching (3-7 mm standard length) (Wexler *et al.* 2001). The maximum age obtained was 4 years (Wild 1986). The age and length data were included as frequencies of ages (in quarters) conditional on length (10-cm class) and not disaggregated by sex, coming from the year 1985 and F18, a dolphin-associated purse-seine fishery that catches fish of a wide range of sizes (Figure 5A), in the models that estimate growth parameters. Figure 7 shows the age frequency conditional on length class and the fixed assumption for growth (see 3.1.1).

2. ASSUMPTIONS AND PARAMETERS

An integrated statistical age-structured stock assessment modelling framework was used to carry out the benchmark assessment of yellowfin tuna in the EPO for 2019 (*Stock Synthesis* Version V3.30.15;_2020_03_26, Methot and Wetzel 2013). Two subsequent unreleased versions (V3.30.15.03-opt and V3.30.15.04-safe, provided by Rick Methot, NOAA Fisheries) were used to estimate the variability of $F_{current}/F_{MSY}$, $F_{current}/F_{LIMIT}$ and dynamic S_0 . Francis weights, recruitment deviation bias corrections and other auxiliary quantities and graphs were computed using the R library *r4ss* (version 1.38.0), the set of packages from *tidyverse* (1.3.0) and original code available from the IATTC repository [IATTCassessment](#).

The model period is 1984-2019. The start year differs from the previous stock assessments, which started in 1975, because data from the purse-seine fishery before 1984 with spatial information necessary to standardize the index and length frequencies are limited. The time step of the model is a quarter, 30 age classes are defined, from 0 quarters to 29+ quarters (7.25 years). The population size structure was defined in 2-cm intervals from 2 to 200+ cm. The model is structured by sex, with sex-specific natural mortality. The size compositions are defined using 2-cm intervals, from 20 to 198+ cm, for the fisheries, and 10-cm intervals, from 20 to 190 cm, for the 'survey'. The models are fitted to catches (with high precision), relative abundance indices, and size composition data. Models that estimated growth were also fitted to conditional age-at-length data. The observed total catches were assumed to be unbiased and relatively precise and were fitted assuming a lognormal error distribution with standard error (SE) of 0.01.

3.1 Biological and demographic information

3.1.1 Growth

The average length-at-age is assumed to follow a Schnute-Richards curve (Richards 1959; Schnute 1981) reparameterized with L_1 , L_2 , a_1 , a_2 as implemented in *Stock Synthesis* (Methot and Wetzel 2013,):

$$L_a = L_1^b + (L_2^b - L_1^b) \left(\frac{1 - \exp(-K(a - a_1))}{1 - \exp(-K(a_2 - a_1))} \right)^{1/b} \quad (\text{Equation 1})$$

where a_1 is the first age at which the growth follows this curve (zero quarters), and L_1 is the corresponding mean size at that age. L_2 is the average size of fish of age a_2 (29 quarters) K is the growth rate, and b influences the shape of the growth curve.

There is uncertainty in the estimates of growth of yellowfin in the EPO. Wild (1986) estimated an asymptotic length (L_∞) of 188.2 cm, using a Richards curve fitted to otolith age-length data. This estimate, however, is an extrapolation well beyond the maximum age of 4 years in Wild's (1986) study. The limited tagging data available, reliable but restricted in their spatial and temporal distribution, are consistent with

an L_2 of about 172 cm.

Maunder and Aires-da-Silva (2009) estimated the growth within the stock assessment model, these parameter estimates have been used in previous assessment and are used in this assessment, in models that assumed fixed growth. The fixed parameter values are: $L_1 = 18.3686$, $L_2 = 182.307$, $a_1 = 0$ quarters, $a_2 = 29$ quarters, $K = 0.19228628$ quarter⁻¹ $b = -0.542255$.

Misfit of the length-composition data (mainly to the fishery with asymptotic selectivity, F19) from preliminary runs assuming fixed growth indicates that the assumed growth function may not represent growth for the core of the exploited population. Given the uncertainty in the estimates of growth, models were included in the reference set that estimate growth while fitting to the conditional age-at-length data.

The variability of size at age may also be important, as this will determine what sizes are plausible in the population. This assessment assumes a coefficient of variation of 7.5% length at age. This value was set as a compromise between the former stock assessment assumption of 10%, which was considered too high during the external review, and the assumption of 5% used in research models presented at that review, which was considered too low. Figure 7 shows the fixed growth curve and the variability assumption; 95% of five-year-old fish (20 quarters) are between 144 and 193 cm, and 95% of the 7.25-year-olds (a_2) are between 155 and 209 cm.

The weight at age w_a is obtained by replacing the average length at age L_a in the length-weight equation for yellowfin tuna in the EPO (Wild 1986):

$$w_a = 1.387 \times 10^{-5} L_a^{3.086} \text{ (Equation 2)}$$

3.1.2 Natural mortality (M)

For this assessment, as in previous assessments, it is assumed that, as yellowfin grow older, the natural mortality rate (M) changes (Maunder and Aires-da-Silva 2012). Males and females are treated separately in this assessment, and M differs between males and females (Figure 8). The highest M is at age 0, then declines almost linearly until the fish are 10 quarters old, then increases again for females. These values were estimated by fitting to sex ratio-at-length data (Schaefer 1998), considering the values with those estimated for yellowfin in the western and central Pacific Ocean (Hampton 2000; Hampton and Fournier 2001). Maunder and Watters (2001) describe in detail how the age-specific M schedule for yellowfin in the EPO was determined.

The assumed level of M for age 0 has no impact on the assessment results. Recruitment occurs at age 0 in the assessment model. Age 0 is used for convenience, and the assumed M for ages not vulnerable to the fisheries is not intended to represent the actual M , and only arbitrarily scales the recruitment at age 0.

3.1.3 Reproductive biology and recruitment

Yellowfin can spawn almost every day if the water temperatures are in the range of 24 to 30°C, resulting in spawning year-round in lower latitudes and in the summer in higher latitudes (Nishikawa *et al.*, 1985; Schaefer 1998; Itano 2000). An “index” of total egg production (fecundity; O_t^F), rather than spawning biomass (S), is used in the assessment, nevertheless those two notations are used interchangeably. This is obtained from the reproductive biology study by Schaefer (1998) (Figure 9). The fecundity O_t^F at time t at age a is given by:

$$O_t^F = \sum_{a=0}^{29} \frac{p_a f_a d_a N_{a,s=1,t}}{1,000,000} \text{ (Equation 3)}$$

where p_a is the proportion of mature females at age a , f_a is the batch fecundity (the number of migratory-nucleus or hydrated oocytes in an ovary) at age a , d_a is the fraction of females spawning per day at age a . To obtain the p_a , f_a , and d_a , from those quantities estimated at length by Schaefer (1998), the average

length at age (Maunder and Aires-da-Silva 2009) was used in the equations below.

The proportion of mature females at length p_L is:

$$p_L = e^{-(\exp(-0.059347(L-85.901241)))} \quad (\text{Equation 4})$$

where L is the fork length in centimeters.

The batch fecundity for a female of fork length L (in millimeters) is:

$$f_L = 0.0003747 L^{3.180758} \quad (\text{Equation 5})$$

The fraction of females spawning per day d_L at fork length L (in centimeters) is:

$$d_L = 0.742(1 - e^{-0.046(L-54.892)}) \quad (\text{Equation 6})$$

Four recruitments are estimated in a year. Recruitment (age-0 fish) is assumed to follow a Beverton and Holt (1957) stock-recruitment curve. The Beverton-Holt curve is parameterized so that the relationship between spawning biomass (fecundity in this assessment) and recruitment is determined by the average recruitment produced by an unexploited population (virgin recruitment) and steepness (h). Steepness is defined as the proportion of the virgin recruitment that a population produces when reduced to 20% of its virgin state. A steepness of 1.0 implies that the stock may produce recruitments equal to the virgin level, on average, at all levels of spawning biomass, while a steepness of 0.7 indicates that when a stock is at 20% of its virgin spawning biomass, only 70% of the virgin recruitment is produced, on average.

Steepness is a key parameter of a stock assessment, but it is problematic to estimate (Lee *et al.* 2012). For tunas, there is little evidence for any particular value. In previous assessments the base-case model had the assumption of $h = 1$. This assessment incorporates the uncertainty in steepness by including four hypotheses in the reference models, $h = 0.7, 0.8, 0.9, \text{ or } 1.0$.

The recruitment is assumed to vary lognormally around the stock recruitment curve with a standard deviation of 1 on the logarithm of the recruitment deviations. The variability of the recruitments is constrained by a penalty added to objective function. The recruitments are corrected so that the expected values are unbiased. The bias correction is computed using the method of Methot and Taylor (2011).

3.1.4 Movement and stock structure

Yellowfin is widely distributed in the tropical and subtropical waters of the Pacific Ocean. Yellowfin are found principally in the mixed layer at temperatures between 20°C and 30°C but may perform “bounce” dives below the thermocline for foraging during the day (Schaefer *et al.* 2007). Juveniles and small fish may aggregate around floating objects while older fish may be found associated with several species of dolphins. Although considered a highly migratory species, the tagging studies done in the EPO have indicated that yellowfin tuna move in restricted areas mostly within 1000 nautical miles of their tagging locations (Fink and Bayliff, 1970; Bayliff, 1979, 1984, Schaefer *et al.*, 2011; Schaefer *et al.*, 2014). However, the evidence from tagging data is not enough to support neither complete mixing nor spatial (Joseph *et al.* 1964, Schaefer 2009). Genomic studies are promising in detecting stock structure, and in the Pacific Ocean some genomic evidence for heterogeneous structure exists (Grewe *et al.* 2015, Pecoraro *et al.* 2018). No such study has been done within the EPO. While yellowfin tuna in the EPO may be composed of spatially disaggregated units (Schaefer, 2009) the available data is insufficient to estimate movement rates or assist in the delimitation of these units. For this assessment, as in previous assessments, it is assumed that there is a single stock of yellowfin tuna in the EPO.

3.2 Fisheries dynamics

3.2.1 Initial conditions

The model is assumed to start from a non-virgin (fished) equilibrium state, with R_{init} , the initial recruitment as an offset of the virgin recruitment, and F_{init} , the initial fishing mortality, being estimated, with no penalty associated with initial equilibrium catches. F_{init} was assumed to correspond to fishery F16, the purse-seine fishery on dolphins in the NE area. This fishery was chosen because it catches a wide range of sizes, thus it could best represent the equilibrium fishing mortality at age for the stock. Additionally, 16 recruitment (quarter) deviations before the start of the model initial quarter are estimated.

3.2.2 Selectivity

The selectivity was modelled as a function of length and age, except for the “discard” fisheries, for which only selectivity at age was assumed (fixed at 1 for ages 1 to 3 quarters, and 0 for the other ages). For all other fleets, the selectivity at age was fixed to be 1 for all ages, except for age zero, which had selectivity of zero.

Selectivity curves at size were assumed to be dome-shaped for most fleets and were modeled initially using double-normal functions. The preliminary fits to the double-normal function were unsatisfactory and indicated that more flexible selectivity functions should be used. The need for more complex and flexible shapes for selectivity may be because the selectivity encompasses not only the gear selection pattern but also the spatial-temporal availability of fish of different sizes. It is likely that seasonal patterns are present because of oceanographic conditions or movement, and those will be absorbed in the selectivity curves. Therefore, selectivity curves based on cubic splines were adopted (Table 2). The cubic splines' number and location of nodes was initially obtained by fitting spline functions to the “empirical” selectivity (the ratio of the number of fish by length class in the catches to the corresponding number in the population obtained from preliminary population models that assume double-normal selectivity.. The fits were done using an external fit in R (library [freeknotsplines](#)). Initial fitting was performed with the suggested spline configurations and then fine tuning was done. A node at the beginning and another at the end of the size distribution for the fleet was always included to avoid extreme changes in selectivity at the tails of the distribution. The parameter for one of the splines nodes was fixed at an arbitrary value and the values for the other nodes were estimated relative to that fixed parameter within the assessment models.

The selectivity for the longline fisheries were assumed to be asymptotic and fixed to values estimated in preliminary runs. The selectivity for the purse-seine fisheries south of 5°N were set equal to those of the longline fisheries. The composition data for those fisheries were not fit in the reference models, only in the preliminary runs. The selectivity for the longline fisheries in weight was assumed to be the same as that of the corresponding fishery in number.

The basic assumption is that selectivity is time-invariant. There are two periods that seem to depart from this assumption (Figures 5B and 6). The first period was during, and for some quarters after, the peak in the index of abundance in about 2002 (Figure 3). The second period occurred in recent years, since about 2015, when the average size of yellowfin in this fishery was higher. Because the standardized size compositions associated with the index (Figure 6) are derived from dolphin-associated data, those changes in average size are also seen in the index. These uncertainties regarding selectivity were addressed in some of the reference models.

The dolphin-associated purse-seine fishery that catches the largest yellowfin in the core area (F19) was chosen as the fishery with asymptotic selectivity whose length-composition data were included in the objective function. Models with asymptotic selectivity had poor fits to these data. To improve the fit, an alternative hypothesis that this fishery has dome-shaped selectivity was also considered in some of the

reference models.

3.3 Data weighting

Likelihood functions encompass not only the sampling (observation) variability, but also model misspecification and unmodelled process variability. Therefore, the CVs of the index of abundance are set equal to the CVs estimated from the standardization model plus a constant added so that the average CV for a range of years is 0.15 (the average CV for the whole time series about 0.18).

The size-composition data were assumed to have multinomial distributions, with the variance proportional to the sample size. The input sample size for the purse-seine fisheries was equal to the number of wells sampled. The number of fish sampled within a well cannot be used to represent the sample size because fish stored in the same well may come from the same school and thus not be independent samples, and their sizes may be highly correlated (Pennington *et al.* 2002). For the preliminary model runs used to estimate the selectivity of the longline fisheries, the length frequency of the longline fishery was used and the sample size was set equal to the number of fish/100. The Francis method for weighting the size composition data was used (TA1.8 in Francis 2011). A preliminary run was conducted with weighting equal 1 and reweighting factors (“Francis weights”) were computed based on how well the model fitted the size length composition data (Table 2). In addition to that, all length compositions with multimodal distributions were further downweighted by multiplying the Francis weights by 0.5, since it is likely that the multimodal distributions result from processes not modelled explicitly (*e.g.* movement). Similarly, to the index of abundance, the length composition likelihood will also absorb model misspecification and unmodelled process variability. The adjusted sample sizes are shown in Figure 5C.

3.4 Model diagnostics

A suite of approaches was used as diagnostics to determine whether a model fits the data well and is correctly specified:

Index of abundance: The root-mean-square error (RMSE) of the residuals was compared to the input CV to evaluate how well the reference models fit the index of abundance and to evaluate the validity of the variability assumption. The residuals are inspected for trends or patterns that may indicate model misspecification.

Size composition data: Predicted and empirical selectivity curves for every fishery that has composition data were compared. The empirical selectivity of a fishery is defined as the average observed catch at length from the fishery divided by the average predicted population number at length. The empirical selectivity was scaled to a maximum value of 1, unless noted otherwise. If the assessment model fits a fishery’s composition well, the two selectivity curves should be similar. The residuals of the length composition data were inspected for trends over time and across length classes. The effective sample size (McAllister and Ianelli, 1997) implied by the model fit was compared to the input sample size. The effective sample size is the size of the random sample needed, on average, to achieve a fit that is as good as the fit to the composition vector (Methot and Wetzel 2013). The better the fit, the larger the effective sample size. It should be noted that the method of McAllister and Ianelli (1997) generally overestimates the effective sample size due to unaccounted correlations in residuals, but relative patterns among fisheries and over time should be well represented.

Integrated model diagnostics: Age-structured production models (Maunder and Piner 2015), catch-curve analysis (Carvalho *et al.* 2017), likelihood profile on the global scaling parameter (Lee *et al.* 2014, Wang *et al.* 2014) and retrospective analyses (Mohn 1999, Hurtado-Ferro *et al.* 2015) were used to detect model misspecification, influence of different data sets, and other potential issues with the models (Appendix 2).

3. REFERENCE MODELS

This benchmark assessment is the basis for a risk analysis that addresses the uncertainties about several assumptions and explicitly includes uncertainty in the evaluation of stock status and formulation of management advice (SAC-11-08).

The first step to apply the risk analysis framework (SAC-11 INF-F, Maunder *et al.* 2020a) is to list the unresolved issues and uncertainties that need to be accounted for in the management advice. Then several hypotheses are formulated that represent different states of nature that may resolve these issues or represent the uncertainties, and these are arranged in a hierarchy. The most encompassing hypotheses (overarching hypotheses) are at the top of the hierarchy, then other levels unfold nested under the upper levels. The main issues and uncertainties when assessing the stock status of yellowfin tuna include: a) spatial structure; b) inconsistencies between the indices of abundance based on CPUE from the dolphin-associated purse-seine fishery and that based on CPUE from the longline fishery; c) inability of the model to fit the high values in the indices of abundance; d) and misfit to the composition data for the fishery that is assumed to have asymptotic selectivity.

The set of overarching hypotheses (Level 1, Figure 10a) addresses the issue of spatial structure. Although there is some evidence of the existence of northern and southern stocks, the divisions are not clear and mixing between the two potential stocks may be episodic, or the magnitude may vary from year to year. The overarching hypotheses formulated for the spatial structure of yellowfin tuna in the EPO were “High mixing”, “Episodic/high variability mixing”, and “Negligible mixing”. The “High mixing” overarching hypothesis is represented by single-stock models similar to previous assessments. The “Episodic/high variability mixing” overarching hypothesis is represented by single-stock models that are driven by the north or the southern stock data. This means that the model is fitted to data for the north (south) and the selectivity for the fisheries in the south (north) is fixed. The “Negligible mixing” hypothesis is represented by two independent assessments, one for the north and one for the south. Many of these models were developed for the yellowfin tuna [review](#) and this informed the decision to eliminate all hypotheses except the “High mixing” hypothesis from the risk analysis to make it practical to implement. This assessment thus focuses on the hypotheses nested within the overarching “High mixing” hypothesis (Figure 10b).

Under the “High mixing” hypothesis are hypotheses that address the misfit to the index of abundance and the changes in selectivity (Level 2A, Figure 10b), and the misfit to the length-composition data for the fishery with asymptotic selectivity (Level 2B). Models representing different steepness scenarios are added as a third level in the hierarchy (Level 3).

The four hypotheses at Level 2A that address issues with the index misfit and changes in selectivity (and the models used to implement them) are:

Index is proportional to abundance (BASE): This hypothesis is most similar to that of previous assessments of yellowfin tuna in the EPO. The BASE model is the basis for all the other models. This hypothesis assumes that the index is proportional to abundance for the whole time period.

Density dependent catchability (DDQ): This hypothesis postulates that the abundance and the index are non-linearly related. The DDQ model estimates a coefficient c that determines how the catchability is influenced by abundance. It is hypothesized that during periods of high abundance the purse seine fleet that fishes on yellowfin associated with dolphins can more efficiently catch yellowfin tuna and this will allow the models to better fit the high index observations (hyper-depletion, $c > 0$, Methot *et al.* 2020).

Time block in the middle (TBM): During the period of peak index values and shortly after, the fishery associated with dolphins catches smaller yellowfin on average. This hypothesis assumes It is hypothesized that: (a) the purse seine fleet that fishes on yellowfin associated with dolphins can more efficiently catch yellowfin tuna during that period and this will allow the model to better fit the high index observations,

but assumes that at other times this is not the case; and (b) if catchability changes, selectivity is also likely to change. The TBM model estimates a block in catchability and selectivity of the index during the period when there are peak values.

Time block at the end (TBE): This hypothesis postulates that catchability and selectivity changed for the survey and some fisheries during the later period when the size of fish caught by the purse seine fishery associated with dolphins is higher on average. The TBE model estimates a block in selectivity and catchability starting in 2015.

There are three hypotheses at Level 2B that address the misfit to the length composition data:

Fixed growth (BASE): It represents the null hypothesis that growth is well described using the fixed parameters. The model is the same as in 2A.

Estimate growth (GRO): The fixed value used for asymptotic length is higher than largest length from the limited tagging data, but is somewhat consistent with the otolith data (although old fish cannot be aged using the daily increment method). The otolith data comes from before the assessment period and the tagging data is limited in its spatial and temporal distribution. This hypothesis postulates that estimation of growth within the stock assessment model may be appropriate and may allow a better fit of the model to the length composition data for the fishery with asymptotic selectivity. The GRO model also fits to the conditional age-at-length data to inform the estimates of growth.

Dome-shape selectivity (DS): This hypothesis postulates that the selectivity to be dome shape for the main purse seine fishery on yellowfin associated with dolphins. The DS model estimates the parameters of the double normal selectivity curve. This assumption may allow the model to fit the length composition data better.

The combination of these hypotheses comprises the configurations that compose the 12 reference models for the assessment of yellowfin tuna in the EPO (Table 3), each with four assumption for the steepness of stock-recruitment function (h equal to 0.7, 0.8, 0.9, and 1.0), for a total of 48 models.

4. RESULTS

5.1 Model diagnostics

5.1.1 Model convergence

All the 48 model for yellowfin converged (produced positive definite Hessian matrices); 28 had small maximum gradients (< 0.001), and eight (TBM.DS and TBM.GRO, for all four steepness values) had large maximum gradients (>1) (Table 4).

5.1.2 Fit to purse-seine indices of abundance

The RMSE and the negative log-likelihood (NLL) of the purse-seine index of abundance are used to evaluate how well the models fit that data (Table 5, Figure 11). Small values for both RMSE and NLL suggest the assessment fits the data well. The models that best fit the index of abundance were those with the assumption of density-dependent catchability (DDQ), followed by those with the time-block in the middle of the series (TBM) (Figure 11 and Figure A3). Of those, the models that best fit the index and had less residual patterns (Figure A3) were the ones that estimated growth (DDQ.GRO). Changing the steepness did not improve the fit; within a configuration, the models with different steepness fit the index about the same.

The model configurations that showed fewer residual patterns over time were those that assumed a non-linear relationship between the index of abundance and its vulnerable biomass (DDQ, DDQ.GRO and DDQ.DS) (Figure A3). They were followed by the models with a time-block in the middle of the period (TBM, TBM.GRO, TBM.DS). A seasonal pattern was present in the residuals of all models, with positive

residuals in quarters 1 and 2 and negative residuals in quarters 3 and 4 (Figure A4).

5.1.3 Fits to length-frequency data

The results of this section focus on the model with steepness of 1 because the fit of models with the same configuration, except for the steepness assumption, fit the composition data almost identically, with at most one negative log-likelihood unit of difference (Table 6).

For all fisheries and for the survey, the average effective sample size (based on McAllister and Ianelli 1997) is about 5 to 20 times larger than the input adjusted sample size (Appendix 3, Table A1). As expected, the survey length frequencies are fit best by models that have blocks in selectivity. The fit of different model configurations to the length-frequency composition is very similar for all but three fisheries: F3, F18 and F19. F19 has the asymptotic selectivity assumption in three of the models, dome-shaped in three models, and blocks of selectivity (asymptotic and dome-shaped) in six models. The models that fit this data best are TBM.DS. F18 is one of the fisheries with the largest catches during all the assessment time period (Figure 2), and the one that shows an increase in the average size in recent years (Figure 5B). These data are fitted best by the models that consider that the index is non-linearly related to abundance (DDQ, DDQ.DS and DDQ.GRO). Finally, F3 is a floating-object fishery that expanded after the mid-1990s and has a skewed length-frequency distribution towards larger sizes (Figure 5A). The models that fit these data best are those that estimate growth (DDQ.GRO, GRO, TBE.GRO and TBM.GRO).

Another way to visualize how well the models fit the length-composition data is through the empirical selectivity: the two curves should be similar if the model is a good fit to the data. Figure 14 shows this plot for BASE $h = 1$. The fit to most of the length frequency data is similar, except for F19, which is fitted better by the model that assumed dome-shaped selectivity (Figure 15).

Residual plots are shown for the survey, and fisheries F18 and F19 (Figure A5). The negative residuals after 2000 improve using the models with a time block at that time (TBM, TBM.GRO, TBM.DS). The trend towards positive residuals at the end of the time series is reduced by the models that have a time block at the end (TBE, TBE.GRO and TBE.DS). Finally, the trends towards negative residuals for larger sizes for F19 is improved by the models that either used dome-shaped selectivity or estimated growth.

5.1.4 Overall fit

The overall fit was assessed using AIC (*Akaike Information Criterion*). Because some models are also fitted to the conditional age-at-length, the AIC was computed without this component to make it more comparable among models. The comparison is an approximation, however, since the models that do not use the conditional age-at-length are expected to have better AIC scores than the models that use it; this is because the latter models' fit to the data will still be affected by the conditional age-at-length. The models that fit the data best were TBM.DS at any steepness value (Table 6).

5.1.5 Integrated model diagnostics

5.1.5.1 Age-structured production model (ASPM) and catch-curve analyses

The age-structured production models (ASPMs) show different trends to their corresponding reference models (Figures A6-A8). The ASPM for all models starts from a depleted state, then increases to two to three times the virgin biomass, and stabilizes at a spawning biomass ratio (SBR; the ratio of the spawning biomass to the virgin spawning biomass) of 1 during 1990-2019. ASPMs that estimate recruitment deviations (ASPM-R) show population trajectory trends that are more similar to the reference model but at a much lower SBR. Only the ASPM-Rs with density-dependent catchability (DDQ) configurations (DDQ, DDQ.GRO and DDQ.DS) have a positive definite Hessian. Of these, only the DDQ configuration, however, has confidence intervals that overlap with the reference model. These results indicate that information on relative recruitment over time is needed to extract absolute abundance information from the

abundance index.

The catch-curve analyses (CCAs) are aligned with reference models in several periods (Figure A7). Both the CCA based on the standardized length composition corresponding to the index of abundance and the CCA based on the fisheries length compositions show similar results. For BASE, TBM and GRO, the beginning of the series is markedly different for the reference models and the CCAs. The implied index of abundance for the CCAs also shows that, for an index to follow the same pattern that the CCAs are inferring, it should be stable throughout the period, but with large interannual variations (Figure A8). This, in addition to the ASPM results above, indicates that the scale of the models is highly influenced by the composition data, but the index of abundance is needed to inform or constrain the temporal variation and trends.

Dissimilar trajectories of the CCA and the integrated models indicate model misspecifications (Figure A7). The discrepancies observed in the BASE model around 2000 and after 2015 are solved by the TBM and TBE configurations. The DDQ reference models have better overall match with the CCAs. This indicates that the length-composition data and the index tend to support similar trajectories when either the index is assumed to have a non-linear relationship with the vulnerable biomass or when there is a block of catchability and selectivity for the index and the main fisheries (F18 and F19).

Some assumptions seem key to solving model misspecification, while others have no effect: the dome-shape selectivity assumption and the change in catchability seem key, while estimating growth is not important. The TBE and the DS reference models have better correspondence with the CCAs except in 2001-2007, when the CCA predicts smaller biomass than the reference models, and around 2010, when the CCA predicts larger biomass (Figure A7). The TBM solves the 2001-2007 discrepancies. The models that estimate growth have similar patterns to those with fixed growth (and identical otherwise). This indicates that understanding the changes in fisheries strategies (*e.g.* gear, search behavior, market demands for larger fish) are a central part for assessing the yellowfin tuna population.

5.1.5.2 Likelihood profile on R_0

This diagnostic is helpful in determining the relative importance of different data components on the estimates. The likelihood profile on R_0 (in log scale) indicates that the results for all reference models are driven by the length composition data, except for GRO and TBE.GRO (Figure A9). For TBM, and to some extent for TBM.GRO, although the length composition data is the most influential component, it is not in contradiction with the index of abundance. This means that including a time-block when the index indicates large abundance may resolve some model misspecification.

5.1.5.3 Retrospective analyses

The retrospective analyses show the behavior of the models when new data are added. Two model configurations showed instabilities when years of data were sequentially removed (DDQ.DS for both SBR and F , and TBM.DS, for F) (Figures A10 and A11). These issues were not resolved even after several attempts to start the model fits from different starting value. The other model configurations that treated the index as one continuous series linearly related to the biomass (BASE, GRO and DS) showed retrospective pattern, both in the spawning biomass ratio and in the fishing mortality. The other models show no important changes in the results by sequentially removing data for the terminal years, indicating that addressing misspecifications in the observation model for the index of abundance, and in the selectivities for the main fleets and survey, improves the robustness of the models.

The previous assessment was not considered reliable for management advice because it was too sensitive to the addition of new abundance index data from the longline fishery (SAC-10-INF-F). These data are no longer used in the assessment and this may be why most models are now not over-sensitive to new data.

5.1.6 Parameter estimates

5.1.6.1 Initial conditions

All models estimate that the population starts from a depleted state in 1984 (spawning biomass ranging from 14% to 72% of unexploited).

5.1.6.2 Selectivity

In general, estimated selectivities follow the empirical selectivities well (Figure A12), except for fishery F19, for which the fit depends on the model assumptions. Differences in the selectivity of F19 occurs primarily at large sizes (Figure 12). When assuming asymptotic selectivity, the selectivity at sizes >100 cm is estimated to be 1. In comparison, when assuming dome-shaped selectivity, selectivity is estimated to reach the peak (*i.e.*, 1) at around 100 cm before dropping to a final (at L_2) level of about zero (DS runs). The runs that estimate growth predict no or a very low proportion of large fish (>175cm). The models with a time block at the end (TBE) have better correspondence between the estimated selectivity for smaller sizes and the empirical selectivity. This is because during most of the period (1984-2014) the selectivity is dome-shaped (Figure 14). Fit is improved by the models that either used dome-shaped selectivity or estimated growth. However, models that estimated growth predicted no fish in size classes that had fish of those sizes in the observations (Figure 13). The TBM models estimated differences in selectivity for small fish in the index of abundance and for large fish in the fisheries. The TBE models estimated a shift towards larger sizes in the selectivity curve of the F19 fishery after 2015, as expected (Figure 13).

5.1.6.3 Catchability and density dependence

The catchability estimates for the index of abundance at the end of the time-series ranged from 86% to 91% of the catchability value for the earlier period, in the models with catchability block at the end (Table 7). The catchability estimates for the index of abundance in the models with block in the middle of the time series are 164% to 182% of the catchability for the rest of the time period, almost doubling the catchability during that period. The density-dependent parameter ranged from 1.7 to 2.1, in the DDQs models.

5.1.6.4 Growth

The estimated growth curves differ in two main ways from the fixed values (Figure 14): L_1 is about 21 cm instead of 18.4 cm, and L_2 (which ranges from 149.4 to 161.9 cm) is smaller than the fixed value (of 182.3 cm) and. Because L_2 is the average size of the oldest fish, and given the assumed CV = 0.075 of variation of length at age, a fish of age 29 quarters as large as 182.4 cm or more might still be found in the population, but with very small probability (for example, about 16 fish in 10,000 would be that large or larger if $L_2 = 149.4$ cm). Another implication of the estimated growth functions is that fish stop growing at about 4 years of age.

5.2 Stock assessment results

5.2.1 Recruitment

Time series of estimated quarterly age-zero recruits are shown in Figure A13, and the annual recruitment in Figure 15. The recruitment estimates are not sensitive to the value of steepness. All models estimate an initial period of above-average recruitment, culminating in 1999, followed by below-average recruitments. This pattern follows the general trend shown in the index of abundance. All models have an increase in the point estimate of recruitment in the last year, but with a large confidence interval, which is expected since there is not much information in the data about recruitment in the most recent year.

The general patterns vary in their magnitude for the different configurations. For example, the large 1999 recruitment is much larger in the DS model than in the TBM.DS model, where most of the increase in the

index is attributed to a change in catchability, rather than the result of a very strong recruitment. In the TBM.DS model, the 1999 recruitment is comparable to that of 1993. The DDQ models have less pronounced difference between the high and the low recruitment periods, as much of the differences are absorbed in the non-linear relationship between the index and abundance. The DS models have the opposite pattern: recruitment in the high recruitment period is much larger than in the low recruitment period. Models that estimate growth have similar point estimates of recruitment as their fixed counterpart but have an increase in the uncertainty of the estimates, as expected. Assuming a dome-shape selectivity for fishery F19 increased the uncertainty of the recruitment estimates..

Like the recruitment results, spawning biomass estimates are not sensitive to steepness (Figure 16). However, they differ in a key point: the estimate of equilibrium virgin biomass. This translates to differences in the SBRs among the steepness runs (Figure 17): the models with $h = 1$ and $h = 0.7$ have the least depleted and the most depleted series, respectively. All the point estimates of the trajectories with different steepness, however, are within each other's confidence intervals.

All biomass trajectories have declining trends, but they vary in the magnitude of the declines. At one extreme, the TBM, TMB.GRO, and TMB.DS models, which show the most pronounced declines, assume a time-block in the middle of the timeseries (2001-2003) of the index of abundance, which isolates the period of the sudden increase in the abundance index by assuming a different catchability during that time; of these, the model TBM.DS has the largest biomass decline. The models that assume that the increase in the index of abundance in 2001-2003 is real, and not a sudden change in catchability (e.g. BASE, GRO, TBE), estimate lesser declines over time. At the other extreme, the models that assume a non-linear relationship between the index and the biomass (DDQ, DDQ.GRO and DDQ.DS) estimate the least declines. All models show the lowest SBR in mid-2016 and an increasing trend afterwards.

5.2.2 Fishing mortality (F)

Regarding fishing mortality (F), similarities and contrasts among models are also apparent (Figure 18). The main similarities are in the relative magnitude of F between age classes. All models indicate the highest F for fish aged 21+ quarters (5.25+ years), followed by fish aged 11-20 quarters (2.75-5 years). The lowest F at age is on the youngest fish, and is about the same for all models. All models estimate an increase in F for the two oldest age classes over time.

The main difference among models is in the magnitude of F for the oldest age classes. The models with fixed growth have a higher F for fish age 21 quarters or more than those that estimate growth or assume dome-shaped selectivity. This is because the fixed-growth models assume that older fish are larger ($L_2 = 182.6$ cm), and therefore, because fish of these sizes are rare in the observations, the fishing mortality must be high. The models that estimate growth explain the scarcity of those large fish in the observations by decreasing the average size of the older fish (L_2) from the fixed value of 182.6 cm to the estimated 149.4 to 161.9 cm, depending on the model, making the proportion of large fish smaller (given the CV of length at age is 7.5%), as seen in the data. The models that assume dome-shape selectivity imply that the reason there are no large fish in the data is because those sizes are not vulnerable to the purse-seine fisheries that operate north of 5°N (the fisheries that provide the length frequency data used to fit the models). The models that estimate a time block in selectivity of both the index and the F19 fishery, from 2015 on (TBE, TBE.GRO and TBE.DS), isolate the period of large increase in sizes with a different selectivity, assume dome-shape selectivity before 2015, thus explaining the lack of large fish in the past by assuming that those fish were not vulnerable to the purse-seine fishery before 2015. It is intriguing that models with a time block in the middle (TBM, TBM.GRO and TBM.DS) estimate the same low fishing mortality as the TBE ones. By using a time block for the index catchability (2001-2003) and selectivity (2002-2007), the TBM models estimate a biomass trajectory analogous to the TBE models, given that the catches are fixed, that translates into similar fishing mortalities.

5.2.3 Fisheries impacts

This analysis compares the impact on the spawning biomass of fisheries that have different selectivities (Wang *et al.* 2009). The impact for each type of fishery was estimated by projecting the population without their catches and obtaining the resulting spawning biomass. The increased spawning biomass in the absence of the catches of those fisheries relative to the current spawning biomass indicates the impact of those fisheries.

All models estimate similar impacts of the different types of fisheries (Figure 19). The longline and the sorted discard fisheries have the smallest impact, while the purse-seine fisheries associated with dolphins have the greatest impact during most of the 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 be noteworthy, and surpassed that of the unassociated fisheries around 2008; in 2018, it surpassed the impact of the purse-seine fisheries associated with dolphins.

5.2.4 Comparison with the previous assessment

One of the main differences between this assessment and the previous one is on its use of the data available. Previous assessments, including the SAC-10 assessment, were fitted to five indices of abundance, one from the longline fishery and four from the purse-seine fisheries. The longline index was based on standardized CPUE from the Japanese fleet. The purse-seine indices were nominal CPUEs and were limited to certain areas of the EPO. The purse-seine and the longline indices had inconsistencies that were considered a major issue for the previous assessments. A new spatio-temporal modeling framework was developed and applied to the CPUE data to create new indices, but the inconsistencies were not resolved. Standardized length frequencies suggest that the two indices may be indexing different groups of fish. The prominent index values in 2001-2003 seems to occur earlier in the longline index and later in the purse-seine index (opposite to what was expected given the growth and selectivity assumptions of the model), were due mainly to the 1998 cohort (of an important El Niño year) in longline fisheries and to the 1999 cohort (of an equally important La Niña year) in purse-seine fisheries. Why these indices tracked those two cohorts differently is still an unresolved issue, a topic for future research. Also, how (or whether) other cohorts of smaller magnitude may be subject to the same phenomenon is unknown. One of the hypotheses is spatial heterogeneity, which is somewhat addressed in the current assessment as the models are fitted to a purse-seine index of abundance for the EPO north of 5°N and also to the length-composition data from the purse-seine fisheries that operate north of 5°N, but not to the data from fisheries in the south.

Results of the current assessment and the previous differ in the uncertainty of the estimates (Figure A14). The uncertainty in the SAC-10 model was very small, due to (a) the limited combinations of parameter values that allowed the model to fit the contradictory information from the indices, and (b) the greater weighting of length-composition data relative to new models, which implement the Francis method of composition data reweighting. The confidence intervals for any model in the current assessment are much wider than the SAC-10 model.

Another important differences between the SAC-10 model and the current assessment are the fishery definitions and the assumed selectivity for each fishery. The fisheries in the current assessment were defined using a regression tree analysis that maximized the differences in size composition among fisheries and minimize the difference within fisheries (in space and time). Then, splines were used to best characterize the selectivity for each fishery. Splines allow more flexibility in the shapes of the selectivity functions than the double-normal functions used for the SAC-10 model. In the current assessment, the mortality at size was characterized by more flexible curves, which changed the F at age. The SAC-10 model estimated higher F -at-age for ages 10 to 21 quarters (Figure 3 in [SAC-10-07](#)), while this assessment estimates that F for fish age 21+ may be as high or higher (Figure 18) . All models have a lower F for fish younger than 10

quarters, similar to SAC-10, but even lower than SAC-10, due to a larger number of recruits being estimated in all models. This is also related to selectivity functions of the floating-object fisheries, which include a narrower range of lengths than those in the previous assessment, in, corresponding to better fits to the length frequencies, which translate into some ages not being fully selected to the floating-object fisheries, given the assumption of variability of length at age.

The relative impact of different fisheries estimated in this assessment (Figure 19) is similar to the previous assessment (Figure 4 in [SAC-10-07](#)) as well as the tendency of increase in F . The tendency for an increase in the impact of the floating-object fisheries and the decrease of impact of the unassociated fisheries shown in previous assessments, is also estimated for all models of the reference set from this assessment. The overall F has increased in recent years, similar to what was estimated by the SAC-10 assessment.

2. STOCK STATUS

The stock status of yellowfin tuna in the EPO is assessed by considering calculations based on the spawning biomass and the maximum sustainable yield (MSY). Maintaining tuna stocks at levels capable of producing MSY is the management objective specified by the IATTC Antigua Convention.

6.1 Definition of reference points

Resolution [C-16-02](#) defines target and limit reference points, expressed in terms of spawning biomass (S) and fishing mortality (F), for the tropical tuna species: bigeye, yellowfin, and skipjack. They, and the method used to compute them in this document, are described below, as is the harvest control rule (HCR) that implements them.

6.1.1 Limit reference points

The **spawning biomass limit reference point (S_{LIMIT})** is the threshold value of S that should be avoided because further depletion could endanger the sustainability of the stock. The interim S_{LIMIT} adopted by the IATTC in 2014 is the spawning biomass that produces 50% of the virgin recruitment (R_0) if the stock-recruitment relationship follows the Beverton-Holt function with a steepness (h) of 0.75. This spawning biomass is equal to 0.077 of the equilibrium virgin spawning biomass (S_0) (Maunder and Deriso 2014 – [SAC-05-14](#)). The HCR requires action be taken if the probability (P) of the spawning biomass at the beginning of 2020 ($S_{current}$) being below S_{LIMIT} is greater than 10%. Thus, to provide management advice, $S_{current}/S_{LIMIT}$, and the probability of $S_{current} < S_{LIMIT}$ (or $P(S_{current}/S_{LIMIT} < 1)$, which is computed by assuming the probability distribution function for the ratio is normal), are reported (Table 8).

The **fishing mortality limit reference point (F_{LIMIT})** is the threshold value of F that should be avoided because fishing more intensively could endanger the sustainability of the stock. The interim F_{LIMIT} adopted by the IATTC in 2014 is the fishing mortality rate that, under equilibrium conditions, maintains the spawning population at S_{LIMIT} . The HCR requires action to be taken if the probability of the average fishing mortality during 2017-2019 ($F_{current}$) being above F_{LIMIT} is greater than 10%. Thus, to provide management advice, $F_{current}/F_{LIMIT}$, and the probability of this ratio being > 1 (by assuming the probability distribution function for the ratio is normal), are reported (Table 8).

6.1.2 Target reference points

The **spawning biomass target reference point** is the level of spawning biomass that should be achieved and maintained. The IATTC adopted S_{MSY} (the spawning biomass that produces the MSY) in 2014 as the target reference point. The HCR requires that actions taken to achieve S_{MSY} have at least a 50% probability of restoring the spawning biomass to the dynamic MSY level (S_{MSY_d}) within five years or two generations. Here, S_{MSY_d} is equal to $S_{MSY_{d1}}$, which is derived by projecting the population into the future, assuming historical recruitment and a fishing mortality rate that produces MSY, $F = F_{MSY}$. The value of $S_{MSY_{d1}}$ used to compute reference points for yellowfin is the mean S for the last four quarters of the projection. To

provide management advice, $S_{current}/S_{MSY_d1}$, and the probability that this ratio is < 1 (by assuming CV is equal to that of $F_{current}/F_{MSY}$), are included (Table 8).

The dynamic MSY (MSY_d) is the sum of the total catches for the last four quarters of the projection.

The **fishing mortality target reference point** of is the level of fishing mortality that should be achieved and maintained. The IATTC adopted F_{MSY} (the fishing mortality rate that produces the MSY) in 2014 as the target reference point. Thus, to provide management advice, $F_{current}/F_{MSY}$, and the probability that this ratio is > 1 (by assuming the probability distribution function for the ratio is normal), are reported, as is the inverse of $F_{current}/F_{MSY}$ (the F multiplier) (Table 8).

In the **Kobe plot** (Figure 21), the time series of S_{MSY_d} is computed based on two approximations: S_{MSY_d1} as previously defined, and $S_{MSY_d2} = S_{0_d} (S_{MSY}/S_0)$, where S_{0_d} is the dynamic spawning biomass in the absence of fishing and S_{MSY}/S_0 is the depletion level that, under equilibrium, produces the maximum sustainable yield. (The two approximations are weighted as follows to obtain the trajectory of S_{MSY_d} in the Kobe plot:

$$S_{MSY_d}(t) = + (1 - p(t)) S_{MSY_d1}(t) + p(t) S_{MSY_d2}(t) \text{ (Equation 7)}$$

where p increases linearly as a function of year (t) from 0 in the start year to 1 in the end year.

6.1.3 Estimates of stock status

According to the 48 reference models, $S_{current}$ ranged from 49% to 219% S_{MSY_d} (Table 8, Figure 20). The probability that the $S_{current} < S_{MSY_d}$ is 50% or less for 13 of the 48 models. $F_{current}$ ranged from 40% to 168% of F_{MSY} . The probability that $F_{current} > F_{MSY}$ was 50% or more for 14 of the 48 models.

$S_{current}$ ranged from 145% to 345% of S_{LIMIT} . The probability that the spawning biomass at the beginning of 2020 is below S_{LIMIT} ranges from 0 to 2%. The point estimate of $F_{current}$ ranged from 22% to 65% of the F_{LIMIT} . The probability that $F_{current} > F_{LIMIT}$ was estimated to be zero for all models.

Every reference model suggests that a lower steepness value corresponds to more pessimistic estimates of stock status: lower spawning biomass relative to the reference points and higher fishing mortality relative to reference points. However, regardless of what value is assumed for steepness, the BASE models (which assume either fixed growth, a linear relationship between the index of abundance and the vulnerable biomass or no changes in selectivity and asymptotic selectivity for the purse-seine fishery that catches the largest fish), estimate the stock to be below the MSY level ($S < S_{MSY_d}$) and that the fishing mortality was above that level ($F > F_{MSY}$). Conversely, models that assume dome-shaped selectivity for fishery F19 (DS, TBM.DS, TBE.DS, DDQ.DS) estimate the opposite. The stock status at the beginning of 2020 estimated by the remaining models depends on the value assumed for steepness. There is considerable uncertainty associated with those estimates (Figure 20); several models that are in the green quadrant of the Kobe plot (Figure 20) have confidence intervals that include the yellow and red quadrants, implying that those models also provide some support for the hypotheses that the stock is below the MSY level and the fishing mortality is above that level.

All models show a “one-way trip” type of trajectory (Figure 21), gradually moving from high spawning biomass and low fishing mortality to low spawning biomass and high fishing mortality over time (Figure 21). For most models, most of the trajectory stayed in the green quadrant of the Kobe plot. The (equilibrium) MSY and S_{MSY} of yellowfin in the EPO has been stable over time (Figure 22).

The results of all these models are used in a risk analysis ([SAC-11-08](#)) to evaluate the probability of exceeding the reference points specified in the harvest control rule.

3. FUTURE DIRECTIONS

Recommendations of the [external review panel](#), as well as lessons learnt in this benchmark assessment,

will be taken into account in future work. Specifically, the staff plans to focus on:

7.1 Collection of new and updated information

- a. Continue its collection and analysis of purse-seine data (catch, effort, and size-composition)
- b. Continue collaborative work with longline CPCs
- c. Continue tagging and biology studies and analyses

7.2 Refinements to the assessment model and methods

- a. Address uncertainty in spatial/stock structure
- b. Continue research on CPUE and length-frequency standardization methods
- c. Work with purse-seine CPCs to understand changes in fishing strategies to inform selectivity modelling
- d. Continue exploring uncertainty in growth and selectivity
- e. Explore uncertainty in natural mortality
- f. Explore different stock assessment time spans, initial conditions and types of models (monthly/weekly models, depletion models)
- g. Explore other integrated model diagnostics

ACKNOWLEDGEMENTS

Many IATTC and member country staff provided data for the assessment. IATTC staff members, and member country scientists provided advice on the stock assessment, fisheries, and biology of yellowfin tuna. Nick Webb provided editorial assistance and Christine Patnode aided on the figures.

REFERENCES

- Aires-Da-Silva, A., and Maunder, M. N. 2009. Status of yellowfin tuna in the eastern Pacific Ocean in 2007 and outlook for the future. [IATTC Stock Assessment Report 9: 3-94](#)
- Bayliff, W.H. 1979. Migrations of yellowfin tuna in the eastern Pacific Ocean as determined from tagging experiments initiated during 1968-1974. *Inter-Amer. Trop. Tuna Comm., Bull.* 17: 445-506.
- Bayliff, W.H. 1979. Migrations of yellowfin tuna in the eastern Pacific Ocean as determined from tagging experiments initiated during 1968-1974. *Inter-Amer. Trop. Tuna Comm., Bull.* 17: 445-506.
- Bayliff, W.H. 1984. Migrations of yellowfin and skipjack tuna released in the central portion of the eastern Pacific Ocean, as determined by tagging experiments. *Inter-Amer. Trop. Tuna Comm., Intern. Rep.* 18: 107 p
- Beverton, R.J., and Holt, S.J. 1957. On the dynamics of exploited fish populations. Fisheries Investigation Series 2, volume 19, UK Ministry of Agriculture, Fisheries, and Food, London, UK.
- Carvalho, F., Punt, A.E., Chang, Y.-J., Maunder, M.N., and Piner, K.R. 2017. Can diagnostic tests help identify model misspecification in integrated stock assessments? *Fisheries Research* **192**: 28-40.
- Estes, D. H. 1983. Shio-Japanese pioneers in San Diego's Fishery. In: *Boat and Ship building in San Diego; 11th Annual Cabrillo Festival Historic Seminar*. San Diego California, Cabrillo Historical Association. 1: 25-43
- Fink, B.D. and W.H. Bayliff. 1970. Migrations of yellowfin and skipjack tuna in the eastern Pacific Ocean as determined by tagging experiments, 1952-1964. *Inter-Amer. Trop. Tuna Comm., Bull.* 15: 1-227.
- Francis, R.I.C.C. 2011. Data weighting in statistical fisheries stock assessment models. *Canadian Journal of Fisheries and Aquatic Sciences* **68**(6): 1124-1138.

- Grewe, P., Hampton, J. 1998. An assessment of bigeye (*Thunnus obesus*) population structure in the Pacific Ocean, based on mitochondrial DNA and DNA microsatellite analysis. CSIRO Marine Research, Hobart, Australia.
- Hampton, J. 2000. Natural mortality rates in tropical tunas: size really does matter. *Canadian Journal of Fisheries and Aquatic Sciences* 57(5): 1002-1010.
- Hampton, J.A., and Fournier, D.A. 2001. A spatially disaggregated, length-based, age-structured population model of yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. *Mar. Freshwater Res* 52, 937–963
- Hurtado-Ferro, F., Szuwalski, C.S., Valero, J.L., Anderson, S.C., Cunningham, C.J., Johnson, K.F., Licandeo, R., McGilliard, C.R., Monnahan, C.C., Muradian, M.L., Ono, Vert-Pre, K.A, Whitten, A.R., Punt, A.E. 2015. Looking in the rear-view mirror: bias and retrospective patterns in integrated, age-structured stock assessment models *ICES Journal of Marine Science* 72(1), 99–110. doi:10.1093/icesjms/fsu19
- IATTC 2010. The IATTC program for in-port sampling of tuna catches. IATTC Document SAC-01-11. <http://www.iattc.org/Meetings/Meetings2010/Aug/English/SAC-01-11-Port-sampling-program.pdf>
- Itano, D. G., 2000. The reproductive biology of yellowfin tuna (*Thunnus albacares*) in Hawaiian waters and the western tropical Pacific Ocean: project summary. SOEST 00–01, JIMAR Contribution 00–328. Pelagic Fisheries Research Program, JIMAR, University of Hawaii.
- Lee, H. H., Maunder, M.N., Piner, K.R., and Methot, R.D. 2012. Can steepness of the stock–recruitment relationship be estimated in fishery stock assessment models? *Fisheries Research* 125: 254-261.
- Lee, H.H., Piner, K.R., Methot, R.D., and Maunder, M.N. 2014. Use of likelihood profiling over a global scaling parameter to structure the population dynamics model: an example using blue marlin in the Pacific Ocean. *Fish. Res.* 158: 138-146.
- Lennert-Cody, C.E., and Tomlinson, P.K. 2010. Evaluation of aspects of the IATTC port sampling design and estimation procedures for tuna catches. *Inter-Amer. Trop. Tuna Comm., Stock Assessment Report*, 10: 279-309.
- Lennert-Cody, C.E., Minami, M., Tomlinson, P.K., Maunder, M.N. 2010. Exploratory analysis of spatial-temporal patterns in length-frequency data: An example of distributional regression trees. *Fisheries Research* 102: 323-326
- Lennert-Cody, C.E., Maunder, M.N., Tomlinson, P.K., Aires-da-Silva, A., Pérez, A. 2012. Progress report on the development of postratified estimators of total catch for the purse-seine fishery port-sampling data. IATTC Document SAC-03-10. <http://www.iattc.org/Meetings/Meetings2012/May/English/SAC-03-10-Post-stratified-estimators.pdf>
- Matsumoto, T. and Bayliff, W.H. 2008. A review of the Japanese longline fishery for tunas and billfishes in the eastern Pacific Ocean, 1998-2003. *Inter-American Tropical Tuna Commission, Bulletin*, 24 (1): 1-187.
- Maunder, M.N. and G.M. Watters. 2001. Status of yellowfin tuna in the eastern Pacific Ocean. *Inter-Amer. Trop. Tuna Comm., Stock Assess. Rep.* 1: 5-86.
- Maunder, M.N. and Watters, G.M. 2003. A-SCALA: an age-structured statistical catch-at-length analysis for assessing tuna stocks in the eastern Pacific Ocean. *IATTC Bull.* 22: 433-582.
- Maunder, M.N., and Aires-da-Silva, A. 2012. A review and evaluation of natural mortality for the assessment and management of yellowfin tuna in the eastern Pacific Ocean. *Inter-Amer. Trop. Tuna Comm. Document* YFT-01-07.
- Maunder, M.N., and R.B. Deriso. 2014. Proposal for biomass and fishing mortality limit reference points based on reduction in recruitment. *Inter-Amer. Trop. Tuna Comm. 5th Scient. Adv. Com. Meeting*. SAC-05-14.

- Maunder, M.N, Xu, H., Lennert-Cody, C.E., Valero, J.L, Aires-da-Silva, A., Minte-Vera, C. 2020a Implementing reference point-based fishery harvest control rules within a probabilistic framework that considers multiple hypotheses. Inter-Amer. Trop. Tuna Comm., 11th Scient. Adv. Com. Meeting. SAC-09. SAC-11-INF-F
- Maunder, M.N., Thorson, J.T., Xu, H., Oliveros-Ramos, R., Hoyle, S.D., Tremblay-Boyer, L., Lee, H.H., Kai, M., Chang, S.-K., and Kitakado, T. 2020b. The need for spatio-temporal modeling to determine catch-per-unit effort based indices of abundance and associated composition data for inclusion in stock assessment models. Fisheries Research **229**: 105594
- McAllister, M.K. and Ianelli, J.N. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. Can. J. Fish. Aquat. Sc **54**: 284–300.
- Methot, R.D., Taylor, I.G. 2011. Adjusting for bias due to variability of estimated recruitments in fishery assessment models. Can. J. Fish. Aquat. Sci. **68**: 1744–1760
- Methot, R.D., and Wetzel, C.R. 2013. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fisheries Research **142**: 86-99.
- Methot, R.D., and Wetzel, C.R., Taylor, I.G., and Doering, K. 2020. Stock Synthesis User Manual Version 3.30.15. NOAA Fisheries Seattle, WA. NOAA. Processed Report NMFS-NWFSC-PR-2020-05
- Mohn, R. 1999. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. ICES Journal of Marine Science: Journal du Conseil **56**(4): 473-488.
- Nishikawa, Y., Honma, M., Ueyanagi, S., Kikawa, S. 1985. Average distribution of larvae of oceanic species of scombrid fishes, 1956–1981. Far Seas Fisheries Research Laboratories S Series 12, 1–99.
- Pecoraro, C., Babbicci, M., France, R., Rico, C., Papetti, C., Chassot, E., Bodin, N., Cariani, A., Bargelloni, L., Tinti, F. 2018. The population genomics of yellowfin tuna (*Thunnus albacares*) at global geographic scale challenges current stock delineation. Scientific Reports **8**, 13890.
- Pennington, M., Burmeister, L.-M., Hjellvik, V. 2002. Assessing the precision of frequency distributions estimated from trawl-survey samples. Fisheries Bulletin **100**:74-80.
- Richards, F. 1959. A flexible growth function for empirical use. Journal of experimental Botany **10**(2): 290-301.
- Schaefer, K., Fuller, D., Hampton, J., Caillot, S., Leroy, B., Itano, D. 2015. Movements, dispersion and mixing of bigeye tuna (*Thunnus obesus*) tagged and release in the equatorial Central Pacific, with conventional and archival tags. Fisheries Research **161**, 336–355.
- Schaefer, K.M. 1998. Reproductive biology of yellowfin tuna (*Thunnus albacares*) in the eastern Pacific Ocean. Inter-Amer. Trop. Tuna Comm., Bull. **21**: 203-272.
- Schaefer, K.M. 2009. Stock structure of bigeye, yellowfin, and skipjack tunas in the eastern Pacific Ocean. IATTC Stock Assessment Report, 9, Status of the tuna and billfish stocks in 2007: 203-221. <http://www.iattc.org/PDFFiles2/StockAssessmentReports/SAR9-Stock-Structure-SPN.pdf>
- Schaefer, K.M., Fuller, D.W., and Block, B.A. 2007. Movements, behavior, and habitat utilization of yellowfin tuna (*Thunnus albacares*) in the northeastern Pacific Ocean, ascertained through archival tag data. Mar. Biol., **105**: 503-525.
- Schaefer, K.M., Fuller, D.W., and Block, B.A. 2011. Movements, behavior, and habitat utilization of yellowfin tuna (*Thunnus albacares*) in the Pacific Ocean off Baja California, Mexico, determined from archival tag data analyses, including Kalman filtering. Fish. Res. **112**, 22-37.
- Schaefer, K.M., Fuller, D.W., and Aldana, G. 2014. Movements, behavior, and habitat utilization of yellowfin tuna (*Thunnus albacares*) in waters surrounding the Revillagigedo Islands Archipelago Biosphere Reserve, Mexico. Fish. Ocean. **23**: 65-82.

- Schnute, J. 1981. A versatile growth model with statistically stable parameters. *Canadian Journal of Fisheries and Aquatic Sciences* **38**(9): 1128-1140.
- Scott, M.D., Lennert-Cody, C.E., Gerrodette, T., Skaug, H.J., Minte-Vera, C.V., Hofmeister, J., Barlow, J., Chivers, S.J., Danil, K., Duffy, L.M., Olson, R.J., Hohn, A.A., Fiedler, P.C., Ballance, L.T., Forney, K.A., 2016. Data available for assessing dolphin population status in the eastern tropical Pacific Ocean. Workshop on Methods for Monitoring the Status of Eastern Tropical Pacific Ocean Dolphin Populations: [DEK-01](#)
- Shimada, B.M., and Schaefer, M. B.(1956 *A study of changes in fishing effort, abundance, and yield for yellowfin and skipjack tuna in the Eastern Tropical Pacific Ocean*. Inter-Amer. Trop. Tuna Comm Bulletin, 1(7), pp. 347-469.
- 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. [IATTC Special Report 18](#).
- Thorson, J.T., and Barnett, L.A.K. 2017. Comparing estimates of abundance trends and distribution shifts using single- and multispecies models of fishes and biogenic habitat. *ICES Journal of Marine Science* **74**(5): 1311-1321.
- Thorson, J.T., and Haltuch, M.A. 2018. Spatiotemporal analysis of compositional data: increased precision and improved workflow using model-based inputs to stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences* (999): 1-14.
- Tomlinson, P.K. 2002. Progress on sampling the eastern Pacific Ocean tuna catch for species composition and length-frequency distributions. Inter-Amer. Trop. Tuna Comm, Stock Assess. Rep. 2: 339-365.
- Vogel 2014 <http://www.iattc.org/Meetings/Meetings2014/May/English/SAC-05-06-Fishery-in-the-EPO-2013-PRES.pdf>
- Wang, S.-P., Maunder, M.N., Aires-da-Silva, A., Bayliff, W.H. 2009. Evaluating fishery impacts: Application to bigeye tuna (*Thunnus obesus*) in the eastern Pacific Ocean. *Fisheries Research* **99**:106-111.
- Wang, S. P., Maunder, M.N., Piner, K.R., Aires-da-Silva, A., and Lee, H.H.. 2014. Evaluation of virgin recruitment profiling as a diagnostic for selectivity curve structure in integrated stock assessment models. *Fish. Res.* 158: 158-164.
- Wexler, J. B., D. Margulies, S. Masuma, N. Tezuka, K. Teruya, M. Oka, M. Kanematsu, and H. Nikaido. 2001. Age validation and growth of yellowfin tuna, *Thunnus albacares*, larvae reared in the laboratory. Inter-Am. Trop. Tuna Comm. Bull. 22:52–91.
- Wild, A., and Foreman, T. J. 1980. The relationship between otolith increments and time for yellowfin and skipjack tuna marked with tetracycline. Inter-Amer. Trop. Tuna Comm. Bull. 17: 507-560.
- Wild, A., Wexler, J. B., and Foreman, T. J. 1995. Extended studies of increment deposition rates in otoliths of yellowfin and skipjack tunas. *Bull. Mar. Sci.* 57: 555-562.
- Wild, A. 1986. Growth of yellowfin tuna, *Thunnus albacares*, in the eastern Pacific Ocean based on otolith increments. Inter-Amer. Trop. Tuna Comm. Bull. 18: 421-482.
- Xu, H., Lennert-Cody, C.E., Maunder, M.N., and Minte-Vera, C.V. 2019. Spatiotemporal dynamics of the dolphin-associated purse-seine fishery for yellowfin tuna (*Thunnus albacares*) in the eastern Pacific Ocean. *Fish. Res.* 213: 121-131.
- Yamanaka, K. L. 1990. Estimates of age, growth and spawning of yellowfin tuna (*Thunnus albacares*) in the Philippines as determined from examination of increments on sagittal otoliths Tuna Develop. Mgmt. Prog., Coll. Vol. Work. Doc. 3: 27-39.

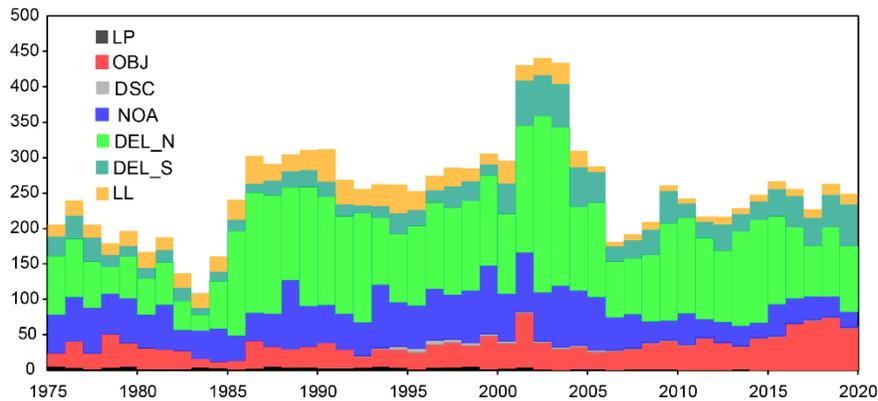
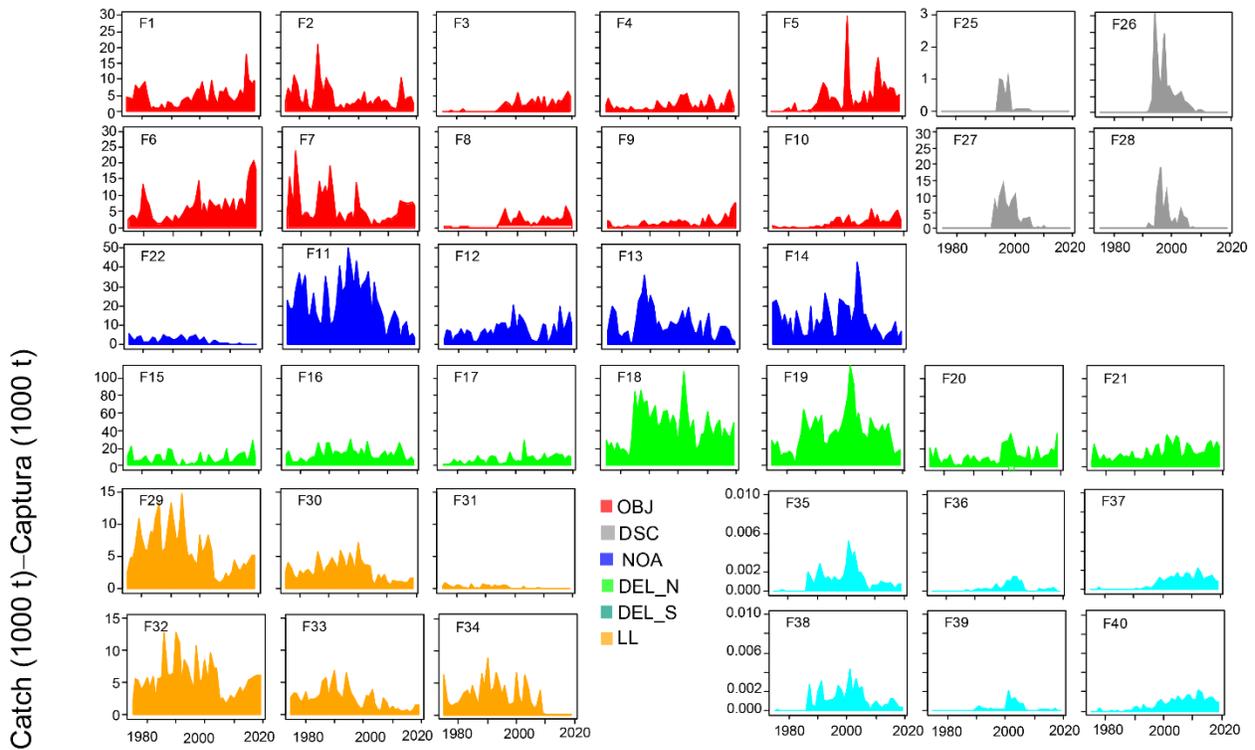


FIGURE 2. Annual catches of yellowfin tuna, in tons, in the EPO, 1975-2019, by fishery (top), and by gear/set type (bottom). Catches for fisheries F29-F35 are recorded in numbers of fish, and are converted to weights by the model. NOTE: The annual y-axis scale varies by color.

FIGURA 2. Capturas anuales de atún aleta amarilla, en toneladas, en el OPO, 1975-2019, por pesquería (arriba) y por tipo de arte/lance (abajo). The fisheries definition is in Table 1.. Las capturas de las pesquerías F29-F35 están registradas en número de peces, y el modelo las convierte en peso. NOTA: La escala anual del eje y varía por color. La definición de las pesquerías está en la Tabla 1.

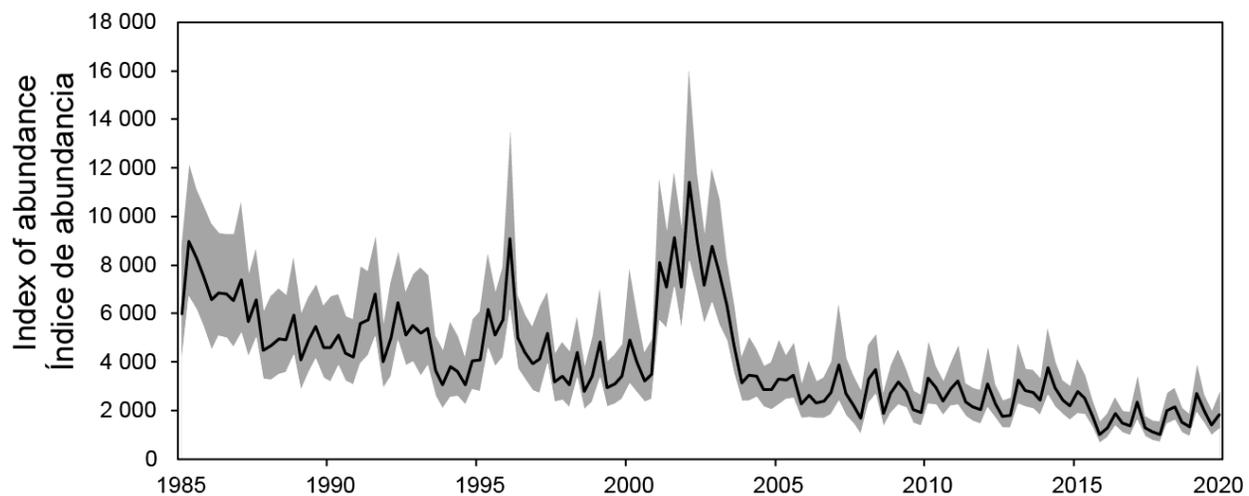


FIGURE 3. Standardized index of abundance used in the stock assessment of yellowfin tuna in the EPO, 1985-2019 (black line), and its associated 95% confidence interval (shading), based on data from the purse-seine fisheries on dolphins north of 5°N (F15-F19).

FIGURA 3. Índice de abundancia estandarizado usado en la evaluación del aleta amarilla en el OPO, 1985-2019 (línea negra), y su intervalo de confianza de 95% asociado (sombreado), basado en datos de las pesquerías cerqueras sobre delfines al norte de 5°N (F15-F19).

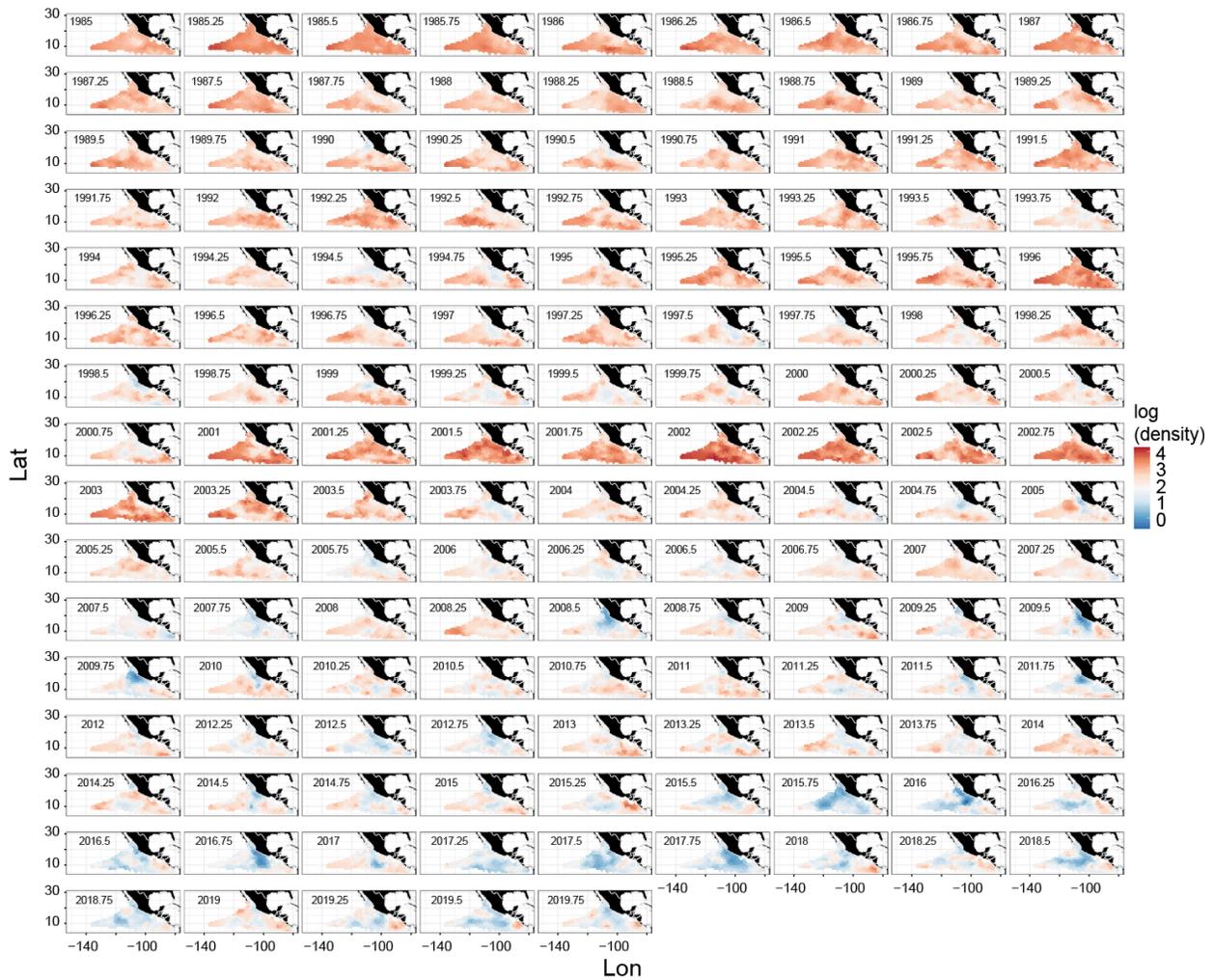


FIGURE 4. Log(density) of yellowfin tuna in the EPO, 1985-2019, by quarter, predicted by the delta-lognormal VAST model.

FIGURA 4. Log(densidad) del aleta amarilla en el OPO, 1985-2019, por trimestre, predicho por el modelo VAST delta-lognormal.

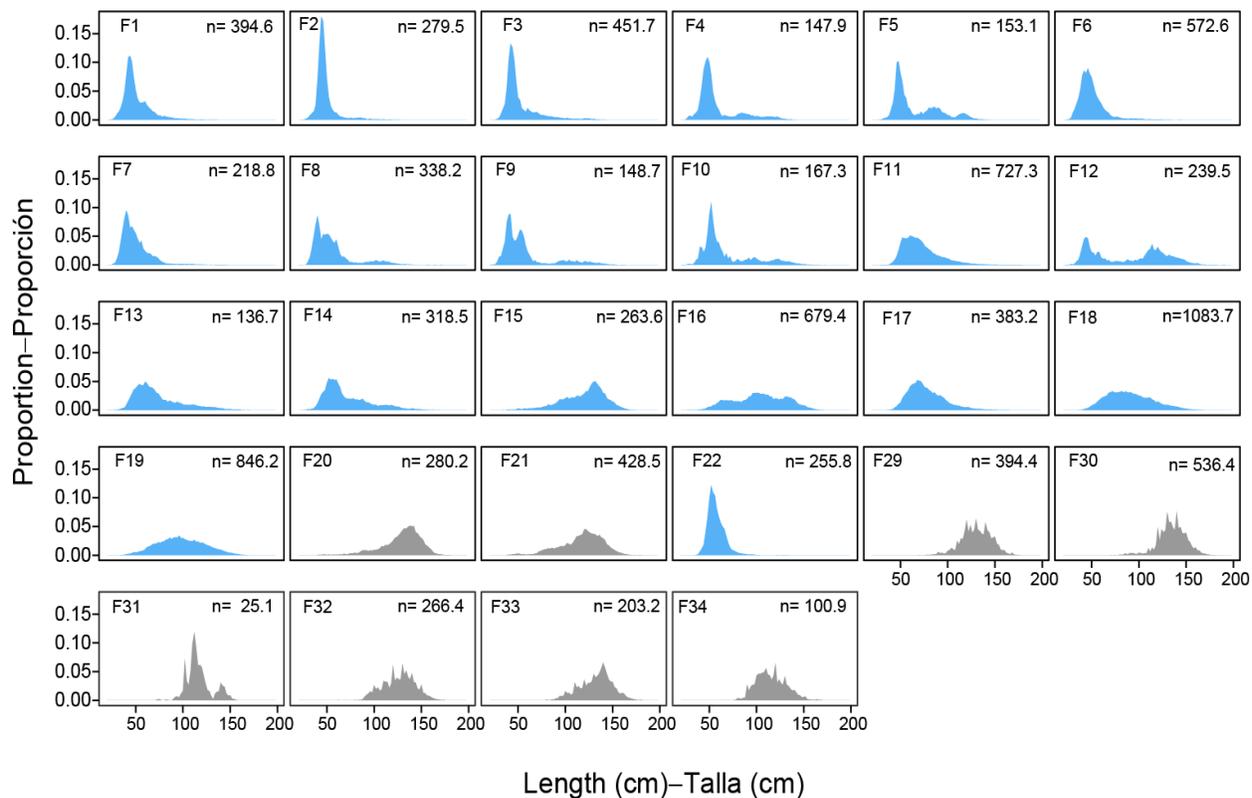


FIGURE 5A. Weighted average observed length composition of yellowfin tuna in the EPO, by fishery; n is the total sample size, adjusted by the weight given to the data in the models. Blue: data fitted by the reference models; grey: data not fitted.

FIGURA 5A. Promedio ponderado de la composición por talla observada del atún aleta amarilla en el OPO, por pesquería; n es el tamaño total de la muestra, ajustado por el peso asignado a los datos en los modelos. Azul: datos ajustados por los modelos de referencia; gris: datos no ajustados.

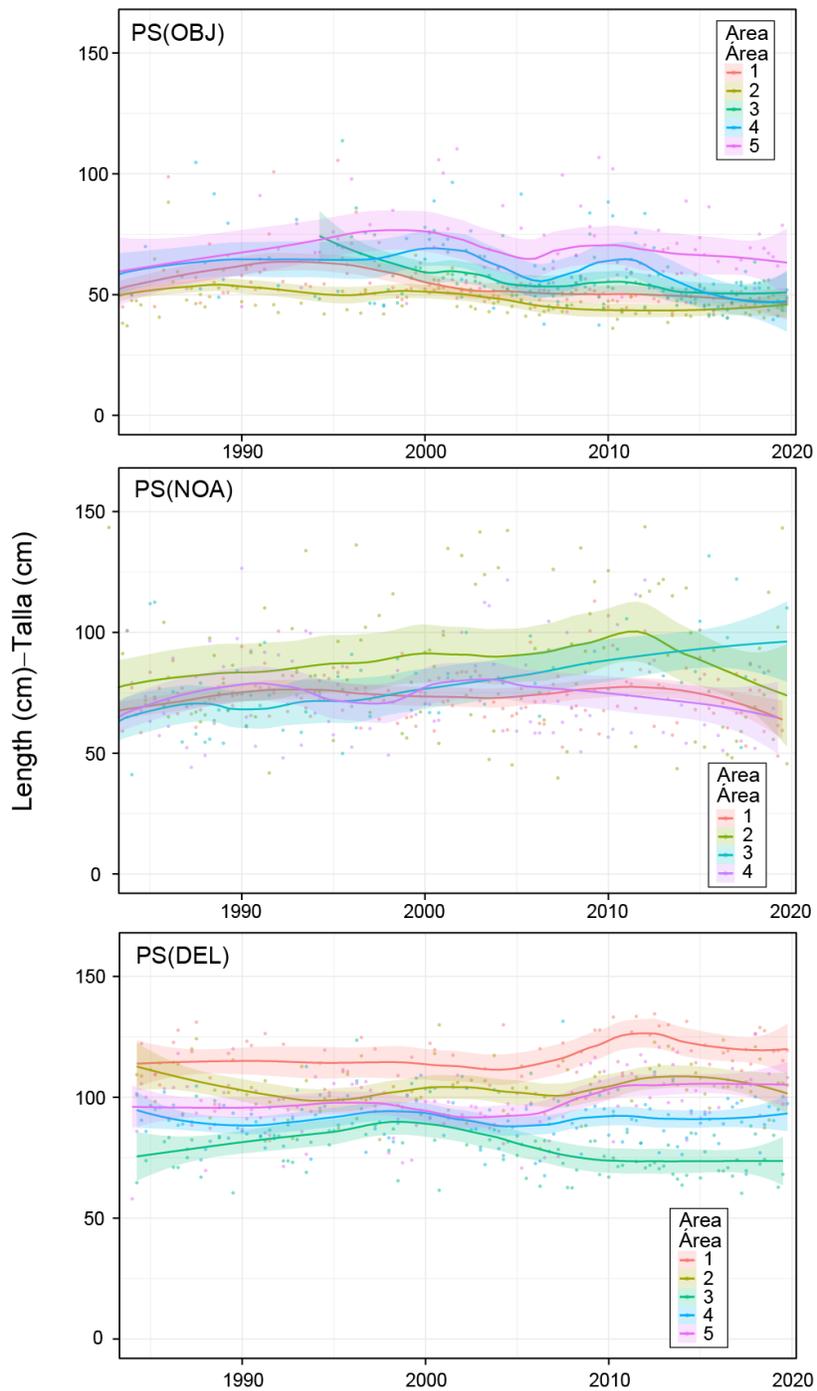


FIGURE 5B. Mean length of yellowfin tuna caught in purse-seine fisheries in the eastern Pacific Ocean, 1985-2019. The colored lines represent the LOESS-smoothed (span = 0.5) values. See Figure 1 and Table 1 for the definition of the areas.

FIGURA 5B. Talla promedio del aleta amarilla capturado en las pesquerías de cerco en el OPO, 1985-2019. Las líneas de colores representan los valores suavizados con LOESS (ancho de banda = 0.5). Ver la definición de las áreas en la Figura 1 y la Tabla 1.

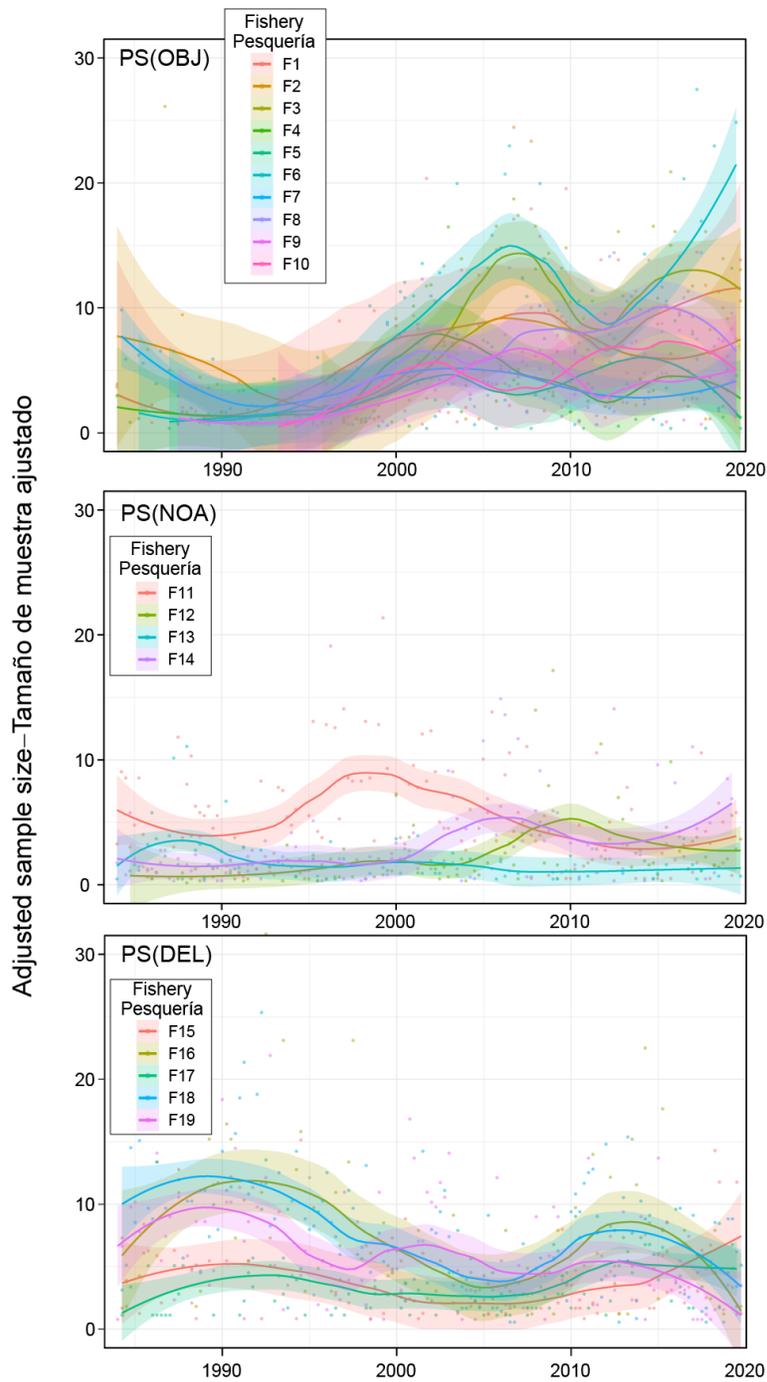


FIGURE 5C. Sample sizes of the length-composition data used in the stock assessment of yellowfin tuna in the EPO, 1985-2019, adjusted using the Francis weighting method. The colored lines represent the LOESS-smoothed values (span = 0.25). See Figure 1 and Table 1 for the definition of the fisheries.

FIGURA 5C. Tamaños de muestra de los datos de composición por talla usados en la evaluación del aleta amarilla en el OPO, 1985-2019, ajustados usando el método de ponderación de Francis. Las líneas de colores representan los valores suavizados con LOESS (ancho de banda = 0.25). Ver la definición de las pesquerías en la Figura 1 y la Tabla 1..

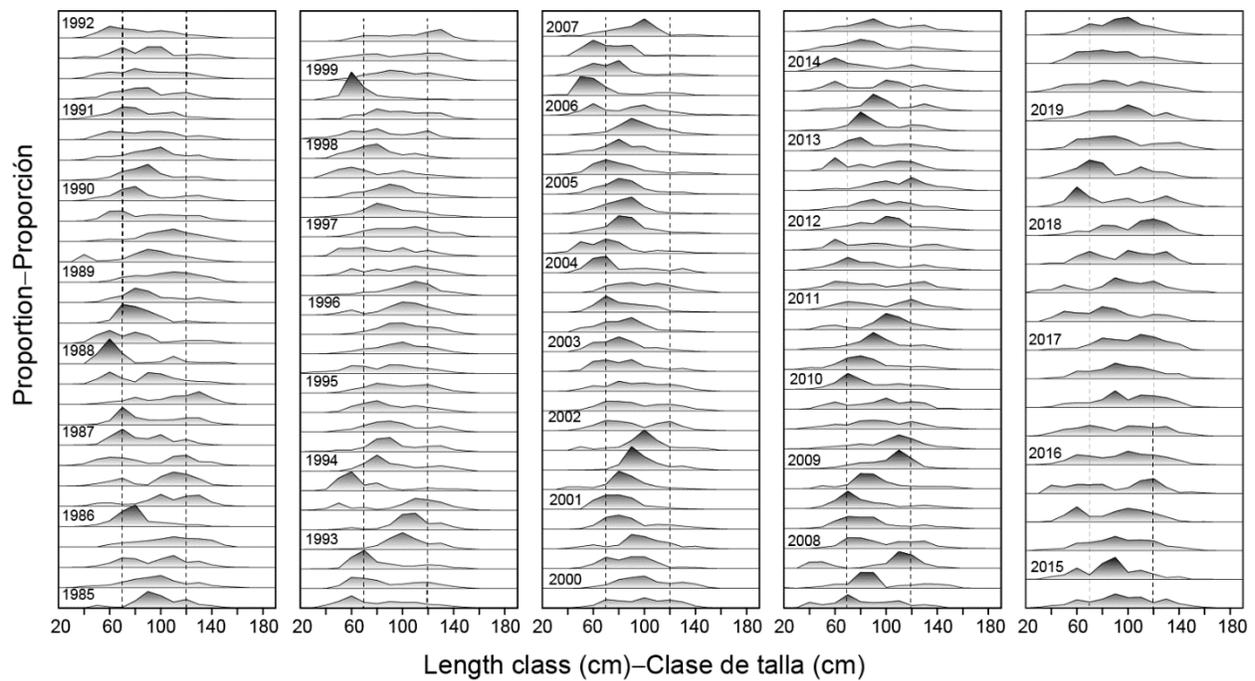


FIGURE 6. Standardized length compositions of yellowfin tuna in the EPO, by quarter, 1985-2019. The dashed vertical lines are at 70 and 120 cm.

FIGURA 6. Composición por talla estandarizada del aleta amarilla en el OPO, por trimestre, 1985-2019. Las líneas de trazos verticales están en 70 y 120 cm.

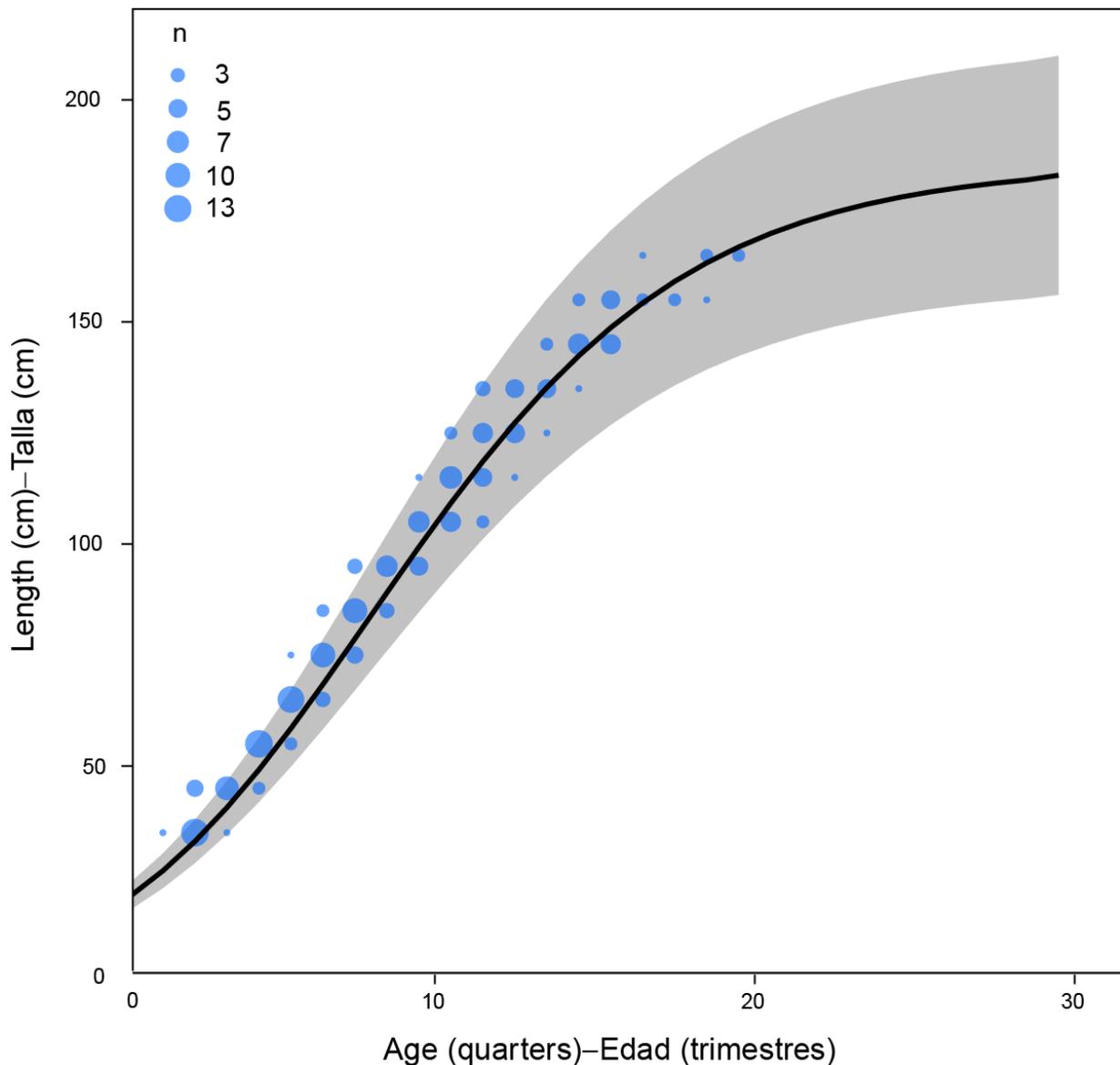


FIGURE 7. Age conditional on length for yellowfin tuna in the EPO, from Wild (1986). The size of the dots represents the number of fish (n) of each age, in quarters, by 10-cm intervals. The black line is the externally-estimated growth assumed in the fixed-growth models. The shaded region represents variation in length-at-age, assuming a CV = 7.5% (mean \pm 1.96 standard deviations).

FIGURA 7. Edad condicional a la talla para el aleta amarilla en el OPO, de Wild (1986). El tamaño de los puntos representa el número de peces (n) de cada edad, en trimestres, por intervalo de 10 cm. La línea negra es el crecimiento, estimado externamente, supuesto en los modelos de crecimiento fijo. La región sombreada representa la variación de la talla por edad, suponiendo un CV = 7.5% (promedio \pm 1.96 desviaciones estándar).

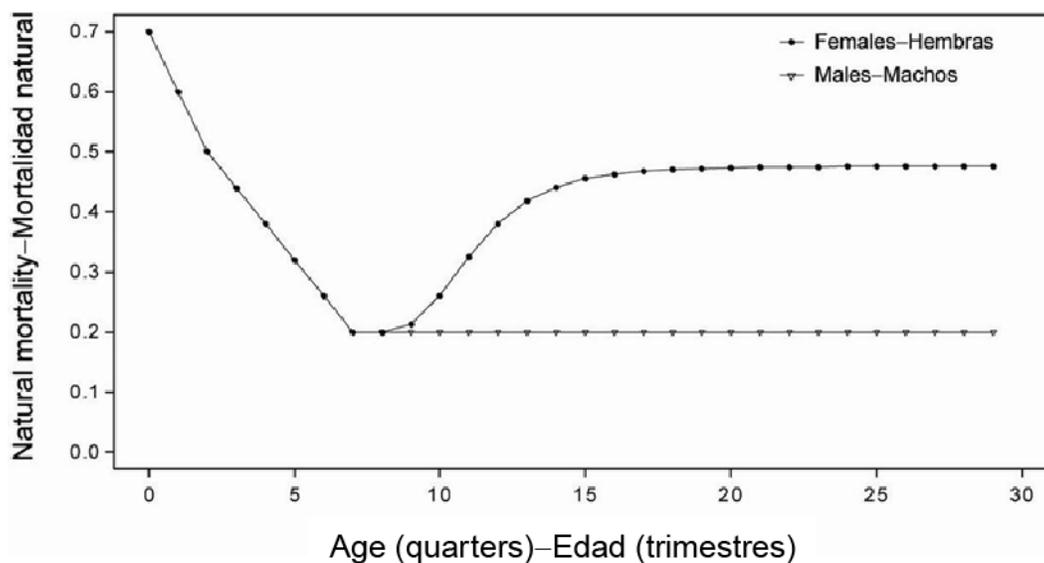


FIGURE 8. Natural mortality rates (M), by age and sex, at quarterly intervals, used for the assessment of yellowfin tuna in the EPO.

FIGURA 8. Tasas de mortalidad natural (M), por edad y sexo, en intervalos trimestrales, usadas para la evaluación del aleta amarilla en el OPO.

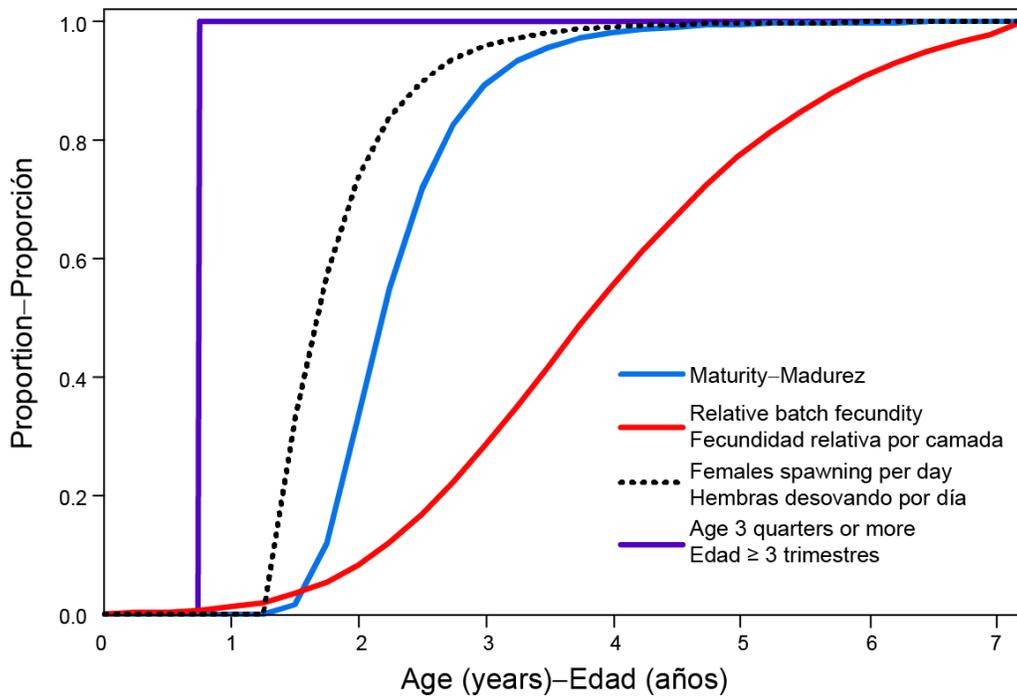


FIGURE 9. Top: Relative contribution of each age to the reproductive output component (scaled to a maximum of one) for yellowfin tuna in the EPO (from Schaefer 1998). **Bottom:** Relative fecundity-at-age curve used to estimate the index of spawning biomass of yellowfin tuna in the EPO.

FIGURA 9. Arriba: Contribución relativa de cada edad al componente de rendimiento reproductivo (escala ajustada a un máximo de uno) del aleta amarilla en el OPO (de Schaefer 1998). **Abajo:** Curva de fecundidad relativa por edad usada para estimar el índice de biomasa reproductora del aleta amarilla en el OPO.

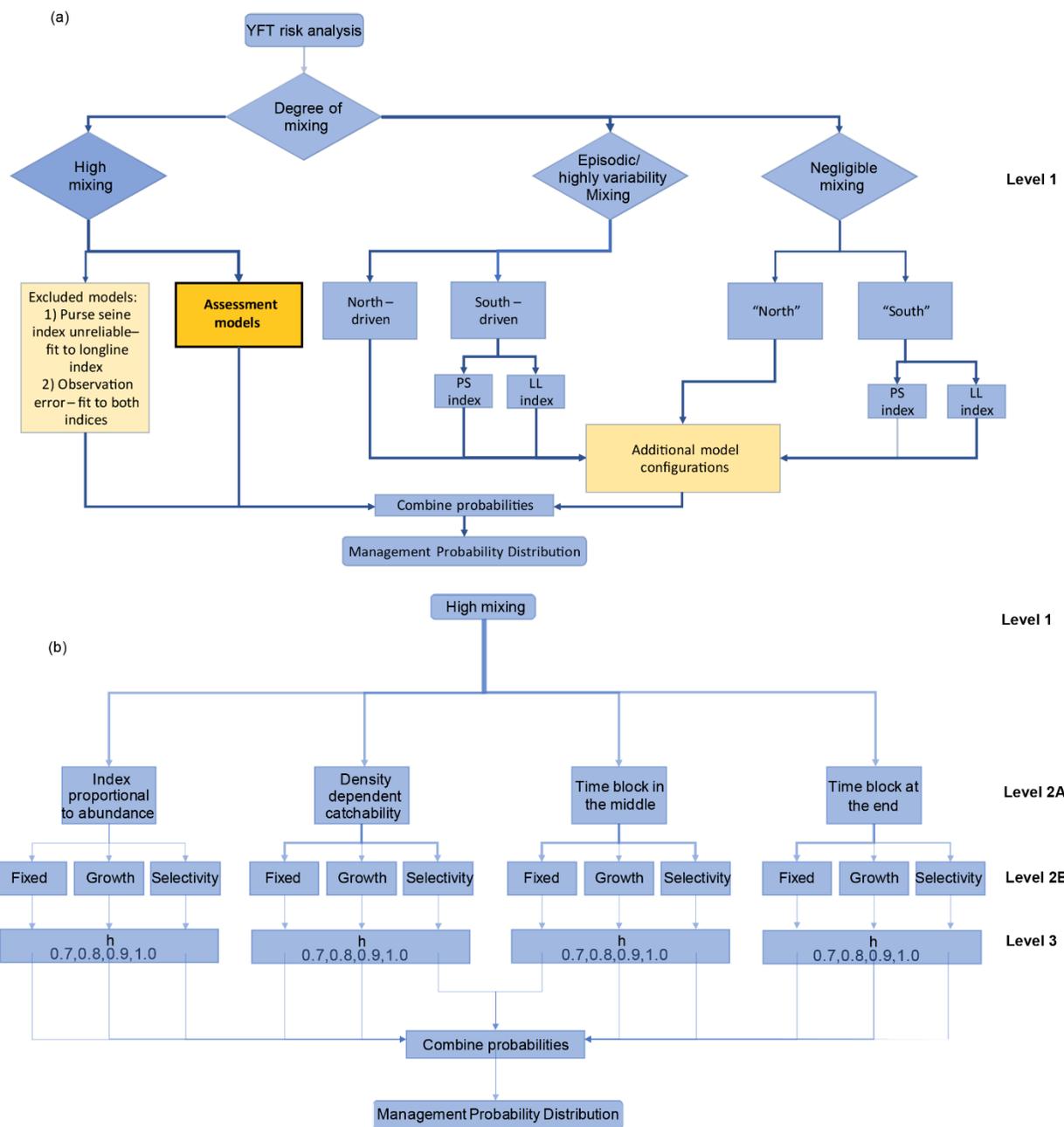


FIGURE 10. Flow chart of hypotheses and models (a) considered and (b) included in the yellowfin risk analysis (see text for details).

FIGURA 10. Diagrama de flujo de hipótesis y modelos (a) considerados y (b) incluidos en el análisis de riesgos del aleta amarilla (ver detalles en el texto).

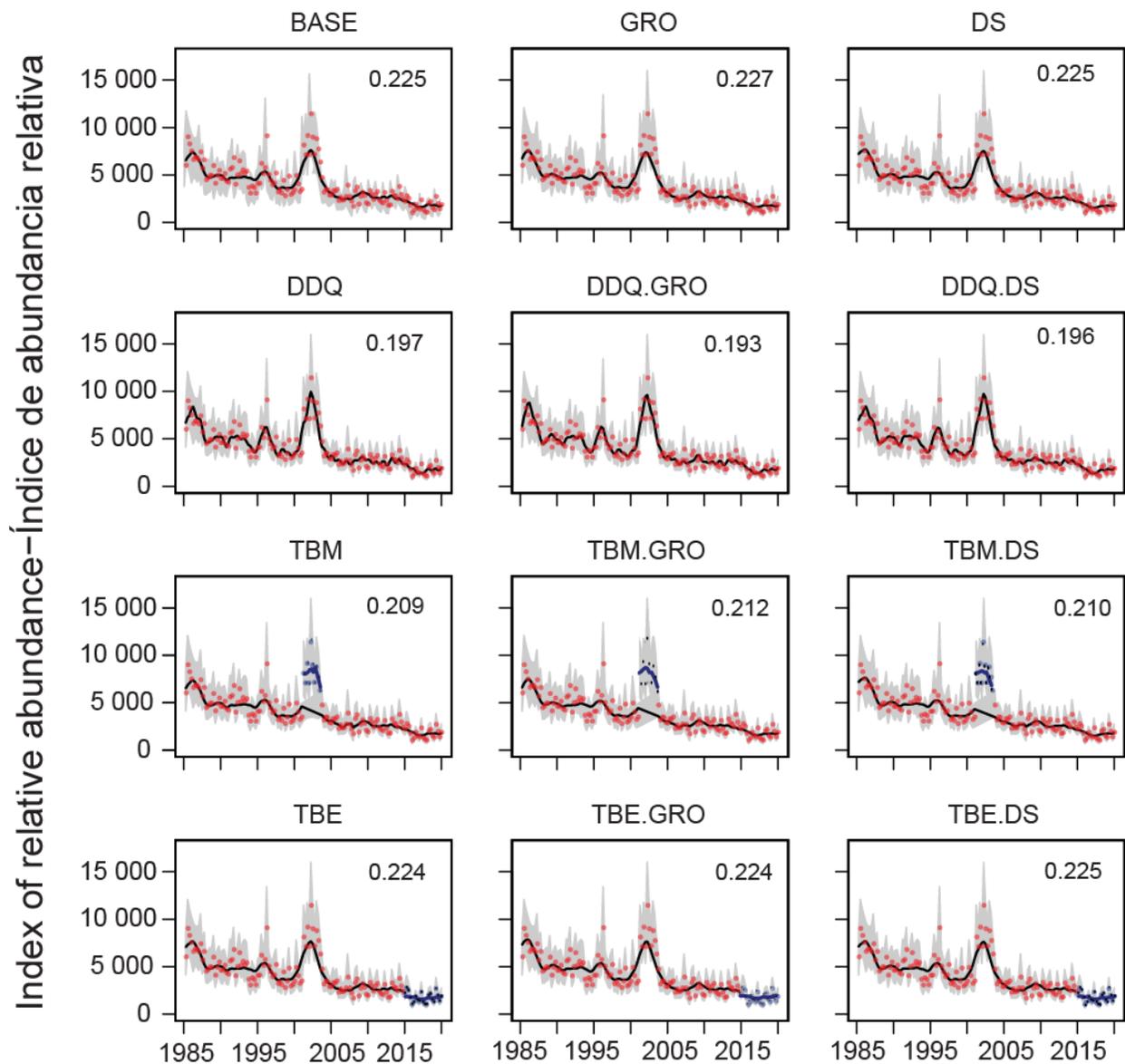


FIGURE 11. Model fits to the CPUE-based indices of abundance for the dolphin-associated fisheries, with steepness fixed at 1.0. The black lines represent the estimated indices, and the shading the approximate 95% confidence intervals (see 3.3). The colored dots indicate the observed CPUE values: blue dots the data corresponding to time blocks (TBM and TBE), and red dots the data outside those time blocks. See Table 3 for details of models.

FIGURA 11. Ajustes del modelo a los índices de abundancia basados en CPUE para las pesquerías asociadas a delfines, con la inclinación fija en 1.0. Las líneas negras representan los índices estimados, y el sombreado los intervalos de confianza de 95% aproximados (ver 3.3). Los puntos de colores indican los valores de CPUE observados: los puntos azules indican los datos correspondientes a los bloques de tiempo (TBM y TBE), y los puntos rojos los datos fuera de esos bloques de tiempo. Ver detalles de los modelos en la Tabla 3.

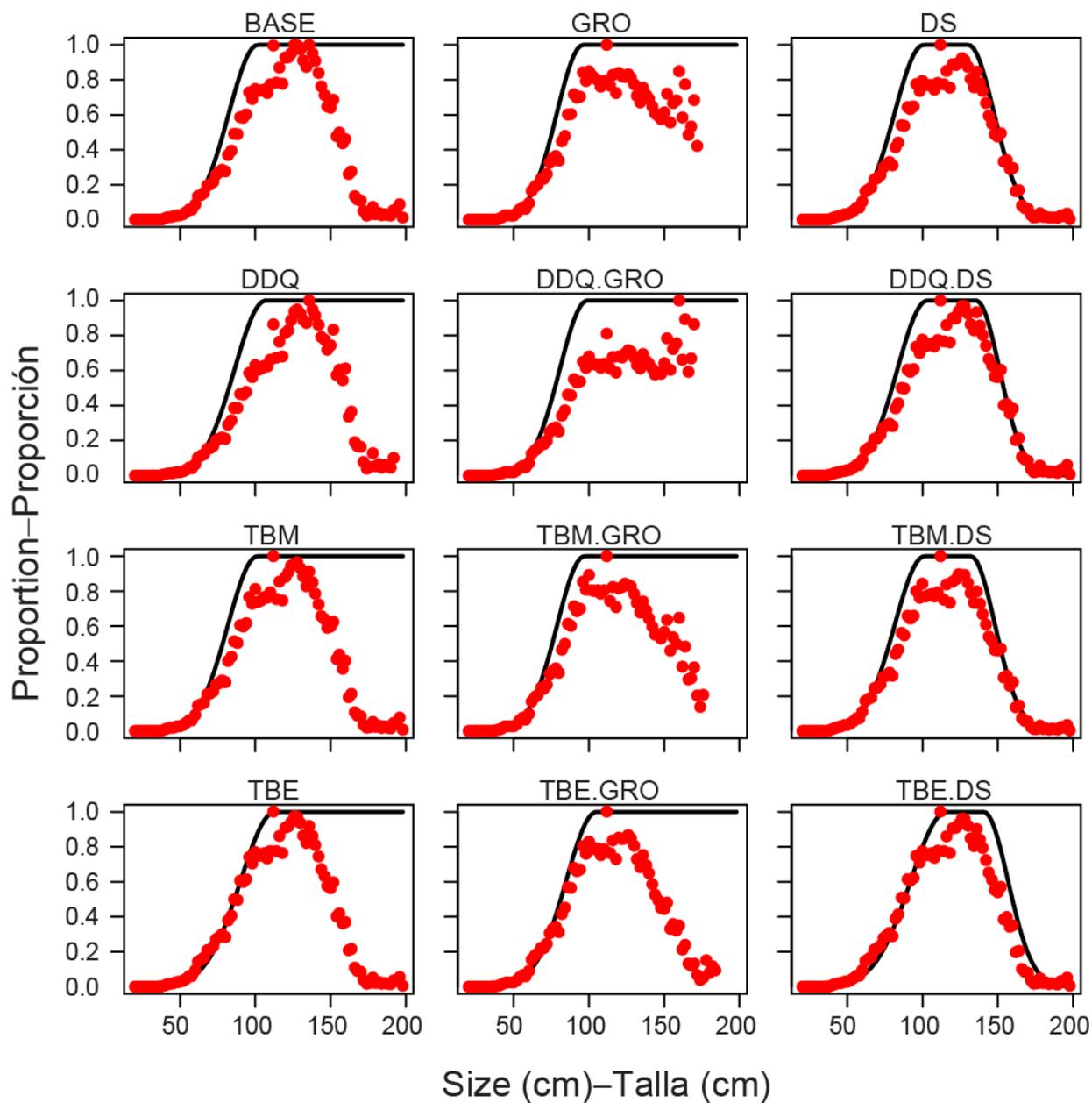


FIGURE 12. Comparison of estimated (black line) and empirical (red dots, see 3.2.2) average selectivity for fishery F19 in each reference model, with steepness = 1.0. The runs without red dots in the largest sizes predicted zero fish of those sizes in the population, even though there were fish in the sample. See model descriptions in Table 3.

FIGURA 12. Comparación de la selectividad promedio estimada (línea negra) y empírica (puntos rojos, ver 3.2.2) para la pesquería F19 en cada modelo de referencia, con inclinación = 1.0. Las ejecuciones sin puntos rojos en los tamaños más grandes predijeron cero peces de esos tamaños en la población, aunque hubo peces en la muestra. Ver descripciones de los modelos en la Tabla 3.

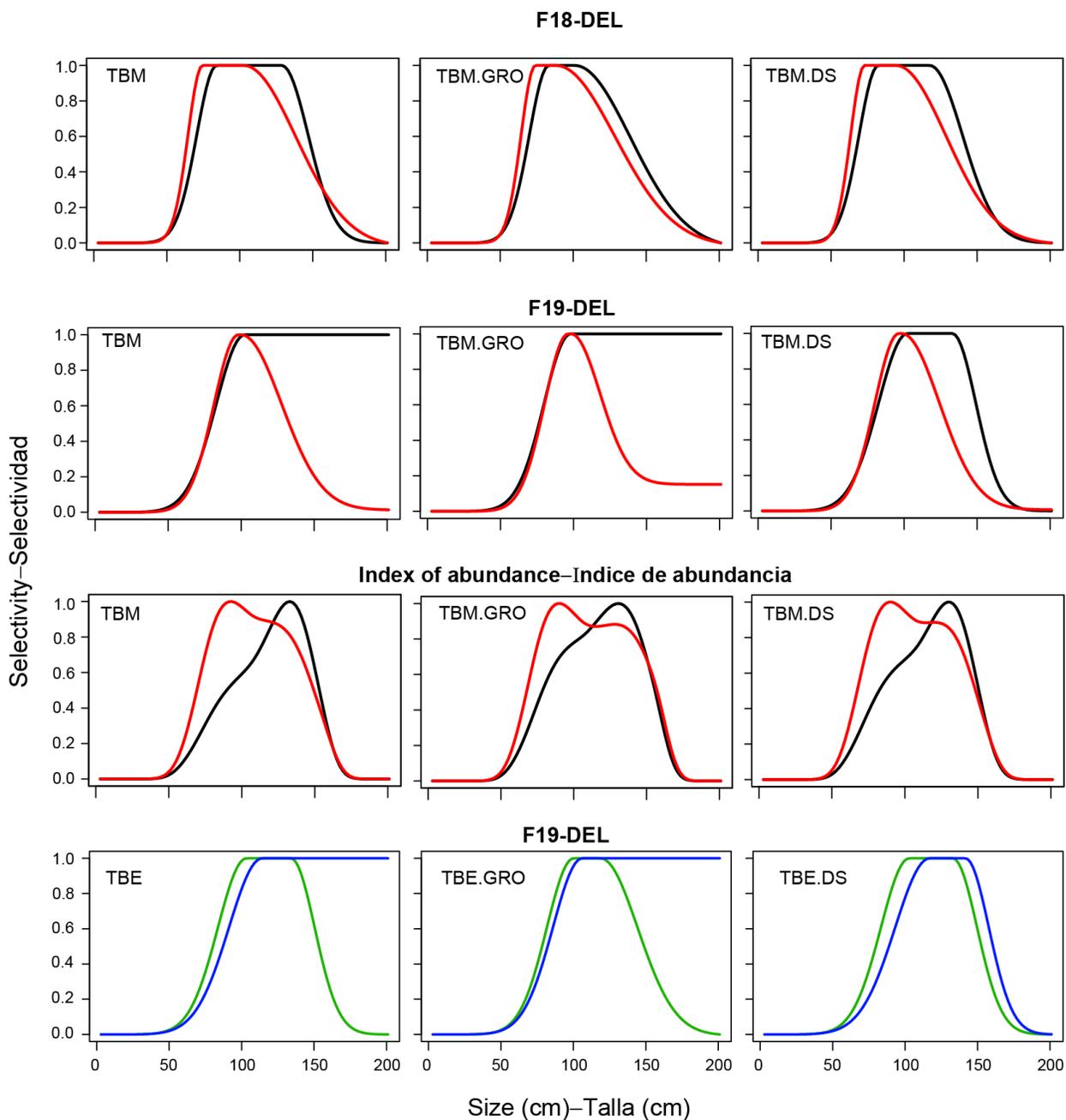


FIGURE 13. Estimated selectivity for the models with blocks in selectivity. The black line represents baseline selectivity, the red line the selectivity for 2002 (quarter 3) to 2007 (quarter 3). The blue and orange lines represent the selectivity for 1984-2014 and 2015-2019, respectively. See model descriptions in Table 3.

FIGURA 13. Selectividad estimada para los modelos con bloques de selectividad. La línea negra representa la selectividad base, la línea roja la selectividad de 2002 (trimestre 3) a 2007 (trimestre 3). Las líneas azul y naranja representan la selectividad de 1984-2014 y 2015-2019, respectivamente. Ver descripciones de los modelos en la Tabla 3.

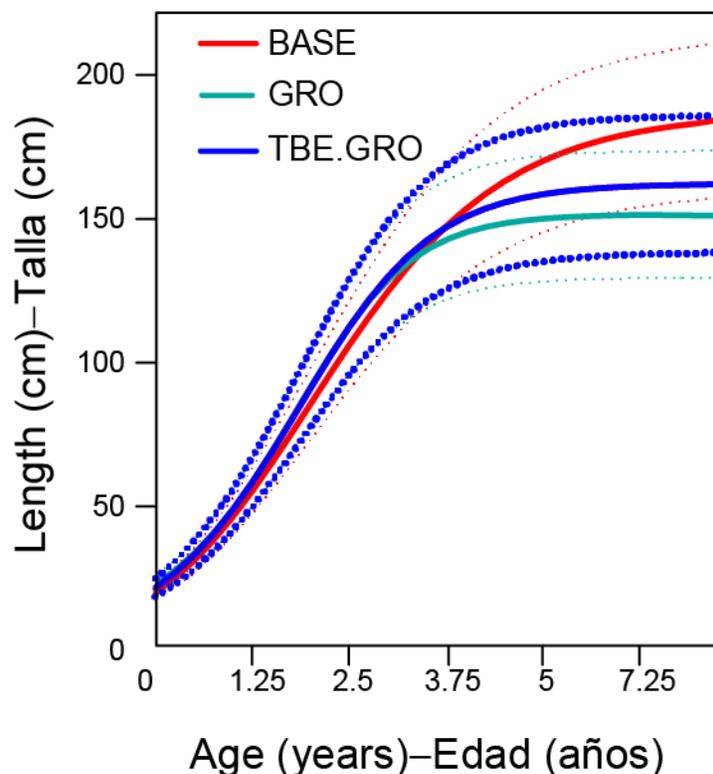


FIGURE 14. Schnute-Richards growth curves for different reference models in the yellowfin tuna benchmark assessment for the EPO. The solid lines represent the mean length-at-age, the dotted lines the variation in length-at-age (mean \pm 1.96 standard deviations) that encompasses 95% of the fish of that age in the population (assuming the length-at-age has a normal distribution). The BASE curves is the fixed assumption. The others are estimated. The TBM.GRO curve (not shown) is similar to the GRO curve. See model descriptions in Table 3.

FIGURA 14. Curvas de crecimiento de Schnute-Richards de modelos de referencia diferentes de la evaluación de referencia del atún aleta amarilla del OPO. Las línea sólidas representan la talla promedio por edad, las líneas de trazos la variación de la talla por edad (promedio \pm 1.96 desviaciones estándar) que comprende el 95% de los peces de esa edad en la población (suponiendo que la talla por edad tiene una distribución normal). La curva BASE es el supuesto de crecimiento fijo. Las otras son estimadas. La curva TBM.GRO (no ilustrada) es similar a la curva de GRO. Ver descripciones de los modelos en la Tabla 3.

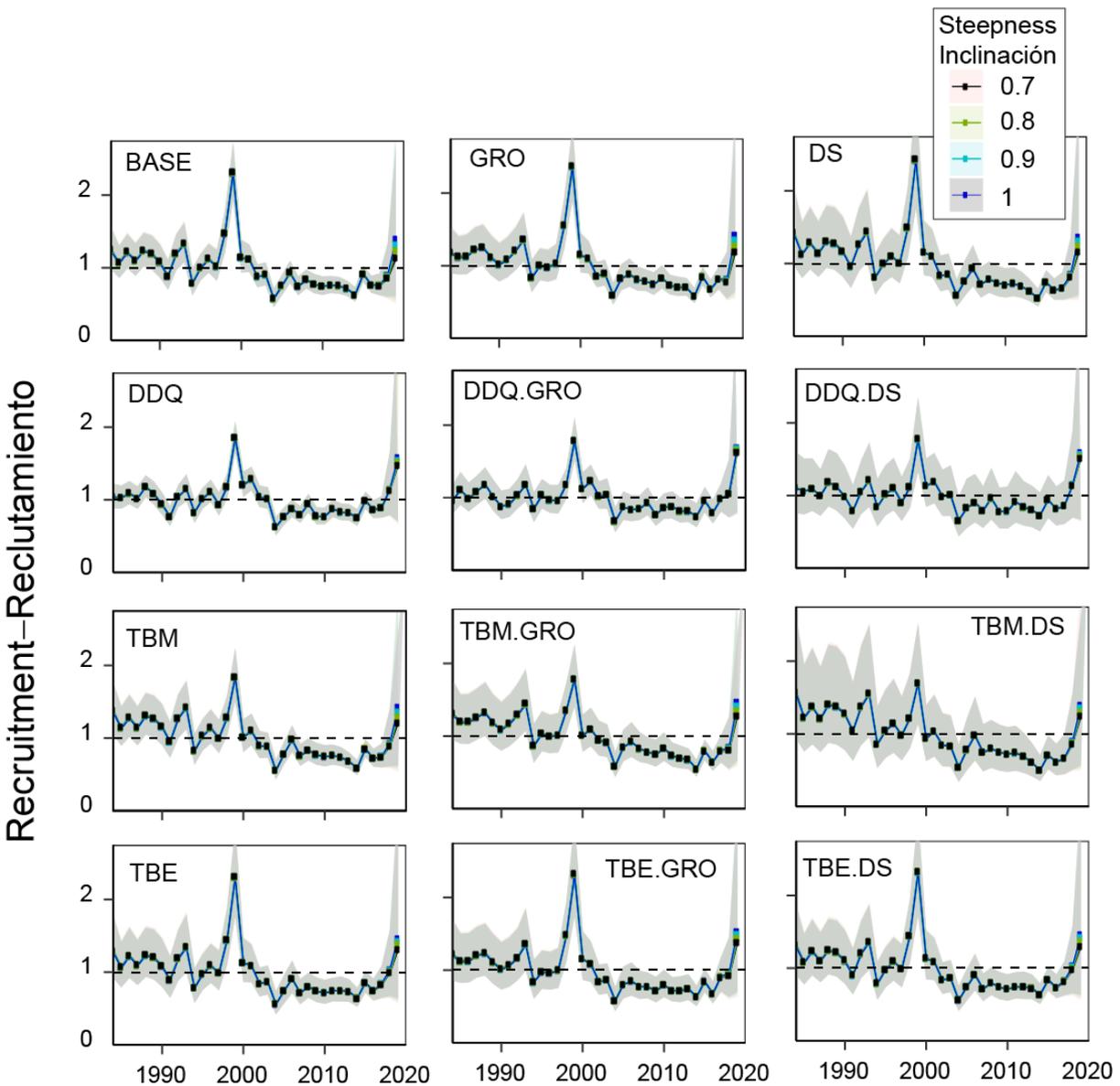


FIGURE 15. Annual relative recruitment of yellowfin tuna to the fisheries of the EPO estimated by the reference models. The lines and dots indicate the maximum likelihood estimates of recruitment, and the shaded areas the approximate 95% confidence intervals around the estimates. The estimates are scaled so that the average recruitment is equal to 1.0 (dashed horizontal line). The results for the four values of the steepness parameter (h) are almost identical. See model descriptions in Table 3.

FIGURA 15. Reclutamiento anual relativo del aleta amarilla en las pesquerías del OPO estimado por los modelos de referencia. Las líneas y puntos indican las estimaciones de máxima verosimilitud (EMV) del reclutamiento, y las áreas sombreadas los intervalos de confianza de 95% aproximados alrededor de las estimaciones. Se ajusta la escala de las estimaciones para que el reclutamiento promedio sea igual a 1.0 (línea de trazos horizontal). Los resultados correspondientes a los cuatro valores del parámetro de inclinación (h) son casi idénticos. Ver descripciones de los modelos en la Tabla 3.

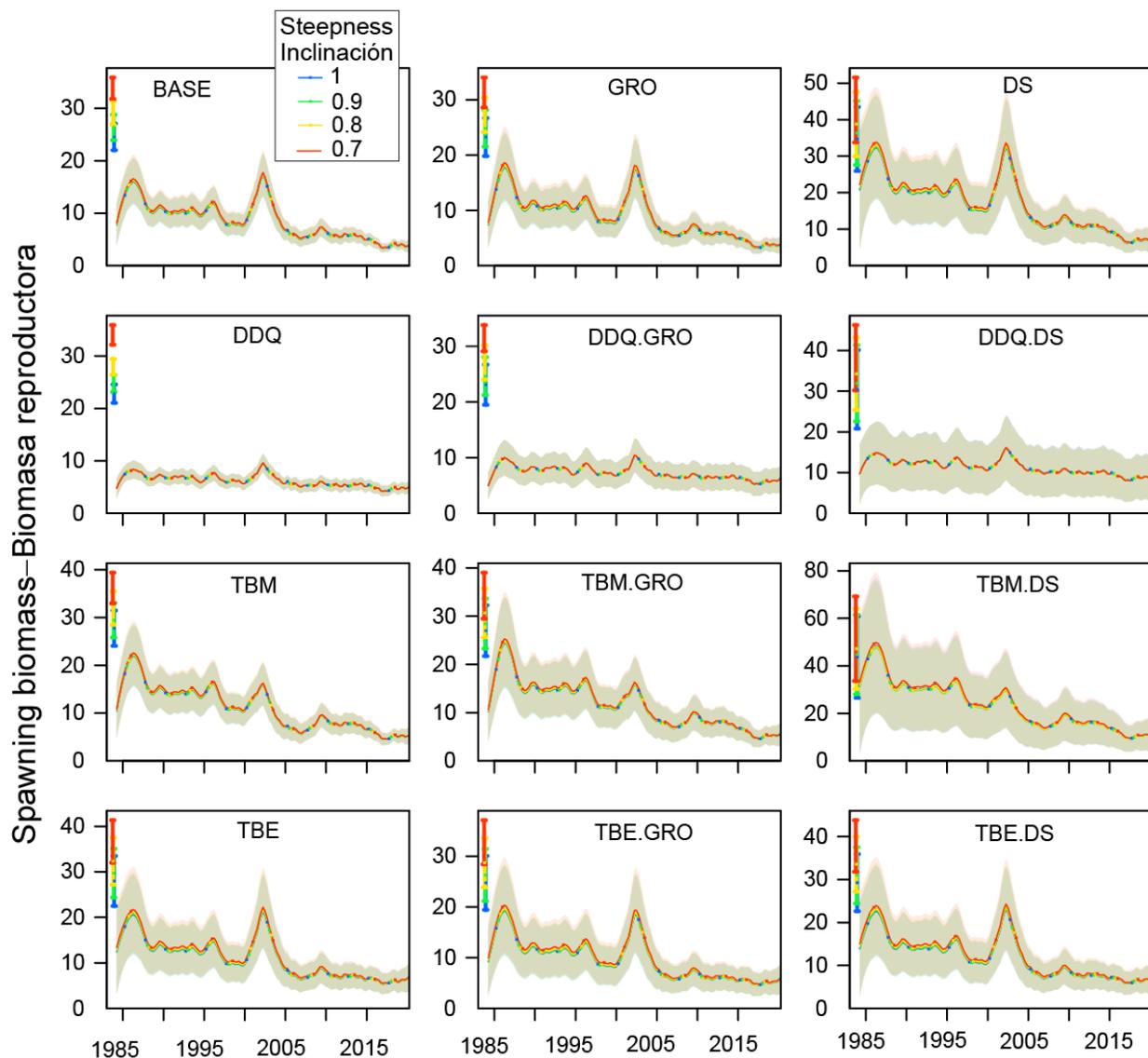


FIGURE 16. Spawning biomass, in thousands of fish, of yellowfin tuna in the EPO, 1985-2019, from the 12 reference models. The solid lines indicate the maximum likelihood estimates for four values of the steepness parameter (h), and the shaded areas the approximate 95% confidence intervals around those estimates. The colored bars and points on the left edge of each panel are the estimates of virgin spawning biomass for each model. See model descriptions in Table 3.

FIGURA 16. Biomasa reproductora, en miles de peces, del aleta amarilla en el OPO, 1985-2019, de los 12 modelos de referencia. Las líneas sólidas indican las estimaciones de máxima verosimilitud correspondientes a cuatro valores del parámetro de inclinación (h), y las áreas sombreadas los intervalos de confianza de 95% aproximados alrededor de esas estimaciones. Las barras y los puntos de color al borde izquierdo de cada panel son las estimaciones de biomasa reproductora virgen para cada modelo. Ver descripciones de los modelos en la Tabla 3

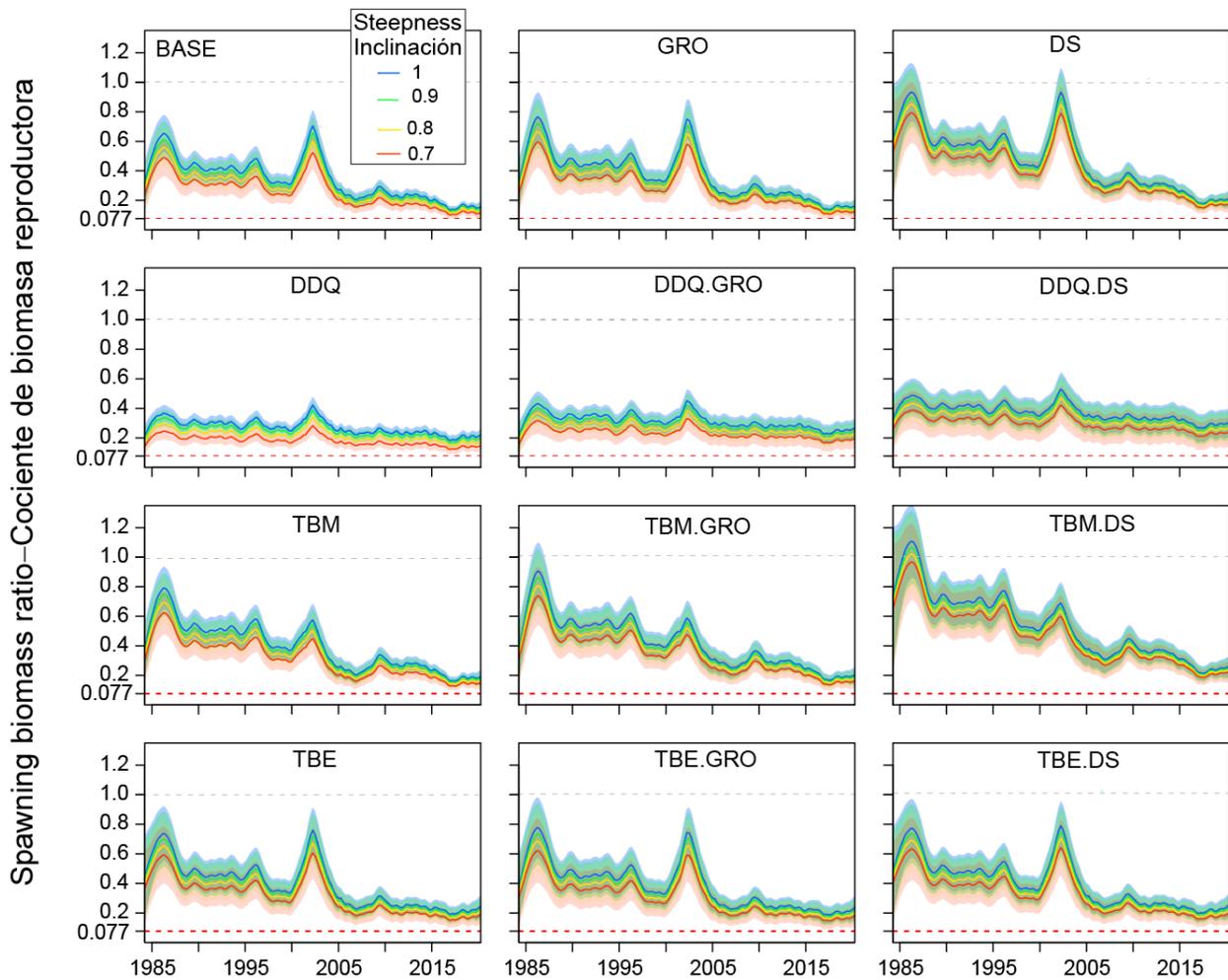


FIGURE 17. Spawning biomass ratios (SBRs) for yellowfin tuna in the EPO, 1985-2019. The solid lines represent the maximum likelihood estimates for four values of the steepness parameter (h), and the shaded areas the approximate 95% confidence intervals around those estimates. The red dashed horizontal line (at 0.077) identifies the SBR at S_{LIMIT} . See model descriptions in Table 3.

FIGURA 17. Cocientes de biomasa reproductora (SBR) del aleta amarilla en el OPO, 1985-2019. Las líneas sólidas representan las estimaciones de máxima verosimilitud correspondientes a cuatro valores del parámetro de inclinación (h). Las áreas sombreadas son los intervalos de confianza de 95% aproximados alrededor de esas estimaciones. La línea de trazos horizontal roja (en 0.077) identifica el SBR en $S_{LÍMITE}$. Ver descripciones de los modelos en la Tabla 3

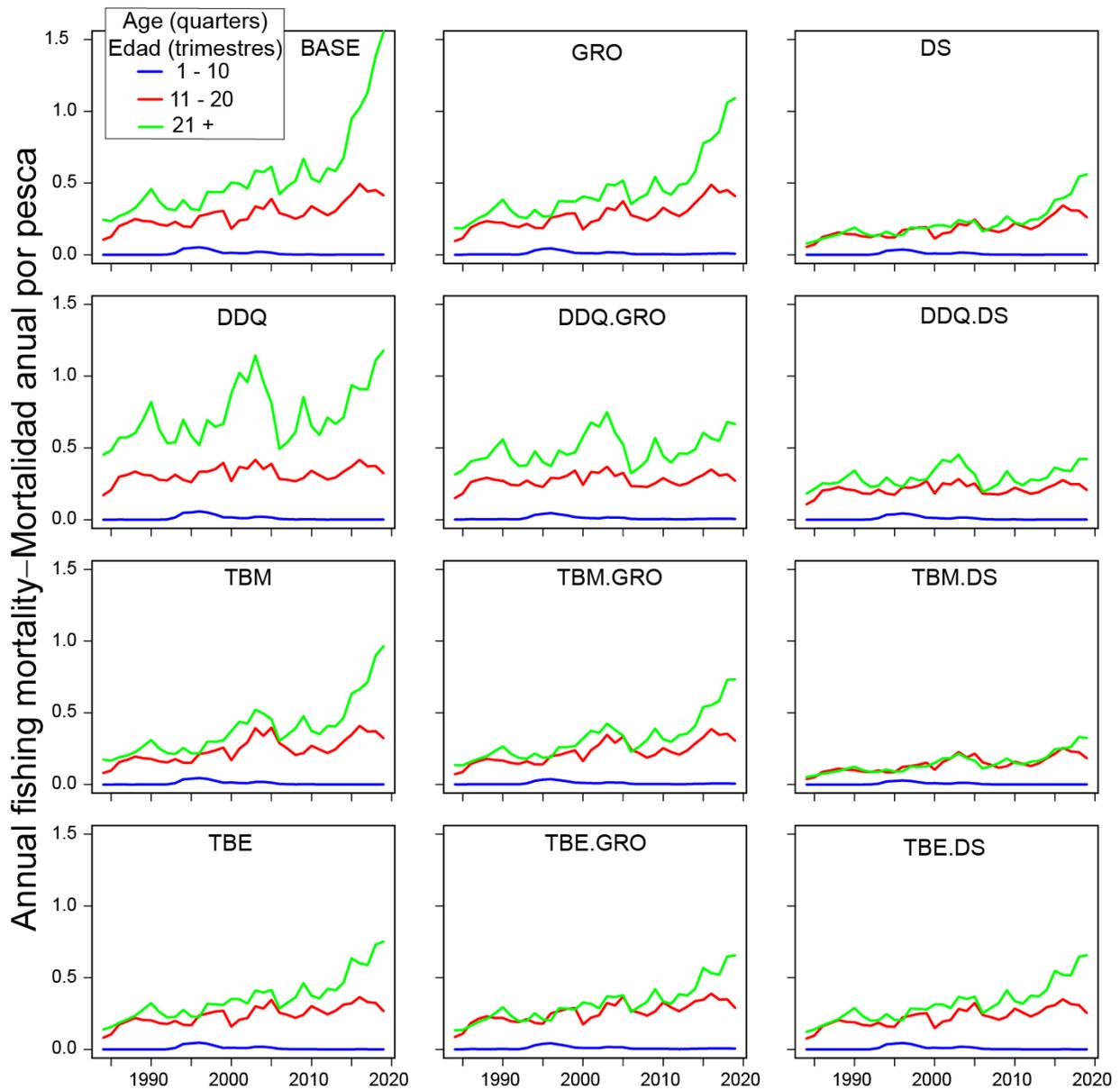


FIGURE 18. Average annual fishing mortality (F) of yellowfin tuna in the EPO, by age group (in quarters), for all gears, estimated by the 12 reference models with steepness = 1.0. See model descriptions in Table 3.

FIGURA 18. Mortalidad por pesca (F) anual promedio del atún aleta amarilla en el OPO, por grupo de edad (en trimestres), por todas las artes, estimada por los 12 modelos de referencia con inclinación = 1.0. Ver descripciones de los modelos en la Tabla 3.

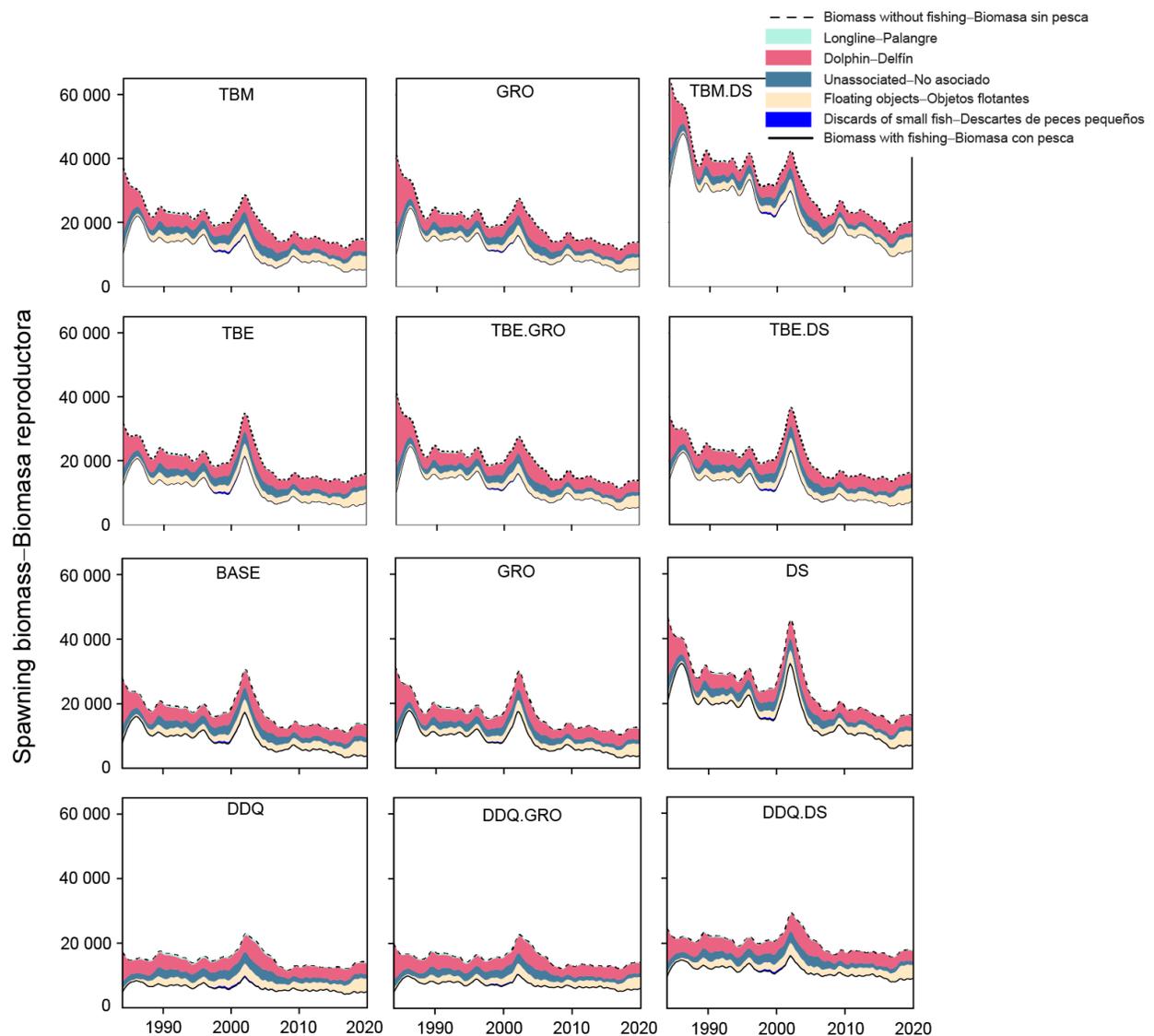


FIGURE 19. Impact of fishing, 1985-2019: trajectory of the spawning biomass (a fecundity index, see text for details) of a simulated population of yellowfin tuna that was never exploited (dashed line) and that predicted by each model, with a steepness of 1.0 (solid line). The shaded areas between the two lines show the portions of the impact attributed to each fishing method. See model descriptions in Table 3.

FIGURA 19. Impacto de la pesca, 1985-2019: trayectoria de la biomasa reproductora (un índice de fecundidad, ver detalles en el texto) de una población simulada de aleta amarilla que nunca fue explotada (línea de trazos) y la trayectoria predicha por cada modelo, con una inclinación de 1.0 (línea sólida). Las áreas sombreadas entre las dos líneas muestran las porciones del impacto atribuido a cada método de pesca. Ver descripciones de los modelos en la Tabla 3.

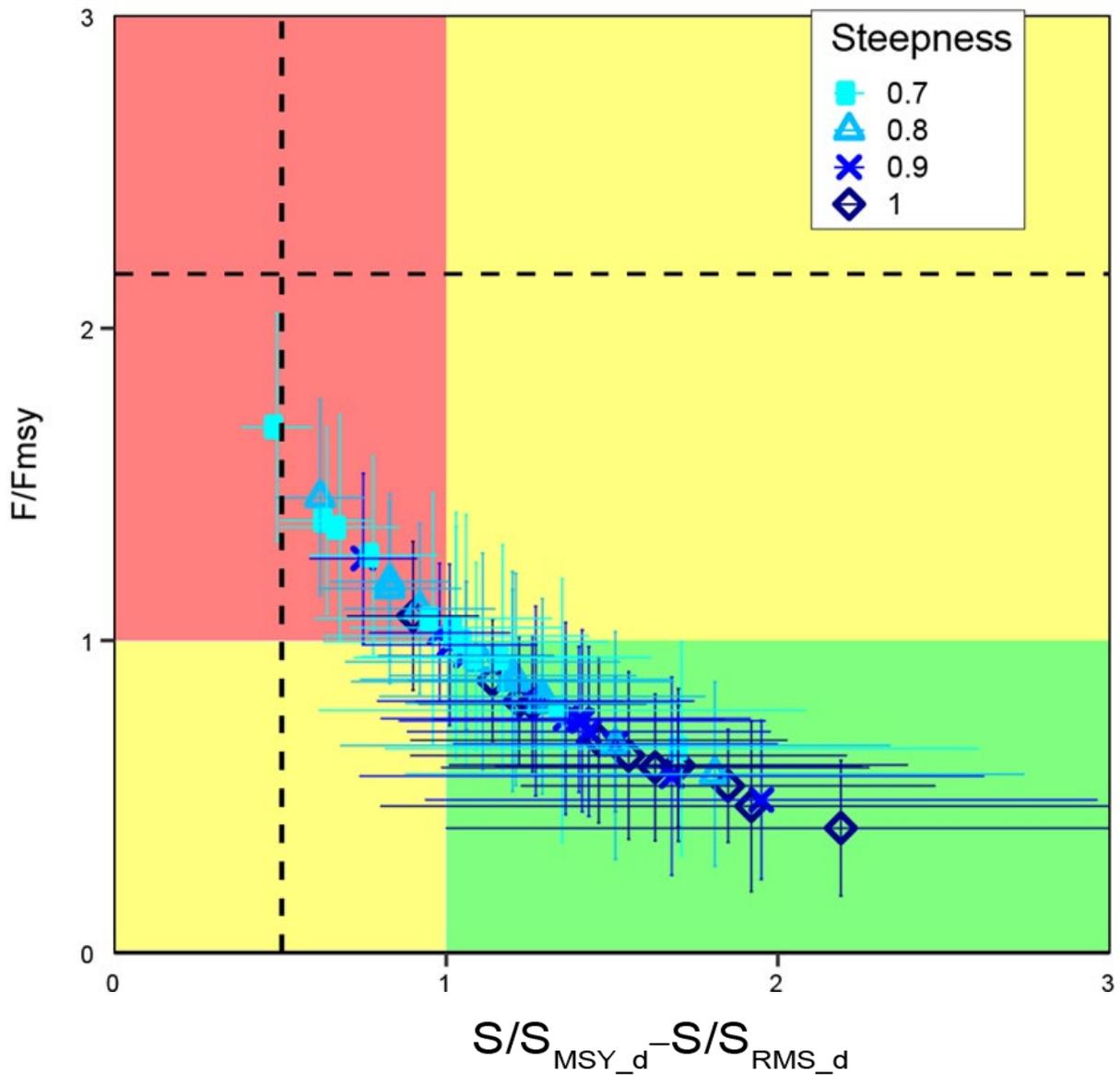


FIGURE 20. Kobe (phase) plot of the estimates of most recent spawning biomass ($S_{current}$) and current (2017-2019) fishing mortality ($F_{current}$) of yellowfin tuna in the EPO relative to their MSY-based reference points (S_{MSY_d} and F_{MSY}), from all models, for four values of the steepness parameter (h). The dashed lines represent the average of all 48 limit reference points. The bars represent the 95% confidence interval of the estimates. See Figure A16 for model names.

FIGURA 20. Gráfica de Kobe (fase) de las estimaciones de la biomasa reproductora (S_{actual}) más reciente y de la mortalidad por pesca (F_{actual}) actual (2017-2019) del atún aleta amarilla en el EPO, relativas a sus puntos de referencia basados en RMS (S_{RMS_d} y F_{RMS}), de todos los modelos, correspondientes a cuatro valores del parámetro de inclinación (h). Las líneas de trazos representan el promedio de los puntos de referencia límite de todos los 48 modelos. Las barras representan el intervalo de confianza de 95% de las estimaciones. Ver los nombres de los modelos en la figura A16.

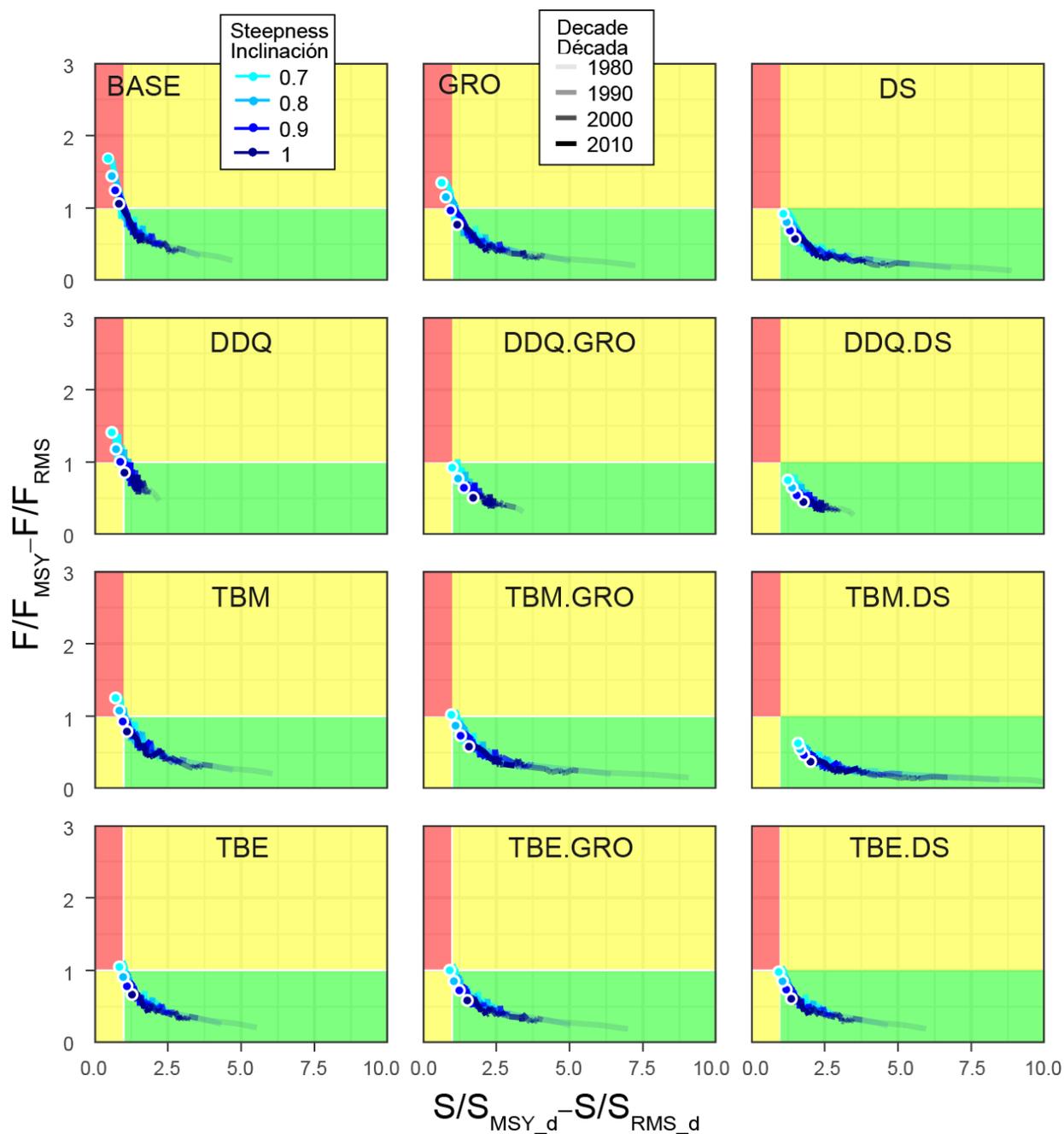


FIGURE 21. Kobe plot of the time series of estimated spawning biomass (S) and fishing mortality (F) relative to their MSY reference points, for each combination of reference model and steepness assumption. Each dot is based on the average F over three years. The white circles represent the most recent estimates. See Table 3 for explanation of model names.

FIGURA 21. Gráfica de Kobe de las series de tiempo de la biomasa reproductora (S) y mortalidad por pesca (F) estimadas con respecto a sus puntos de referencia de RMS, para cada combinación de modelo de referencia y supuesto de inclinación. Cada punto se basa en la F promedio en tres años. Los puntos de color representan las estimaciones más recientes. En la Tabla 3 se explican los nombres de los modelos.

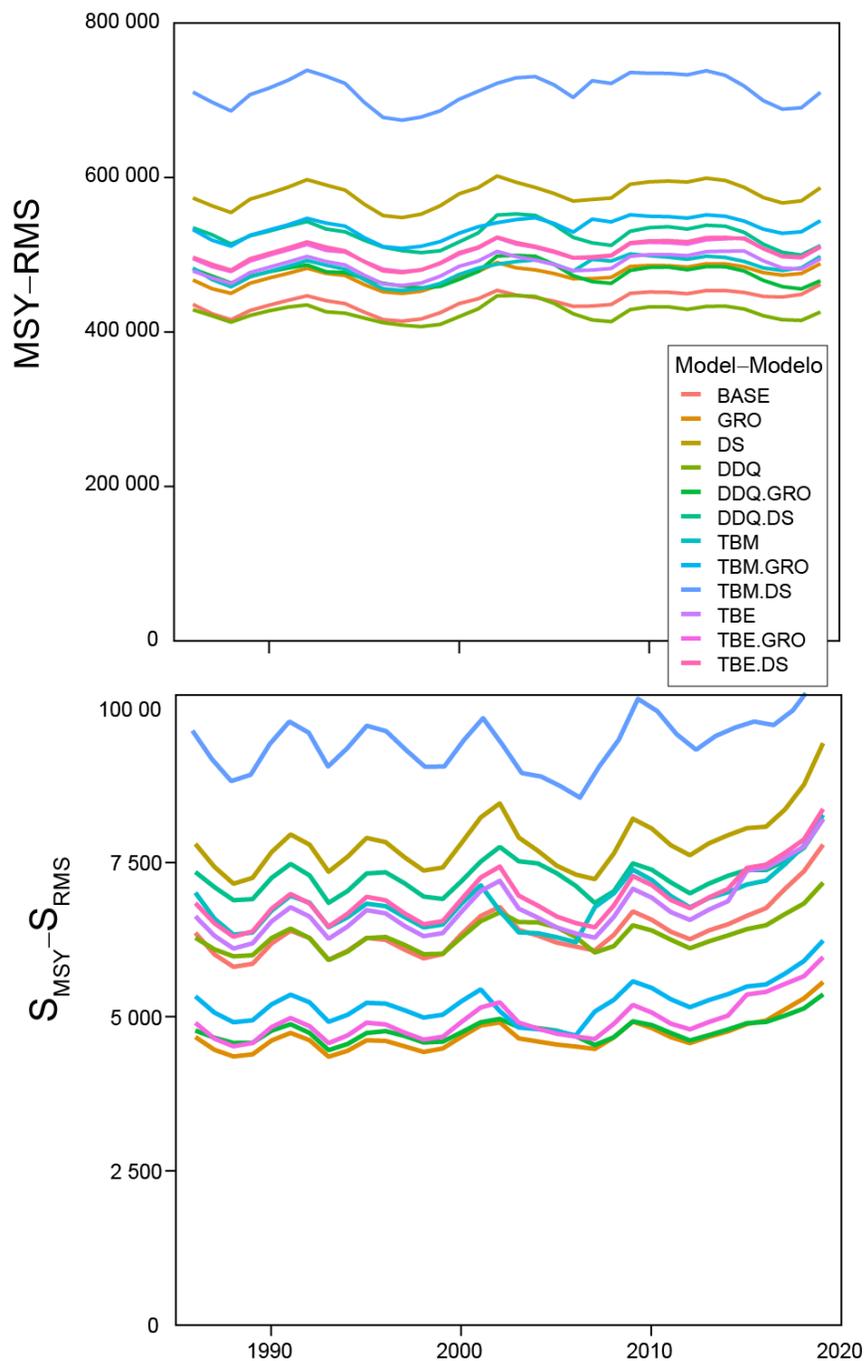


FIGURE 22. Estimates of spawning biomass (S) and the associated maximum sustainable yield in equilibrium (S_{MSY}) using the average age-specific fishing mortality for each year, for all models, with steepness = 1.0. See Table 3 for explanation of model names.

FIGURA 22. Estimaciones del rendimiento máximo sostenible (RMS) y la biomasa reproductora asociada (S_{RMS}) usando el promedio de la mortalidad por pesca por edad para cada año. Solo se muestran las estimaciones para un supuesto de inclinación de 1.0. En la Tabla 3 se explican los nombres de los modelos.

TABLE 2. Selectivity (at length unless noted otherwise) and weighting of composition data specified for the fisheries and surveys in the assessment. The asymptotic curves were modeled using a double-normal function. Selectivity at age was 1.0 for ages 1 to 29 quarters (unless noted otherwise). Sel – selectivity assumption, W – length-frequency data weighting. The number in parenthesis after *splines* is the number of knots. D. normal is the double normal selectivity function. F-A: Fixed-Asymptotic.

TABLA 2. Selectividad (por talla, salvo indicación al contrario) y ponderación de los datos de composición especificados para las pesquerías y estudios en la evaluación. Las curvas asintóticas fueron modeladas con una función doble normal. La selectividad por edad fue 1.0 para las edades 1 a 29 (salvo indicación al contrario). Sel – supuesto de selectividad, W - ponderación de los datos de frecuencia detalladas. El número entre parentesis despues *splines* es el número de nudos. D. normal es la funcion de selectividad doble normal. F-A: Fija-Asintótica.

Fishery - Pesquería	F1	F2	F3	F4	F5	F6
Sel	<i>Splines</i> (12)	<i>Splines</i> (11)	<i>Splines</i> (10)	<i>Splines</i> (10)	<i>Splines</i> (9)	<i>Splines</i> (9)
W	Francis/2	Francis	Francis/2	Francis/2	Francis/2	Francis
Fishery - Pesquería	F7	F8	F9	F10	F11	F12
Sel	<i>Splines</i> (6)	<i>Splines</i> (12)	<i>Splines</i> (14)	<i>Splines</i> (11)	<i>Splines</i> (10)	<i>Splines</i> (8)
W	Francis/2	Francis/2	Francis/2	Francis/2	Francis	Francis/2
Fishery - Pesquería	F13	F14	F15	F16	F17	F18
Sel	<i>Splines</i> (8)	<i>Splines</i> (9)	<i>Splines</i> (7)	<i>Splines</i> (10)	<i>Splines</i> (7)	Table - Tabla 3
W	Francis/2	Francis/2	Francis	Francis/2	Francis	Francis
Fishery - Pesquería	F19	F20	F21	F22		
Sel	D normal	F-A	F-A	<i>Splines</i> (9)		
W	Francis	0	0	Francis		
Fishery - Pesquería	F25	F26	F27	F28	F29	F30
Sel	Ages Edades 3-5	Ages Edades 3-5	Age Edades 3-5	Ages Edades 3-5	F-A	F-A
W	-	-	-	-	0	0
Fishery - Pesquería	F31	F32	F33	F34	F35	F36
Sel	F-A	F-A	F-A	F-A	= F29	= F30
W	0	0	0	0	0	0
Fishery - Pesquería	F37	F38	F39	F40		
Sel	= F31	= F32	= F33	= F34		
W	0	0	0	0		
Survey - Estudio	S41	S23	S24	S42	S43	
Sel	<i>Splines</i> (5)	D. normal	D. normal	D. normal	D. normal	
W	Francis	0	0	0	0	

TABLE 3. Reference models for the benchmark assessment of the yellowfin tuna in the EPO. The model input files and results for this assessment are available in [html and pdf formats](#).

TABLA 3. Modelos de referencia para la evaluación de referencia del atún aleta amarilla en el OPO. Los archivos de entrada del modelo y los resultados de esta evaluación están disponibles en [formato html y pdf](#).

	Model	Growth	Catchability of index	Index selectivity	Selectivity F19	Selectivity F18	Auxiliary data
1.	BASE	Fixed	Constant	Constant	Asymptotic	Dome-shape (11-knot spline)	
2.	GRO	Estimated	Constant	Constant	Asymptotic		age-at-length
3.	DS	Fixed	Constant	Constant	Dome-shape		
4.	DDQ	Fixed	Density-dependent	Constant	Asymptotic		
5.	DDQ.GRO	Estimated	Density-dependent	Constant	Asymptotic		age-at-length
6.	DDQ.DS	Fixed	Density-dependent	Constant	Dome-shape		
7.	TBM	Fixed	Blocks: Baseline, 2001-2003.Q2	block: 1984 – 2002.Q2 2002.Q3-2007.Q3	Blocks: Asymptotic (baseline) Dome-shape (2002.Q3- 2007.Q3)	Double-normal Block: Dome-shape (baseline) Dome-shape (2002.Q3- 2007.Q3)	
8.	TBM.GRO	Estimated	Blocks: Baseline, 2001-2003.Q2	block: 1984 – 2002.Q2 2002.Q3-2007.Q3			age-at-length
9.	TBM.DS	Fixed	Blocks: Baseline, 2001-2003.Q2	block: 1984 – 2002.Q2 2002.Q3-2007.Q3			
10.	TBE	Fixed	Blocks: Baseline, 2001-2003.Q2	Constant	Blocks: Dome-shape (1984-2014), Asymptotic (2015-2019)	Dome-shape (11-knot spline)	
11.	TBE.GRO	Estimated	Blocks: Baseline, 2001-2003.Q2	Constant			age-at-length
12.	TBE.DS	Fixed	Blocks: Baseline, 2001-2003.Q2	Constant			

TABLE 4. Maximum gradients of the models.

TABLA 4. Gradientes máximos de los modelos.

Steepness- Inclinación (h)	1	2	3	4	5	6	7	8	9	10	11	12
	BASE	GRO	DS	DDQ	DDQ.GRO	DDQ.DS	TBM	TBM.GRO	TBM.DS	TBE	TBE.GRO	TBE.DS
1.0	0.0013	0.0001	0.0001	0.0002	0.0000	0.0002	0.0022	3.5	10.5	0.0002	0.0001	0.0001
0.9	0.0001	0.0001	0.0001	0.0006	0.0109	0.0030	0.0005	3.5	1.1	0.0000	0.0041	0.0001
0.8	0.0001	0.0132	0.0001	0.0002	0.0000	0.0016	0.0002	11.2	1.3	0.0005	0.0000	0.0004
0.7	0.0002	0.0004	0.0001	0.0285	0.0013	0.0006	0.0010	4.4	9.8	0.0016	0.0008	0.0014

TABLE 5 Root mean square error (RMSE) and negative log-likelihood (NLL) for the index of abundance for all model runs.

TABLA 5. La raíz del error cuadrático medio (RECM) y log-verosimilitud negativa (LVN) para el índice de abundancia en todas las ejecuciones de los modelos.

Steepness- Inclinación (h)	1	2	3	4	5	6	7	8	9	10	11	12
	BASE	GRO	DS	DDQ	DDQ.GRO	DDQ.DS	TBM	TBM.GRO	TBM.DS	TBE	TBE.GRO	TBE.DS
RECM												
1.0	0.23	0.23			0.19	0.20	0.21	0.21	0.21	0.22	0.22	0.22
0.9	0.23	0.23	0.22	0.22	0.20	0.20	0.21	0.21	0.21	0.22	0.22	0.22
0.8	0.23	0.23	0.23	0.20	0.19	0.20	0.21	0.21	0.21	0.22	0.22	0.22
0.7	0.23	0.23	0.23	0.20	0.19	0.20	0.21	0.21	0.21	0.22	0.22	0.22
Differences to the lowest NLL - Diferencias de LVN al más bajo (164.85)												
1.0	32.3	34.8	32.3	4.5	0.0	2.5	16.0	17.3	16.2	31.8	32.1	32.0
0.9	32.3	34.8	32.3	4.5	0.1	2.5	16.0	17.3	16.2	31.8	32.1	32.0
0.8	32.4	34.9	32.4	4.5	0.1	2.6	16.0	17.4	16.3	31.9	32.2	32.0
0.7	32.5	35.2	32.5	4.5	0.1	2.6	16.0	17.4	16.3	32.0	32.4	32.2

TABLE 6. Number of estimated parameters, negative log-likelihood (NLL): (1) without the conditional age-at-length data, (2) only the age-at-length data and Akaike information criterion (AIC) estimated excluding the conditional age-at-length data, reported as the difference to the model with the lowest value (Δ AIC).

TABLA 6. Número de parámetros estimados, log-verosimilitud negativa (LVN): (1) sin los datos de edad condicional a la talla y (2) solo los datos de edad condicional a la talla y criterio de información de Akaike (AIC) estimado excluyendo los datos de edad condicional a la talla, presentado como diferencia al modelo con el menor valor (Δ AIC).

Steepness- Inclinación	Number of estimated parameters - Número de parámetros estimados											
	BASE	GRO	DS	DDQ	DDQ.GRO	DDQ.DS	TBM	TBM.GRO	TBM.DS	TBE	TBE.GRO	TBE.DS
1.0	332	336	334	333	337	335	339	343	341	341	345	343
0.9	332	336	334	333	337	335	339	343	341	341	345	343
0.8	332	336	334	333	337	335	339	343	341	341	345	343
0.7	332	336	334	333	337	335	339	343	341	341	345	343
	1. NLL, without age-at-length data - LVN, sin datos de edad por talla											
1.0	2134.1	2112.3	2127.0	2107.3	2071.6	2079.7	2087.4	2069.8	2058.9	2127.0	2035.8	2125.9
0.9	2133.7	2111.8	2126.6	2106.7	2071.6	2079.6	2086.8	2069.2	2058.3	2126.6	2035.5	2125.5
0.8	2133.4	2111.5	2126.2	2106.1	2071.6	2079.6	2086.3	2068.7	2057.6	2126.2	2035.2	2125.1
0.7	2133.4	2111.5	2126.1	2105.6	2071.6	2079.6	2086.0	2068.4	2056.8	2126.1	2035.1	2124.9
	2. NLL, only age-at-length data - LVN, solo datos de edad por talla											
1.0	N/A	56.0	N/A	N/A	55.8	N/A	N/A	50.3	N/A	N/A	N/A	55.4
0.9	N/A	56.0	N/A	N/A	55.8	N/A	N/A	50.3	N/A	N/A	N/A	55.4
0.8	N/A	56.0	N/A	N/A	55.8	N/A	N/A	50.2	N/A	N/A	N/A	55.3
0.7	N/A	55.8	N/A	N/A	55.8	N/A	N/A	50.2	N/A	N/A	N/A	55.2
	Δ AIC, without age-at-length data - AIC, sin datos de edad por talla (AIC min = 4795.7)											
1.0	136.6	100.9	87.0	55.3	21.5	33.7	57.1	30.0	4.1	140.2	131.7	142.0
0.9	135.7	99.8	85.7	55.2	21.5	33.6	56.0	28.8	2.8	139.4	130.9	141.2
0.8	135.1	99.2	84.5	55.1	21.5	33.5	54.9	27.7	1.4	138.8	130.4	140.5
0.7	135.2	99.4	83.5	55.2	21.6	33.5	54.3	27.0	0.0	138.5	130.3	140.1

TABLE 7. Ratio between catchabilities of the index of abundance in the block ($q_{2015-2019}$ or $q_{2001-2003.Q2}$) and the baseline (q) and non-linearity coefficient c , for each model and steepness value of the stock-recruitment function (h) (see Table 3 and 5.1.6.c) .

TABLA 7. Razón entre las capturabilidades del índice de abundancia en el bloque ($q_{2015-2019}$ or $q_{2001-2003.Q2}$) y la línea de base (q) y el coeficiente de no linealidad c , para cada modelo y valor de la inclinación de la función población reclutamiento (h) (vea la Tabla 3 y 5.1.6.c) .

h	$q_{2015-2019}/q$		
	TBE	TBE.GRO	TBE.DS
1.0	0.91	0.92	0.86
0.9	0.92	0.92	0.86
0.8	0.92	0.93	0.87
0.7	0.92	0.93	0.87
	$q_{2001-2003.Q2}/q$		
	TBM	TBM.GRO	TBM.DS
1.0	1.64	1.82	1.74
0.9	1.64	1.81	1.74
0.8	1.65	1.81	1.74
0.7	1.65	1.82	1.75
	c		
	DDQ	DDQ.GRO	DDQ.DS
1.0	1.7	2.2	2.1
0.9	1.7	2.2	2.1
0.8	1.7	2.2	2.0
0.7	1.7	2.2	2.0

TABLE 8. Management table for yellowfin tuna in the EPO, 2020. Respectively, $S_{current}$ and S_{MSY_d} are the spawning biomass at the beginning of 2020 and at dynamic MSY level; $F_{current}$ and F_{MSY} are the fishing mortality during 2017-2019 and at MSY; and S_{LIMIT} and F_{LIMIT} are the limit reference points for spawning biomass and fishing mortality. $C_{current}$ is the total catch of yellowfin in 2019, in metric tons, and MSY_d is the dynamic MSY, for each reference model and steepness value of the stock-recruitment function (h). (See 6.1 for the definitions of the quantities reported).

TABLA 8. Tabla de ordenación para el atún aleta amarilla en el OPO, 2020. Respectivamente, S_{actual} y S_{RMS_d} son la biomasa reproductora a principios de 2020 y en RMS dinámico; F_{actual} y F_{RMS} son la mortalidad por pesca durante 2017-2019 y en RMS; y $S_{LÍMITE}$ y $F_{LÍMITE}$ son los puntos de referencia límite de biomasa reproductora y mortalidad por pesca. C_{actual} es la captura total (en toneladas) de aleta amarilla en 2019 y RMS_d es el RMS dinámico, para cada modelo de referencia y valor de la inclinación de la función población reclutamiento (h). (Vea 6.1 para la definición de las cantidades presentadas).

	1	2	3	4	5	6	7	8	9	10	11	12
	BASE	GRO	DS	DDQ	DDQ.GRO	DDQ.DS	TBM	TBM.GRO	TBM.DS	TBE	TBE.GRO	TBE.DS
$h = 1.0$												
MSY	461,752	488,404	586,672	425,788	466,324	511,876	497,760	543,960	710,188	494,796	509,932	510,824
MSY_d	257,732	263,175	290,662	271,054	299,762	319,271	269,331	288,203	353,699	290,869	300,961	297,008
$C_{current}/MSY_d$	0.97	0.95	0.87	0.92	0.83	0.79	0.93	0.87	0.72	0.86	0.83	0.85
S_{MSY}/S_0	0.32	0.24	0.27	0.31	0.23	0.27	0.30	0.23	0.26	0.29	0.24	0.29
$S_{current}/S_0$	0.15	0.16	0.21	0.22	0.27	0.30	0.19	0.20	0.26	0.24	0.24	0.25
$S_{current}/S_{LIMIT}$	2.00	2.09	2.71	2.84	3.45	3.93	2.47	2.62	3.37	3.17	3.05	3.26
$p(S_{current} < S_{LIMIT})$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$F_{current}/F_{LIMIT}$	0.40	0.40	0.27	0.33	0.28	0.22	0.33	0.31	0.20	0.28	0.30	0.27
$p(F_{current} > F_{LIMIT})$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$S_{current}/S_{MSY_d}$	0.90	1.26	1.63	1.14	1.85	1.92	1.22	1.70	2.19	1.46	1.70	1.55
$p(S_{current} < S_{MSY_d})$	0.84	0.07	0.03	0.14	0.00	0.05	0.08	0.01	0.03	0.06	0.02	0.05
$F_{current}/F_{MSY}$	1.08	0.80	0.59	0.87	0.53	0.47	0.81	0.60	0.40	0.68	0.60	0.63
$p(F_{current} > F_{MSY})$	0.74	0.03	0.00	0.10	0.00	0.00	0.03	0.00	0.00	0.01	0.00	0.00
$h = 0.9$												
MSY	468,040	481,752	573,148	436,744	459,168	501,548	496,352	528,252	677,592	493,256	501,144	506,556
MSY_d	260,403	252,946	267,120	267,881	276,496	293,116	259,476	263,425	308,512	276,548	278,752	279,319
$C_{current}/MSY_d$	0.96	0.98	0.94	0.93	0.90	0.86	0.97	0.95	0.82	0.91	0.90	0.90
S_{MSY}/S_0	0.35	0.29	0.31	0.35	0.28	0.31	0.33	0.28	0.30	0.33	0.29	0.32
$S_{current}/S_0$	0.14	0.15	0.20	0.20	0.25	0.29	0.18	0.19	0.25	0.23	0.22	0.24
$S_{current}/S_{LIMIT}$	1.86	1.95	2.59	2.61	3.23	3.73	2.32	2.47	3.24	2.97	3.05	3.06
$p(S_{current} < S_{LIMIT})$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$F_{current}/F_{LIMIT}$	0.47	0.46	0.31	0.38	0.32	0.25	0.38	0.36	0.23	0.33	0.30	0.31
$p(F_{current} > F_{LIMIT})$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$S_{current}/S_{MSY_d}$	0.75	1.01	1.43	0.98	1.51	1.68	1.06	1.40	1.95	1.27	1.41	1.36
$P(S_{current} < S_{MSY_d})$	1.00	0.47	0.06	0.57	0.02	0.08	0.33	0.04	0.03	0.13	0.07	0.10

	1	2	3	4	5	6	7	8	9	10	11	12
	BASE	GRO	DS	DDQ	DDQ.GRO	DDQ.DS	TBM	TBM.GRO	TBM.DS	TBE	TBE.GRO	TBE.DS
$F_{\text{current}}/F_{\text{MSY}}$	1.26	0.99	0.71	1.02	0.67	0.57	0.95	0.75	0.49	0.81	0.74	0.75
$p(F_{\text{current}} > F_{\text{MSY}})$	0.97	0.46	0.02	0.59	0.00	0.00	0.34	0.02	0.00	0.10	0.04	0.06
$h = 0.8$												
MSY	483,904	485,012	565,840	462,136	463,640	498,952	502,580	521,748	658,140	499,520	502,460	509,704
MSY_d	269,568	251,063	249,703	271,954	261,577	272,308	254,710	248,137	276,016	268,398	266,591	267,751
$C_{\text{current}}/\text{MSY}_d$	0.92	0.99	1.01	0.92	0.95	0.93	0.98	1.01	0.92	0.94	0.94	0.94
S_{MSY}/S_0	0.37	0.32	0.34	0.37	0.32	0.34	0.36	0.32	0.33	0.36	0.32	0.35
S_{current}/S_0	0.13	0.14	0.19	0.18	0.23	0.27	0.16	0.18	0.24	0.21	0.20	0.22
$S_{\text{current}}/S_{\text{LIMIT}}$	1.68	1.79	2.44	2.31	2.93	3.46	2.13	2.30	3.07	2.74	2.63	2.83
$p(S_{\text{current}} < S_{\text{LIMIT}})$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$F_{\text{current}}/F_{\text{LIMIT}}$	0.55	0.53	0.36	0.45	0.37	0.29	0.44	0.42	0.26	0.38	0.41	0.37
$p(F_{\text{current}} > F_{\text{LIMIT}})$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$S_{\text{current}}/S_{\text{MSY}_d}$	0.62	0.83	1.29	0.83	1.29	1.51	0.92	1.20	1.81	1.11	1.21	1.20
$p(S_{\text{current}} < S_{\text{MSY}_d})$	1.00	0.94	0.12	0.97	0.08	0.11	0.75	0.15	0.04	0.30	0.19	0.21
$F_{\text{current}}/F_{\text{MSY}}$	1.46	1.17	0.82	1.19	0.80	0.66	1.10	0.89	0.57	0.93	0.88	0.87
$p(F_{\text{current}} > F_{\text{MSY}})$	1.00	0.86	0.13	0.93	0.06	0.04	0.77	0.21	0.00	0.35	0.24	0.23
$h = 0.7$												
MSY	518,192	502,584	566,512	521,896	488,020	508,960	521,792	526,380	650,584	518,396	517,428	524,164
MSY_d	289,293	256,702	235,527	291,255	254,438	255,332	255,934	238,816	248,957	266,352	262,019	261,308
$C_{\text{current}}/\text{MSY}_d$	0.86	0.97	1.07	0.86	0.98	0.99	0.98	1.04	1.02	0.94	0.96	0.96
S_{MSY}/S_0	0.40	0.35	0.37	0.40	0.35	0.37	0.38	0.35	0.36	0.38	0.35	0.38
S_{current}/S_0	0.11	0.12	0.17	0.15	0.19	0.24	0.15	0.16	0.22	0.19	0.18	0.20
$S_{\text{current}}/S_{\text{LIMIT}}$	1.45	1.58	2.26	1.90	2.53	3.10	1.90	3.21	2.89	2.44	2.35	2.55
$p(S_{\text{current}} < S_{\text{LIMIT}})$	0.02	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$F_{\text{current}}/F_{\text{LIMIT}}$	0.65	0.62	0.42	0.54	0.44	0.34	0.52	0.33	0.30	0.45	0.47	0.43
$p(F_{\text{current}} > F_{\text{LIMIT}})$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$S_{\text{current}}/S_{\text{MSY}_d}$	0.49	0.68	1.17	0.64	1.09	1.35	0.78	1.03	1.71	0.96	1.03	1.06
$p(S_{\text{current}} < S_{\text{MSY}_d})$	1.00	1.00	0.23	1.00	0.31	0.17	0.99	0.43	0.06	0.59	0.44	0.39
$F_{\text{current}}/F_{\text{MSY}}$	1.68	1.36	0.94	1.38	0.95	0.78	1.27	1.04	0.65	1.07	1.01	1.00
$p(F_{\text{current}} > F_{\text{MSY}})$	1.00	0.98	0.38	0.99	0.37	0.15	0.95	0.60	0.02	0.64	0.53	0.49

APPENDIX 1

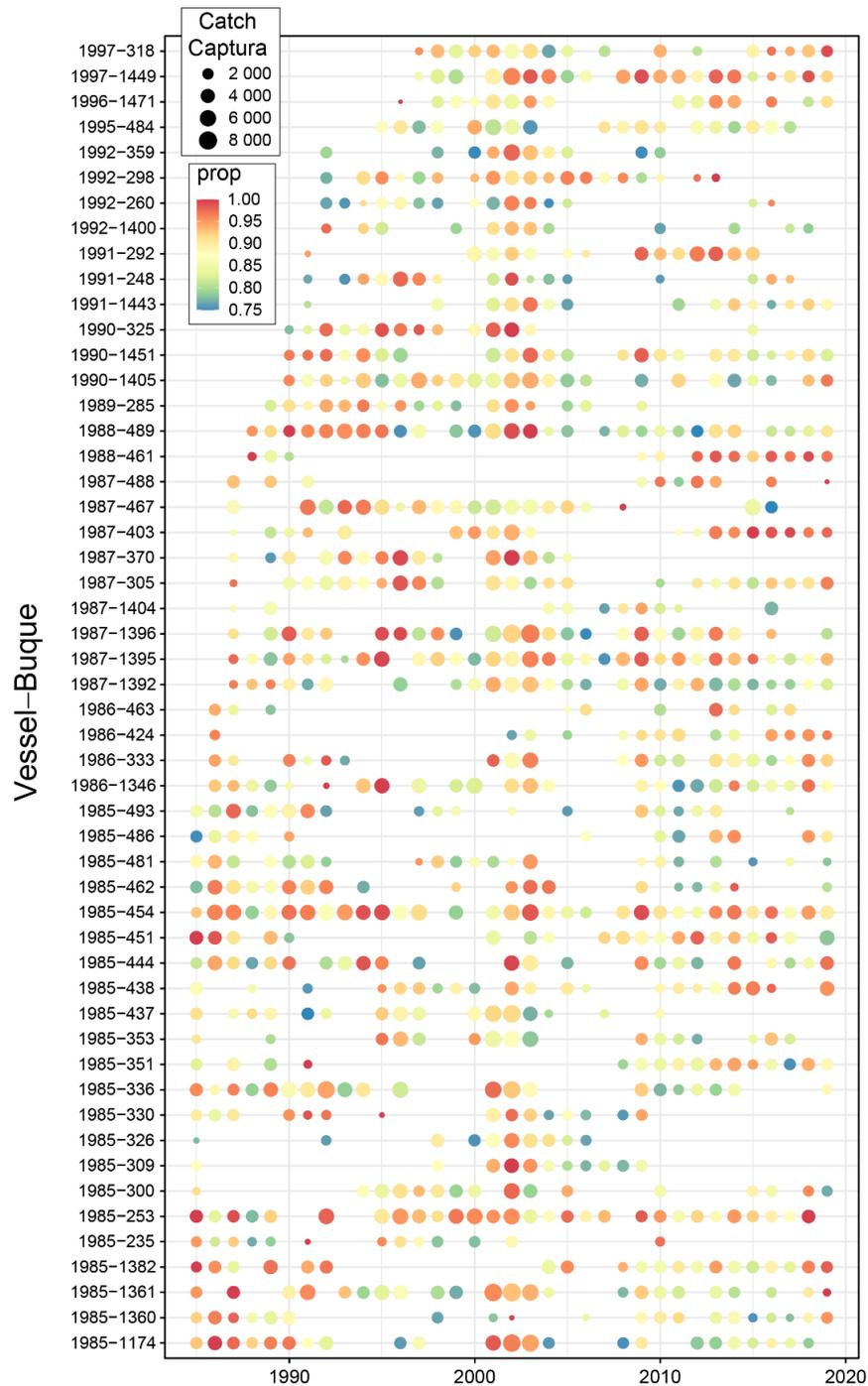


FIGURE A1. Vessels included in the standardization of CPUE to obtain the index of abundance coming from the dolphin-associated fisheries. Vessels (indicated by their codes) are shown on the y-axis. The size of the dot represents the annual catch, and the color the annual proportion of sets on dolphins.

FIGURA A1. Buques incluidos en la estandarización de la CPUE para el índice de abundancia proveniente de las pesquerías asociada a delfines. Los buques (indicados por sus códigos) se muestran en el eje y. El tamaño del punto representa la captura anual, y el color la proporción anual de lances sobre delfines.

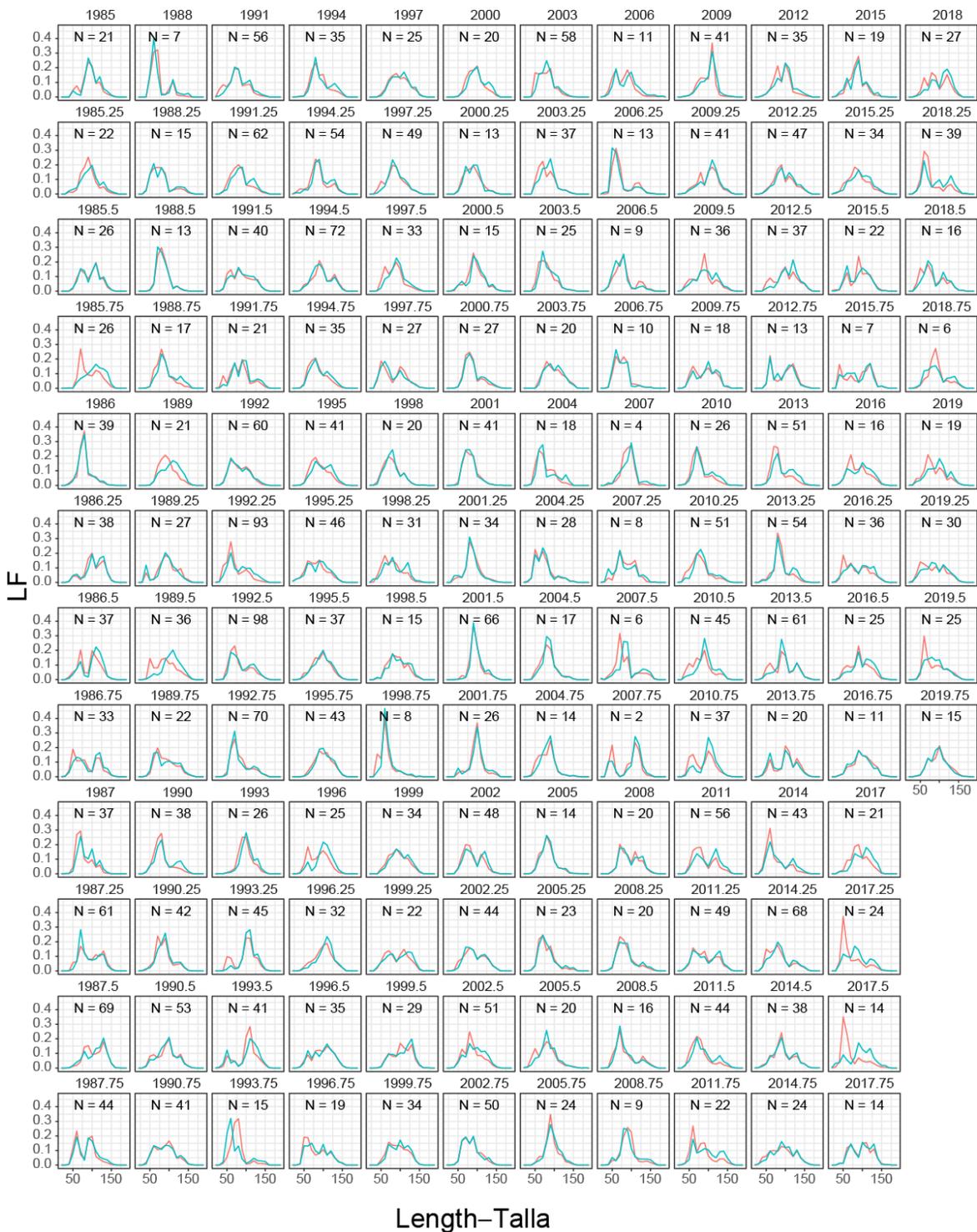


FIGURE A2. Comparison of nominal (red lines) and standardized (blue lines) length frequencies used to represent the index of abundance.

FIGURA A2. Comparación de las frecuencias de talla nominales (líneas rojas) y estandarizadas (líneas azules) usadas para representar el índice de abundancia.

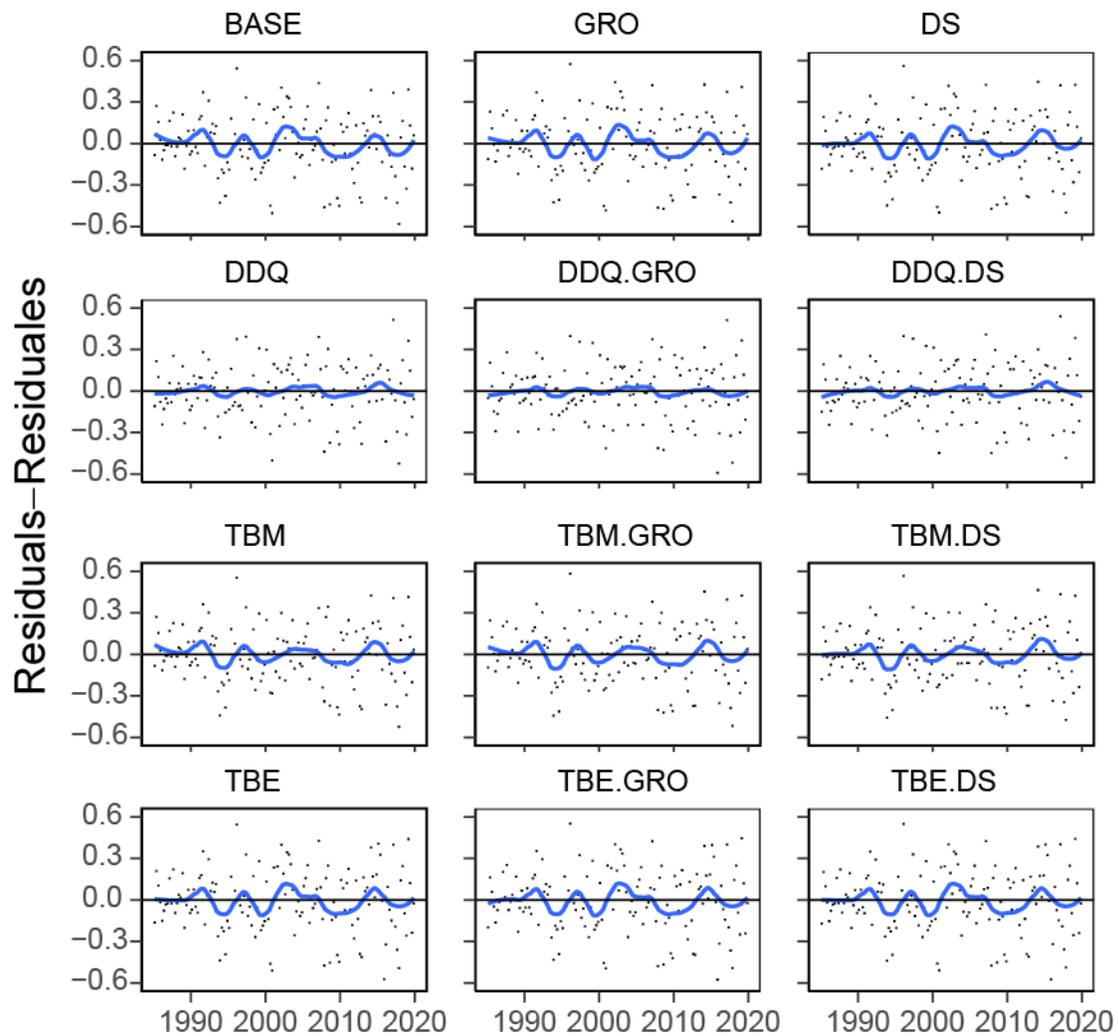


FIGURE A3. Residuals ($\log(\text{observed index}) - \log(\text{expected index})$) for the twelve model configurations with steepness $h = 1$. The lines were built using the *R stats::loess* function for fitting a local polynomial smoother with span = 0.25.

FIGURA A3. Residuales ($\log(\text{índice observado}) - \log(\text{índice esperado})$) para las 12 configuraciones de los modelos con inclinación $h = 1$. Las líneas se construyeron usando la función de *R stats::loess* para ajustar un suavizador polinomial local con ancho de banda = 0.25.

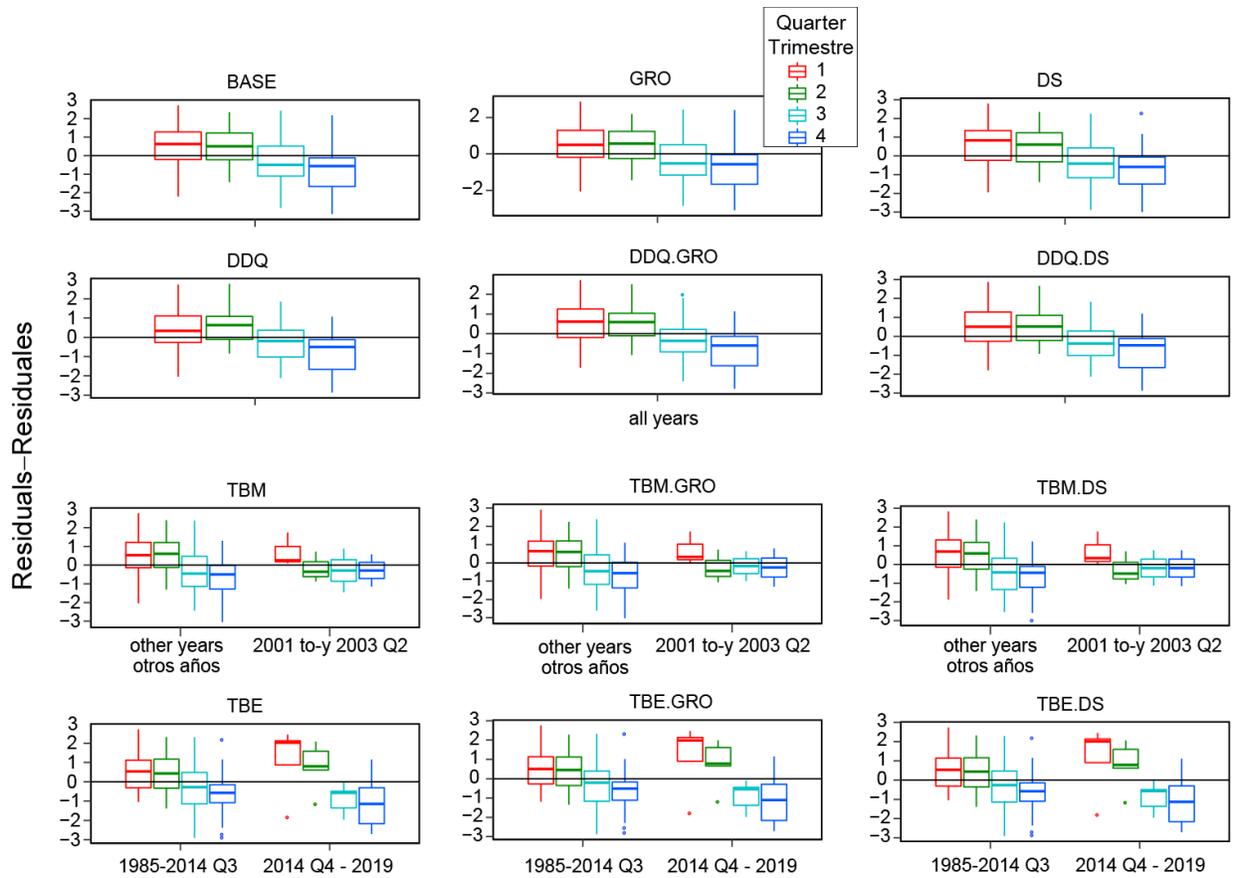


FIGURE A4. Residual ($\log(\text{observed index}) - \log(\text{expected index})$) plots for the twelve model configurations with steepness = 1.

FIGURA A4. Gráficas de residuales ($\log(\text{índice observado}) - \log(\text{índice esperado})$) para las 12 configuraciones de los modelos con inclinación = 1.

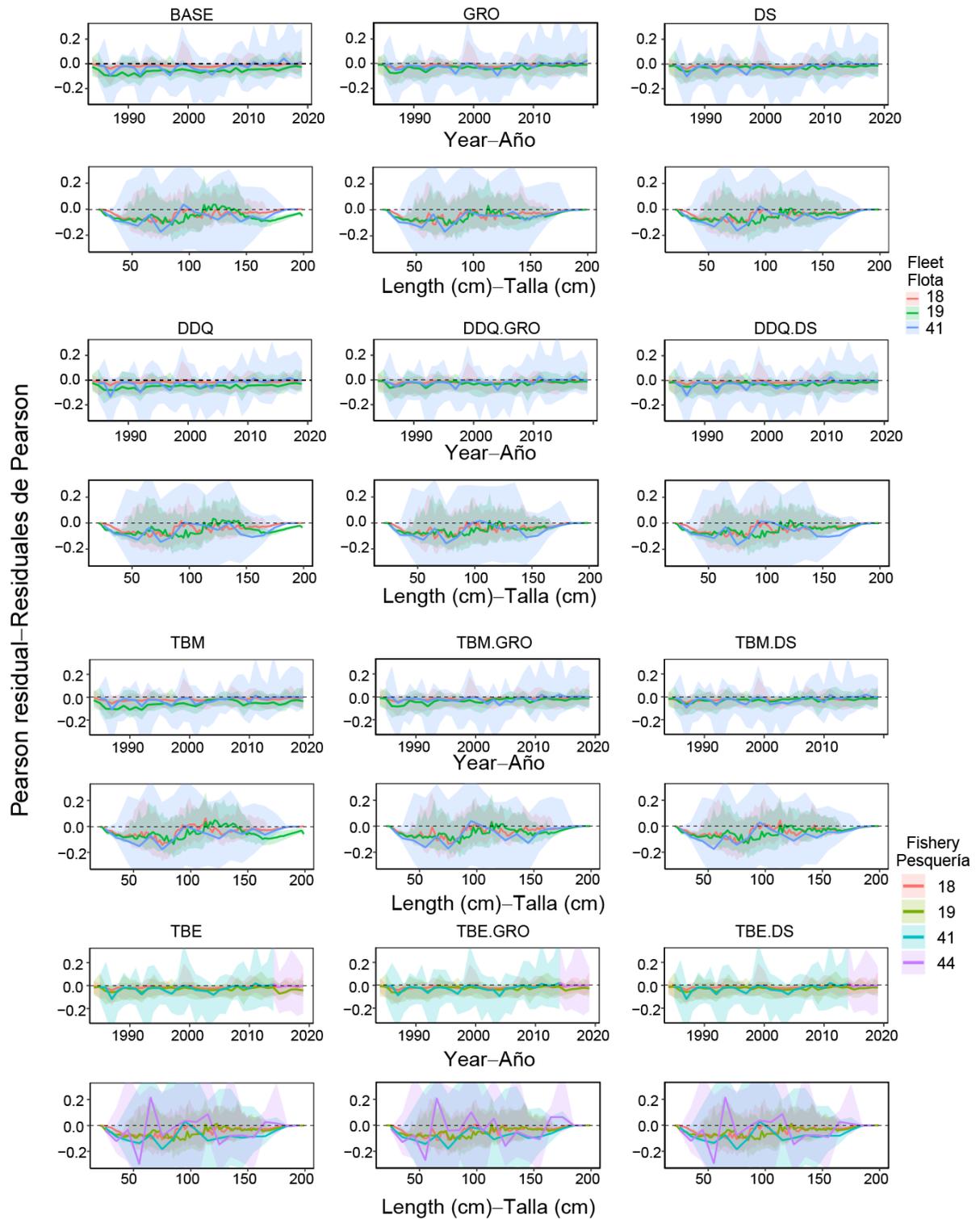


FIGURE A5. Residual plots for the survey (41 – baseline period and 44 - block) and fisheries F18 and F19. The lines are the median residuals and the shaded area are the 25th and 75th percentiles.

FIGURA A5. Gráficas de residuales para el estudio (41 y 44) y las pesquerías F18 y F19. Las líneas representan los residuales medianos y el área sombreada los percentiles de 25 y 75.

APPENDIX 2.

Integrated models diagnostics

Age-structured production model (ASPM): this diagnostic (Maunder and Piner, 2015) may be used to: (i) evaluate model misspecification, (ii) ascertain the influence of composition data on the estimates of absolute abundance and trends in abundance, and (iii) check whether catch alone can explain the trends in the indices of abundance. The ASPM diagnostic is computed as follows: (i) run the base case model; (ii) fix selectivity parameters at the maximum likelihood estimate (MLE) from the base case model, (iii) turn off the estimation of all parameters except the scaling parameters (R_0), and set the recruitment deviates to zero; (iv) fit the model to the indices of abundance only; (v) compare the estimated trajectory to that of the base case. There is evidence of the existence of a production function if the ASPM can fit well the index of abundance that have good contrast (*i.e.* those that have declining and/or increasing trends), it is also likely that the index, in combination with the catches, provides information on absolute abundance (Maunder and Piner 2015). When the catches cannot explain the changes in the indices, the ASPM will fit the index poorly. This can have several causes: (i) the stock is recruitment-driven; (ii) the stock has not yet declined to the point where catch is a major factor influencing abundance, (iii) the base-case model is incorrect, or (iv) the indices of relative abundance are not proportional to abundance. Checking whether the stock is recruitment-driven involves estimating recruitment deviations when fitting the model (ASPM-R). If this is still not able to capture the population trajectory estimated in the integrated model, it can be concluded that the information about scale in the integrated model is coming from the length composition data. Large confidence intervals on the abundance estimated by the ASPM also indicate that the index of abundance has little information on absolute abundance.

Catch-curve analysis (CCA) is done by fitting the integrated model only to the length composition data, and estimating all parameters except the auxiliary parameters associated with the index (Carvalho *et al.* 2017). The decline in the logarithm of the proportion of catch-at-age with age (the catch curve) provides information on fishing mortality (since the natural mortality assumed to be known), and when combined with catch data provides information on abundance. The CCA is used to verify whether the temporal trend implied by the size composition data is consistent with that coming from the index of abundance. If the two trends are similar, then there is more confidence that the estimated abundance trend is accurate. Two variants of the CCA were used, one that is fit only to data from the fisheries and other that is fit only to the survey data.

Likelihood profile on the global scaling parameter: A likelihood profile of the average recruitment in an unfisher (virgin) population in logarithm scale, $\ln R_0$, is used to determine whether information about absolute biomass scaling is consistent among data sets (*e.g.*, Francis, 2011;; Lee *et al.*, 2014; Wang *et al.*, 2014 2). The profile is done by fixing $\ln R_0$ to a range of values around the maximum likelihood estimate (MLE) and estimating all other parameters, then obtaining the contribution of each data set and penalty components to the likelihood conditioned of the value of $\ln R_0$. The profile quantifies how the fit to each data component is degraded by changing the population scale. The data with large amount of information on population scale will show loss of fit (smaller likelihood, or larger negative-log likelihood) as population scale is changed from its best estimate (Lee *et al* 2001). If different data components favor different values for $\ln R_0$, there is contradictory information among them, conditioned on the model, thus pointing to potential model misspecification.

Retrospective analyses: these analyses are useful for determining how consistent a stock assessment method is from one year to the next (Mohn, 1999). The analysis is generally done by eliminating data for the last time step, then repeating the model fit without changing the method and assumptions, removing the last and the second last, running the model again and so on, until a desired amount of data is

cumulatively removed. This shows the effect on the resulting estimated quantities of including more data. Inconsistencies in the results of this progressive removal of data are a signal of inadequacies in the assessment models. The assessment model has a quarterly time step, but new data are updated annually (four quarters at once). Thus, the retrospective analysis was done by removing whole years of data at once.

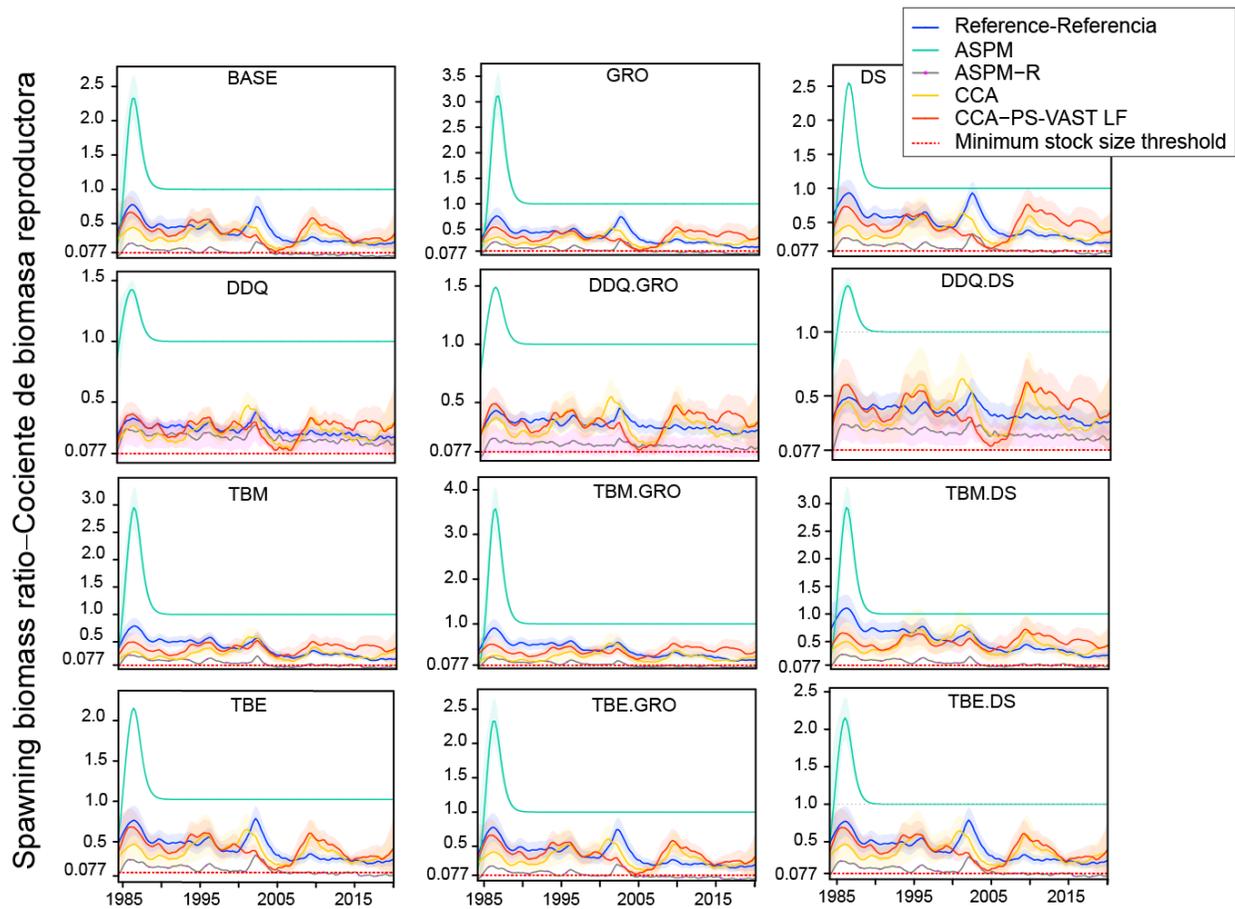


FIGURE A6. Spawning biomass ratios of yellowfin tuna in the EPO, 1985-2019, estimated by the 12 reference models and the corresponding diagnostic models (ASPM, ASPM-R, CCA, CCA-PS-VAST LF). The red dotted line at 0.077 indicates $S = S_{LIMIT}$.

FIGURA A6. Cocientes de biomasa reproductora del aleta amarilla en el OPO, 1985-2019, estimados por los 12 modelos de referencia y los modelos diagnósticos correspondientes (ASPM, ASPM-R, ACC, CCA-PS-VAST LF). La línea de trazos roja en 0.077 indica $S = S_{LÍMITE}$.

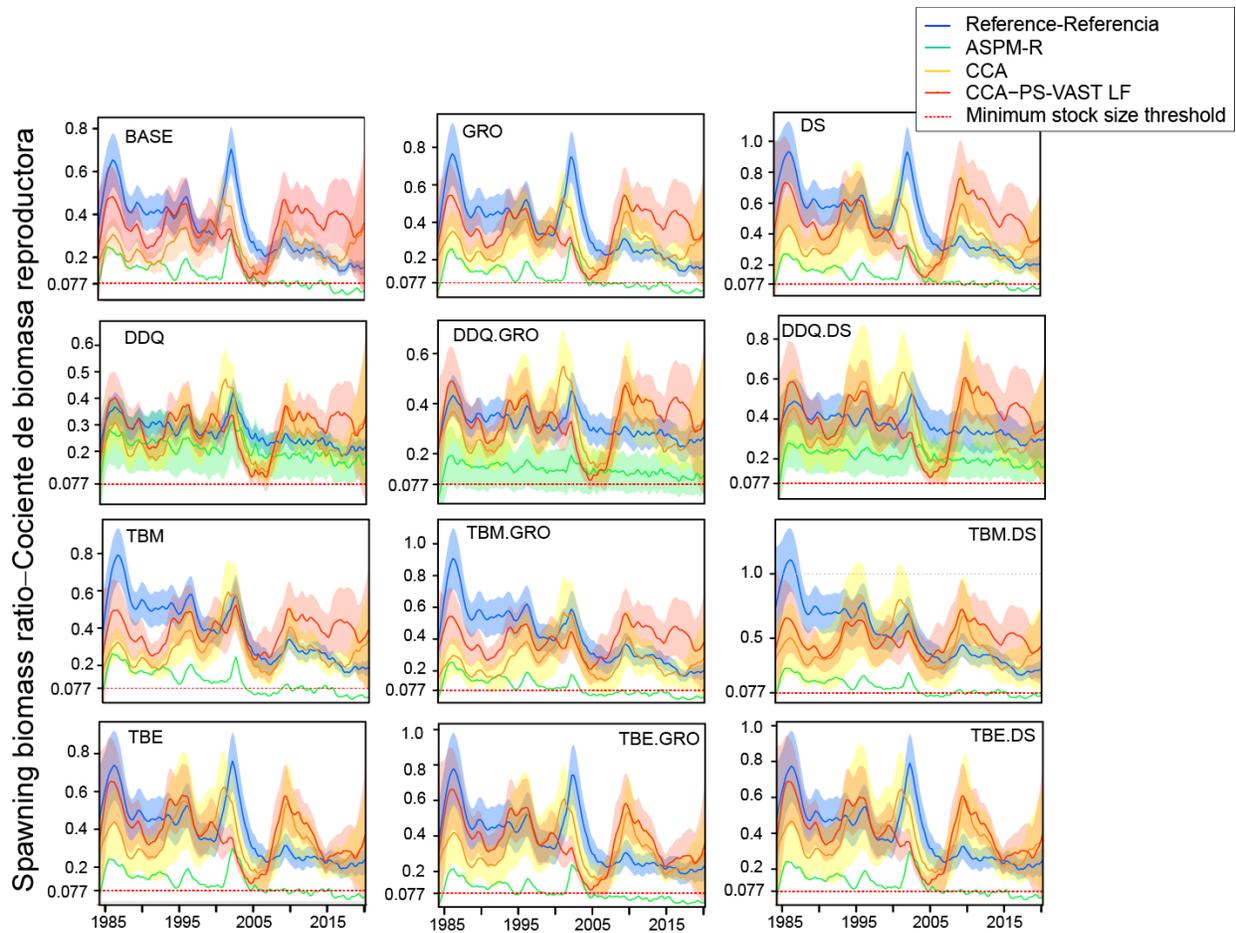


FIGURE A7. Spawning biomass ratios of yellowfin tuna in the EPO, 1985-2019, estimated by the 12 reference models and the corresponding diagnostic models (ASPM-R, CCA, CCA-PS-VAST LF). The lines represent the maximum likelihood estimate (MLE), and the shaded areas the confidence intervals (CI). Models without CI did not produce a positive definite hessia (all ASPM-R except DDQ, DDQ.GRO, DDQ.DS).. The red dotted line at 0.077 indicates $S=S_{LIMIT}$.

FIGURA A7. Cocientes de biomasa reproductora del aleta amarilla en el OPO, 1985-2019, estimados por los 12 modelos de referencia y los modelos diagnósticos correspondientes (ASPM-R, ACC, ACC-PS-VAST LF). Las líneas representan la estimación de máxima verosimilitud (EMV) y las áreas sombreadas los intervalos de confianza (IC). Los modelos sin IC no produjeron una matriz Hessiana positiva definit (todos los ASPM-R excepto DDQ, DDQ.GRO, DDQ.DS). La línea de trazos roja en 0.077 indica $S= S_{LÍMITE}$.

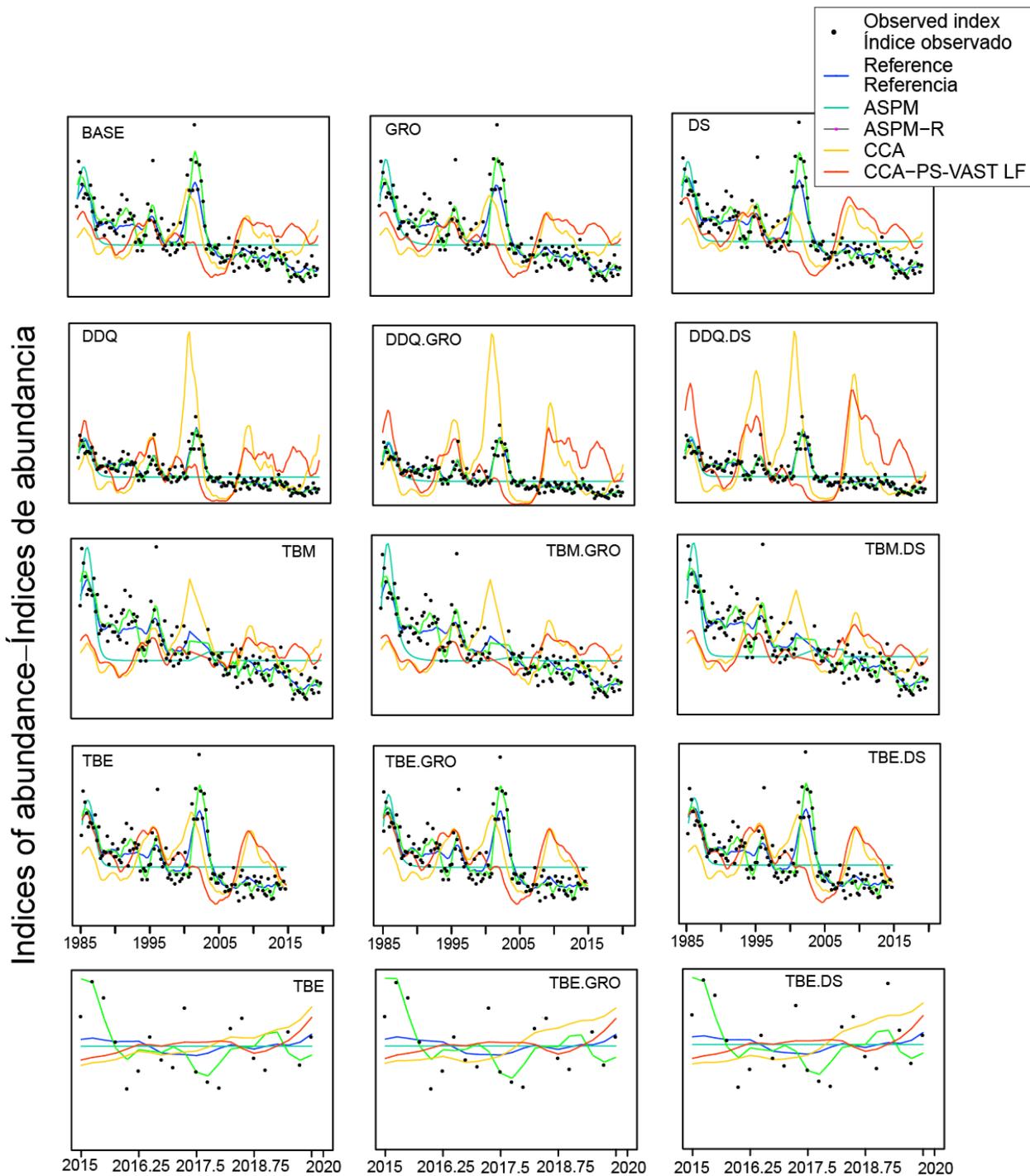


FIGURE A8. Abundance index of yellowfin tuna in the EPO, 1985–2019, estimated by each reference model and the corresponding diagnostic models. The lines represent the maximum likelihood estimates (MLE) or the predictions (for CCA, CCA-PS-VAST LF), the dots are the observed values .

FIGURA A8. Índice de abundancia del atún aleta amarilla en el OPO estimados por cada modelo de referencia y los modelos diagnósticos correspondientes. Las líneas representan la estimación de máxima verosimilitud (EMV) o los valores predichos (para CCA, CCA-PS-VAST LF), los puntos son los valores observados

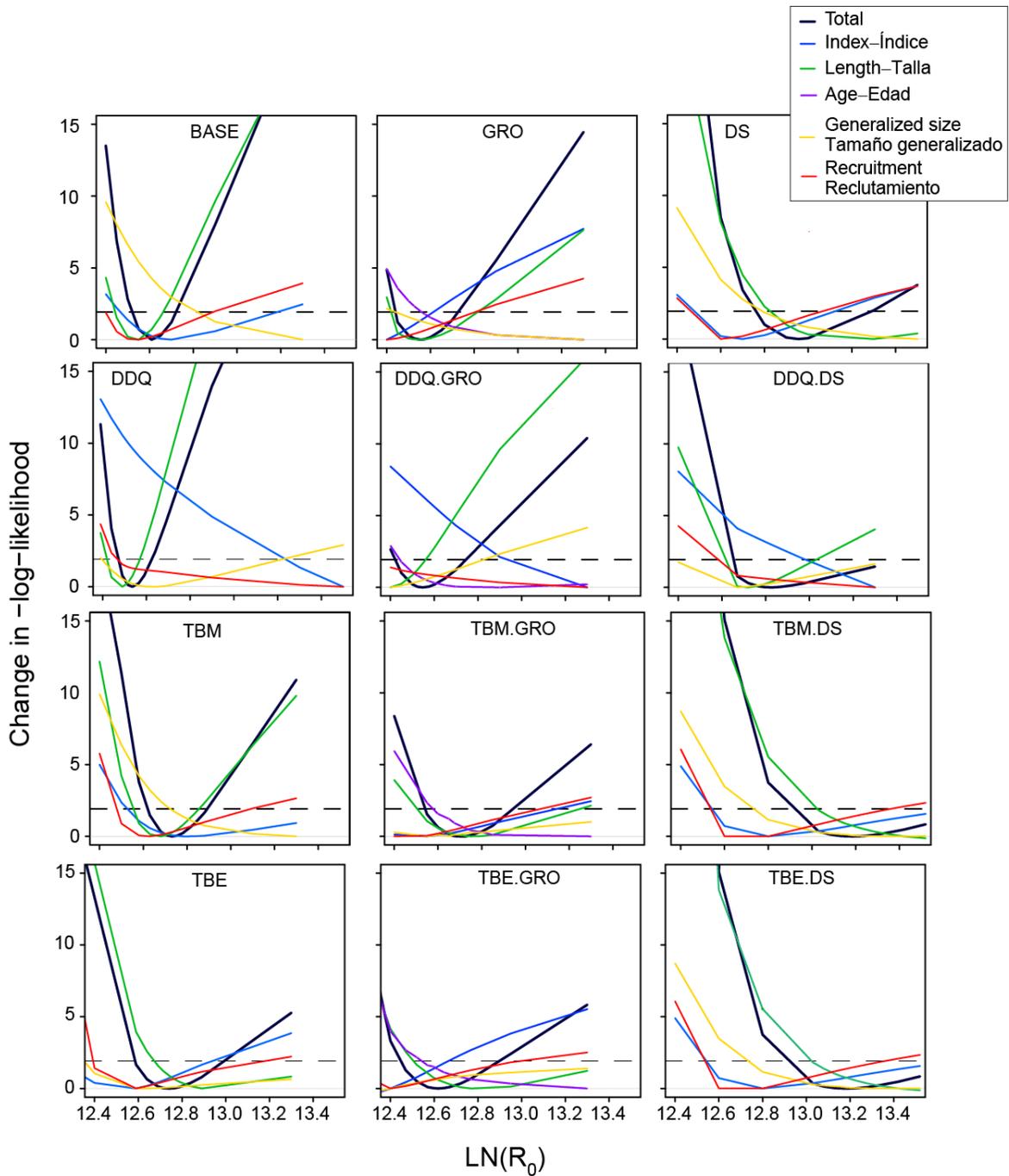


FIGURE A9. Comparison of the likelihood profile for $\ln R_0$ (scaling parameter) for the twelve reference models for yellowfin tuna in the EPO. The two point where the dotted line crossed the total likelihood line are the lower and upper values of the 95% confidence interval (obtained from the chi-square distribution with one degree of freedom).

FIGURA A9. Comparación del perfil de verosimilitud para $\ln R_0$ (parámetro de escala) para los 12 modelos de referencia para el aleta amarilla en el OPO. Los dos puntos donde la línea de trazos cruza la línea de la verosimilitud total son el valor mínimo y máximo del intervalo de confianza del 95% (obtenido de la distribución chi-cuadrada con un grado de libertad)

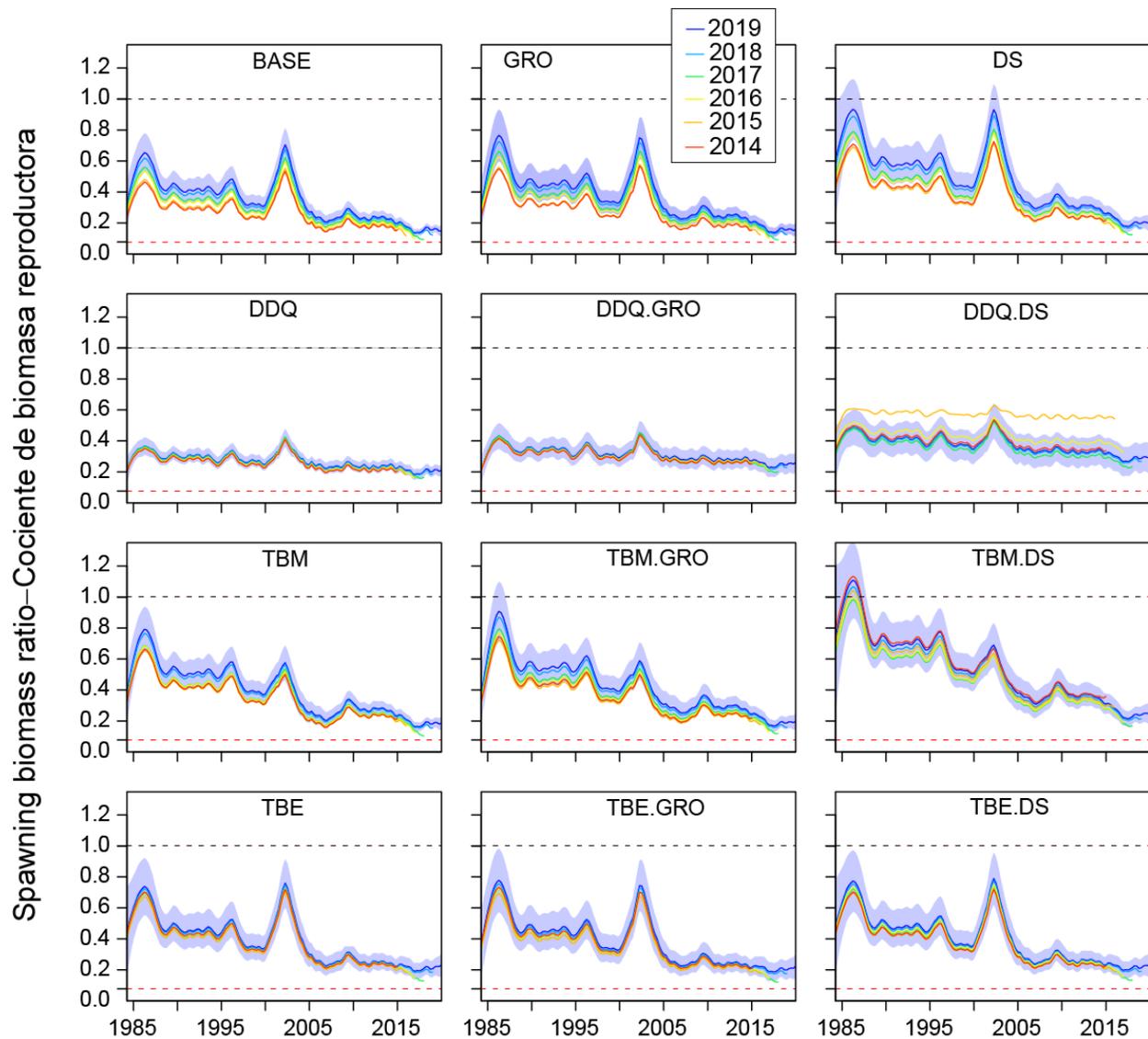


FIGURE A10. Retrospective patterns of the spawning biomass ratio (SBR) for the 12 reference models for yellowfin tuna in the EPO. The black and red dashed lines indicate $SBR = 1$ ($S=S_0$) and $SBR = 0.077$ ($S=S_{LIMIT}$), respectively.

FIGURA A10. Patrones retrospectivos del cociente de biomasa reproductora para los 12 modelos de referencia para el atún aleta amarilla en el OPO. Las líneas de trazos negra y roja indican $SBR = 1$ ($S=S_0$) y $SBR = 0.077$ ($S=S_{LIMITE}$), respectivamente.

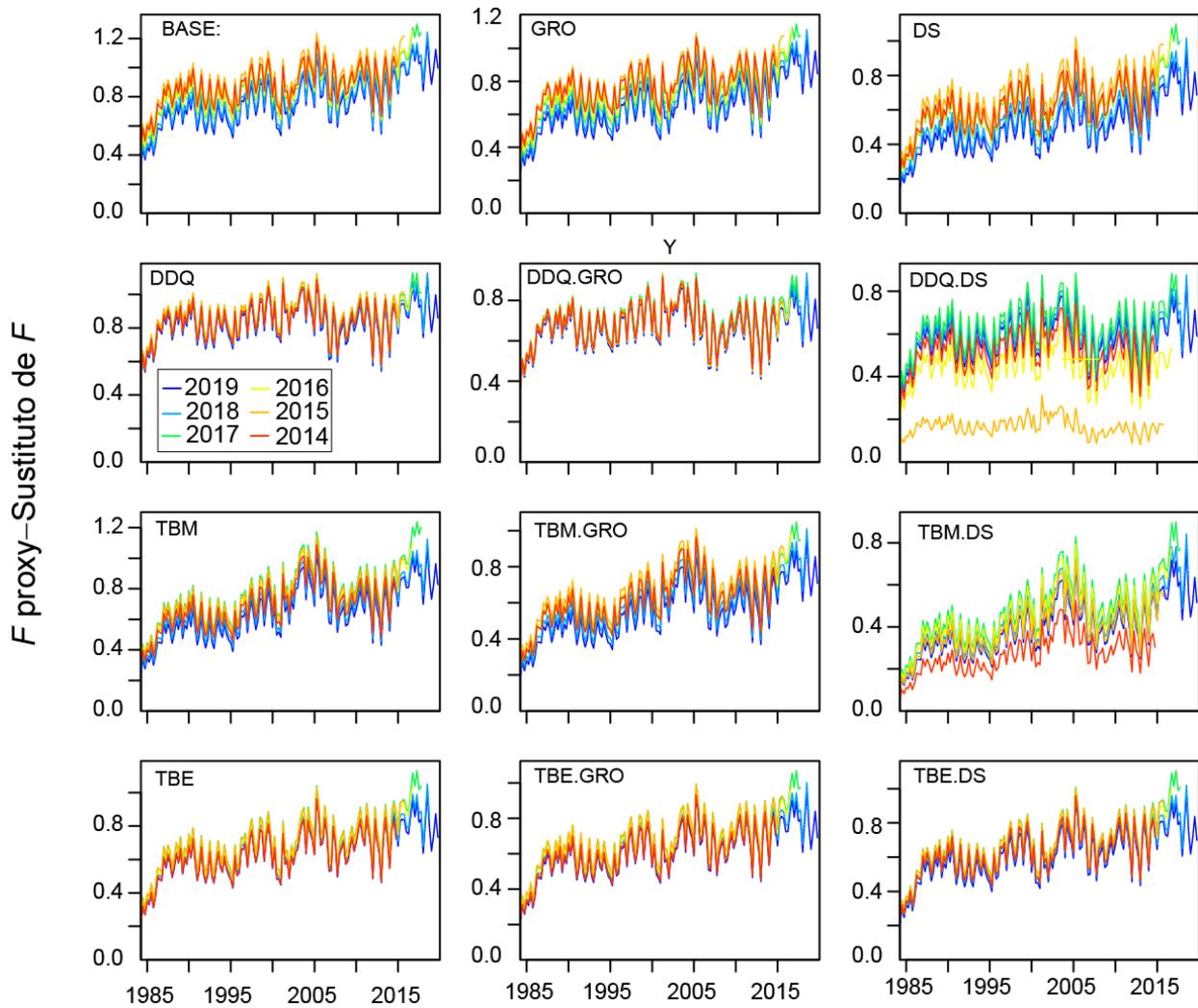


FIGURE A11. Retrospective patterns of fishing mortality proxy $[(1-SBR)/(1-SBR_{MSY})]$, 1984-2019, for the 12 reference models for yellowfin tuna in the EPO.

FIGURA A11. Patrones retrospectivos del sustituto de mortalidad por pesca $[(1-SBR)/(1-SBR_{RMS})]$, 1984-2019, para los 12 modelos de referencia para el aleta amarilla en el OPO.

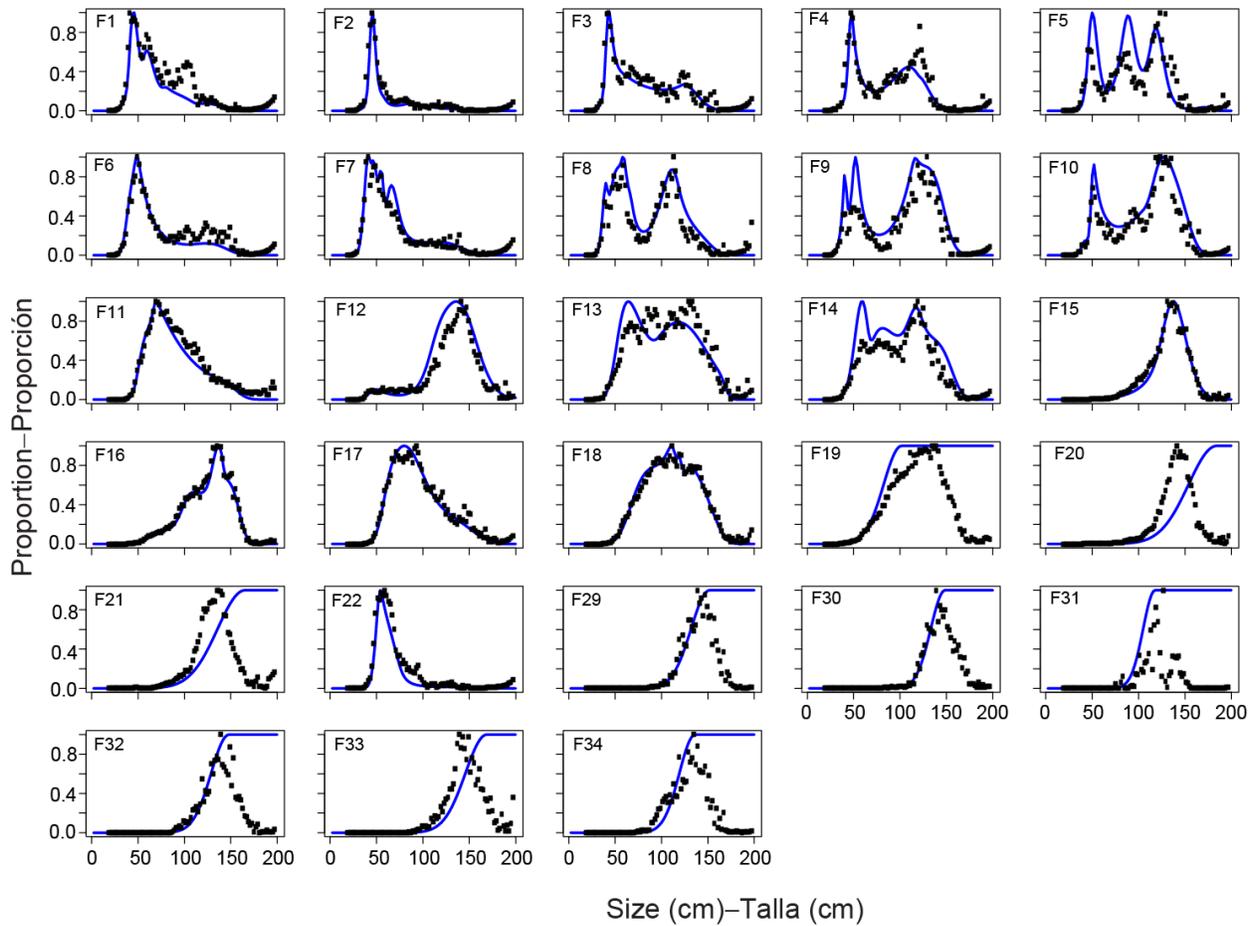


FIGURE A12. Estimated (purple line) and empirical (black dots) average selectivity, by fishery, for the BASE model with $h = 1$ (see text for details).

FIGURA A12. Selectividad promedio estimada (línea morada) y empírica (puntos negros), por pesquería, para el modelo BASE con $h = 1$ (ver detalles en el texto).

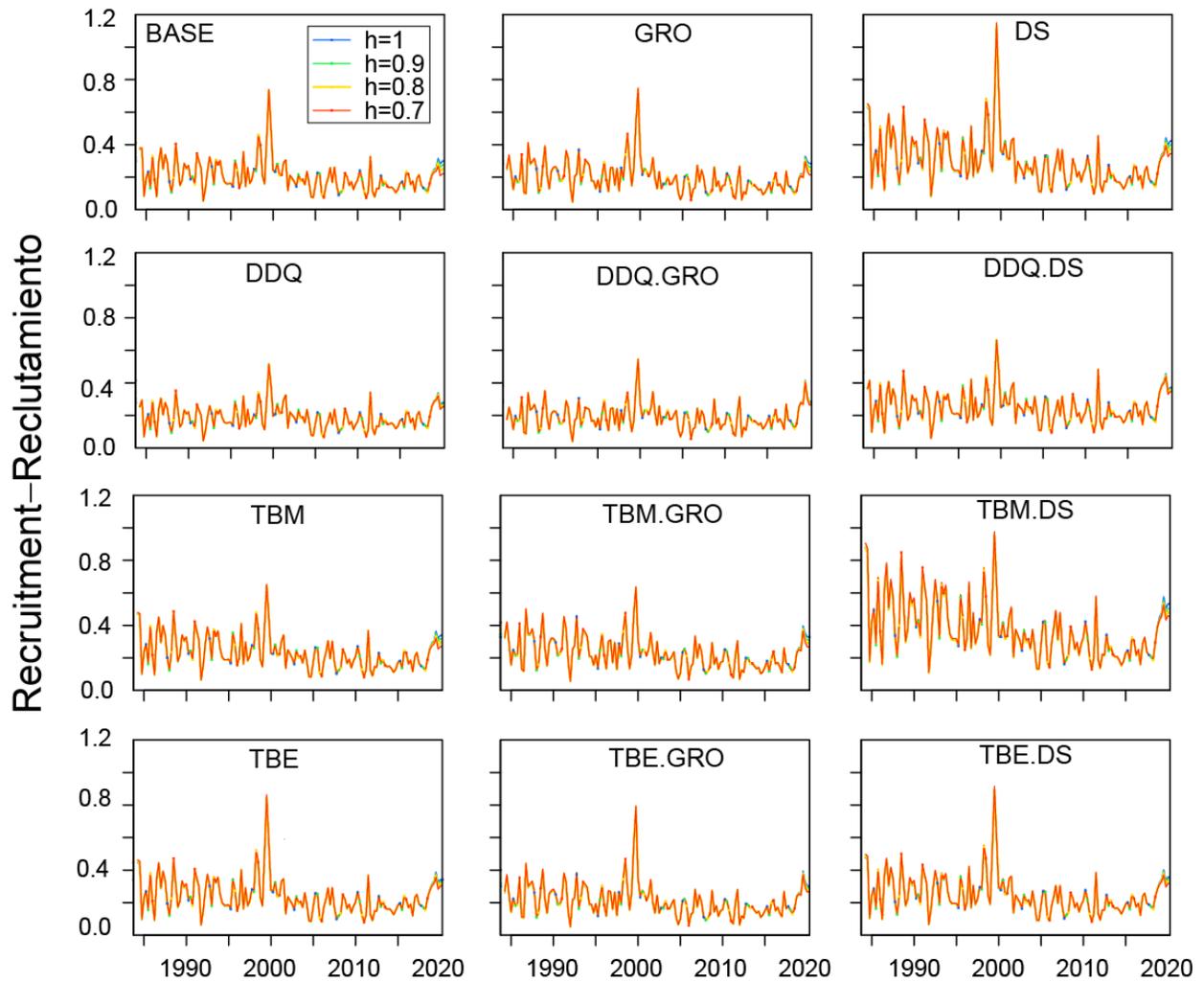


FIGURE A13. Quarterly recruitment of yellowfin tuna to the fisheries of the EPO, 1984-2019, estimated by the 48 models.

FIGURA A13. Reclutamiento trimestral del aleta amarilla en las pesquerías del OPO estimado los 48 modelos.

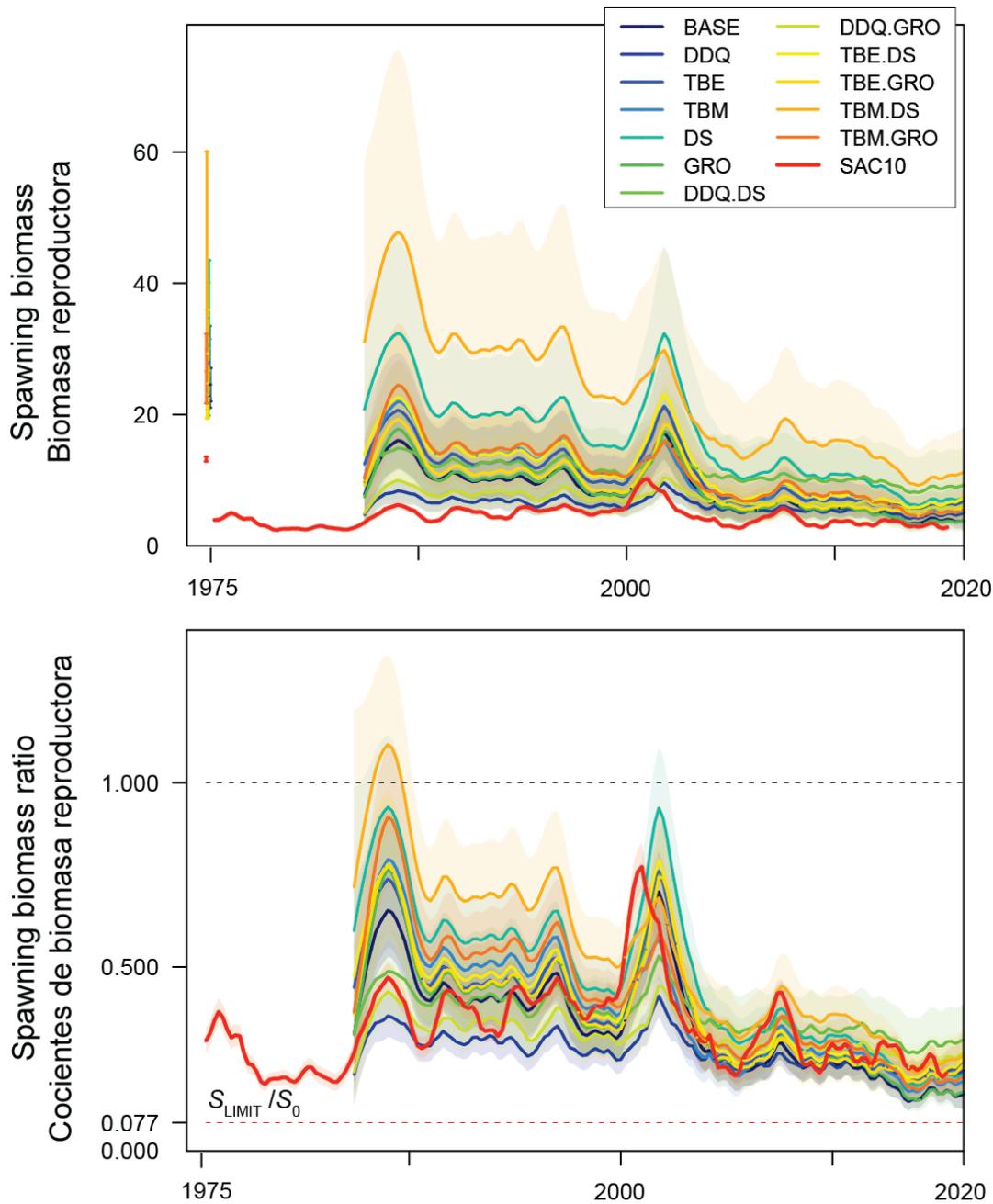


FIGURE A14. Spawning biomass (top) and spawning biomass ratios (SBRs; bottom) of yellowfin tuna in the EPO, for the 12 reference models with $h = 1.0$ (2020) and for the previous assessment (SAC-10, 2019; red line). The solid lines represent the maximum likelihood estimates, and the shaded areas their approximate 95% confidence intervals. Top: the colored bars and points on the left edge are the estimates of virgin spawning biomass for each model. Bottom: the red dashed horizontal line (at 0.077) identifies the $S=S_{LIMIT}$. **FIGURE A14.** Biomasa reproductora (arriba) y cociente de biomasa reproductora (abajo) del aleta amarilla en el OPO para los 12 modelos de referencia con $h = 1.0$ (2020) y para la evaluación previa (SAC 10, 2019; línea roja). Las líneas representan las estimaciones de máxima verosimilitud y las áreas sombreadas indican los intervalos de confianza aproximados de 95% alrededor de esas estimaciones. Arriba: Las barras y los puntos de color al principio del panel son las estimaciones de biomasa reproductora virgen para cada modelo. La línea horizontal punteada roja (en 0.077) identifica el $S=S_{LÍMITE}$.

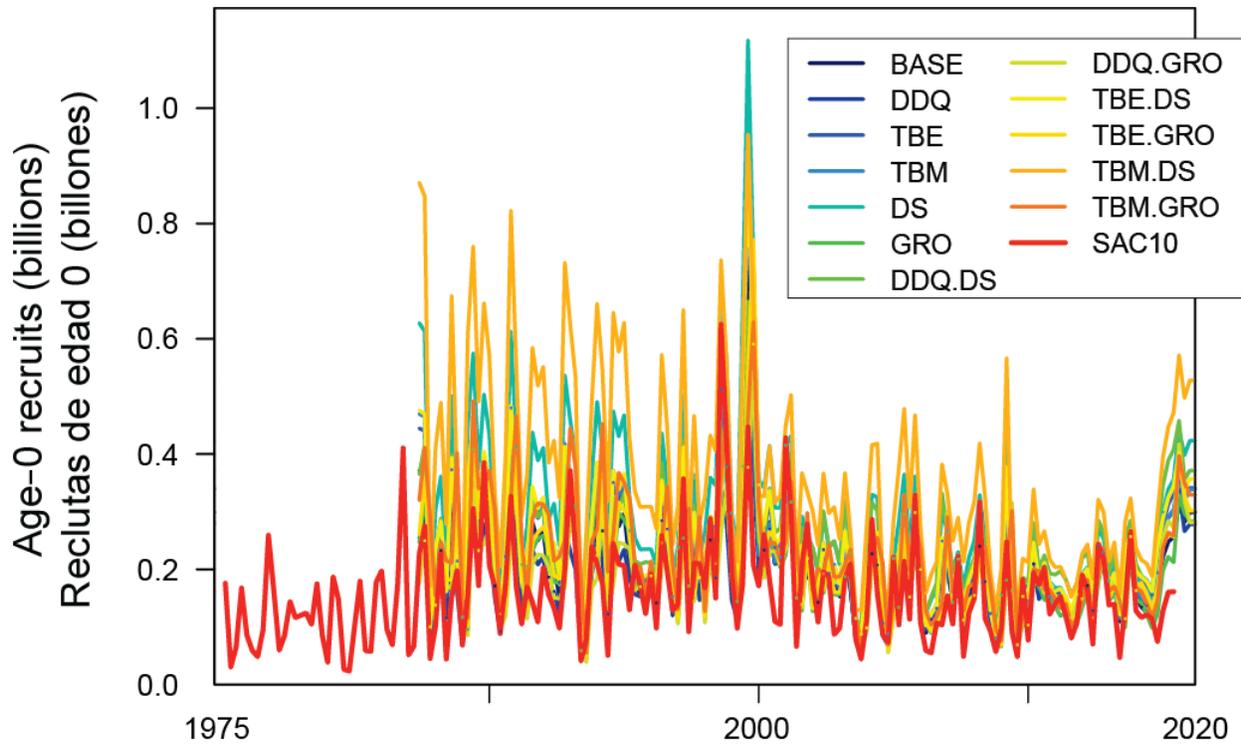


FIGURE A15. Quarterly recruitment of yellowfin tuna to the fisheries of the EPO, by the 12 reference models with $h = 1.0$ (1984- 2020) and by the previous assessment model (SAC-10,1975-2019).

FIGURA A15. Reclutamiento trimestral estimado de aleta amarilla en las pesquerías del OPO de los 12 modelos con inclinación $h = 1.0$ (1984- 2020) y por el del modelo de la evaluación anterior (SAC-10, 1975-2019).

APPENDIX 3

TABLE A1. Average adjusted input sample size (n adj) and average effective sample size for each fishery and for the survey, by model, with $h = 1$. The data in the shaded cells were not used in the models. Data in boldface indicate the fisheries with the largest ranges in effective sample size.

TABLA A1. Tamaño promedio de la muestra de entrada ajustado (n adj) y tamaño promedio efectivo de la muestra para cada pesquería y para el estudio, por modelo, con $h = 1$. Los datos en las casillas sombreadas no fueron usados en los modelos. Datos en negritas indican las pesquerías con los mayores rangos de tamaño de muestra.

Survey	Effective n													Range
	n adj	BASE	DDQ	DDQ.DS	DDQ.GRO	DS	GRO	TBE	TBE.DS	TBE.GRO	TBM	TBM.DS	TBM.GRO	
Survey	11.8	59	62	63	63	61	60	64	64	65	64	64	63	7
Fisheries														
F1	7.3	63	64	63	61	62	61	62	62	61	63	63	61	3
F2	6.1	36	35	35	35	36	35	36	36	35	36	36	35	1
F3	9.4	58	59	59	72	58	70	59	59	70	59	59	72	14
F4	4.0	33	33	33	32	33	32	33	33	32	33	33	32	2
F5	3.6	40	39	38	38	40	40	41	41	40	40	39	39	3
F6	9.4	83	83	82	87	83	88	83	83	88	83	83	87	6
F7	4.1	38	37	37	35	38	36	39	39	36	38	38	35	4
F8	6.8	55	55	54	52	55	53	55	55	52	55	55	53	3
F9	3.8	30	30	30	31	30	32	30	30	31	30	30	31	2
F10	4.8	25	26	26	25	25	24	25	25	25	26	26	25	1
F11	5.3	63	61	61	61	63	62	63	63	62	63	62	62	2
F12	2.3	25	25	25	24	25	25	25	25	25	25	25	25	1
F13	2.0	33	33	33	34	34	34	34	34	34	34	34	34	1
F14	3.0	22	21	22	21	22	21	22	22	21	22	22	21	1
F15	4.1	108	107	106	105	109	107	109	109	109	109	109	108	4
F16	7.5	113	113	113	114	114	115	114	114	116	114	115	115	3
F17	3.8	53	53	53	53	53	54	53	53	54	53	53	53	1
F18	7.8	159	161	162	162	160	160	160	160	160	155	157	155	8
F19	6.2	105	110	111	111	110	107	110	110	111	108	114	111	9
F20	4.2	61	69	48	101	44	103	52	50	66	53	40	92	63
F21	4.5	45	47	39	50	37	50	40	39	41	41	35	47	15
F22	5.2	45	45	45	50	45	50	45	45	50	45	45	49	5
F29	6.9	44	49	41	55	38	54	42	41	45	42	37	52	18
F30	8.9	58	67	55	77	49	76	54	53	62	55	47	74	31
F31	2.8	18	19	17	18	16	18	17	17	17	18	17	18	2
F32	5.0	47	52	45	52	42	49	45	44	45	45	40	48	12
F33	3.4	31	35	28	44	26	41	28	28	31	28	25	38	19
F34	2.0	23	26	23	24	21	23	22	22	22	23	21	23	4