DOCUMENT SAC-11-06 REV

BIGEYE TUNA IN THE EASTERN PACIFIC OCEAN, 2019: BENCHMARK ASSESSMENT

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

CONTENTS

---

[1] Postponed until a later date to be determined

EXECUTIVE SUMMARY

1. This year's benchmark assessment of bigeye tuna in the eastern Pacific Ocean is the basis for a risk analysis used to provide management advice (SAC-11-08). 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 a total of fourteen reference models, each with four steepness assumptions (0.7, 0.8, 0.9, and 1.0).

3. The fourteen reference models included in this assessment are developed within a hierarchical framework, and combine components that address three major uncertainties in the previous assessment: a) the apparent regime shift in recruitment (R shift), b) the misfit to the length-composition data for the longline fishery that is assumed to have asymptotic selectivity, and c) the steepness of the stock-recruitment relationship.

4. The 44 converged model runs made for this assessment indicate that:

    a. At the beginning of 2020, the spawning biomass (S) of bigeye ranged from 14% to 212% of the level at dynamic MSY (S_{MSY_d}); 26 runs suggest that it was below that level.

    b. During 2017-2019, the fishing mortality (F) of bigeye ranged from 51% to 223% of the level at MSY (F_{MSY}); 26 runs suggest that it was above that level.

    c. At the beginning of 2020, the spawning biomass (S) of bigeye ranged from 51% to 532% of the limit reference level (S_{LIMIT}); five runs suggest that it was below that limit.

    d. During 2017-2019, the fishing mortality of bigeye ranged from 32% to 114% of the limit reference level (F_{LIMIT}); three runs 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, all short-term models and two environment models (Models *Env-Fix* and *Env-Mrt*) estimate that, regardless of what value is assumed for steepness, S is below S_{MSY_d} and F is above F_{MSY} at the beginning of 2020, whereas Models *Gro* and *Sel* estimate the opposite. The stock status at the beginning of 2020 estimated by the remaining models depends on the value assumed for steepness.

6. 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 assessment[2] of bigeye tuna (*Thunnus obesus*) in the eastern Pacific Ocean (EPO), conducted using an integrated statistical age-structured stock assessment modeling platform (Stock Synthesis 3.30.15). 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.

---

[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.

1.1. Background

In recent years, some problems and sources of uncertainty had arisen in the staff's assessment of bigeye, culminating in the suspension of the staff's previous 'best assessment' approach and the adoption of the work plan to improve its stock assessments. The main problem was the apparent 'recruitment shift' ('*R* shift') in the mid-1990s, when the assessment estimated that the average recruitment doubled at the same time as the purse-seine catches of bigeye, mostly smaller fish, increased from 10,000 to nearly 50,000 metric tons (t) in three years with the rapid expansion of the fishery on fish-aggregating devices (FADs) in the equatorial EPO, while longline catches of large bigeye halved. Although it is possible that recruitment did in fact increase, this result appears anomalous, and various hypotheses have been proposed to explain it (see Aires-da-Silva *et al.* (2010), Valero *et al.* (2019), and Punt *et al.* (2019) for details).

Another issue in the assessment of bigeye has been the systematically poor fits of the model to longline length-composition data. In particular, the model predicts higher proportions of larger (older) bigeye than those observed in the length-composition data for the longline fishery, which is assumed to have an asymptotic selectivity. While this could be due to observation error, it could also be the result of mis-specified assumptions in the model about parameters such as growth, natural mortality and/or selectivity.

One major source of uncertainty, and potentially also the explanation for other issues identified in the model, is the possibility of the stock having a spatial structure that is not captured in the model. However, preliminary modeling work did not find that stock structure in the EPO could explain the *R* shift, and there were insufficient data to model interactions with the central Pacific Ocean (CPO) (Valero *et al.* 2019). Therefore, the assessment assumes, as before, that there is a single stock of bigeye in the EPO, with minimal net movement of fish between the EPO and the CPO.

There is also uncertainty about the nature of the stock-recruitment relationship (a measure of the degree to which the recruitment to a stock is determined by the stock's spawning biomass, expressed as steepness (*h*)). The staff's previous assessments of bigeye have consistently presented analyses of the sensitivity of the results to different assumptions about this *h* parameter, but only to show the impact of these assumptions on estimated management quantities, and the uncertainty around *h* was not explicitly incorporated in the management advice.

More recently, 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-09 INF-B). This may 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.

As a result of these issues and uncertainties, in 2018 the staff concluded that the results of its stock assessment of bigeye were not reliable enough to be used as a basis for management advice to the Commission (IATTC-94-03), and therefore implemented the work plan to improve the stock assessments for tropical tunas. This benchmark assessment not only addressed the existing issues, but also included reviews by external experts of the assessments of both bigeye and yellowfin tunas and the development of a new approach to standardizing the longline catch per unit effort (CPUE) and associated length-composition data, using spatio-temporal models. Neither external reviews identified a particular replacement for the current base-case models, but both suggested a variety of alternatives for the staff to consider regarding natural mortality schedules, growth models, selectivity curves, and estimation procedures.

The staff's approach to the stock assessment of tropical tunas could be improved by incorporating model uncertainty to derive information for management advice.

1.2. The new approach

This 2020 benchmark assessment of bigeye in the EPO, and the companion assessment of yellowfin tuna (SAC-11-07), represent a new approach to stock assessments by the staff. Previously, it used a '*best as-sessment*' approach, which bases the evaluation of stock status on a single 'base-case' model: the new assessments are based on '*risk analysis*' methodology, which uses 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 different results, thus effectively incorporating uncertainty into the management advice as it is formulated[3]. This change, which represents a paradigm shift at IATTC, both for the staff's work and the Commission's decision-making regarding the conservation of tropical tunas, also allows the staff to evaluate explicitly the probability statements in the IATTC harvest control rule for tropical tunas established in Resolution C-16-02.

This new approach to formulating management advice for tropical tunas includes the following four components:

1. Two **stock assessment reports**, for bigeye (this document) and yellowfin (SAC-11-07), presenting the results from all reference models for each species (model fits, diagnostics, stock status);
2. A **risk analysis** (SAC-11 INF-F), assessing the consequences of using each model 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 (bigeye, skipjack, yellowfin); and
4. The **staff's recommendations** (SAC-11-15) for the conservation of tropical tunas, based on the risk analysis above.

2. DATA

2.1. Fisheries and 'surveys'

The 2nd external review of the bigeye assessment did not consider the development of a spatial model for the EPO a high priority in the short term. Accordingly, this assessment uses an 'areas-as-fleets' approach, in which the geographical areas (Figure 1) are implicitly modeled as separate fleets in a one-area model.

There are no fishery-independent surveys of tuna abundance and size composition in the EPO: all the available data are obtained from the fishery, directly or indirectly. However, the philosophy of how to treat fishery-derived indices of abundance has changed for this assessment. Although still based on CPUE data, they are disconnected from the fisheries structure, total catch, and catch composition, and are therefore more like true surveys. They are called "surveys" to be consistent with Stock Synthesis terminology, and also because they are modeled as having data, such as indices of abundance and size compositions, but no catch, as a real survey would.

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

---

[3] See SAC-11 INF-F (Maunder *et al*. 2020) for a description of the technical details of the risk analysis, using bigeye as a case study.
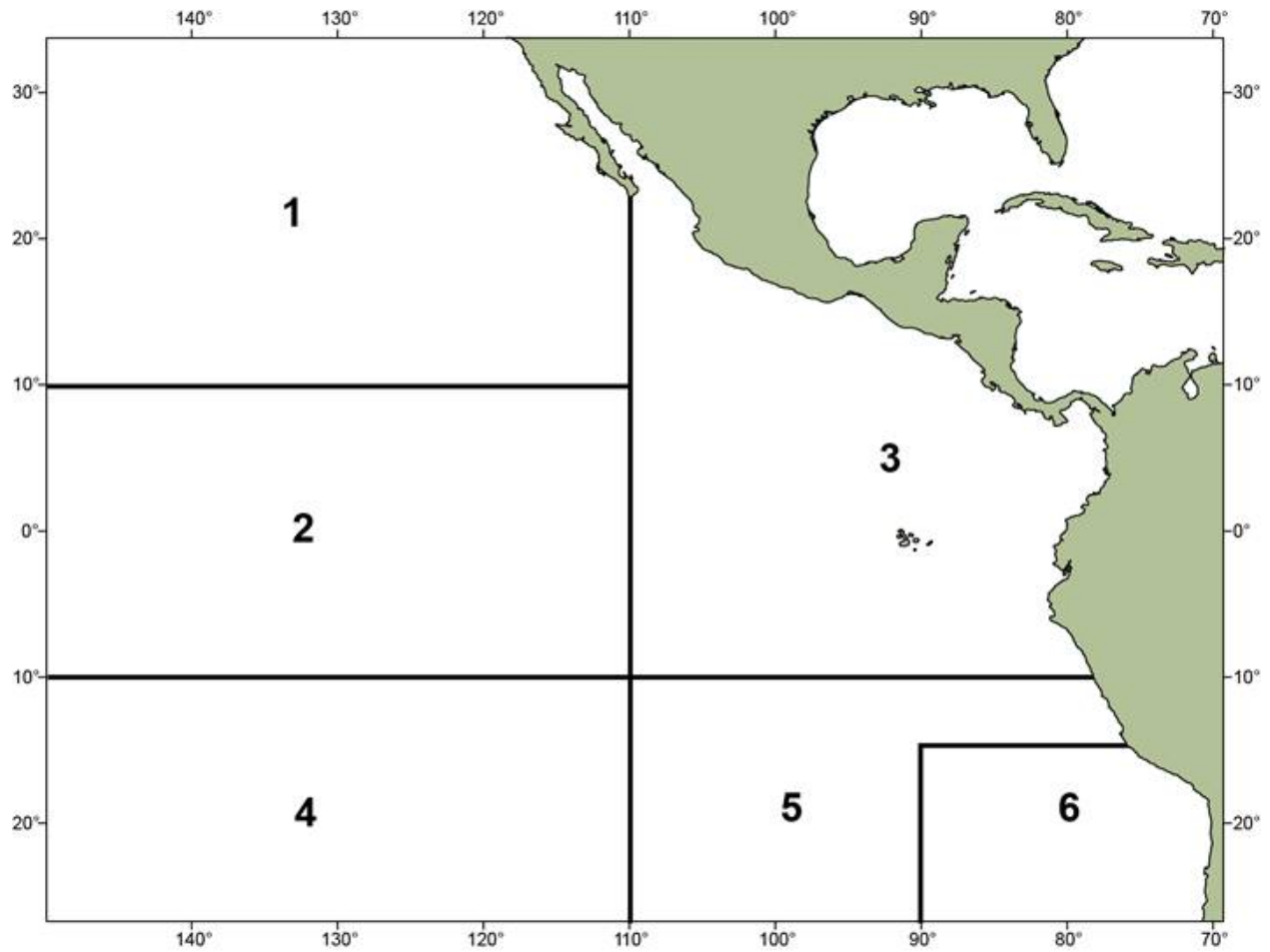
FIGURE 1. Spatial extent of the fisheries defined for the stock assessment of bigeye tuna in the EPO. The fisheries are summarized in Table 1.

FIGURA 1. Extensión espacial de las pesquerías definidas para la evaluación del atún patudo en el OPO. En la Tabla 1 se resumen las pesquerías.

TABLE 1. Fishery and "survey" fleets defined for the stock assessment of bigeye tuna in the EPO. PS = purse-seine; LL = longline; OBJ = sets on floating objects; NOA = sets on unassociated fish; DEL = sets on dolphins. See Figure 1 for map of areas.

Fleet Number	Gear	Set type	Years	Area	Catch data	Unit
FISHERIES						
1	LL	-	1979-present	1	Retained catch only	1,000s
2			Block 1: 1979-1993 Block 2: 1994-present	2		
3				3		
4				4		
5				5		
6				6		
7	LL	-	1979-present	1	Retained catch only	tons
8			Block 1: 1979-1993 Block 2: 1994-present	2		
9				3		
10				4		
11				5		
12				6		
13	PS	OBJ	1979-present	2	Retained catch + discards (inefficiencies)	tons
14				3		
15				4		
16				5		
17				6		
18	PS	OBJ	1979-present	2-6	Discards (size-sorting)	tons
19	PS	NOA+DEL	1979-present	2	Retained catch + discards (all)	tons
20				3		
21				4		
22				5		
23				6		
SURVEYS						
24	LL	-	1979-1992	2-6	No catches	-
25			1995-present			

2.1.1. Fisheries

Twenty-three fisheries are defined for the stock assessment of bigeye tuna in the EPO, classified by gear (purse-seine/longline/pole-and-line), purse-seine set type (floating object/unassociated/dolphin), area of operation (Figure 1), and unit of catch (number/weight) (Table 1).

These new spatial definitions are based on the results of a regression tree analysis that uses both CPUE and length-frequency data to investigate the stock structure of bigeye in the EPO (Lennert-Cody *et al*. 2010; Document [WSBET-02-02](#)). The analysis synthesizes spatial structure in these data types across different gears to provide one set of spatial boundaries for all fisheries in the assessment, which necessarily represent a compromise between those that would be best for each gear type. 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 area. As in the previous assessment (Xu *et al*. 2018), since the catches of bigeye by pole-and-line (LP) vessels and by purse-seine vessels setting on dolphins (DEL) and unassociated tunas (NOA) are small relative to other fisheries, for convenience, they are pooled (Fisheries 19-23).

For the longline fishery, an abrupt increase in the number of hooks between floats (HBF) around 1993-1994 resulted in deeper longline sets, which implies different selectivity curves after 1993; therefore, the ten longline fisheries in Areas 2-6[4] were divided into two time blocks (1979-1993 and 1994-2019; Table 1) with different selectivity parameters.

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. Two longline 'surveys' are defined for this assessment of bigeye in the EPO based on the time of operation, 'early' (1979-1992) and 'late' (1995-2019) (Table 1). No surveys are assigned to Area 1 because the indices of abundance and composition data from this area are likely not representative of the "core" equatorial region of the bigeye distribution, all of which lies outside Area 1 (Figure 1). Catchability and selectivity are estimated separately for the two surveys. The coefficient of variation (CV) of the late index of abundance is fixed while that of the early index is estimated. See section 2.3 for information about how CVs are specified for survey indices of abundance.

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

---

[4] In Area 1, the length-composition sampling before 1993 was too limited to enable different selectivities to be estimated for the two time blocks

information on species composition (see section 2.3.4).

For this assessment, EPO total catches by species were estimated by catch stratum (area, month, set type, and vessel carrying capacity) and then aggregated across catch strata to obtain quarterly estimates for each fishery. 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 or observer or logbook data, as applicable. To correct for underestimated bigeye catches, an adjustment 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 individual 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 in the EPO by the fleets of distant-water CPCs (China, Chinese Taipei, French Polynesia, Japan, and Korea) and coastal CPCs (principally Mexico and the United States).

For this assessment, these data are aggregated in line with the new fishery definitions based on area of operation (Figure 1 and Table 1). Because two longline fisheries are defined for each area, the catches are entered in their original units (number or weight), and the conversion between number and weight is done internally in the assessment model.

Updated and new catch data for the longline fisheries (Fisheries 1-12), available to the IATTC staff on 10 April 2020, were incorporated into the current assessment. New or updated catch data were available for Vanuatu (2018-2019), Chinese Taipei (2016-2019), French Polynesia (2018-2019), China (2018-2019), Japan (2016-2019), Korea (2018-2019), the United States (2014-2018) and Panama (2019). 2019 data are from monthly reports. If catch data for a recent year or years were unavailable, catches were set equal to the last year for which data were available. For fleets that reported catch aggregated by year and 5° cell, the data were disaggregated, using the proportion of catches by quarter and area for the closest year for which data 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 catch by fishery 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.

For the purse-seine fishery, the amount of bigeye discarded, regardless of the reason, is estimated with information collected by AIDCP 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. Total catch by Fisheries 13-17 (OBJ) represents retained catch plus discards resulting from inefficiencies in the fishing process. Fishery 18 represents discards resulting from catch sorting in Fisheries 13-17, which are treated separately, following the rationale of Maunder and Watters (2001), and are assumed to be composed only of fish that are 2-4 quarters old. In Fisheries 19-23 (NOA and DEL), total catch represents retained catch plus some discards resulting from inefficiencies in the fishing process and from sorting the catch, although the latter is infrequent in these fisheries.

Discards by the longline fisheries cannot be estimated with the minimal data available, so it is assumed that the retained catch represents the total catch (Table 1).

2.2.4. Catch trends

To compare the catch trends of the purse-seine and longline fisheries (Figure 2), longline catches in number are converted to weight inside the assessment model. The catch trends for bigeye have several important features:

1. Longline fisheries dominate bigeye catches before 1993; since then the OBJ fishery has become the main fishery for bigeye.
2. The total annual catch has been relatively stable since 1997.
3. The two tropical areas in the EPO, Areas 2 and 3, are the most important fishing grounds for both longline and purse-seine fisheries.
4. Longline catches in Areas 2 and 3 have decreased since 1979. In Area 2 the purse-seine catch increased continuously since the expansion began in 1993, while in Area 3 it increased during 1993-2000 and decreased notably thereafter.
5. Longline catches in Area 1 are much smaller than in the other areas in the EPO, supporting the exclusion of Area 1 from the surveys.

2.3. Indices of abundance

While both purse-seine and longline indices of abundance are available for bigeye in the EPO, this assessment includes only the longline indices, which primarily inform the trend of adult abundance. Standardizing purse-seine indices of abundance is notoriously challenging, because the relationship between abundance and fishing effort, and how it evolves with technological advancements over time, is unclear, and at least some of the covariates that impact the catchability of the purse-seine fishery (*e.g.*, soak time, FAD density, presence of echosounder) are not available for the entire period of interest. The staff is investigating novel methods for producing reliable purse-seine indices of abundance using FAD data, which might potentially inform the trend of juvenile abundance in future assessments of bigeye.

2.3.1. Data selection

The longline indices of abundance developed for this assessment are based on newly available aggregated catch and effort data for the Japanese commercial fleet, at a resolution of 1° cell x month x vessel. Only the data for 1979-1992 ('early') and 1995-2019[5] ('late') were used for standardizing these indices, for two reasons: 1) vessel identification, a catchability covariate included in the index standardization procedure, was not available before 1979; 2) the sharp increase in HBF in 1993-1994 may cause changes in both survey catchability and selectivity.

---

[5] Since catch and effort information for the fourth quarter of the terminal year (2019) is incomplete, it was removed from the dataset, as in previous assessments.

The 'core' fishing ground, defined as all 1° x 1° cells in the EPO with at least 80 quarters of observations during 1979-2019 (Figure A1), was used for the index standardization. This selection is due primarily to the concern that the marked contraction of the Japanese longline fishery in recent years, in terms of fishing effort (Figure 3a) and area of operation (Figure 3b), may bias the indices for those years. Using the catch and effort dataset for the core fishing ground only when fitting a spatiotemporal model reduces the potential impact of spatial imputation of catch rates for unfished locations on the accuracy of standardized indices from such models. To allow the standardization model to accurately estimate vessel effects on catchability over an extended period of time, vessels with missing identification or less than 40 quarters of observations are filtered out of the dataset. The number of vessels selected by this criterion increased approximately linearly from 20 in 1979 to a historical high level (150) in 1992 before dropping also approximately linearly to less than 20, the lowest level, in 2019 (Figure 3c).

2.3.2. Standardization procedure

A spatiotemporal delta-generalized linear mixed model (VAST; Thorson and Barnett 2017) is used to estimate the longline indices of abundance for bigeye in the EPO. This approach models encounter probability and positive catch rate separately to deal with zero-inflated catch observations, uses area-weighting, and can impute catch rates in unfished locations based on estimated spatial correlation patterns, which are assumed to be constant over both space and time. The early and late indices of abundance used in the assessment are standardized by two separate VAST models, both of which use the *logit* and *log* link functions for the linear predictors of encounter probability and positive catch rate, respectively. Both predictors include an intercept term (year-quarter effect), spatial term, spatiotemporal term, and two catchability covariates (vessel identification as a factor term and HBF as a linear term) to account for the effects of vessel and gear on fishing efficiency (for details see Xu *et al*. 2019). The two VAST models group all 1° cells in the core fishing ground into 200 spatial knots to improve computational efficiency and estimate the standardized index of abundance and associated CV for every quarter in the early and late periods (Figure 3). The code used to standardize the two indices of abundance can be found here.

The CVs of the indices are estimated in VAST based on the quality of catch rate data in each quarter, including spatial coverage and sample size. The low spatial coverage and sample size in recent years (Figure 3) result in high CVs being estimated for those years (Figure 4). In the previous assessments of bigeye (Aires-da-Silva *et al*. 2015a, Xu *et al*. 2018), a time-invariant CV of 0.15 was used for longline indices of abundance. In this assessment, however, the CV of the early survey index is the sum of the CV estimated by VAST and a constant estimated by the assessment model depending on the degree of consistency between this index with other data inputs; whereas the CV of the late survey index is fixed at the sum of the CV estimated by VAST and a constant that scales the mean CV in 1995-2014 to be 0.15 (Figure 4).

It is worth noting that the late longline index increased markedly after the first quarter of 2018, which is likely a result of the strong El Niño in 2016 (Figure 4). The previous two El Niño events of similar magnitude (1983 and 1998) induced concurrent high recruitment and higher values of the index of abundance, the latter peaking around three years later (Figure 4). There is also the possibility of a positive impact of El Niño on longline catchability, which merits further investigation in future assessments.

2.4. Size-composition data

2.4.1. Fisheries

2.4.1.a 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.

2.4.1.b Longline

The length-composition data for longline fisheries in this assessment are based on 1) new monthly 1° x 1° length-frequency data for the Japanese commercial fleet; 2) new monthly 1° x 1° catch and effort data for individual Japanese commercial vessel; and 3) 5° x 5° quarterly longline catch data reported by CPCs. The length-composition data should be representative of longline catches, so the raw length-frequency observations, filtered to remove any not precise to within 2 cm, are raised to the corresponding fishery catch in the same location and time strata, as follows:

1. Raise Japanese monthly 1° x 1° length-frequency data to Japanese total catch in the same strata;
2. Aggregate the raised data from the previous step to quarterly 5° x 5° catch;
3. Raise the aggregated length-frequencies from the previous step to the total catch of all CPCs in the same strata;
4. Aggregate the raised length-frequencies from the previous step for each longline fishery (Fisheries 1-12).

The length compositions are aggregated by 2-cm intervals from 20 cm to 198 cm, 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, to very low values in the most recent quarters (Figure 5). The sample sizes are scaled within the assessment using the Francis (2011) approach (Table 2). The longline fishery in Area 2 (Fishery 2) catches the highest proportion of large (150+ cm) bigeye in the late period (Figure 6) and is therefore assumed in most reference models to have an asymptotic selectivity (see section 4).

2.4.1.c Trends in sample size and mean length

Since the expansion of the OBJ fisheries, their sample sizes, especially those in Areas 2 and 3, have increased pronouncedly (Figure 5 top), to the point that most length-composition data for those fisheries in recent years are from those two areas. The longline fisheries show a different trend: before the expansion of the OBJ fisheries, the longline length-composition data were predominantly from Area 2, but since then the sample sizes for all longline fisheries, including in that area, have fallen to a very low level in recent years (Figure 5 bottom).

The average length of bigeye caught in most OBJ and longline fisheries in the EPO does not show a strong temporal pattern (Figure 6 top). One exception is the OBJ fishery in Area 6, where the average size of bigeye caught tripled before 2000 and halved thereafter. However, the very low sample sizes for this fishery (Figure 5 top) suggest that this large temporal variation may be primarily an artifact of sampling error. Overall, the long-term average length composition of each main longline fishery (in Areas 2-5) shifted to larger bigeye from the early to the late period (Figure 7), which could be caused by the increase in HBF over time.

2.4.2. Surveys

2.4.2.a Data selection

The length-composition data for longline indices of relative abundance in this assessment are based on 1) new monthly 1° x 1° length-frequency data, in 1- or 2-cm intervals, for the Japanese commercial fleet; and 2) new monthly 1° x 1° catch and effort data for individual Japanese commercial vessel. Unlike the length-composition data from the fisheries that are representative of catches, the length compositions of surveys are representative of abundance (Maunder *et al*. 2020), so the raw length-frequency observations are raised to the catch rate obtained from the catch and effort data in the same location and time strata. To match the indices of abundance to which the survey length compositions correspond, the raw length frequency observations in only the core fishing ground (Figure A1) are selected for the standardization of survey length compositions.

2.4.2.b Standardization procedure

Survey selectivity is assumed to be time-invariant in this assessment. To ensure that this assumption is approximately accurate, VAST is also used to standardize the length-composition data for surveys, following the procedure described in Thorson and Haltuch (2018). Standardizing length-specific catch-rate data using spatiotemporal models like VAST is computationally intensive and time-consuming, so the input composition data were aggregated into 10-cm intervals, from 20 cm to 190 cm. The early and late longline survey length compositions are standardized by two separate VAST models, both of which use the *logit* and *log* link functions for the linear predictors of encounter probability and positive catch rate, respectively. Both predictors for each length interval include an intercept term (year-quarter effect), spatial term, and spatiotemporal term, all of which are assumed to be independent among length intervals, due mainly to the infeasibly long computation time when estimating covariances among length intervals. VAST groups all 5° x 5° cells in the core fishing ground into 50 spatial knots to further improve computational efficiency. The two VAST models provide standardized length compositions for every quarter in the early and late periods. The input sample size of the survey length compositions is computed as the total number of fish sampled divided by 100. The sample sizes are scaled within the assessment model using the Francis (2011) approach.

2.5. Age-at-length data

Age-at-length data derived from otolith readings (Schaefer and Fuller 2006) were integrated into some, but not all, reference models to provide information on mean length-at-age and variation in length-at-

age. These data consist of age estimates from counts of daily increments on otoliths and length measure-ments of 254 bigeye tuna caught in 2002 by floating-object fisheries (Figure 8). The otoliths were collected by length-stratified sampling and were therefore included in the model as age conditioned on length. Age-at-length data derived from otolith readings are available for fish up to four years of age only because otolith daily increments for large (older) fish are very difficult to interpret (Schaefer and Fuller 2006).

3. ASSUMPTIONS AND PARAMETERS

3.1. Biological and demographic information

3.1.1. Growth

As with many tuna species, specifying growth in the bigeye stock assessment for the EPO presents some challenges. Age-at-length data derived from otolith readings are available for fish up to four years of age only (Schaefer and Fuller 2006). This is a narrow spectrum of ages of the longevity of at least 15-16 years estimated from tagging studies (Langley *et al*. 2008). Otolith daily increments for large (old) fish are very difficult to interpret. Bigeye growth estimates from tagging studies are available, but again these are mostly limited to juvenile ages (Schaefer and Fuller 2006). Acquiring tag-recapture information for older fish is problematic, since they are difficult to catch for tagging, and few tag recoveries from larger fish are available from the longline fisheries.

Following the recommendations of the first [external review of the IATTC staff's assessment of bigeye](#) in 2010, a transition was made from the traditional von Bertalanffy growth model to a more flexible Richards (1959) growth model. In Stock Synthesis, the Richards growth curve for the mean length (*L*) at age (*a*) is re-parameterized according to Schnute (1981):

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

Where a_1 and a_2, the first and last ages at which the fish follow the Richards curve, are specified to be 0 and 40 quarters, respectively. Therefore, L_1 and L_2 are the mean length at age 0 and 40 quarters, respec-tively. *K* represents the growth rate, and *b* influences the shape of the Richards growth curve. Previous sensitivity analyses (Aires-da-Silva and Maunder 2010, Aires-da-Silva *et al*. 2015a, Hampton and Maunder 2005) have shown that the results of the bigeye stock assessment are highly sensitive to L_2. Based on a recently-available length-increment tagging dataset that includes large bigeye in the EPO, Aires-da-Silva *et al*. (2015b) developed an integrated model that incorporates both otolith age-at-length data and length-increment tagging data into the estimation of growth parameters. The integrated model was found to improve the estimates of growth parameters, especially L_2, than the model based on otolith age-at-length only. Aires-da-Silva *et al*. (2015b) estimated that, for bigeye in the EPO, L_1 and L_2 are 21.88 cm and 196.34 cm, respectively; *K is* 0.11 quarter^{-1}, and *b* is 0.23.

Another important component of growth used in age-structured statistical catch-at-length models is the variation of length at age, which can be just as influential as the mean length at age. For bigeye in the EPO, the standard deviation of length at age is assumed to be proportional to mean length at age. Information on the variability of length at age can be obtained from the otolith data with the method of Aires-da-Silva *et al*. (2015), or by integrating the otolith data into the stock assessment. In this assessment, the standard deviations of length at age 0 and 40 quarters are, respectively, estimated internally by the assessment model and fixed at the value (8.878 cm) estimated externally by Aires-da-Silva *et al*. (2015b). The reason for estimating the standard deviation of length at age 0 internally by the assessment model is that the externally-estimated value (2.962 cm) appears to be too small, as suggested by the poor fit of the length-composition data from the floating-object fisheries at small sizes (figure not shown).

The following weight-length relationship, from Nakamura and Uchiyama (1966), was used to convert length to weight in the current stock assessment:

$$w = 3.661 \times 10^{-5} \times l^{2.90182} \quad (Equation\ 2)$$

where *w* is weight in kilograms and *l* is length in centimeters.

3.1.2. Natural mortality (*M*)

Age-specific vectors of natural mortality (*M*) are assumed for bigeye in the EPO. This assessment uses sex-specific models and natural mortality schedules are provided for each sex (Figure 3.2 in SAC-01-08a). *M* at age 0 is assumed to be 0.25 for both sexes, and *M* is assumed to decrease to 0.1 at 5 quarters of age. As in previous assessments, it is assumed that female *M* increases after the fish mature. These age-specific vectors of *M* are based on fitting to the estimates of age-specific proportions of females, maturity at age, and *M* of Hampton (2000).

Different levels of *M* had a large influence on the absolute population size and the population size relative to that corresponding to the maximum sustainable yield (MSY) (Watters and Maunder 2001). Harley and Maunder (2005) performed a sensitivity analysis to assess the effect of increasing *M* for bigeye younger than 10 quarters. In addition, the effect on the bigeye stock assessment of assuming alternative scenarios of juvenile *M* has been evaluated (Document SARM-9-INF-B). The management quantities showed little sensitivity when higher levels of *M* were assumed for fish 0-5 quarters of age, but greater sensitivity to the assumption made about the older early ages (5-12 quarters) included in the early high levels of *M*. However, the high levels of *M* assumed for bigeye 5-12 quarters old (60-120 cm) seem unrealistic.

An ongoing investigation of *M* for bigeye, based on an integrated analysis which includes tagging and sex-ratio data, indicates levels of *M* for adult bigeye higher than previously assumed (Maunder and Aires-da-Silva 2010). However, these estimates are highly uncertain, and strongly dependent on the assumptions made about tag-reporting rates by longliners.

3.1.3. Recruitment and reproduction

It is assumed that bigeye can be recruited to the fishable population during every quarter of the year. Recruitment may occur continuously throughout the year, because individual fish can spawn almost every day if ambient water temperature is in the appropriate range (Kume 1967, Schaefer 2006).

Stock Synthesis allows a Beverton and Holt (1957) stock-recruitment relationship to be specified. The Beverton-Holt curve is parameterized so that the relationship between spawning biomass (*S*, biomass of mature females) and recruitment (modeled in Stock Synthesis as the number of age-0 fish) is determined by estimating the average recruitment produced by an unexploited population (virgin recruitment) and *h*, defined as the fraction of virgin recruitment that is produced if spawning biomass is reduced to 20% of its unexploited level. It controls how quickly recruitment decreases when the spawning biomass is reduced and can vary between 0.2 (recruitment is a linear function of spawning biomass) and 1.0 (recruitment is independent of spawning biomass). In practice, it is often difficult to estimate *h* because of a lack of contrast in spawning biomass and because of other factors, like environmental influences, can cause recruitment to be extremely variable. Previous assessments assumed that *h* is 1 since there is no evidence that recruitment is related to spawning stock size for bigeye in the EPO. If steepness is estimated as a free parameter in the model, it is estimated to be 1. However, simulation analyses have shown *h* is frequently estimated to be 1 even when the true steepness is lower (Lee *et al*. 2012). In this assessment, four values of steepness (0.7, 0.8, 0.9, and 1.0) are evaluated. In addition to the assumptions required for the stock-recruitment relationship, a constraint on quarterly recruitment deviates with a standard deviation of 0.6 is applied. It is also important to note that the method proposed by Methot and Taylor was used to provide bias adjustment for recruitment.

Reproductive inputs are based on the results of Schaefer (2006) and data provided by Dr. N. Miyabe of the National Research Institute of Far Seas Fisheries (NRIFSF) of Japan. Information on age at length (Schaefer and Fuller 2006) was used to convert proportion mature at length into an age-at-maturity schedule (SAC-01-08a, Figure 3.3 and Table 3.1).

3.1.4. Movement

Movement patterns of tagged bigeye in the EPO and adjacent areas are available from Schaefer and Fuller (2009) and Schaefer et al. (2015). With respect to juvenile fish, Minte-Vera *et al*. (2019) reported movement in an eastward direction at rates between 16% to 23% per quarter, using tag-recoveries from conventional tags and advection-diffusion analysis of archival tag data. Movement information for older fish is very limited. Valero *et al*. (2019b) implemented a series of exploratory spatially-structured models with and without movement for bigeye in the EPO and adjacent areas. Although some spatial configurations with movement reduced the R shift, there are uncertainties in the movement rates of juvenile bigeye (which seem too high, based on assessment model results), and those for adults are unknown, both in direction and rate. While some combinations of estimated juvenile rates and assumed rates and directions of adult movement reduce the R shift, they resulted in accumulation of fish to the coastal areas, and lead to either unrealistically large differences in longline catchability between areas, or lack of model convergence when mirroring catchability among areas (Valero *et al*., 2019a, b).

The current assessment does not consider movement explicitly; rather, it is assumed that the population is randomly mixed at the beginning of each quarter of the year. The IATTC staff is studying the movement of bigeye within the EPO, using data recently collected from conventional and archival tags, and these studies indicate movement of juvenile bigeye into the EPO from the CPO. The studies may eventually provide information useful for stock assessment. The spatial definition of the fisheries implicitly accommodates some forms of movement through different selectivities. It is worth noting that natural mortality in this closed-population assessment model can also account for the east-west movement across the 150°W management boundary. Moving westward (eastward) across the management boundary is interpreted by the model as an increased (decreased) natural mortality rate.

3.1.5. Stock structure

Minte-Vera *et al*. (2019) provide a review of information available for identifying the stock structure of bigeye within the EPO. Tagging analysis and multivariate regression-tree analysis of size distribution and CPUE trends (Lennert-Cody *et al*., 2010; Minte-Vera *et al*, 2019) have been used to investigate spatial structure in the EPO. Information on the stock structure is limited, in part because stock identity is from fishery-dependent sources and information from other disciplines is limited or not available (see Punt *et al*. 2019). Information and analyses currently available suggest a single genetic population of bigeye tuna across the Pacific Ocean, with some regional sub-population structure, with considerable connectivity between the EPO management unit and the CPO (Schaefer, 2009; WPFMC, 2014; Schaefer *et al*., 2015, Minte-Vera *et al*., 2019; Punt *et al*. 2019). Several evaluations of potential spatial structure in the bigeye stock assessment have been conducted (Aires-da-Silva and Maunder 2010b; Valero *et al*. 2018, 2019b). Exploratory assessment models for the central EPO area removed the R shift and indicated a more depleted stock (Valero *et al*. 2018). However, spatially disaggregating the entire EPO did not remove the R shift, which remained in all spatially-structured EPO-wide exploratory assessments that did not have major diagnostic problems (Aires-da-Silva and Maunder, 2010; Valero *et al*., 2018; Valero *et al*. 2019a; Valero *et al*., 2019b), suggesting that spatial structure alone will not resolve the problems associated with the R shift. Further spatially-structured frameworks will be considered in future stock assessments.

Despite tagging data showing movement of juvenile bigeye from the CPO into the EPO, for this assessment it is assumed that there are two stocks, one in the EPO (the focus of the assessment) and the other in the

CPO, and that there is no net exchange of fish between these regions. The IATTC staff periodically conducts a Pacific-wide assessment of bigeye in collaboration with scientists of the Oceanic Fisheries Programme of the Secretariat of the Pacific Community and of the NRIFSF. This work may help indicate how the assumption of a single stock in the EPO is likely to affect the interpretation of the results obtained from Stock Synthesis.

4. REFERENCE MODELS

In this report, the latest version (3.30.15) of Stock Synthesis (Methot and Wetzel 2013) was used to assess the status of bigeye in the EPO. Stock Synthesis is an age-structured, statistical stock assessment model framework that allows models of varying complexity to be implemented and fitted to many different types of data. Previous assessments of bigeye have relied on the results of a single model configuration (base case) for evaluating stock status. As described in section 3, uncertainties about several assumptions have remained in the assessment and the staff has developed a risk analysis framework to explicitly include those in the evaluation of stock status and formulation of management advice.

The first step in the implementation of the risk analysis is to establish the plausible hypotheses (reference models) that define the states of nature associated with the main existing sources of uncertainty. These models are designed to address three major uncertainties in the previous assessment: (1) the apparent regime shift in recruitment (R shift); (2) the misfit to the length-composition data for the longline fishery that is assumed to have asymptotic selectivity (Fishery 2); and (3) the steepness of the stock-recruitment relationship (h). The reference models were developed within a hierarchical framework, and combine components that address all three issues. The models that address the R shift assumed that the shift was either (i) real, or (ii) an artifact of model misspecification. Several models were eliminated from consideration based on previous work, and others after initial runs. The remaining models are divided into those that assume that: (1) the R shift is real and is estimated as a recruitment regime parameter (for 1979-1993) in the model (**Environment** (**Env**); (2) the R shift is an artifact of model mis-specification, but the cause is unknown, so the early data are excluded by starting the model in 2000 (**Short** (**Srt**)); and (3) the R shift is an artifact of model mis-specification, and the process is known. This last group includes models that estimate (a) pre-adult (13-quarter-old) M as a proxy for pre-adult movement (**Movement** (**Mov**)); (b) the growth curve and the variation of length at age (**Growth** (**Gro**)); (c) a dome-shaped selectivity curve for Fishery 2 (**Selectivity** (**Sel**)), and (d) adult (26-quarter-old) M (**Mortality** (**Mrt**)).

The **Env** and **Srt** models are combined with components that address the misfit to the composition data for Fishery 2. These include models that (1) ignore the issue (**Fixed** (**Fix**)); (2) estimate the growth curve and the variation of length at age (**Growth** (**Gro**)); (3) estimate a dome-shape selectivity curve for Fishery 2 (**Selectivity** (**Sel**)), and (4) estimate adult(26-quarter-old) M (**Mortality** (**Mrt**)). The rest of the models that use known processes (movement, growth, selectivity, and M) to explain the R shift also use those processes to address the misfit to the longline composition data.

Two other models were investigated but were subsequently eliminated for use in management advice based on diagnostics. The first (**Catch curve** (**Ind**)) assumed that the index of abundance was not representative of the stock and was therefore not used in the assessment; the second (**Down-weighted** (**Cmp**)) assumed that the longline length-composition data were not representative, and therefore are down-weighted by a factor of 10 in the model. Neither was run in combination with the components that addressed the R shift and the misfit to the composition data.

This benchmark assessment for bigeye is based on fourteen reference models (Table 2), of which twelve models are selected based on model diagnostics. Each model was run with four different values of h (0.7, 0.8, 0.9, and 1.0). However, to reduce the number of analyses and subsequent figures, most results and diagnostics in this report are presented for h = 1.0 only.

The naming convention uses the name of the model used to address the *R* shift, followed by the name of the model used to address the misfit to the composition data. For example, the model that uses the short-term model to address the *R* shift and the growth curve to address the misfit to the composition data is **Short-Growth (Srt-Gro)**, while the model that uses growth to address the *R* shift is simply **Growth (Gro)**.

All reference models of bigeye have a quarterly time step from the first quarter of 1979/2000 (model-specific) to the last quarter of 2019 (Table 2). They include 40 population age bins from 0 to 39+ quarters and 111 population length bins from 2 to 220+ cm with an interval of 2 cm. Also, they are sex-structured models with sex-specific natural mortality. They are fitted to indices of relative abundance and size compositions (and also age compositions for the three models which estimate growth; see Table 2 for more details) by finding a set of population dynamics and fishing parameter estimates that maximize the penalized log-likelihood, given the amount of catch taken by each fishery. The penalized log-likelihood is the sum of the log-likelihood of catches (no initial equilibrium catches), indices of abundance, length compositions, and recruitment deviates. Observed catches are assumed to be unbiased and relatively precise, following a lognormal error distribution with standard error of 0.01. Many aspects of the underlying assumptions have been described in Section 3, but the reference models also include the following important assumptions, unless noted otherwise:

1. Bigeye tuna are recruited to the discard fishery (Fishery 18) one quarter after hatching, and the discard fishery catches fish of the first few age classes only (fully selected between 1 and 3 quarters of age);
2. In the second time block (1994-2019), the size-based selectivity curve of the longline fisheries associated with Area 2 (Fisheries 2 and 8) is asymptotic;
3. The fisheries that show more than one obvious peak in the aggregated size compositions (Fisheries 1, 16, and 19; see Figure 9) have 5-knot spline selectivity (Table 3);
4. One unassociated fishery with negligible catches (Fishery 22) shares the selectivity with another unassociated fishery (Fishery 19) that has similar aggregated length comps (Figure 9) because the selectivity of this fishery cannot be estimated in the model due to the lack of associated composition data (Figure 5). Accordingly, the size compositions of Fishery 22 do not fit to the assessment model (Table 2). Besides, two unassociated fisheries with negligible catch and no size composition information (Fisheries 21 and 23) also share the selectivity with Fishery 19.

The following parameters were estimated in this stock assessment unless noted elsewhere:
1. Variability of length at age 0.
2. Recruitment in every quarter from the first quarter of 1979/2000 through the fourth quarter of 2019.
3. Virgin recruitment.
4. Catchability coefficients of the two surveys, which are assumed to be different.
5. An additional standard deviation of the early longline index of abundance.
6. Selectivity parameters for fisheries and surveys. In this assessment, the double-normal selectivity option is chosen for all asymptotic and dome-shaped selectivity curves, and the 5-knot spline selectivity option is chosen for all spline selectivity curves. The number of parameters estimated for the asymptotic, dome-shaped, and spline selectivity curves is 2, 4, and 5, respectively (Table 3).
7. Initial population size and age structure. One initial recruitment regime parameter and two initial fishing mortality parameters, one for the combined purse-seine fisheries and one for the combined longline fisheries, are estimated. Also, deviates for the youngest 16 age classes are estimated.

The following parameters are assumed to be known unless noted otherwise:
1. Age-specific maturity curve (Table 3.1 and Figure 3.3 in SAC-01-08a);

2. Selectivity curves for the discard fishery;
3. Individual growth except for variability of length at age 0;
4. *M*.

4.1. Assessment results

4.1.1. Model run convergence

Of the 48 model runs for bigeye, 44 converged with small maximum gradients (mostly smaller than 0.001) and positive definite Hessians (Table 4). The results of the four runs that did not converge or produce positive definite Hessians (Model *Env-Fix* with $h = 0.9$, 0.8, or 0.7 and Model *Srt* with $h = 0.7$) are not shown in this section.

4.1.2. Parameter estimates

The estimated *M,* Richards growth curve, and dome-shaped selectivity of Fishery 2 are shown in Figure 10. The difference in the growth curve is larger for old bigeye (Figure 10a). Models *Gro* and *Env-Gro* both estimate that L_2 is slightly less than 170 cm, which is smaller than the value estimated by Model *Srt-Gro* (184.5 cm) and especially the value fixed in other models (196.3 cm). Also, the standard deviations of length at age 40 quarters that estimated by Models *Gro* and *Env-Gro* (around 13 cm) are much larger than either the value estimated by Model *Srt-Gro* (10 cm) or fixed in other models (8.9 cm).

The difference in the selectivity of Fishery 2 is larger for large sizes (>170 cm; Figure 10b). When assuming asymptotic selectivity, the selectivity at large sizes is estimated to be 1. In comparison, when assuming dome-shaped selectivity, selectivity is estimated to reach the peak (*i.e.*, 1) at around 160 cm before dropping to a final (at L_2) level of less than 0.1 for Models *Env-Sel* and *Sel* and 0.24 for Model *Srt-Sel*. Namely, the dome-shaped selectivity estimated at L_2 is less than 25% of the asymptotic selectivity estimated at L_2.

The models which estimate *M* suggest much larger values than that fixed in other models (Figure 10c). Specifically, the *M* for females older than 25 quarters is fixed at 0.14 while was estimated to be 0.19 by Model *Srt-Mrt*, 0.28 by Model *Env-Mrt*, and 0.33 by Model *Mrt*. The *M* for 12-quarter-old female is fixed at 0.1 while was estimated by Model *Mov* to be 0.25. It is important to note that the large estimated *M* can be equivalently interpreted by those reference models as bigeye moving from the EPO to the CPO (westward across the 150° management boundary).

4.1.3. Recruitment

The time series of estimated annual (Figure 11) and quarterly (Figure A2) recruitment of bigeye from the 44 converged model runs have several important features: (1) recruitment estimates are similar within the range of steepness investigated in this assessment; (2) recruitment estimates before 1993 are more uncertain, as the composition data from floating-object fisheries that catch small- and medium-size bigeye were very limited (Figure 6); (3) a pronounced regime shift in recruitment can be identified in model *Env-Fix*. The regime shift, which coincided with the expansion of the fisheries in association with floating objects, is characterized by low recruitments during the early period (1975-1993) followed by relatively large recruitments during the late period (1994-2019). To quantify the extent of the *R* shift, the ratio of the median recruitment in 1994-2019 to that in 1979-1993 was calculated (Table 5); it was >2 in Models *Env-Fix* and *Ind*, meaning that the median recruitment in the late period is more than twice that in the early period. While the *R* shift is also present in the other reference models, it is notably reduced, with none of the ratios of the two median recruitments being greater than 1.5 (Table 5). Since Model *Ind* did not reduce the *R* shift, it was eliminated for use in management advice.

4.1.4. Spawning biomass

In general, the estimates of spawning biomass and spawning biomass ratio (SBR; the ratio of the spawning

biomass of the current stock to that of the unfished stock) are similar within the range of steepness values investigated in this assessment (Figures 12 and 13). However, the estimates of the two quantities are noticeably more sensible to the assumed steepness than is the estimate of recruitment. The time series of SBR from the 12 reference models show large among-model differences. Model *Env-Fix* estimates the lowest level of terminal SBR (0.075), while the other three *Env-* models estimate much higher levels of terminal SBR (0.170 to 0.291). In comparison, the four Srt- models also estimate relatively low SBR in the terminal year, ranging from 0.094 (*Srt-Fix*) to 0.146 (*Srt-Gro*). The remaining models have the most optimistic SBR among the 12 models, estimating a terminal SBR of 0.279 (*Mrt*) to 0.201 (*Gro*).

The uncertainties associated with both spawning biomass (Figure 12) and SBR (Figure 13) are estimated to be relatively large before 1990, due mainly to the lack of composition data from both purse-seine and longline fisheries (Figure 5). It is apparent in results that allowing for additional model flexibility results in increased estimation uncertainty. The models which estimate growth, selectivity, or natural mortality show larger uncertainty in the estimates of spawning biomass and SBR than do the counterpart models where the parameters of those processes are fixed.

4.1.5.   Fishing mortality (*F*)

There have been important changes in the levels of fishing mortality (*F*) caused by the fisheries that catch bigeye in the EPO. All reference models show that the *F* of bigeye less than 9 quarters old increased substantially from around zero before 1993 to the historically high level in recent years, as a result of the continuous expansion of the purse-seine fishery on FADs (Figure 14). In comparison, the *F* of bigeye more than 12 quarters old remained generally stable since 1993. All models but Model *Mov* suggest that the *F* of bigeye less than 9 quarters old is notably lower than that of bigeye above that age.

Fishing has reduced the spawning biomass of bigeye in the EPO. This conclusion is drawn from the result of a simulation in which the spawning biomass of bigeye in the EPO is projected, in the absence of fishing, over the historical period of the assessment using the time series of estimated recruitment deviates. To compare the impact of different fisheries on the stock, the simulations were run with each gear excluded in turn (see Wang *et al*. (2009) for details of the simulation methodology). The fishery impact plot on which the simulations are based showed that the longline fishery had the greatest impact on the stock before 1997, but with the decrease in longline effort and the expansion of the floating-object fishery, the impact of the purse-seine fishery on the spawning population of bigeye is currently far greater than that of the longline fishery (Figure 15). The discards of small bigeye in the floating-object fishery have a small, but detectable, impact on the depletion of the stock.

The 12 reference models tell very different stories about the relative impact of longline and purse-seine fisheries on the population of bigeye in the EPO. Model *Env-Fix* suggests that longline fisheries dramatically reduced the bigeye population before 1997, while other reference models which cover 1979-2019 suggest that longline fisheries caused only a mild reduction in spawning biomass (Figure 15). Model *Env-Fix* also suggests the greatest impact of purse-seine fisheries on the population: the current population would be more than seven times larger in the absence of purse-seine (predominantly floating-object) fisheries (Figure 15).

4.2. Diagnostics

4.2.1.   Fit to longline indices of abundance

The fit to longline indices of abundance is similar among reference models (Figure A3). The predicted index does not follow closely to the observed one since 2015. Particularly, the longline fishery observed large interannual variations in population abundance since 2015 but the predicted abundance trend is very flat. This is a major concern of this assessment, which is nevertheless not surprising considering that

the confidence interval of longline indices since 2015 is wide due to the contraction of the Japanese long-line fishery (Figure 3). In other words, longline indices of abundance since 2015 are relatively down-weighted and therefore have less influence on assessment results and fits.

The root-mean-square error (RMSE) of the Pearson residual of the late longline index of abundance is used to evaluate how well the reference models fit the longline indices of abundance (Table 6). The smaller the RMSE, the better the assessment model fits to the longline index. Model *Env* has the worst fit to the late longline index of abundance (RMSE = 1.18). The two medium-term models which estimate growth (Models *Env-Gro* and *Gro*) have the best fit to the late longline index of abundance (RMSE = 1.05). RMSE is >1 for all reference models, indicating that the CVs of the late longline index of abundance are expected to be greater than that currently specified.

4.2.2. Fit to fisheries composition data

To visualize how well each reference model fits to the composition data for fisheries, the 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 from the assessment model. To facilitate the comparison, the empirical selectivity is smoothed using a Lowess smoother before being scaled to have a maximum value of 1. If the assessment model fits a fishery's composition well, the two selectivity curves should follow closely. In general, Model *Env-Fix* fits well to the composition data of most fisheries, with the no-table exception of the longline fishery in Area 2 (Fishery 2), which is assumed in this reference model to be asymptotic since 1994 (Figure A4).

Reference models were also compared with respect to the fit to the aggregated composition data of Fish-ery 2 since 2000 (Figure A5). All models except Models *Env-Gro* and *Gro* suggest that the selectivity of Fishery 2 should be dome-shaped instead of asymptotic. In fact, Models *Env-Gro* and *Gro* suggest that the selectivity of this fishery may not be asymptotic either: it is expected to increase towards the maximum length interval without an asymptotic plateau. Model *Cmp* has the largest discrepancy between the pre-dicted and empirical selectivity curves. Since model *Ind* did not reduce the misfit of these longline com-position data more than Model *Env-Fix*, it was eliminated for use in management advice.

4.2.3. Retrospective analysis

Retrospective analyses are useful for determining how consistent a stock assessment model is from one year to the next (Mohn 1999): inconsistencies can often signal inadequacies in the model. They are usually carried out by progressively eliminating the last year's data from the analysis without changing the method and assumptions, thus showing the effect on the resulting estimated quantities of including more data. Estimates of population attributes and management quantities for the most recent years are often uncertain and biased. As noted in previous assessments, retrospective bias does not necessarily indicate the magnitude and direction of the bias in the current assessment, only that the model may be mis-specified.

In this report, the retrospective analysis was conducted by removing the last year's data five times. In general, most reference models do not show pronounced retrospective patterns in SBR, as the five retro-spective estimates for the last year are distributed either close to, or on both sides of, the current assess-ment's estimates (Figure A6). Among the 12 reference models, Models *Env-Fix*, *Env-Mrt*, and *Env-Sel* have the most obvious retrospective pattern in SBR, which are largest in the estimates for early years, however. These results suggest that the sensitivity of results to the inclusion of new data from the index of abun-dance based on longline CPUE data, which was the main reason that the previous assessment was not considered reliable enough for management advice, has been reduced.

4.2.4. Age-structured production model

The age-structured production model (ASPM) method proposed by Maunder and Piner (2014) is a diagnostic tool to evaluate whether an assessment model is correctly specified. The ASPM is built by fixing all selectivity parameters at the values estimated by the reference model and removing all composition likelihood components from the total model likelihood. The results, particularly spawning biomass, from the ASPM with zero recruitment deviates are then compared with those from the reference assessment model. If the ASPM is not able to mimic indices of abundance, it could be because the stock is recruitment-driven, the reference model is not correctly specified, or indices of abundance are not proportional to population abundance (Carvalho *et al.* 2017, Maunder and Piner 2014).

The SBR for each reference model is compared to those estimated by the corresponding ASPM with and without recruitment deviates being estimated. In terms of SBR, the ASPM that does not estimate recruitment deviates is significantly different from the reference model, while the ASPM that estimates recruitment deviates (ASPM-R) is more similar to the reference model (Figure A6). The comparison underlines the importance of considering recruitment variation to understanding the population dynamics of bigeye in the EPO. For the reference models of which the ASPM-R converge with positive definite Hessian, their estimated 95% confidence interval of the last year's SBR overlaps with that estimated by the corresponding ASPM-R. It suggests that the most recent SBR estimated by those reference models is primarily determined by indices of abundance or the information is similar between the index and the composition, which, as argued by Francis (2011), is a key indicator of good model performance.

4.2.5. Catch-curve analysis

Catch-curve analysis (CCA) is used in the stock assessment as a diagnostic tool to verify whether the temporal trend in composition data is consistent with that in indices of abundance (Carvalho *et al.* 2017). If the two trends are similar, then there is more confidence that the estimated abundance trend is accurate. This method was first introduced by Carvalho *et al.* (2017) to evaluate the relationship between composition data and fishing mortality.

In general, the trends of SBR estimated by the reference model and the corresponding catch-curve analysis are similar during 1990-2010, but very different before 1990 and after 2010 (Figure A7). Since the composition data for both purse-seine and longline fisheries are very limited before 1990 (Figure 5), the quality of composition data is expected to be too low to accurately inform the trend of population abundance in that period. The similar abundance trends during 1990-2010 imply that the composition data and indices of abundance are, to a certain extent, consistent with each other. The increasingly diverging abundance trends after 2010 imply that the selectivity of some fisheries that catch bigeye in the EPO changed, growth changed, or that the index of abundance is no longer representative.

4.2.6. R_0 likelihood profile

Virgin recruitment (R_0), defined as the equilibrium recruitment in the absence of fishing, is a key parameter in the stock-recruitment relationship that scales the absolute abundance. By running the reference model several times with R_0 fixed at a range of values around the maximum likelihood estimate, the profile of model likelihood (*i.e.*, the total negative log-likelihood and its components) against R_0 is referred to as the R_0 likelihood profile (Wang *et al.* 2009). The R_0 likelihood profile is a diagnostic tool widely used to compare the influence of composition data and indices of relative abundance on absolute abundance.

The R_0 likelihood profile suggests that, in most reference models, the composition data of fisheries are more influential on R_0 than are indices of abundance (Figure A8). A comparison of the R_0 likelihood profile for composition data across fleets shows clearly that its shape is primarily determined by contradictory R_0 information from the longline (Fishery 2) and floating-object (Fishery 13) fisheries in Area 2 (not showed).

4.3. Comparison to the previous assessment

The transition from the previous to the current assessment models, following the recommendations of the second external review, is described in detail in document SAC-11 INF-C.

There are substantial differences between the time series of recruitment (Figure 16) and SBR (Figure 17) estimated by this assessment and the previous one (Xu *et al*. 2018). These differences are primarily due to revised assumptions (mainly fishery and survey definitions, data weighting for indices of abundance and length compositions, modeling of catchability and selectivity, and estimating growth or *M* in some models), improved longline catch and effort data (1° x 1° x month resolution with vessel identification), and improved methodology (how to compute survey indices of abundance and length compositions for both surveys and fisheries) following the recommendation of the second external review. The SBR results from the 12 current reference models bracket those from the previous assessment: specifically, half the reference models (*Gro*, *Env-Gro*, *Sel*, *Mov*, *Mrt*, and *Env-Sel*) are more optimistic, and the other half (*Env-Fix*, *Env-Mrt*, *Srt-Fix*, *Srt-Mrt*, *Srt-Gro*, and *Srt-Sel*) more pessimistic, than the previous base case model.

4.4. Summary of results from assessment models

There are several important features in the time series of estimated bigeye recruitment. First, recruitment estimates are not sensitive to the assumed steepness of the stock-recruitment relationship, at least within the range of 0.7-1.0. Second, recruitment estimates are more uncertain before 1993 when the floating-object fisheries had very limited composition data. Third, most reference models, especially Model *Env-Fix*, indicate a regime shift in recruitment: recruitment estimated by those models tends to be greater after 1993 than before. Lastly, most reference models estimate above-average annual recruitments during 2017-2019, especially in 2017, regardless of the assumed steepness. However, these recruitment estimates are very uncertain, because recently recruited bigeye are represented in only a few length-frequency samples.

The time series of estimated SBR from the 12 reference models are similar with respect to interannual variations and long-term trends but differ notably with respect to the absolute scale. Most reference models estimate that the spawning population has generally declined since 1979 and reached one of its lowest levels in 2019 due to fishing. Some reference models suggest that in 2019 the spawning biomass was depleted to less than 10% of the unfished level, while others suggest that it was above 30% of that level. The SBR is more sensitive than recruitment to the assumed value of *h* within the range of 0.7-1.0.

There have been pronounced changes in the amount of fishing mortality caused by the fisheries that catch bigeye in the EPO. On average, the fishing mortality of juvenile bigeye (less than 9 quarters old) has increased substantially since 1993, as a result of the expansion of the purse-seine fishery on floating objects, and reached a historically high level in recent years, while that of adult bigeye (more than 12 quarters old) has increased to a much lesser extent. For 2019, all reference models except Model *Mov* estimate that the fishing mortality of juvenile bigeye is notably higher than that of adult bigeye.

5. STOCK STATUS

The status of the stock of bigeye 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 Antigua Convention.

5.1. 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

5.1.1. Limit reference points

The spawning biomass limit reference point (S_{LIMIT}) is the threshold 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 S that produces 50% of the virgin recruitment if the stock-recruitment relationship follows the Beverton-Holt function with a steepness of 0.75. This spawning biomass is equal to 0.077 of the equilibrium virgin spawning biomass (Maunder and Deriso 2014). 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 this ratio being < 1 (by assuming the probability distribution function for the ratio is normal), are included in the management table.

The fishing mortality limit reference point (F_{LIMIT}) is the threshold of fishing mortality 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 S 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 included in the management table.

5.1.2. Target reference points

The spawning biomass target reference point is the level of spawning biomass that should be achieved and maintained. In 2014 the IATTC adopted S_{MSY} (the spawning biomass that produces the MSY) 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 current dynamic MSY level (S_{MSY_d}) within five years or two generations. Here, S_{MSY_d} is derived by projecting the population into the future under historical recruitment (bias adjusted) and a fishing mortality rate that produces MSY. The current value of S_{MSY_d} used to compute reference points for bigeye is the last quarter's S in the projection period. To provide management advice, $S_{current}/S_{MSY\_d}$, and the probability that this ratio is < 1 (by assuming the probability distribution function for the ratio is normal with a CV equal to that of $F_{current}/F_{MSY}$), are included in the management table.

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 included in the management table, as is the inverse of $F_{current}/F_{MSY}$ (F multiplier).

In the Kobe trajectory plot, the time series of S_{MSY_d} is computed based on two approximations: (1) S_{MSY_d1} = 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 MSY; (2) S_{MSY_d2}, which is derived by projecting the population into the future under historical recruitment (bias adjusted) and $F = F_{MSY}$. The two approximations are weighted as follows to obtain the trajectory of $S_{MSY\_d}(t)$ in the Kobe plot:

$$S_{MSY\_d}(t) = p(t) \, S_{MSY\_d1}(t) + (1 - p(t)) \, S_{MSY\_d2}(t) \quad (Equation\ 1)$$

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

The dynamic MSY (MSY_d; total fishery catches in the last four quarters of the projection) in the management table is also derived from the projection for S_{MSY_d}.

5.2. Estimates of stock status

According to the forty-four converged model runs included in this assessment, the spawning biomass of bigeye at the beginning of 2020 ranged from 14% (Model *Srt-Fix* with a steepness of 0.8) to 212% (Model *Gro* with a steepness of 1.0) of the spawning biomass at dynamic MSY (Figure 18 and Table 7). Twenty-six of the forty-four runs suggest that the spawning biomass of bigeye at the beginning of 2020 is lower than the MSY level. The fishing mortality of bigeye in 2017-2019 ranged from 51% (Model *Gro* with a steepness of 1.0) to 223% (Model *Srt-Fix* with a steepness of 0.8) of the fishing mortality at MSY. Twenty-six of the forty-four runs suggest that the fishing mortality of bigeye in 2017-2019 is higher than the MSY level (Figure 18 and Table 7). These interpretations, however, are subject to large uncertainty, as indicated by the wide confidence intervals around the most recent estimate in the Kobe plot (Figure 18).

According to the forty-four converged model runs included in this assessment, the spawning biomass of bigeye at the beginning of 2020 ranged from 51% (Model *Srt-Fix* with a steepness of 0.8) to 532% (Model *Gro* with a steepness of 0.7) of the spawning biomass at the limit level (Table 7). Five of the 44 runs suggest that the spawning biomass of bigeye at the beginning of 2020 is lower than the limit reference level. The Fishing mortality of bigeye in 2017-2019 ranged from 32% (Model *Gro* with a steepness of 1.0) to 114% (Model *Srt-Fix* with a steepness of 0.8) of the fishing mortality at the limit level. Three of the 44 runs suggest that the fishing mortality of bigeye in 2017-2019 is higher than the limit reference level (Figure 18 and Table 7).

Every reference model included in this assessment suggests that lower steepness values correspond to more pessimistic estimates of stock status: lower *S* and higher *F* relative to the reference points (Figure 19). However, all short-term models and two environment models (Models *Env-Fix* and *Env-Mrt*) estimate that, regardless of what value is assumed for steepness, *S* is below S_{MSY_d} and *F* is above F_{MSY} at the beginning of 2020. Conversely, Models *Gro* and *Sel* estimate the opposite. The stock status at the beginning of 2020 estimated by the remaining models depends on the value assumed for steepness.

The MSY of bigeye in the EPO could be maximized if the age-specific selectivity pattern were similar to that of the longline fisheries, because they catch larger individuals that are close to the critical weight (the weight at which it should ideally be caught to maximize yield per recruit). Before the expansion of the floating-object fishery that began in 1993, the MSY was greater than the current level (Figure 20 top). Also, S_{MSY} was lower before the expansion of the floating-object fishery in comparison to the current level (Figure 20 bottom).

6. FUTURE DIRECTIONS

6.1. Collection of new and updated information

The IATTC staff intends to continue its collection of catch, effort, and size-composition data from the fisheries that catch bigeye tuna in the EPO. Updated and new data will be incorporated into the next stock assessment. In particular, the staff will continue to collaborate with Asian CPCs to compile longline catch, effort, and size-composition data, which are highly influential in the assessment results.

The staff will also continue tagging studies to improve the understanding of the biology of bigeye in the EPO, especially the growth and movement of adult bigeye that significantly impact assessment results, as shown in this report.

6.2. Refinements to the assessment model and methods

The IATTC staff will continue developing the assessment model for bigeye tuna in EPO. Much of the progress will depend on how the Stock Synthesis software is modified in the future and the availability of data. The following changes would be desirable for future assessments:

1. Explore time-varying selectivity for some fisheries;
2. Conduct a Pacific-wide stock assessment;
3. Explore purse-seine indices of abundance;
4. Explore spatial stock assessment models;
5. Explore alternative assumptions on stock structure, especially for fisheries using spline selectivity;
6. Integrate the tagging growth-increment data into the stock assessment model;
7. Explore alternative model initial conditions and model time spans.

ACKNOWLEDGEMENTS

REFERENCES

Aires-da-Silva, A., and Maunder, M. 2010. Sensitivity analysis of bigeye stock assessment to alternative growth assumptions. Document BET-01-03, External review of IATTC bigeye tuna assessment. La Jolla, CA.

Aires-da-Silva, A., Minte-Vera, C., and Maunder, M.N. 2015a. Status of bigeye tuna in the eastern Pacific Ocean in 2015 and outlook for the future. Inter-Amer. Trop. Tuna Comm., Stock Assess. Rep.

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

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.

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.

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., and Maunder, M. 2005. Comparison of Pacific-wide, western and central Pacific, and eastern Pacific assessments of bigeye tuna.

IATTC 2010. The IATTC program for in-port sampling of tuna catches. IATTC Document SAC-01-11.

Kume, S. 1967. Distribution and migration of bigeye tuna in the Pacific Ocean. Rept. of Nankai Reg. Fish. Res. Lab. 25: 75-80.

Langley, A., Hampton, J., Kleiber, P., and Hoyle, S. 2008. Stock assessment of bigeye tuna in the western and central Pacific Ocean, including an analysis of management options. WCPFC SC3 SA WP-1. Port Moresby, Papua New Guinea 11: 22.

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.

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., 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., Maunder, M.N., Tomlinson, P.K., Aires-da-Silva, A., Pérez, A. 2012. Progress report on the development of post-stratified estimators of total catch for the purse-seine fishery port-sampling data. IATTC Document SAC-03-10.

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., and Aires-da-Silva, A. 2010. Investigation of catch-per-unit-of-effort data used in the eastern Pacific Ocean bigeye assessment model. Document BET-01-04, External review of IATTC bigeye tuna assessment. La Jolla, CA.

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., and Piner, K.R. 2014. Contemporary fisheries stock assessment: many issues still remain. ICES Journal of Marine Science 72(1): 7-18.

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. 2020. 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.

Methot, R.D., and Taylor, I.G. 2011. Adjusting for bias due to variability of estimated recruitments in fishery assessment models. Canadian Journal of Fisheries and Aquatic Sciences 68(10): 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.

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.

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., and Itano, D. 2015. Movements, dispersion, and mixing of bigeye tuna (Thunnus obesus) tagged and released in the equatorial Central Pacific Ocean, with conventional and archival tags. Fisheries research 161: 336-355.

Schaefer, K.M. 2006. Reproductive biology of bigeye tuna Thunnus obesus in the eastern and central Pacific Ocean. Inter-Amer. Trop. Tuna Comm. Bull. 23: 3-31.

Schaefer, K.M., and Fuller, D.W. 2006. Estimates of age and growth of bigeye tuna (Thunnus obesus) in the eastern Pacific Ocean based on otolith increments and tagging data. Inter-American Tropical Tuna Commission.

Schnute, J. 1981. A versatile growth model with statistically stable parameters. Canadian Journal of Fisheries and Aquatic Sciences 38(9): 1128-1140.

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.

Valero, J. L., Aires-da-Silva, A. and Maunder, M. N. 2018. Exploratory spatial stock assessment of Bigeye tuna (Thunnus obesus) in the EPO. Inter-Amer. Trop. Tuna Comm., 9th Scient. Adv. Com. Meeting. SAC-09.

Valero, J. L., Maunder, M., Xu, H., Minte-Vera, C. V., Lennert-Cody, C., Aires-da-Silva, A. 2019. Summary of modeling work on evaluating bigeye tuna recruitment shift hypotheses. Inter-Amer. Trop. Tuna Comm., 10th Scient. Adv. Com. Meeting. SAC-10 INF-G.

Valero, J. L., Maunder, M., Xu, H., Minte-Vera, C. V., Lennert-Cody, C., Aires-da-Silva, A. 2019. Spatial stock assessment model options for bigeye tuna (Thunnus obesus) in the EPO and beyond. Review of the stock assessment of bigeye tuna in the eastern Pacific Ocean. La Jolla, California (USA), 11-15 March 2019. WSBET-02-09.

Valero, J. L., Maunder, M., Xu, H., Minte-Vera, C. V., Lennert-Cody, C., Aires-da-Silva, A. 2019. Investigating potential causes of misspecification-induced regime shift in recruitment in the EPO bigeye tuna (Thunnus obesus) assessment. Review of the stock assessment of bigeye tuna in the eastern Pacific Ocean. La Jolla, California (USA), 11-15 March 2019. WSBET-02-08.

Wang, S.-P., Maunder, M.N., Aires-da-Silva, A., and Bayliff, W.H. 2009. Evaluating fishery impacts: application to bigeye tuna (Thunnus obesus) in the eastern Pacific Ocean. Fisheries Research 99(2): 106-111.

Watters, G.M., and Maunder, M.N. 2001. Status of bigeye tuna in the eastern Pacific Ocean. Inter-Amer. Trop. Tuna Comm., Stock Assessment Report 1: 109-211.

Xu, H., Minte-Vera, C.V., Maunder, M.N., and Aires-da-Silva, A. 2018. Status of bigeye tuna in the eastern Pacific Ocean in 2017 and outlook for the future. Inter-Amer.Trop. Tuna Comm., 9th Scient. Adv. Com. Meeting: SAC-09-05.

Xu, H., et al. (2019). "Spatiotemporal dynamics of the dolphin-associated purse-seine fishery for yellowfin tuna (Thunnus albacares) in the eastern Pacific Ocean." Fisheries Research 213: 121-131.

FIGURES
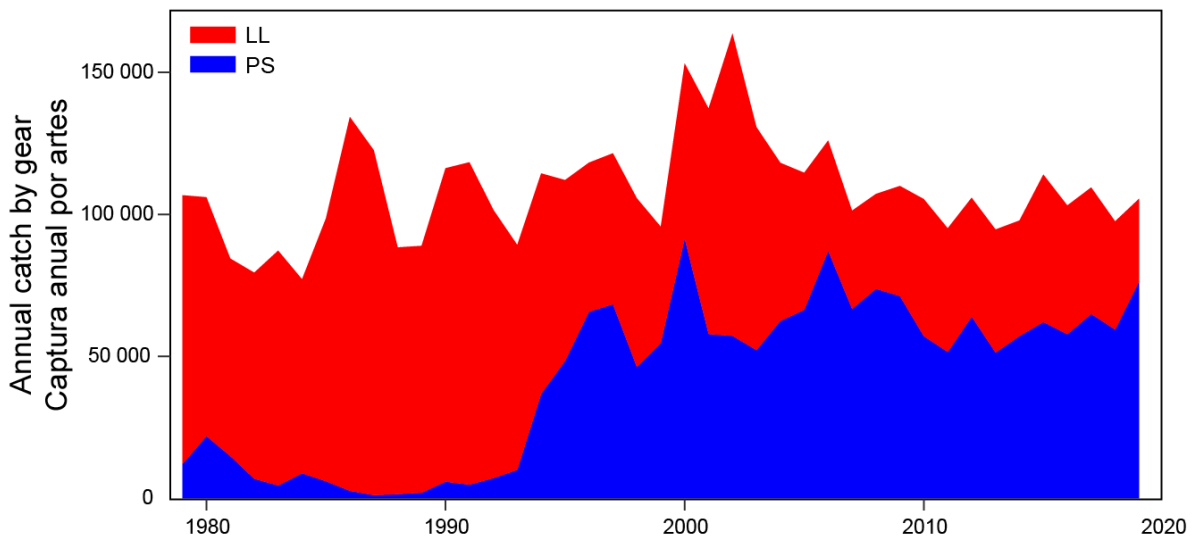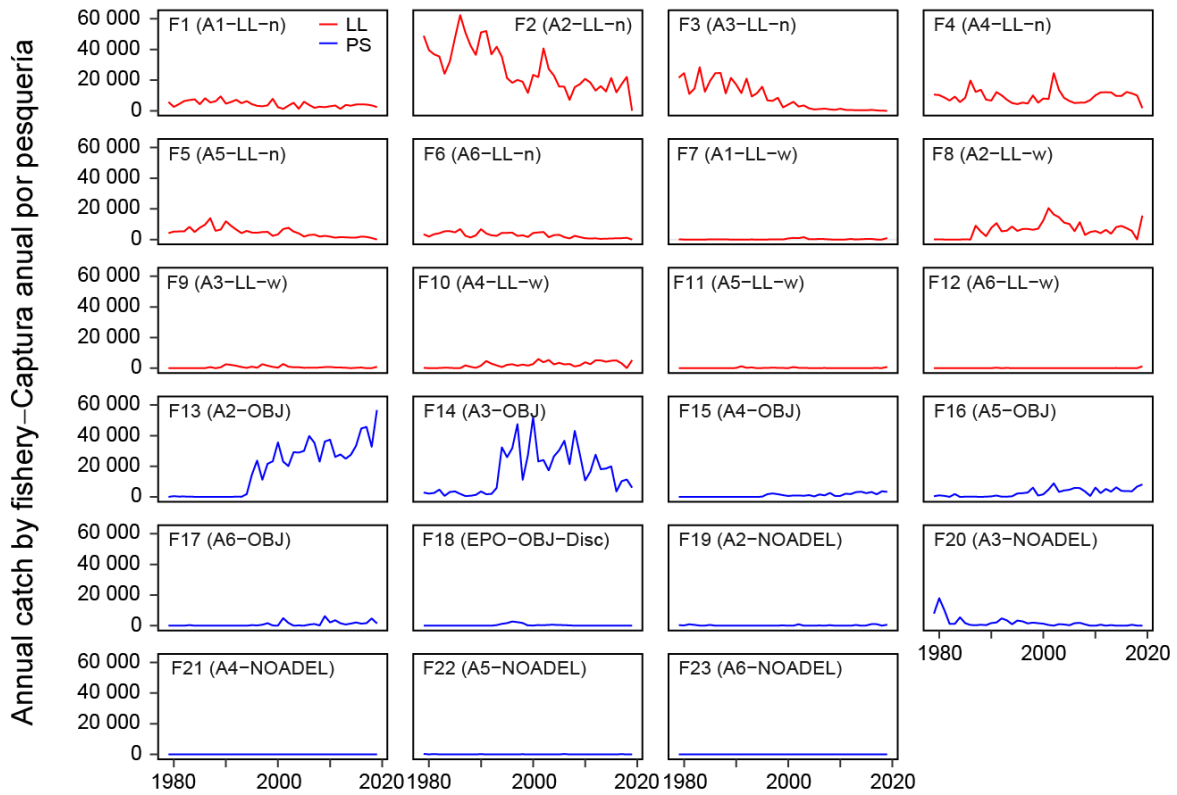
FIGURE 2. Annual catches (metric tons) of bigeye in the eastern Pacific Ocean by fishery (top) and gear (bottom) in 1979-2019. Red: longline; blue: purse seine. To facilitate the comparison, the catches of Fisheries 1-6 are converted by the stock assessment model from number to weight.

FIGURA 2. Capturas anuales (toneladas) de patudo en el Océano Pacífico por pesquería (arriba) y por arte de pesca (abajo) en 1979-2019. Rojo: palangre; azul: cerco. Para facilitar la comparación, las capturas de las pesquerías 1-6 son convertidas por el modelo de evaluación de número a peso.
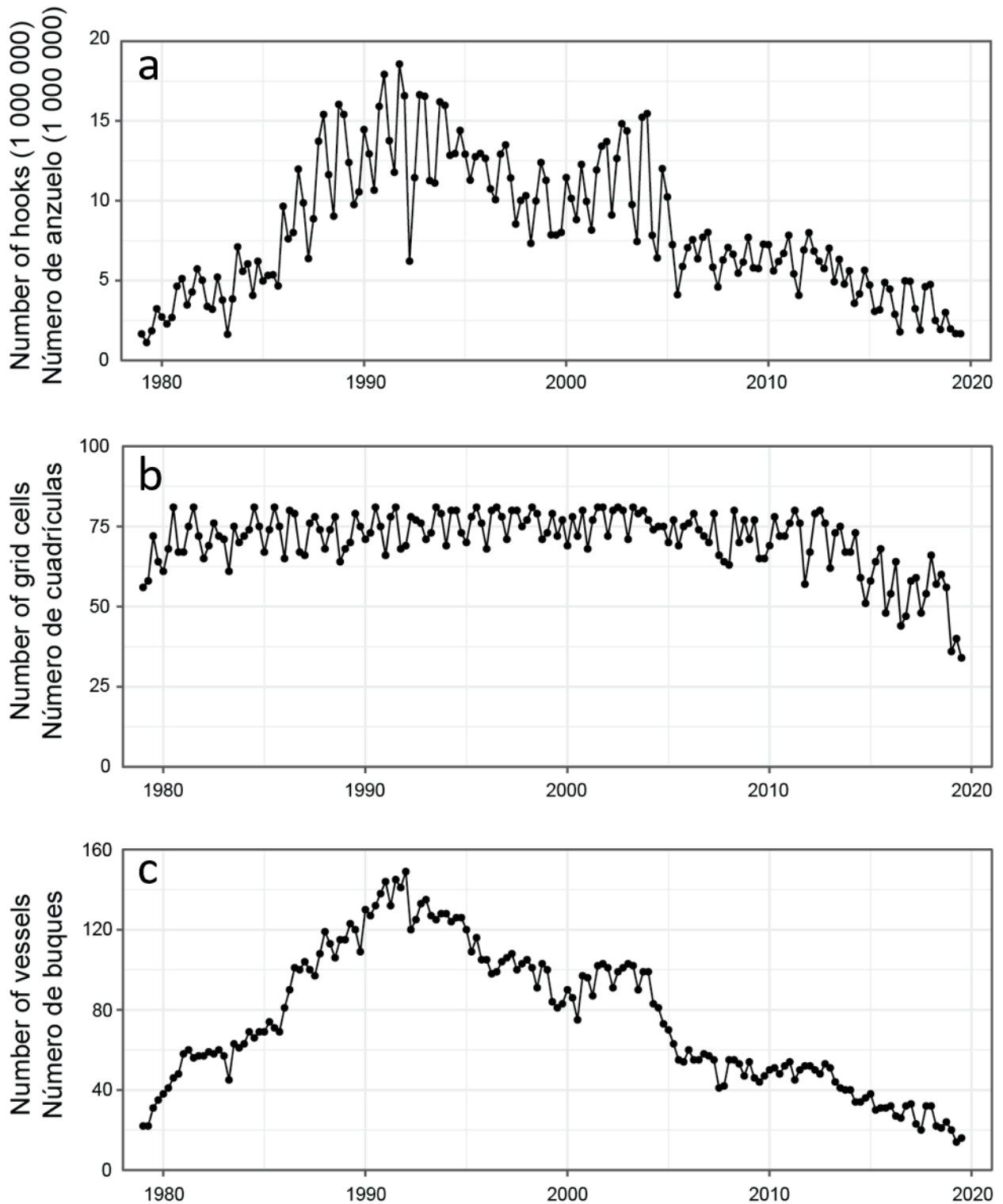
FIGURE 3. Total number of hooks (a), 1°x 1° grid cells (b), and vessels (c) covered by the Japanese longline catch and effort dataset used in the standardization of survey indices of abundance.

FIGURA 3. Número total de anzuelos (a), de cuadrículas de 1° x 1° (b), y buques (c) cubiertos por el conjunto de datos de captura y esfuerzo de palangre de Japón usado en la estandarización de los índices de abundancia de los estudios.
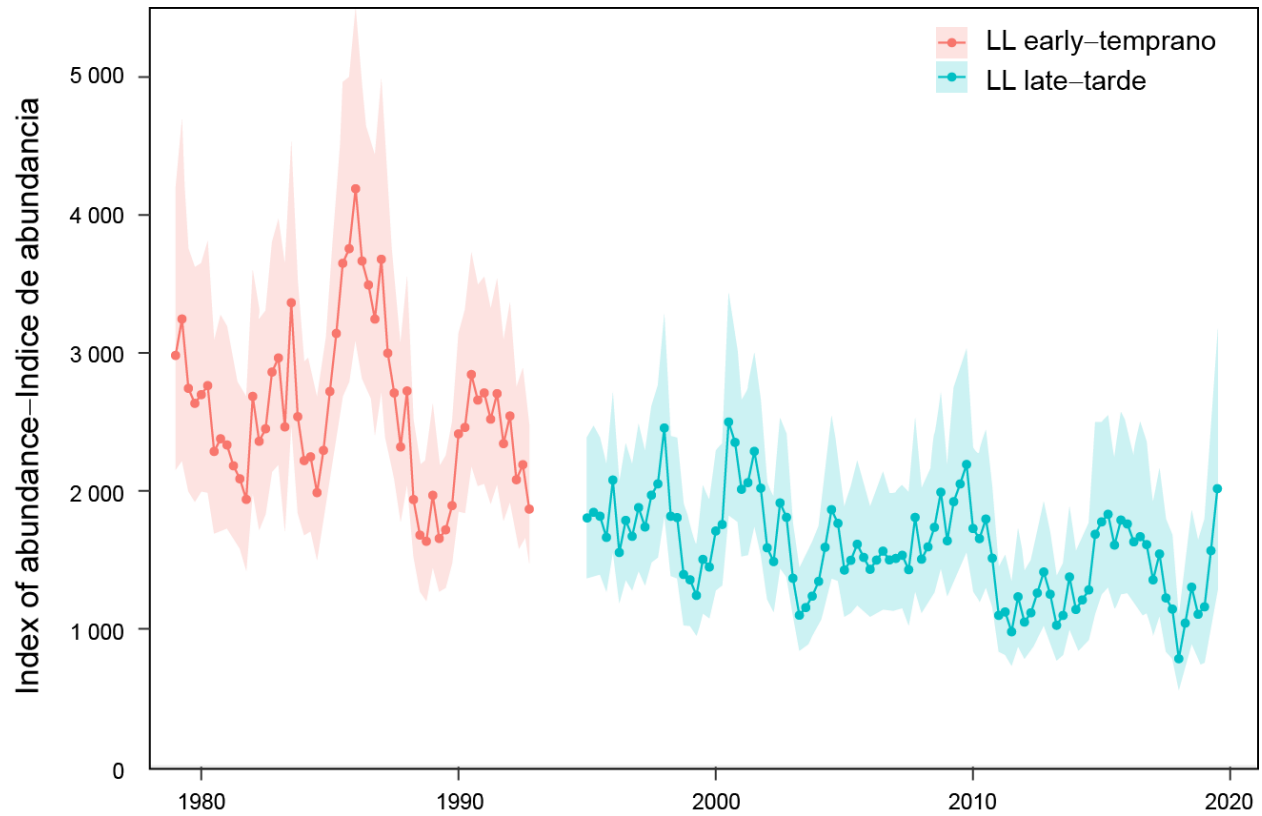
FIGURE 4. The standardized longline indices of abundance and the associated 95% confidence interval used for the stock assessment of bigeye tuna in the eastern Pacific Ocean.

FIGURA 4. Los índices de abundancia de palangre estandarizados y el intervalo de confianza de 95% asociado usados para la evaluación del patudo en el Océano Pacífico oriental.
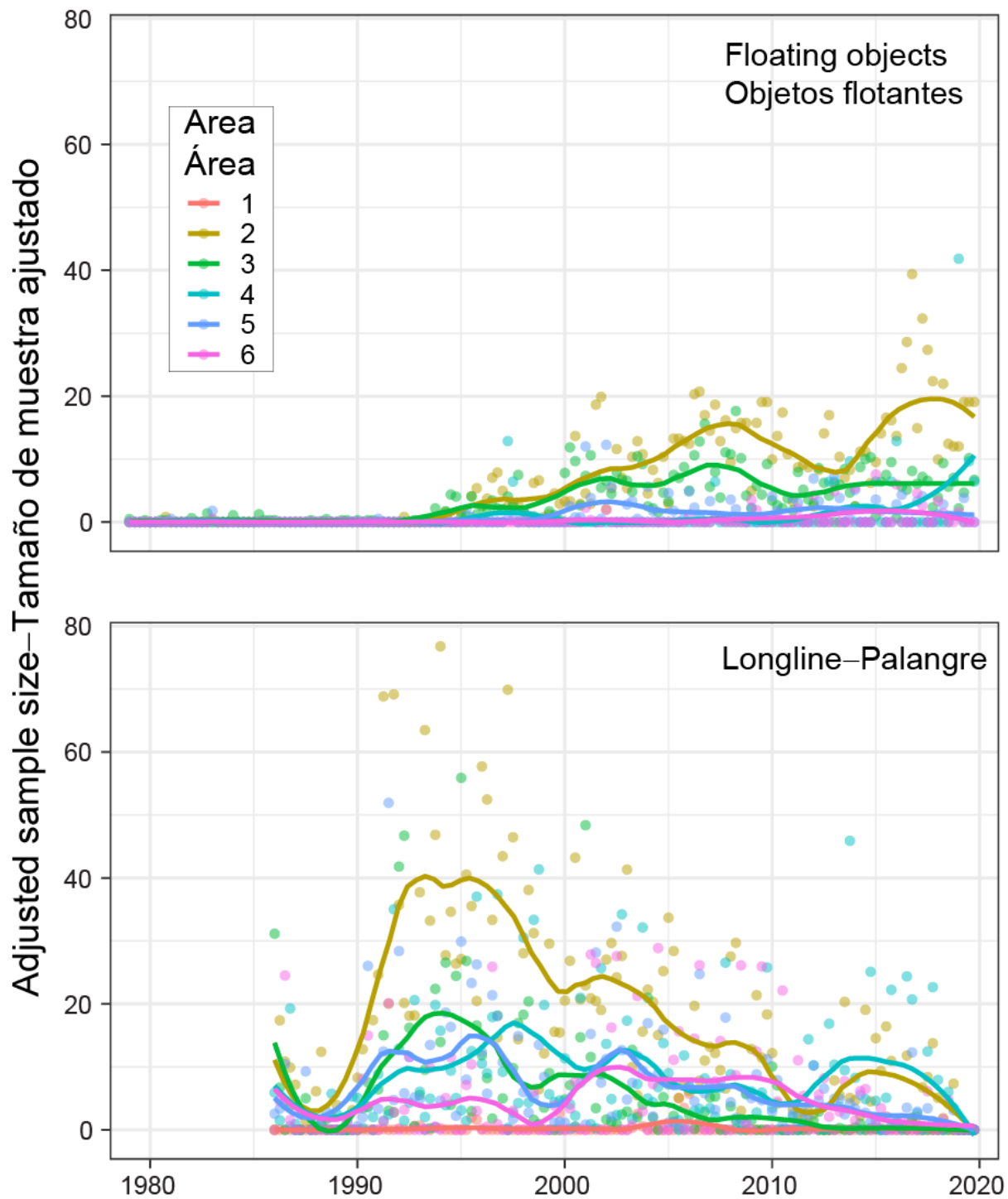
FIGURE 5. The adjusted (based on the Francis weighting method) sample sizes of floating-object (top) and longline (bottom) compositions used in the stock assessment of bigeye tuna in the eastern Pacific Ocean. The color lines are the Lowess smoothed (span=0.25) values. See Figure 1 for map of areas.

FIGURA 5. Los tamaños de muestra ajustados (basados en el método de ponderación de Francis) de las composiciones de objetos flotantes (arriba) y palangre (abajo) usados en la evaluación del patudo en el Océano Pacífico oriental. Las líneas de colores son los valores suavizados con *lowess* (lapso = 0.25). Ver la Figura 1 para consultar el mapa de las áreas.
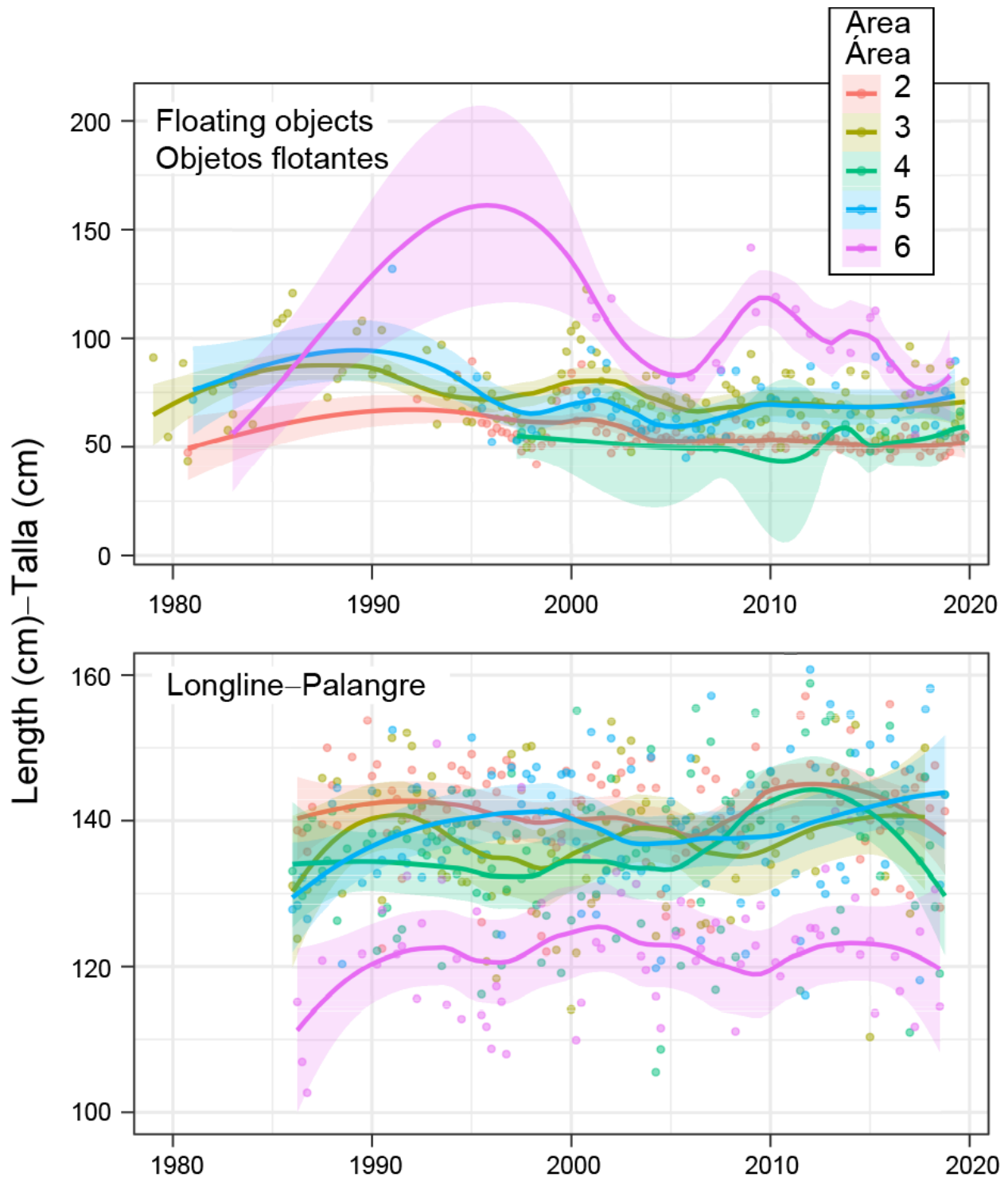
FIGURE 6. The mean length of bigeye caught in purse-seine (top) and longline (bottom) fisheries in the eastern Pacific Ocean. The color lines are the Lowess smoothed (span=0.5) values. See Figure 1 for map of areas.

FIGURA 6. Talla promedio del patudo capturado en las pesquerías de cerco (arriba) y de palangre (abajo) en el Océano Pacífico oriental. Las líneas de colores son los valores suavizados con *lowess* (lapso=0.5). Ver la Figura 1 para consultar el mapa de las áreas.
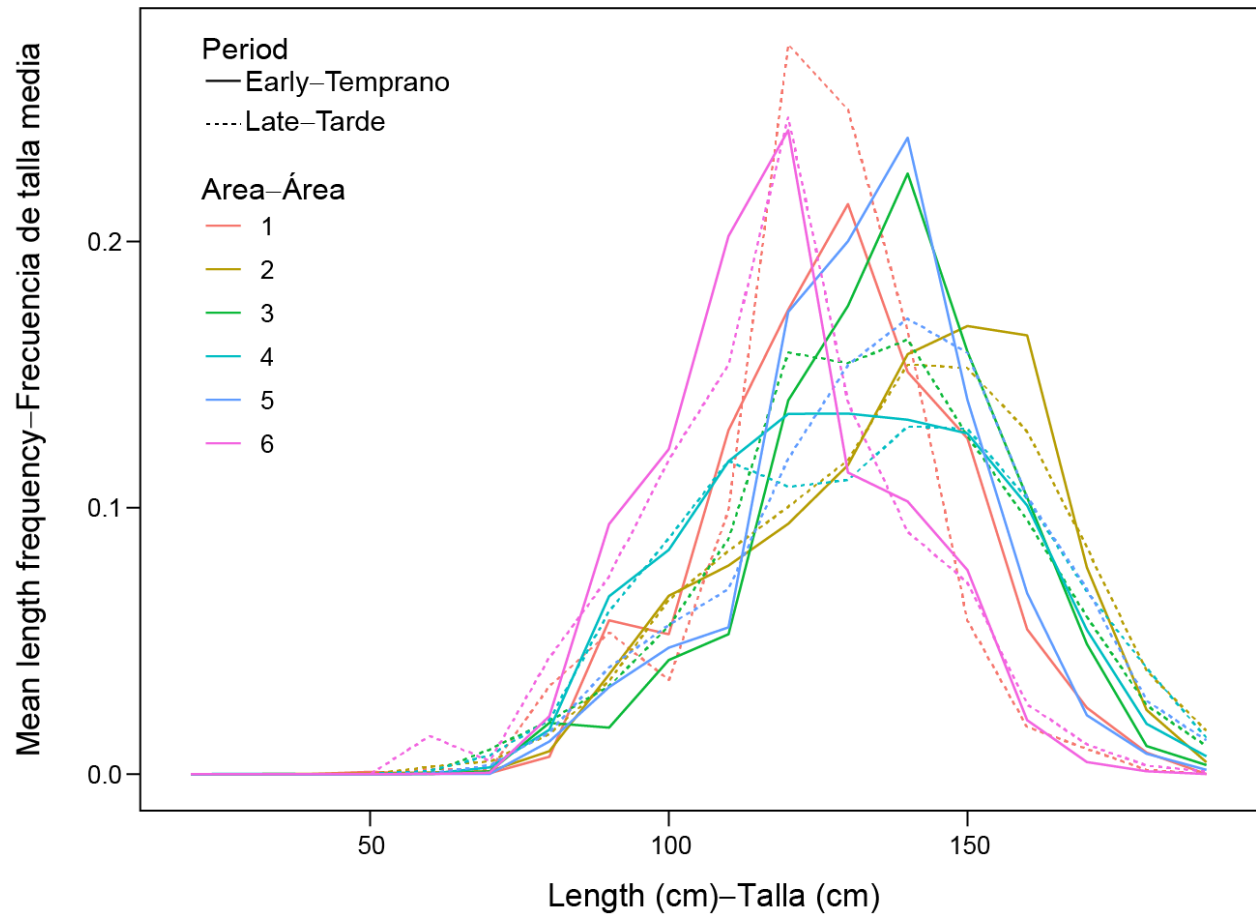
FIGURE 7. Comparison of the mean length compositions of bigeye tuna in longline fisheries by area (defined in Figure 1) and time block (1979-1993 (early) and 1994-2019 (late)).

FIGURA 7. Comparación de las composiciones por talla promedio del patudo en las pesquerías palangreras por área (definidas en la Figura 1) y bloque de tiempo (1979-1993 (temprano) and 1994-2019 (tardío)).
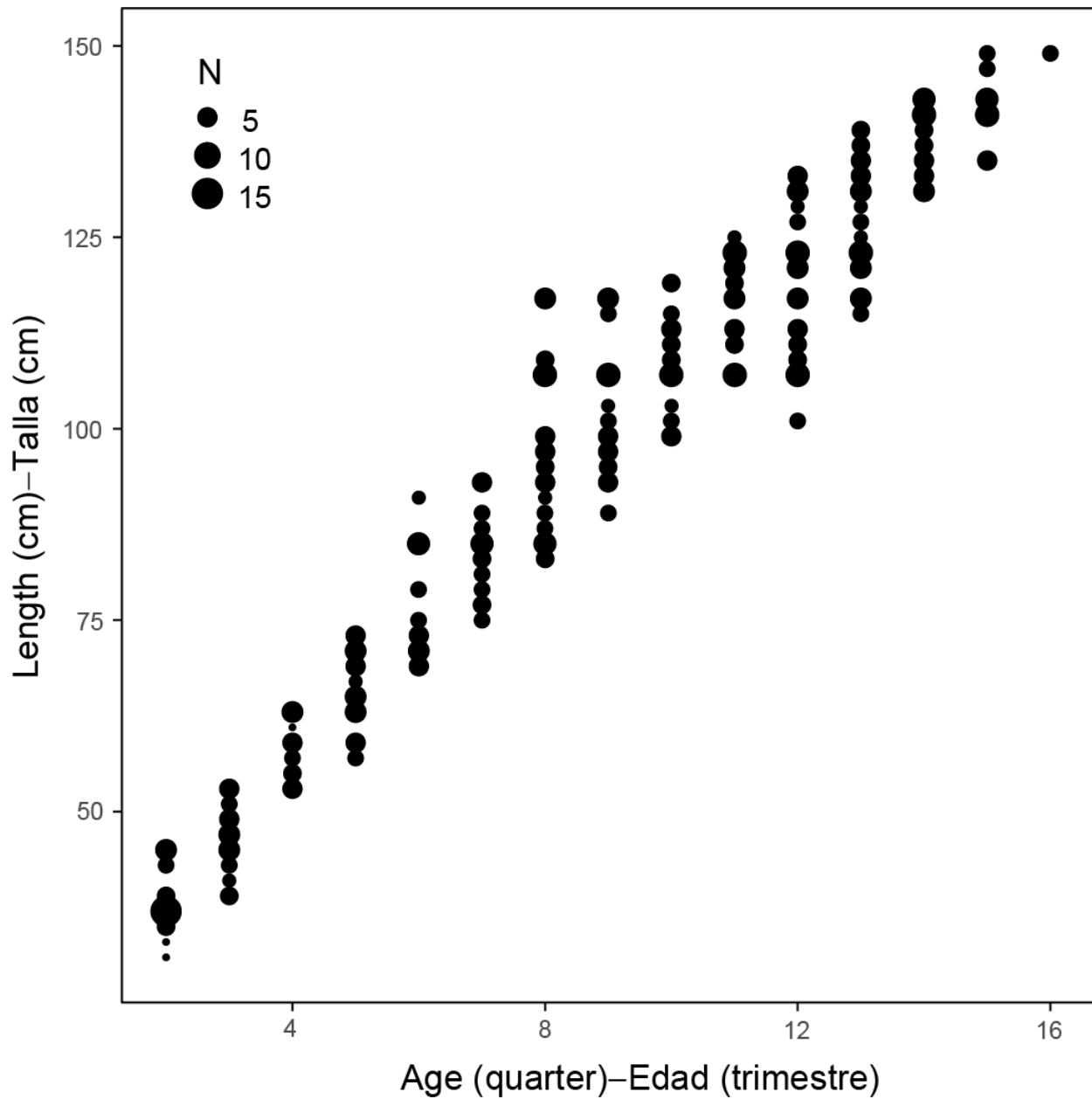
FIGURE 8. Age conditional on length for bigeye tuna in the EPO. The size of black dots represents the number of fish (N) for each age by 10 cm intervals.

FIGURA 8. Edad condicionada a la talla para el patudo en el OPO. El tamaño de los puntos negros representa el número de peces (N) para cada edad a intervalos de 10 cm.
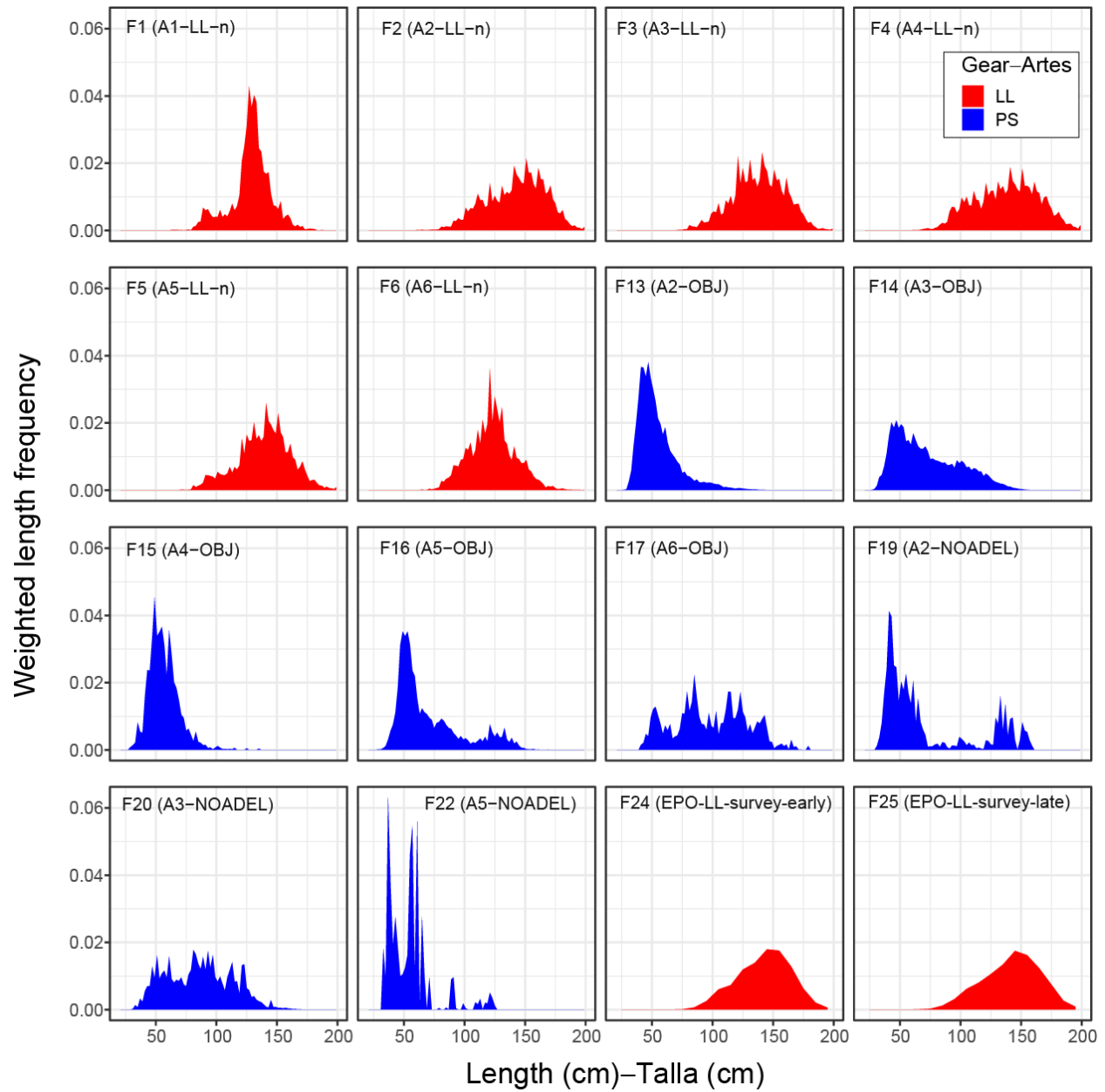
FIGURE 9. Sample-size weighted length frequency of bigeye observed by each fishery and survey. Red: longline; blue: purse seine.

FIGURA 9. Frecuencia de talla ponderada por tamaño de muestra del patudo observada por cada pesquería y estudio. Rojo: palangre; azul: cerco.
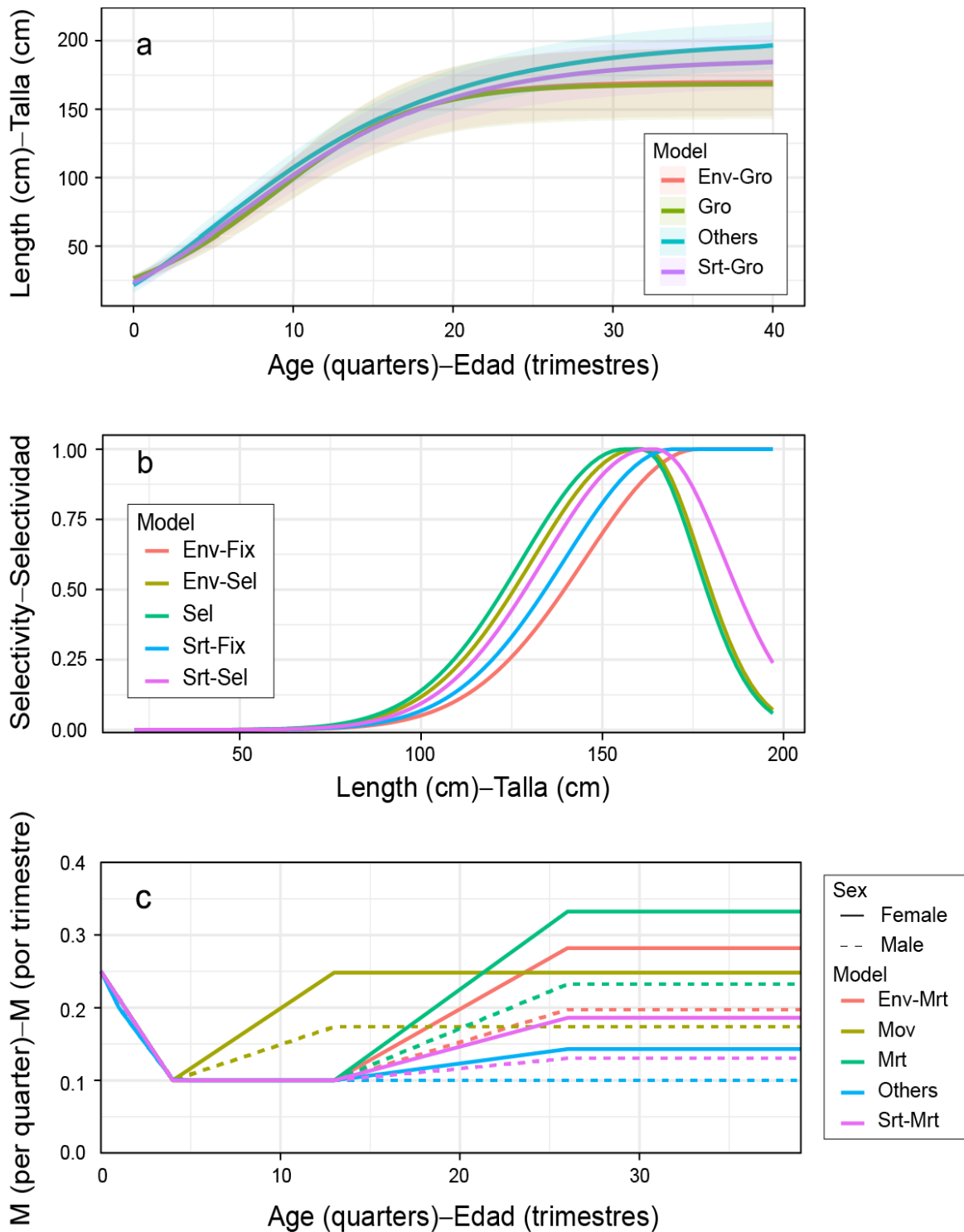
FIGURE 10. Comparison of estimated growth curve and the associated variability (a), selectivity of Fishery 2 in the second time block (b), and sex-specific natural mortality (c) by reference models. The meaning of model name can be found in Table 2.

FIGURA 10. Comparación de la curva de crecimiento estimada y la variabilidad asociada (a), la selectividad de la Pesquería 2 en el segundo bloque de tiempo (b), y la mortalidad natural por sexo (c) por modelo de referencia. En la Tabla 2 se explican los nombres de los modelos.
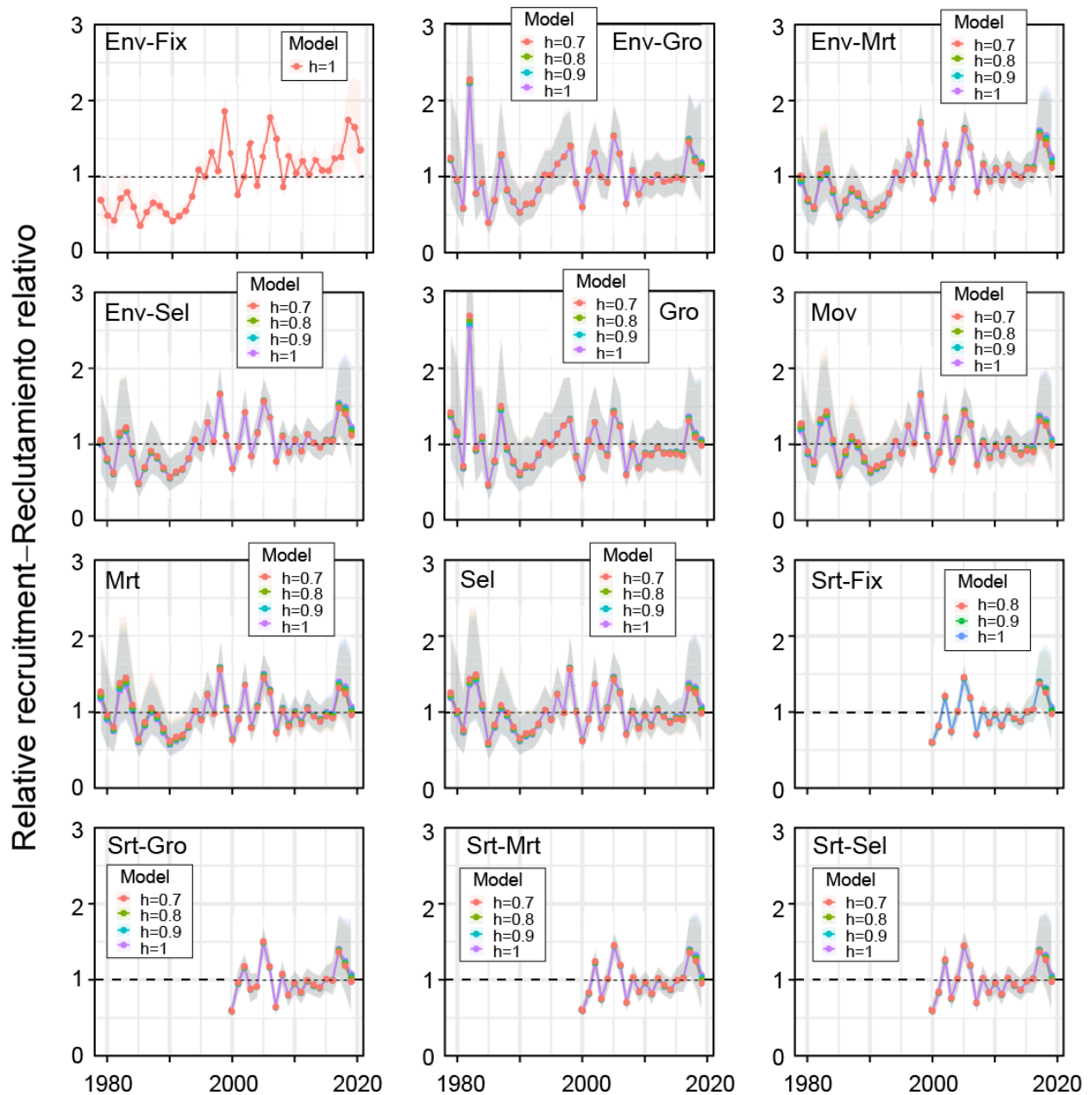
FIGURE 11. Comparison of estimated relative annual recruitment of bigeye tuna in the eastern Pacific Ocean from each reference model under different assumptions on the steepness of the Beverton-Holt stock-recruit relationship (*h*). The shaded areas represent the 95% confidence interval. See Table 2 for explanation of model names.

FIGURA 11. Comparación del reclutamiento anual relativo estimado del patudo en el Océano Pacífico oriental de cada modelo de referencia bajo diferentes supuestos de inclinación de la relación población-reclutamiento de Beverton-Holt (*h*). Las áreas sombreadas representan el intervalo de confianza de 95%. En la Tabla 2 se explican los nombres de los modelos.
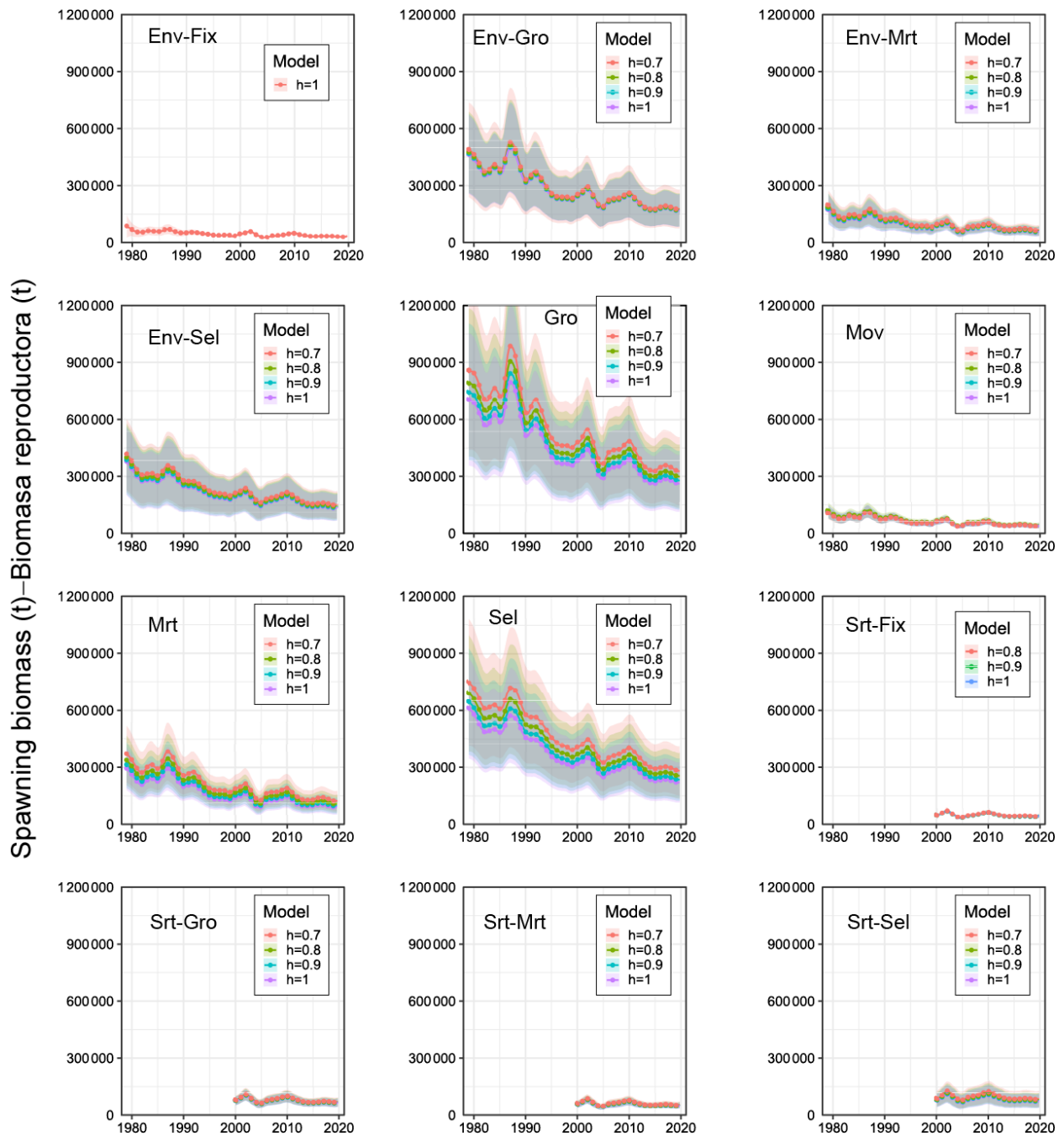
FIGURE 12. Comparison of estimated spawning biomass of bigeye tuna in the eastern Pacific Ocean from each reference model under four different assumptions on the steepness of the Beverton-Holt stock-recruit relationship (*h*). The shaded areas represent the 95% confidence interval. See Table 2 for explanation of model names.

FIGURA 12. Comparación de la biomasa reproductora estimada del patudo en el Océano Pacífico oriental de cada modelo de referencia bajo cuatro diferentes supuestos de inclinación de la relación población-reclutamiento de Beverton-Holt (*h*). Las áreas sombreadas representan el intervalo de confianza de 95%. En la Tabla 2 se explican los nombres de los modelos.
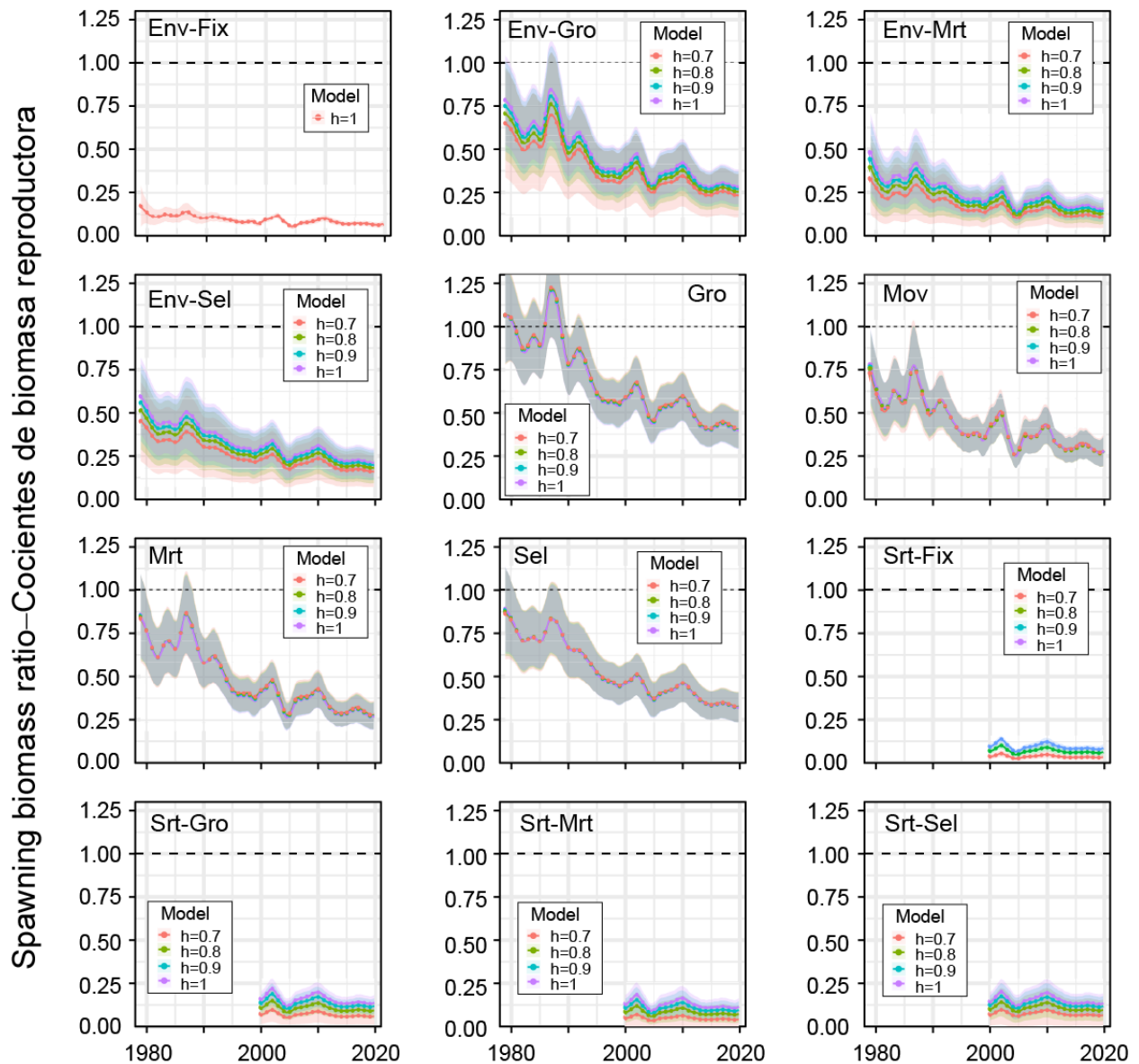
FIGURE 13. Comparison of estimated spawning biomass ratio of bigeye tuna in the eastern Pacific Ocean from each reference model under different assumptions on the steepness of the Beverton-Holt stock-recruit relationship (*h*). The shaded areas represent the 95% confidence interval. See Table 2 for explanation of model names.

FIGURA 13. Comparación del cociente de biomasa reproductora estimado del patudo en el Océano Pacífico oriental de cada modelo de referencia bajo diferentes supuestos de inclinación de la relación población-reclutamiento de Beverton-Holt (*h*). Las áreas sombreadas representan el intervalo de confianza de 95%. En la Tabla 2 se explican los nombres de los modelos.
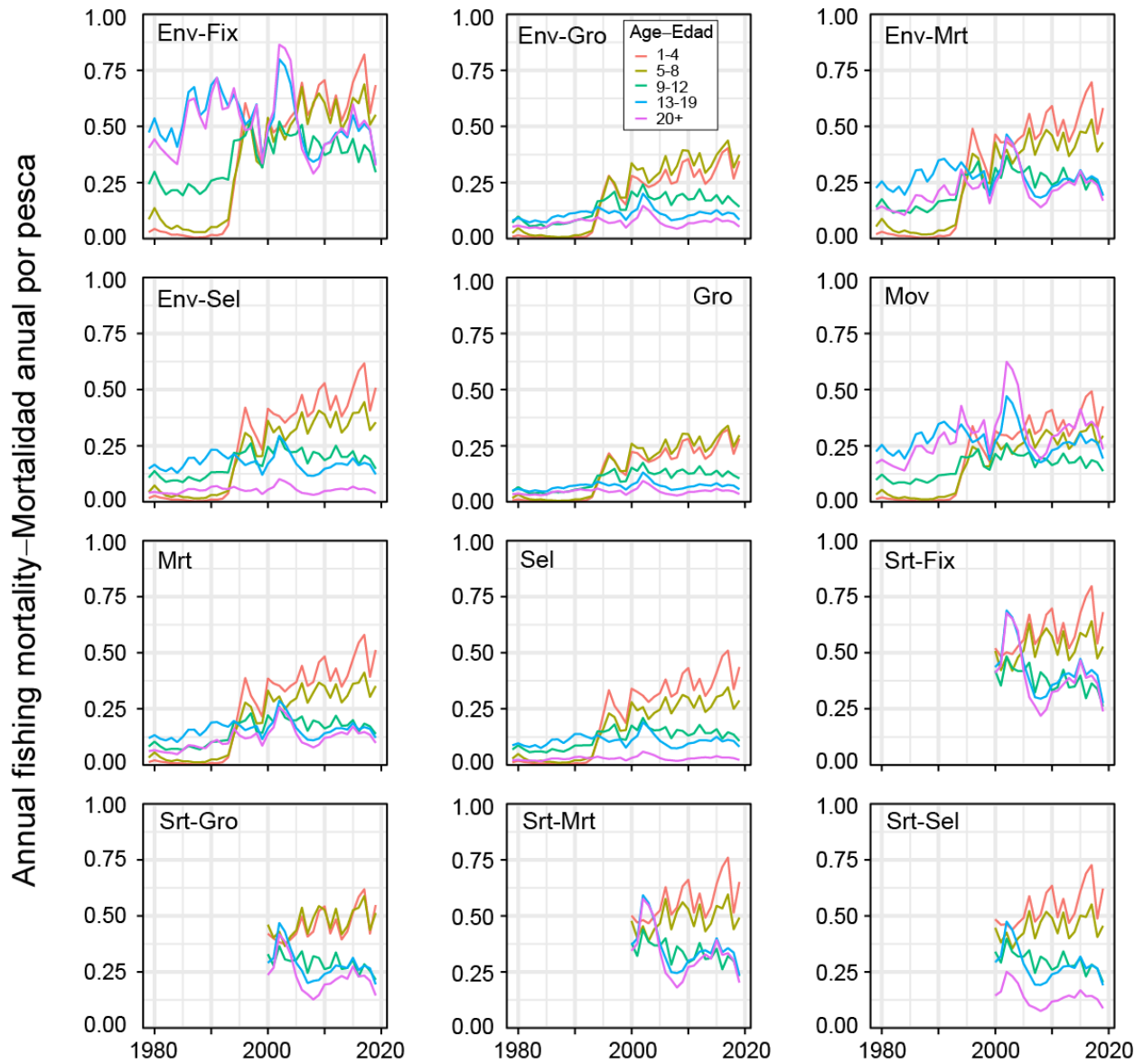
FIGURE 14. Comparison of average annual fishing mortality, by age groups (quarters), of bigeye tuna in the eastern Pacific Ocean. See Table 2 for explanation of model names.

FIGURA 14. Comparación de la mortalidad anual por pesca promedio, por grupo de edad (trimestres), del patudo en el Océano Pacífico oriental. En la Tabla 2 se explican los nombres de los modelos.
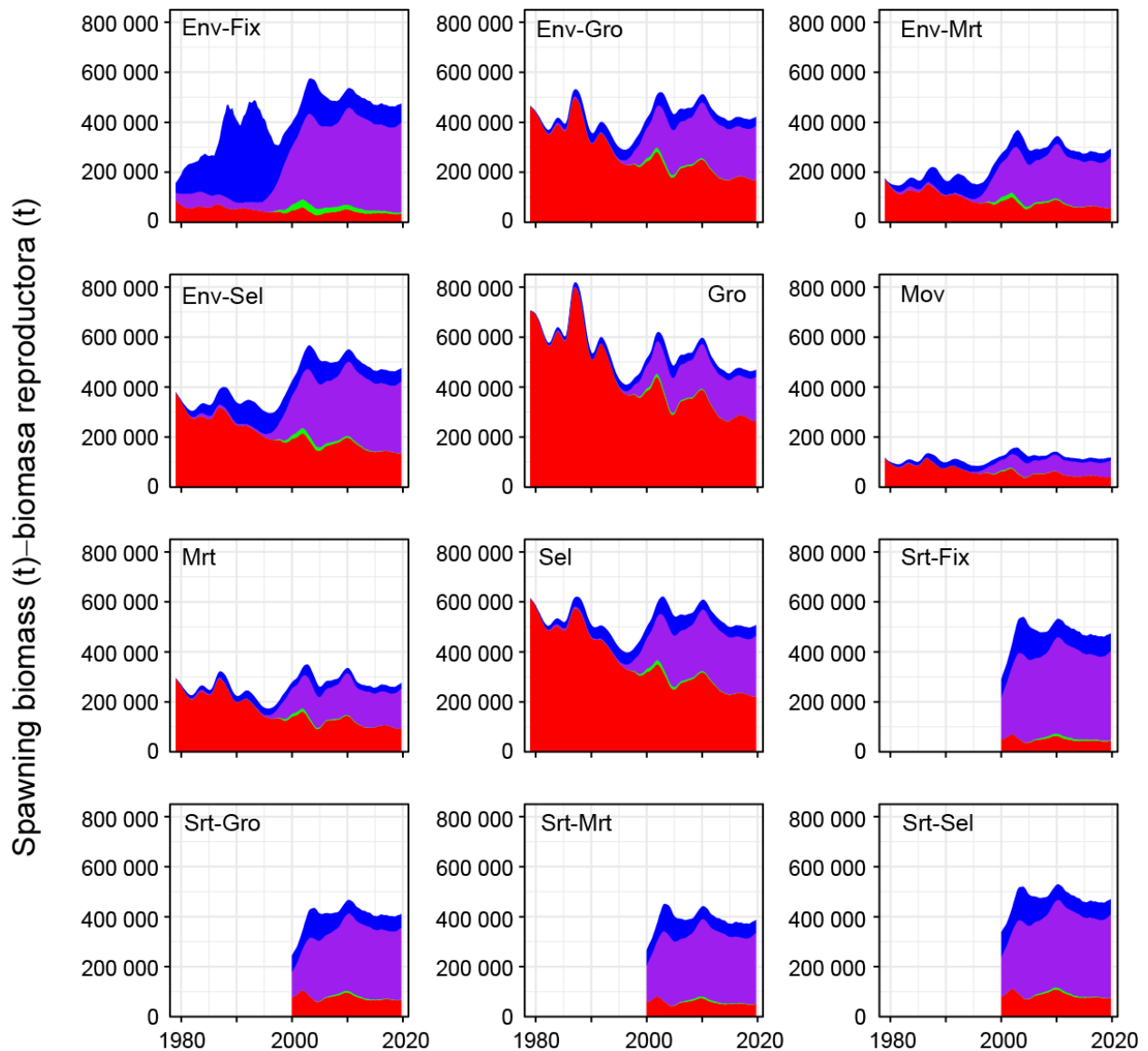
FIGURE 15. Comparison of spawning biomass trajectory of a simulated population of bigeye tuna in the eastern Pacific Ocean that was never exploited (top line) and that predicted by the stock assessment model (bottom line). The shaded green, purple, and blue areas between the two lines show the portions of the impact attributed to the discard fishery, purse-seine fisheries, and longline fisheries, respectively. See Table 2 for explanation of model names.

FIGURA 15. Comparación de la trayectoria de la biomasa reproductora de una población simulada de patudo en el Océano Pacífico oriental que nunca fue explotada (línea superior) y la trayectoria predicha por el modelo de evaluación (línea inferior). Las áreas sombreadas en verde, morado y azul entre las dos líneas muestran las porciones del impacto atribuido a la pesquería de descarte, las pesquerías cerqueras, y las pesquerías palangreras, respectivamente. En la Tabla 2 se explican los nombres de los modelos.
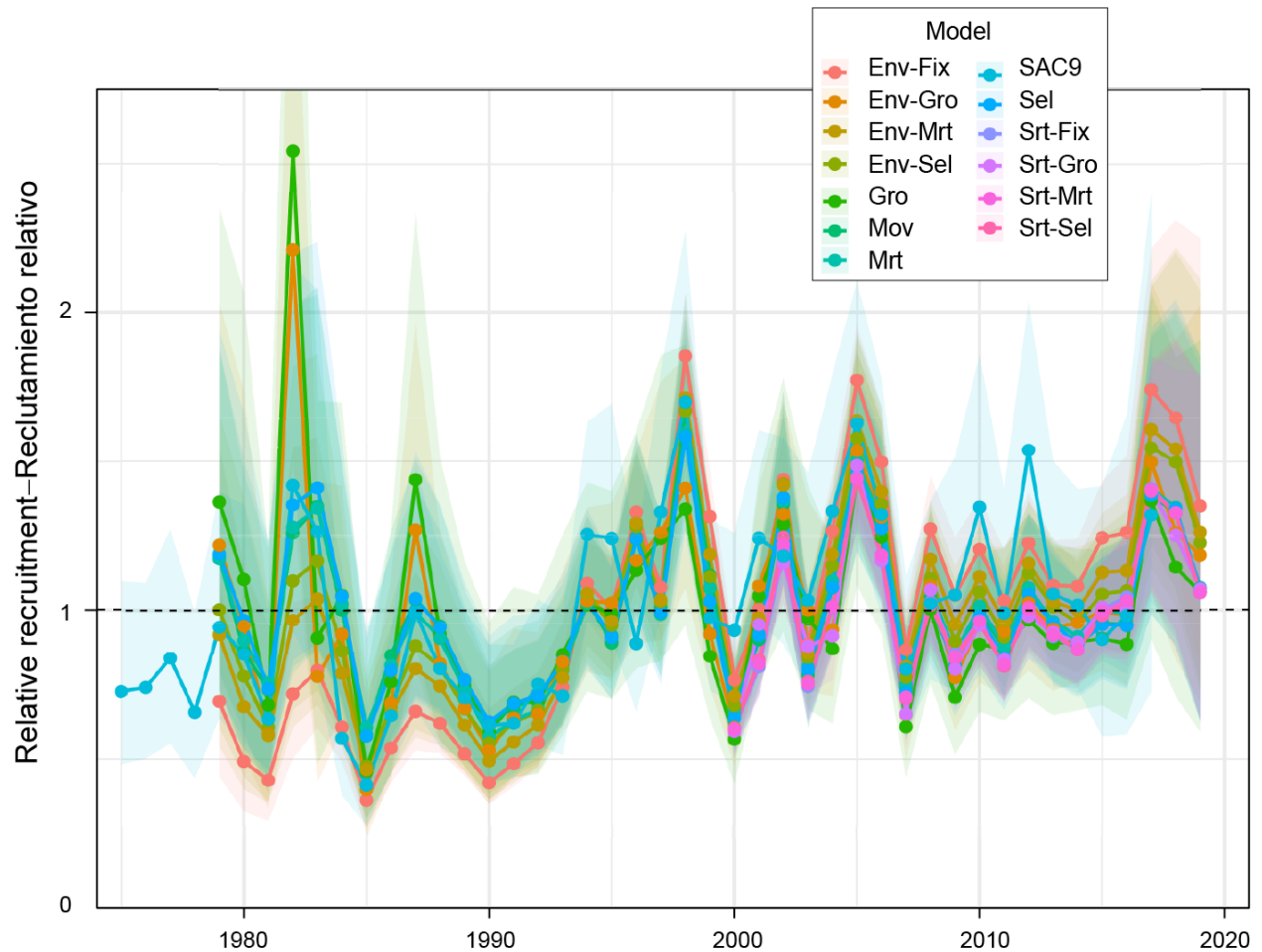
FIGURE 16. Comparison of annual relative recruitment estimates for bigeye tuna in the eastern Pacific Ocean from the twelve SAC-11 reference models (only the estimates that correspond to steepness = 1.0 are shown) and the SAC-9 base-case model. The shaded areas represent the 95% confidence intervals. See Table 2 for explanation of model names.

FIGURA 16. Comparación de las estimaciones de reclutamiento anual relativo del patudo en el Océano Pacífico oriental de los 12 modelos de referencia de SAC-11 (solo se muestran las estimaciones correspondientes a la inclinación = 1.0) y el modelo de caso base de SAC-9. Las áreas sombreadas representan el intervalo de confianza de 95%. En la Tabla 2 se explican los nombres de los modelos.
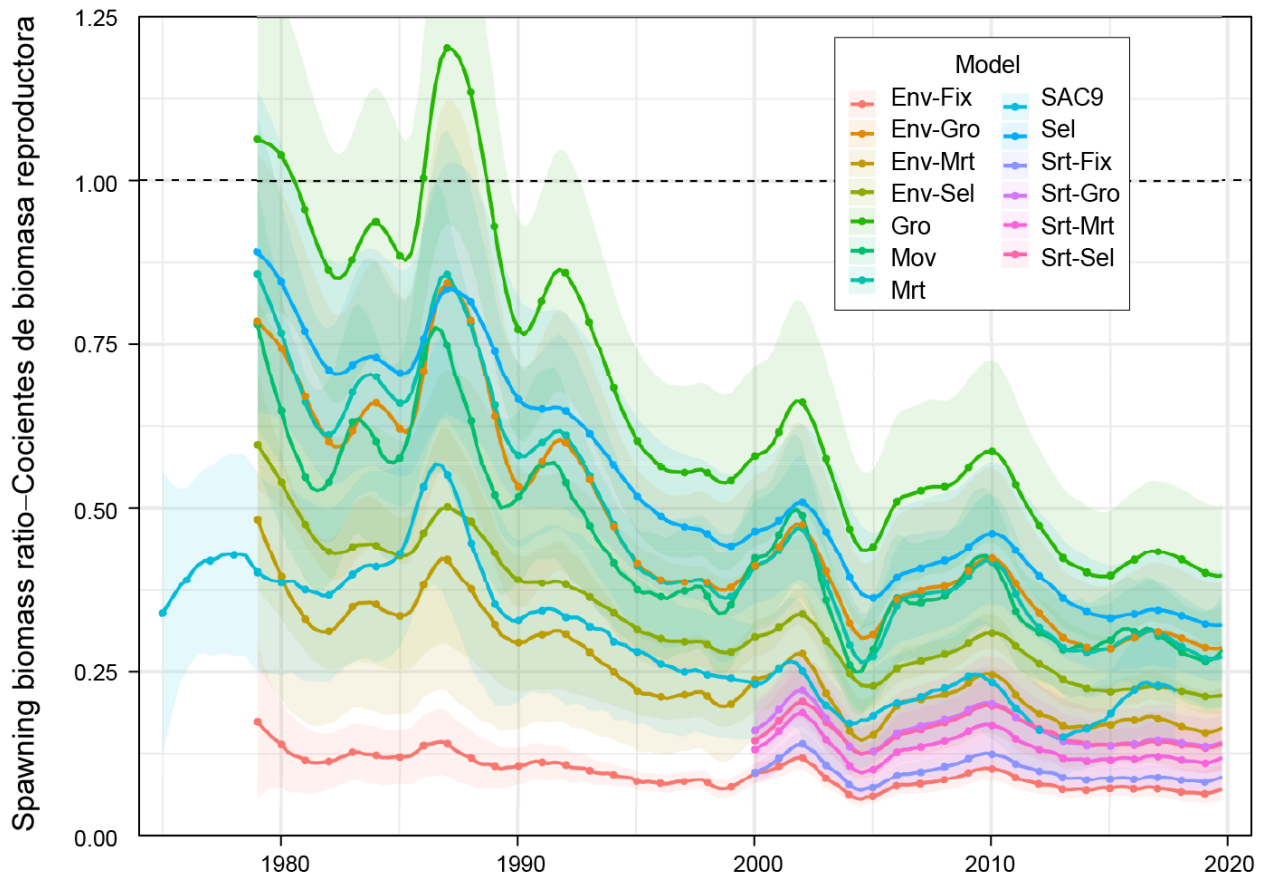
FIGURE 17. Comparison of estimated spawning biomass ratio for bigeye tuna in the eastern Pacific Ocean, from the twelve SAC-11 reference models (only the estimates that correspond to steepness = 1.0 are shown) and the SAC-9 base case model. The shaded areas represent the 95% confidence intervals. See Table 2 for explanation of model names.

FIGURA 17. Comparación del cociente de biomasa reproductora estimado para el patudo en el Océano Pacífico oriental de los 12 modelos de referencia de SAC-11 (solo se muestran las estimaciones correspondientes a la inclinación = 1.0) y el modelo de caso base de SAC-9. Las áreas sombreadas representan los intervalos de confianza de 95%. En la Tabla 2 se explican los nombres de los modelos.
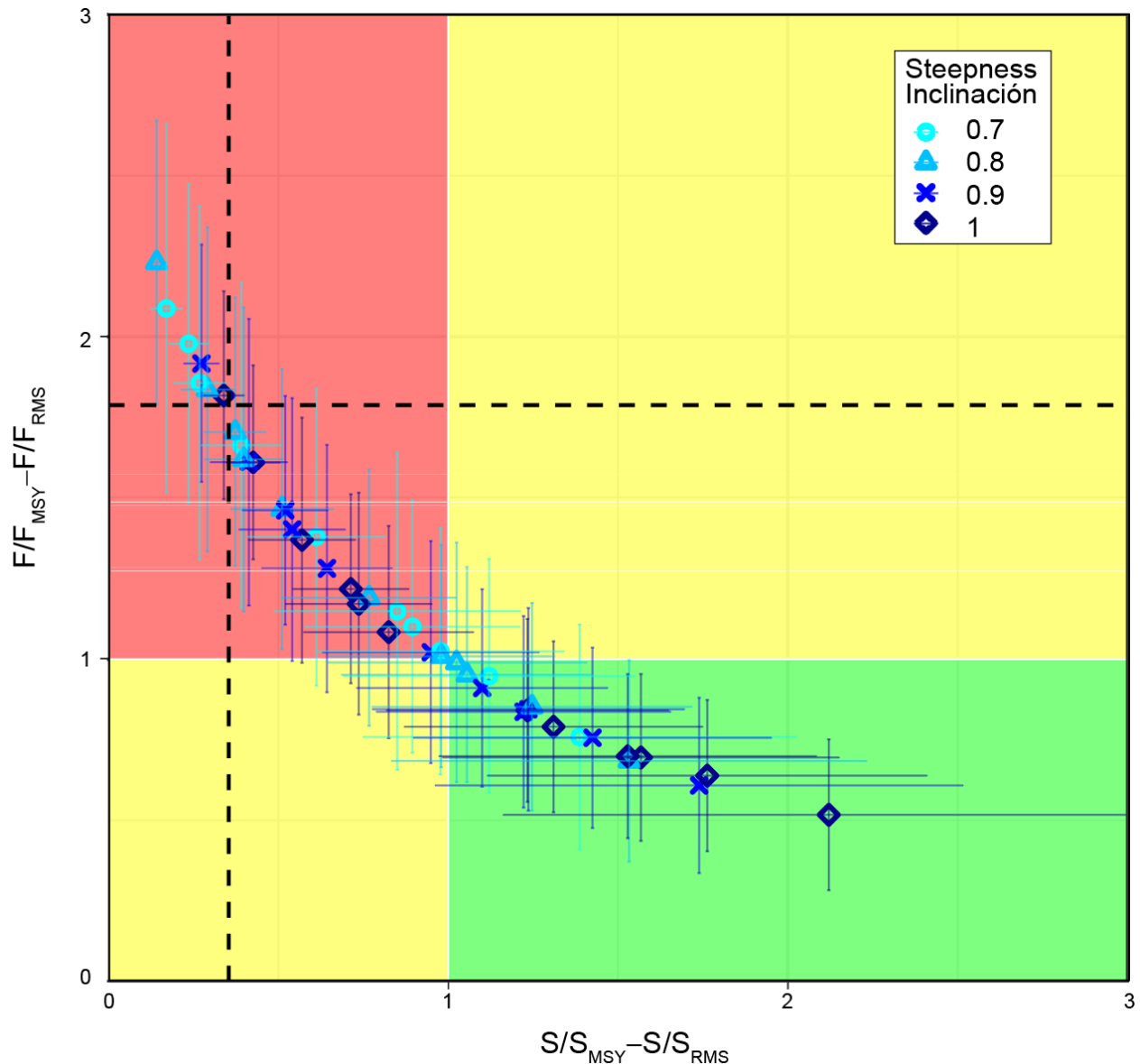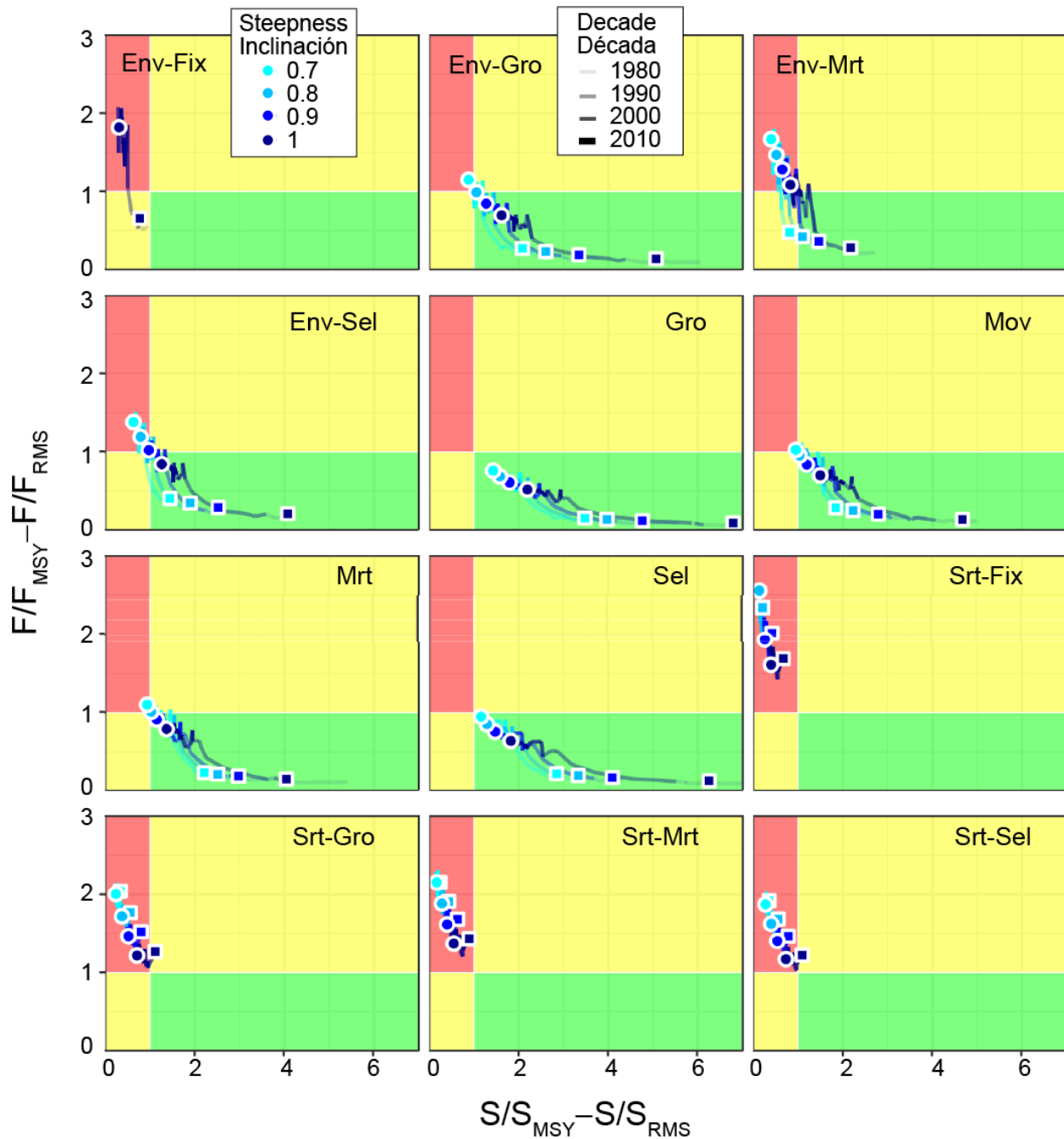
FIGURE 18. Kobe plot of the most recent estimates of spawning biomass (*S*) and fishing mortality (*F*) relative to their MSY reference points (S_{MSY_d} and F_{MSY}) estimated by the 44 converged reference model runs (see Table 4). Each dot is based on the average *F* over the most recent three years. The dashed lines represent the limit reference points averaged for the 44 converged reference model runs. The error bars represent the 95% confidence interval of the estimates.

FIGURA 18. Gráfica de Kobe de las estimaciones más recientes de biomasa reproductora (*S*) y mortalidad por pesca (*F*) con respecto a sus puntos de referencia de RMS (S_{RMS_d} y F_{RMS}) estimados por las 44 ejecuciones convergentes de los modelos de referencia (ver Tabla 4). Cada punto se basa en la *F* promedio de los últimos tres años. Las líneas punteadas representan los puntos de referencia límite promediados para las 44 ejecuciones convergentes de los modelos de referencia. Las barras de error representan el intervalo de confianza de 95% de las estimaciones.

FIGURE 19. 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 colored dots and squares represent the most recent and initial estimates, respectively. See Table 2 for explanation of model names.

FIGURA 19. 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 y cuadrados coloreados representan las estimaciones más recientes y las iniciales, respectivamente. En la Tabla 2 se explican los nombres de los modelos.
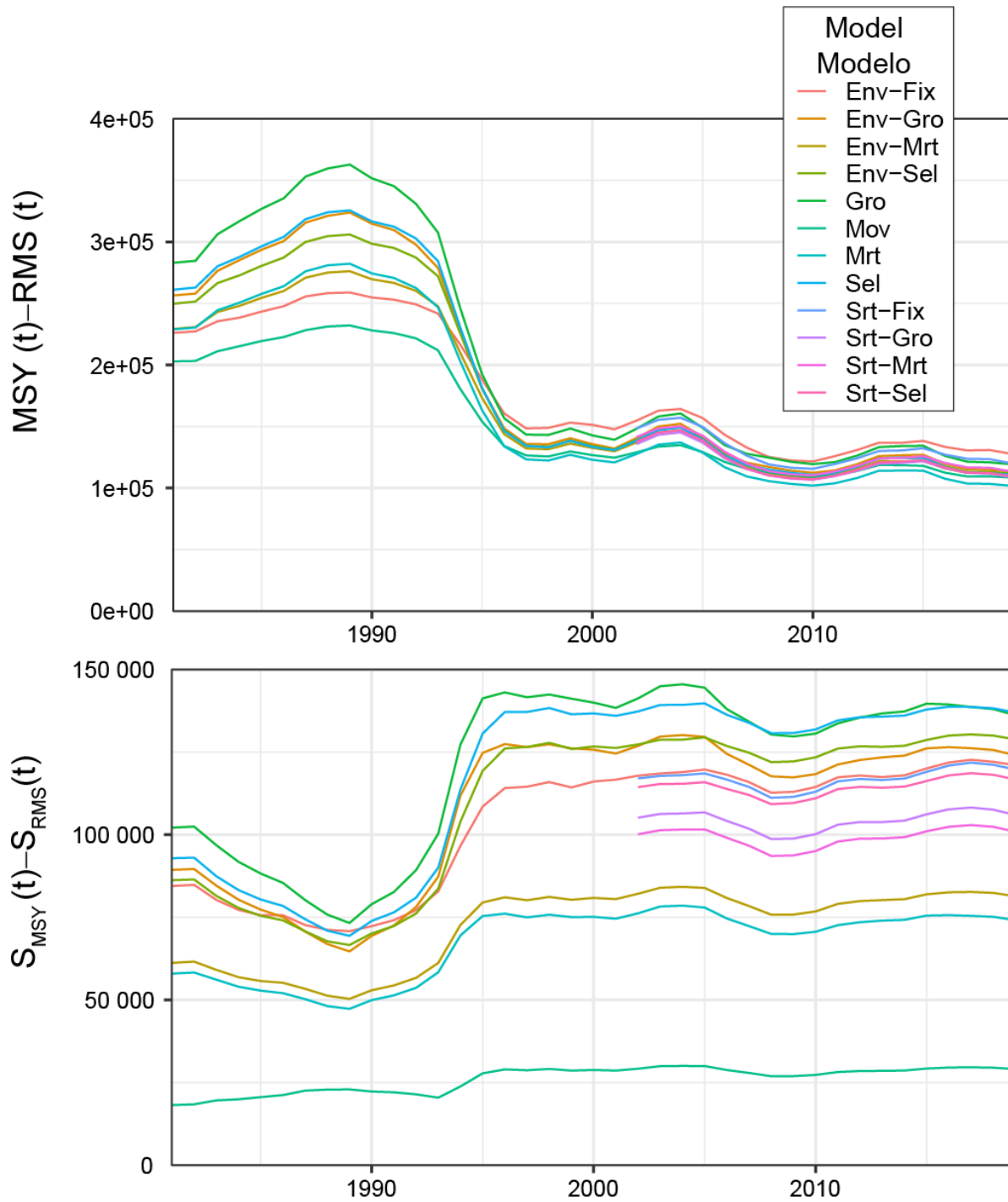
FIGURE 20. Estimates of maximum sustainable yield (MSY) and the associated spawning biomass level (S_{MSY}) using the average age-specific fishing mortality for each year. Only estimates for a steepness assumption of 1.0 are shown. See Table 2 for explanation of model names.

FIGURA 20. Estimaciones del rendimiento máximo sostenible (RMS) y el nivel de biomasa reproductora asociado (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 2 se explican los nombres de los modelos.

TABLES

TABLE 2. The 14 reference models included in the assessment of bigeye tuna in the EPO in 2019. Regime: whether a regime parameter for recruitment is included in the model; Growth: growth parameters estimated in the model; M: natural mortality rate fixed/estimated for some ages; Indices: whether longline indices of abundance are included in the model; LL comps weight: weighting of longline composition data; Selectivity: assumed selectivity curve for Fishery 2 in the second time block (1994-2019); Additional data: additional data fitted in the model. All model input files and output results for this benchmark assessment are available in <u>html and pdf formats</u>.

	Model	Years	Regime	Growth	M	Indices	LL comps weight	Selectivity	Additional data
1.	Env-Fix	1979-2019	Y	std of L(0 quarter)	Fix	Y	Francis	Asymptotic	
2.	Env-Gro		Y	All six parameters	Fix	Y	Francis	Asymptotic	Age-at-length
3.	Env-Mrt		Y	std of L(0 quarter)	Est (26-quarter-old)	Y	Francis	Asymptotic	
4.	Env-Sel		Y	std of L(0 quarter)	Fix	Y	Francis	Dome-shape	
5.	Gro		N	All six parameters	Fix	Y	Francis	Asymptotic	Age-at-length
6.	Mov		N	std of L(0 quarter)	Est (13-quarter-old)	Y	Francis	Asymptotic	
7.	Mrt		N	std of L(0 quarter)	Est (26-quarter-old)	Y	Francis	Asymptotic	
8.	Sel		N	std of L(0 quarter)	Fix	Y	Francis	Dome-shape	
9.	Srt-Fix	2000-2019	N	std of L(0 quarter)	Fix	Y	Francis	Asymptotic	
10.	Srt-Gro		N	All six parameters	Fix	Y	Francis	Asymptotic	Age-at-length
11.	Srt-Mrt		N	std of L(0 quarter)	Est (26-quarter-old)	Y	Francis	Asymptotic	
12.	Srt-Sel		N	std of L(0 quarter)	Fix	Y	Francis	Dome-shape	
13.	Ind	1979-2019	N	std of L(0 quarter)	Fix	N	Francis	Asymptotic	
14.	Cmp		N	std of L(0 quarter)	Fix	Y	Francis/10	Asymptotic	

TABLE 3. Selectivity profiles and weighting of composition data specified for the fisheries and surveys in the assessment.

Fishery	F1	F2	F3	F4	F5	F6
Selectivity	5-knot spline	Dome/asymptotic	Dome	Dome	Dome	Dome
Data weighting	Francis/10	Francis	Francis	Francis	Francis	Francis
Fishery	**F7**	**F8**	**F9**	**F10**	**F11**	**F12**
Selectivity	Mirror F1	Mirror F2	Mirror F3	Mirror F4	Mirror F5	Mirror F6
Data weighting	-	-	-	-	-	-
Fishery	**F13**	**F14**	**F15**	**F16**	**F17**	**F18**
Selectivity	Dome	Dome	Dome	Dome	5-knot spline	Specified ramp
Data weighting	Francis	Francis	Francis	Francis	Francis/2	-
Fishery	**F19**	**F20**	**F21**	**F22**	**F23**	
Selectivity	5-knot spline	Dome	Mirror F19	Mirror F19	Mirror F19	
Data weighting	Francis/2	Francis	-	0	-	
Survey	**S24**	**S25**				
Selectivity	Dome	Dome				
Data weighting	Francis	Francis				

TABLE 4. Maximum gradient of the reference models under different assumptions of steepness (h). See Table 2 for explanation of model names. NA: run does not converge/Hessian is not positive definite.

	1	2	3	4	5	6	7	8	9	10	11	12
Steepness (h)	Env-Fix	Env-Gro	Env-Mrt	Env-Sel	Gro	Mov	Mrt	Sel	Srt-Fix	Srt-Gro	Srt-Mrt	Srt-Sel
1.0	0.0002	0.0001	0.0000	0.0001	0.0000	0.0001	0.0001	0.0010	0.0001	0.0001	0.0001	0.0001
0.9	NA	0.0000	0.0002	0.0001	0.0001	0.0001	0.0001	0.0009	0.0001	0.0003	0.0001	0.0000
0.8	NA	0.0001	0.0000	0.0001	0.0000	0.0001	0.0012	0.0011	0.0056	0.0012	0.0097	0.0000
0.7	NA	0.0021	0.0001	0.0001	0.0001	0.0002	0.0001	0.0002	NA	0.0024	0.0025	0.0000

TABLE 5. Ratio of the median recruitment in 1994-2019 to that in 1979-1993. See Table 2 for explanation of model names. NA: run does not converge/Hessian is not positive definite.

1	2	3	4	5	6	7	8	9	10	11	12	13
Env-Fix	Env-Gro	Env-Mrt	Env-Sel	Gro	Mov	Mrt	Sel	Srt-Fix	Srt-Gro	Srt-Mrt	Srt-Sel	Ind
2.4	1.5	1.8	1.6	1.2	1.3	1.4	1.3	NA	NA	NA	NA	2.1

TABLE 6. Root-mean-square-error of the Pearson residual of the late longline index of abundance. See Table 2 for explanation of model names. NA: run does not converge/Hessian is not positive definite.

1	2	3	4	5	6	7	8	9	10	11	12
Env-Fix	Env-Gro	Env-Mrt	Env-Sel	Gro	Mov	Mrt	Sel	Srt-Fix	Srt-Gro	Srt-Mrt	Srt-Sel
1.18	1.05	1.10	1.07	1.05	1.10	1.08	1.07	1.13	1.07	1.11	1.08

TABLE 7. Management table for bigeye in the EPO. $S_{current}$, and S_{MSY_d}: spawning biomass, in metric tons, at the beginning of 2020 and at dynamic MSY; $F_{current}$ and F_{MSY}: fishing mortality during 2017-2019 and at MSY; S_{LIMIT} and F_{LIMIT}: limit reference points for spawning biomass and fishing mortality; $C_{current}$: total catch of bigeye in 2019, in metric tons; MSY_d: dynamic MSY. See Table 2 for explanation of model names. NA: run does not converge/Hessian is not positive definite.

	1	2	3	4	5	6	7	8	9	10	11	12
	Env-Fix	Env-Gro	Env-Mor	Env-Sel	Gro	Mov	Mrt	Sel	Srt-Fix	Srt-Gro	Srt-Mor	Srt-Sel
h = 1.0												
MSY	127799	113455	110552	111846	119351	108523	101765	110682	120059	108865	113475	109538
MSY_d	129052	121604	116490	119092	133713	121906	115944	124758	124983	114926	119835	117053
$C_{current}$/MSY_d	0.98	0.81	0.92	0.86	0.67	0.78	0.80	0.74	0.95	0.92	0.93	0.91
S_{MSY}/S_0	0.24	0.21	0.23	0.20	0.21	0.20	0.22	0.20	0.24	0.22	0.23	0.21
$S_{current}/S_0$	0.07	0.29	0.17	0.22	0.40	0.30	0.28	0.32	0.09	0.15	0.12	0.14
$S_{current}/S_{LIMIT}$	0.97	3.77	2.21	2.83	5.21	3.85	3.62	4.20	1.22	1.89	1.62	1.87
p($S_{current}<S_{LIMIT}$)	0.59	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.00	0.03	0.01
$F_{current}/F_{LIMIT}$	0.96	0.42	0.62	0.52	0.32	0.43	0.47	0.40	0.87	0.70	0.76	0.70
p($F_{current}>F_{LIMIT}$)	0.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.00	0.01	0.00
$S_{current}/S_{MSY\_d}$	0.34	1.57	0.82	1.23	2.12	1.53	1.31	1.76	0.43	0.71	0.57	0.74
p($S_{current}<S_{MSY\_d}$)	1.00	0.03	0.91	0.14	0.01	0.03	0.08	0.01	1.00	1.00	1.00	0.99
$F_{current}/F_{MSY}$	1.82	0.70	1.08	0.84	0.52	0.70	0.79	0.64	1.61	1.22	1.37	1.17
p($F_{current}>F_{MSY}$)	1.00	0.01	0.69	0.14	0.00	0.01	0.06	0.00	1.00	0.93	0.97	0.84
h = 0.9												
MSY	NA	108924	113270	109872	113442	104656	98568	105546	148278	116160	125985	116490
MSY_d	NA	114717	117480	114770	125898	115780	111485	117756	153133	121070	131548	122610
$C_{current}$/MSY_d	NA	0.87	0.96	0.92	0.71	0.83	0.83	0.78	0.95	0.96	0.95	0.94
S_{MSY}/S_0	NA	0.25	0.27	0.25	0.25	0.25	0.26	0.25	0.27	0.26	0.27	0.26
$S_{current}/S_0$	NA	0.28	0.16	0.20	0.40	0.30	0.28	0.32	0.07	0.13	0.10	0.13
$S_{current}/S_{LIMIT}$	NA	3.60	2.02	2.66	5.24	3.85	3.63	4.21	0.91	1.63	1.36	1.64
p($S_{current}<S_{LIMIT}$)	NA	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.76	0.02	0.14	0.05
$F_{current}/F_{LIMIT}$	NA	0.47	0.70	0.58	0.35	0.47	0.51	0.44	1.00	0.79	0.85	0.78
p($F_{current}>F_{LIMIT}$)	NA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.49	0.02	0.10	0.02
$S_{current}/S_{MSY\_d}$	NA	1.24	0.64	0.95	1.74	1.22	1.10	1.42	0.27	0.52	0.41	0.54
p($S_{current}<S_{MSY\_d}$)	NA	0.16	1.00	0.62	0.03	0.16	0.30	0.06	1.00	1.00	1.00	1.00
$F_{current}/F_{MSY}$	NA	0.84	1.28	1.02	0.61	0.84	0.91	0.76	1.92	1.46	1.61	1.40
p($F_{current}>F_{MSY}$)	NA	0.17	0.92	0.55	0.00	0.14	0.28	0.04	1.00	0.99	1.00	0.97
h = 0.8												

|	1	2	3	4	5	6	7	8	9	10	11	12
	Env-Fix	Env-Gro	Env-Mor	Env-Sel	Gro	Mov	Mrt	Sel	Srt-Fix	Srt-Gro	Srt-Mor	Srt-Sel
MSY	NA	106154	120141	110677	108327	101648	96061	101490	241701	133839	152708	132739
MSY_d	NA	110601	123572	114334	119873	111876	108504	112833	243925	137731	157935	137613
C_{current}/MSY_d	NA	0.91	0.97	0.95	0.74	0.86	0.85	0.81	0.95	0.96	0.96	0.96
S_{MSY}/S_0	NA	0.29	0.30	0.28	0.29	0.28	0.29	0.28	0.30	0.30	0.30	0.29
S_{current}/S_0	NA	0.26	0.14	0.19	0.41	0.29	0.28	0.33	0.04	0.10	0.08	0.10
$S_{\text{current}}/S_{\text{LIMIT}}$	NA	3.38	1.79	2.44	5.27	3.80	3.65	4.23	0.51	1.31	1.06	1.34
p($S_{\text{current}}<S_{\text{LIMIT}}$)	NA	0.00	0.04	0.01	0.00	0.00	0.00	0.00	1.00	0.15	0.43	0.20
$F_{\text{current}}/F_{\text{LIMIT}}$	NA	0.54	0.78	0.66	0.38	0.52	0.55	0.47	1.14	0.90	0.96	0.88
p($F_{\text{current}}>F_{\text{LIMIT}}$)	NA	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.89	0.19	0.37	0.17
$S_{\text{current}}/S_{MSY\_d}$	NA	1.02	0.51	0.77	1.53	1.06	0.98	1.25	0.14	0.37	0.29	0.40
p($S_{\text{current}}<S_{MSY\_d}$)	NA	0.45	1.00	0.96	0.07	0.38	0.55	0.15	1.00	1.00	1.00	1.00
$F_{\text{current}}/F_{\text{MSY}}$	NA	0.99	1.47	1.19	0.68	0.95	1.01	0.85	2.23	1.70	1.84	1.62
p($F_{\text{current}}>F_{\text{MSY}}$)	NA	0.48	0.98	0.83	0.02	0.39	0.52	0.18	1.00	1.00	1.00	1.00
h = 0.7												
MSY	NA	105345	134746	115476	103461	101146	93917	97757	NA	183663	232971	174695
MSY_d	NA	108968	137903	118340	114615	110076	106251	108659	NA	181896	231763	174218
C_{current}/MSY_d	NA	0.94	0.97	0.96	0.77	0.88	0.86	0.84	NA	0.97	0.96	0.97
S_{MSY}/S_0	NA	0.32	0.32	0.31	0.32	0.32	0.32	0.31	NA	0.32	0.33	0.32
S_{current}/S_0	NA	0.24	0.12	0.17	0.41	0.30	0.28	0.33	NA	0.07	0.05	0.07
$S_{\text{current}}/S_{\text{LIMIT}}$	NA	3.10	1.50	2.15	5.32	3.90	3.68	4.24	NA	0.87	0.65	0.95
p($S_{\text{current}}<S_{\text{LIMIT}}$)	NA	0.01	0.14	0.02	0.00	0.00	0.00	0.00	NA	0.66	0.84	0.55
$F_{\text{current}}/F_{\text{LIMIT}}$	NA	0.62	0.88	0.75	0.41	0.54	0.59	0.52	NA	1.04	1.08	0.99
p($F_{\text{current}}>F_{\text{LIMIT}}$)	NA	0.00	0.18	0.02	0.00	0.00	0.00	0.00	NA	0.61	0.71	0.48
$S_{\text{current}}/S_{MSY\_d}$	NA	0.85	0.39	0.61	1.39	0.98	0.89	1.12	NA	0.24	0.17	0.27
p($S_{\text{current}}<S_{MSY\_d}$)	NA	0.79	1.00	1.00	0.12	0.55	0.74	0.29	NA	1.00	1.00	1.00
$F_{\text{current}}/F_{\text{MSY}}$	NA	1.15	1.66	1.38	0.76	1.02	1.10	0.95	NA	1.98	2.09	1.86
p($F_{\text{current}}>F_{\text{MSY}}$)	NA	0.72	1.00	0.95	0.09	0.55	0.69	0.39	NA	1.00	1.00	1.00
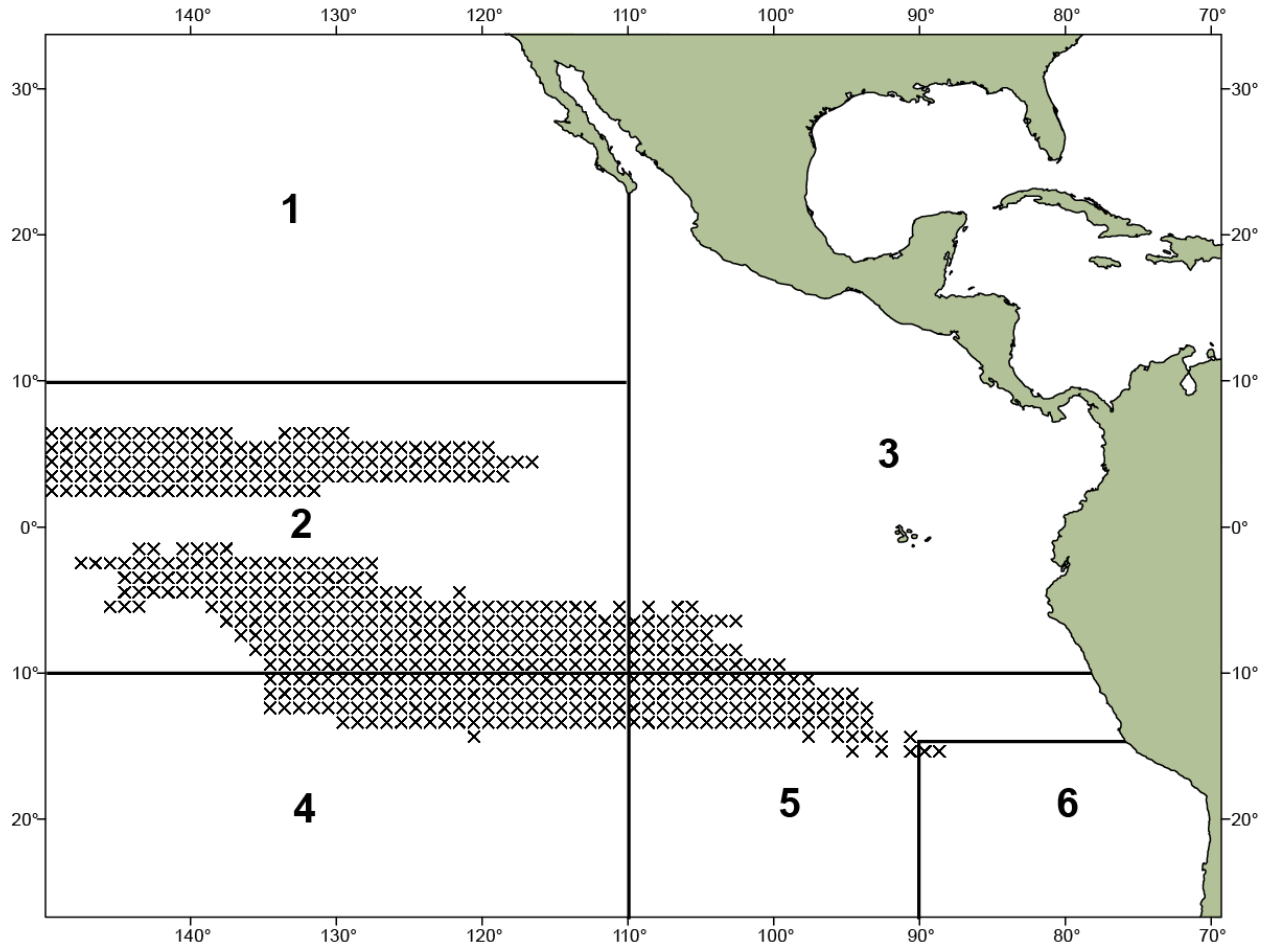
FIGURE A1. The core fishing ground defined for the Japanese commercial longline fleet in the EPO.

FIGURA A1. El caladero central definido para la flota palangrera comercial japonesa en el OPO.
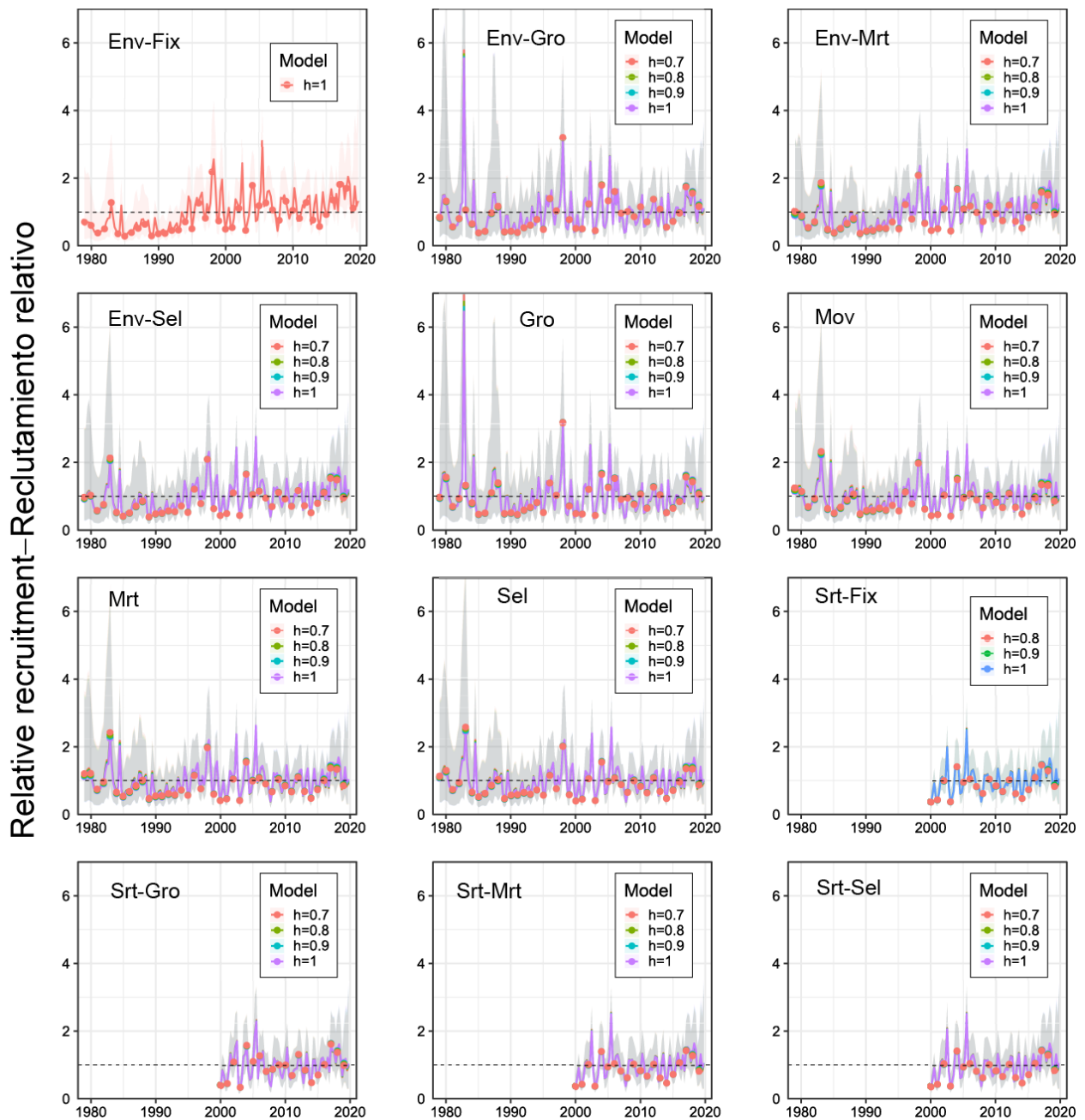
FIGURE A2. Comparison of estimated relative quarterly recruitment of bigeye tuna in the eastern Pacific Ocean from each reference model under different assumptions on the steepness of the Beverton-Holt stock-recruit relationship (*h*). The shaded areas represent the 95% confidence interval. See Table 2 for explanation of model names.

FIGURA A2. Comparación del reclutamiento trimestral relativo estimado del patudo en el Océano Pacífico oriental de cada modelo de referencia bajo diferentes supuestos de inclinación de la relación población-reclutamiento de Beverton-Holt (*h*). Las áreas sombreadas representan el intervalo de confianza de 95%. En la Tabla 2 se explican los nombres de los modelos.
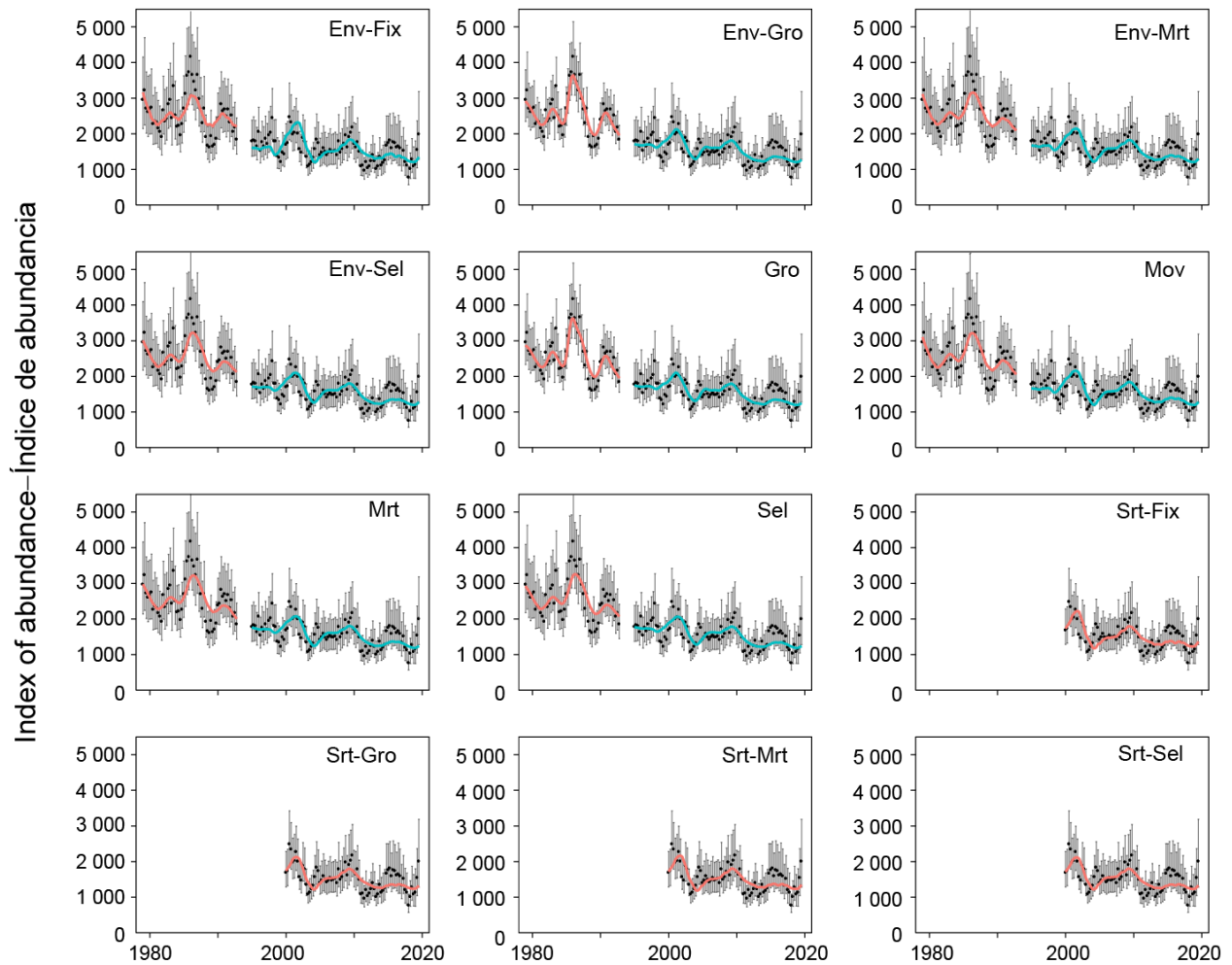
FIGURE A3. Fit of the reference models to longline indices of abundance. The black dots and error bars represent the observed indices and their 95% confidence interval. The solid color lines are the values predicted by the reference model. See Table 2 for explanation of model names.

FIGURA A3. Ajuste de los modelos de referencia a los índices de abundancia de palangre. Los puntos negros y las barras de error representan los índices observados y su intervalo de confianza de 95%. Las líneas sólidas son los valores predichos por el modelo de referencia. En la Tabla 2 se explican los nombres de los modelos.
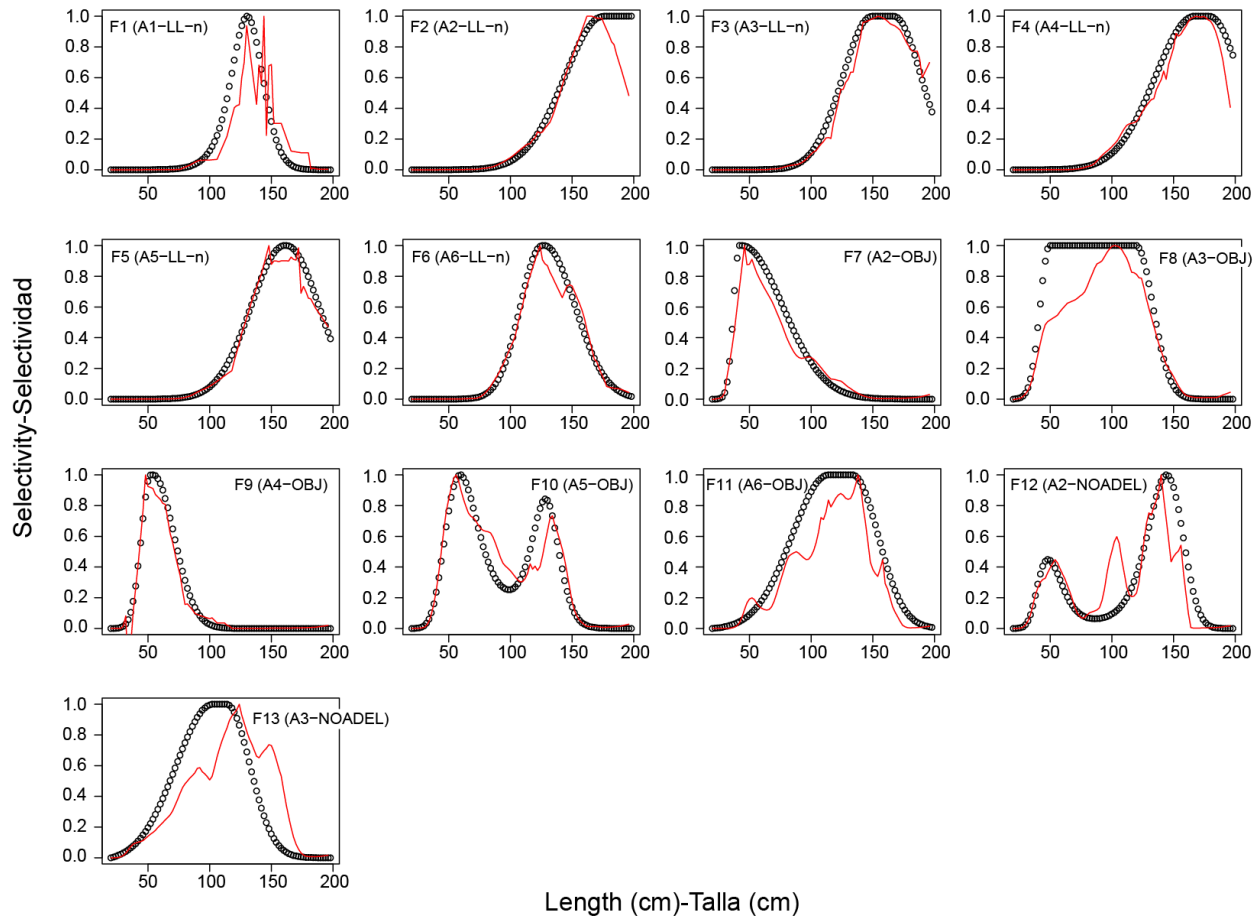
FIGURE A4. Comparison of estimated (black dots) and empirical (red lines) selectivity for every fishery in model Env-Fix that has length-composition data. The Lowess smoother is applied to the empirical selectivity. See Table 2 for explanation of model names.

FIGURA A4. Comparación de la selectividad estimada (puntos negros) y empírica (líneas rojas) para cada pesquería en el Modelo *Env-Fix* que tiene datos de composición por talla. El suavizador *lowess* se aplica a la selectividad empírica. En la Tabla 2 se explican los nombres de los modelos.
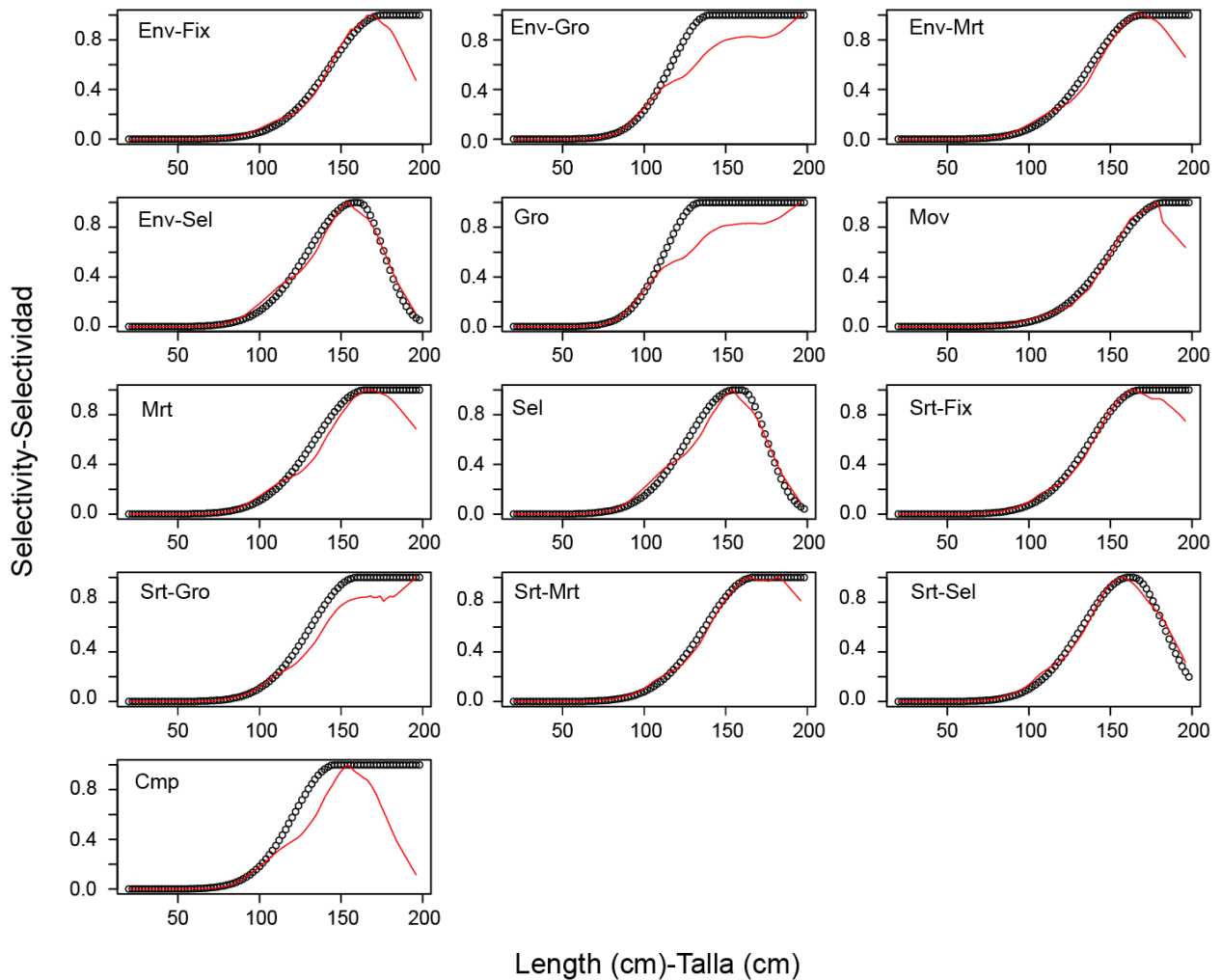
FIGURE A5. Comparison of estimated (black dots) and empirical (red lines) selectivity for Fishery 2 among reference models. The Lowess smoother is applied to the empirical selectivity. See Table 2 for explanation of model names.

FIGURA A5. Comparación de la selectividad estimada (puntos negros) y empírica (líneas rojas) para la Pesquería 2 entre los modelos de referencia. El suavizador *lowess* se aplica a la selectividad empírica. En la Tabla 2 se explican los nombres de los modelos.
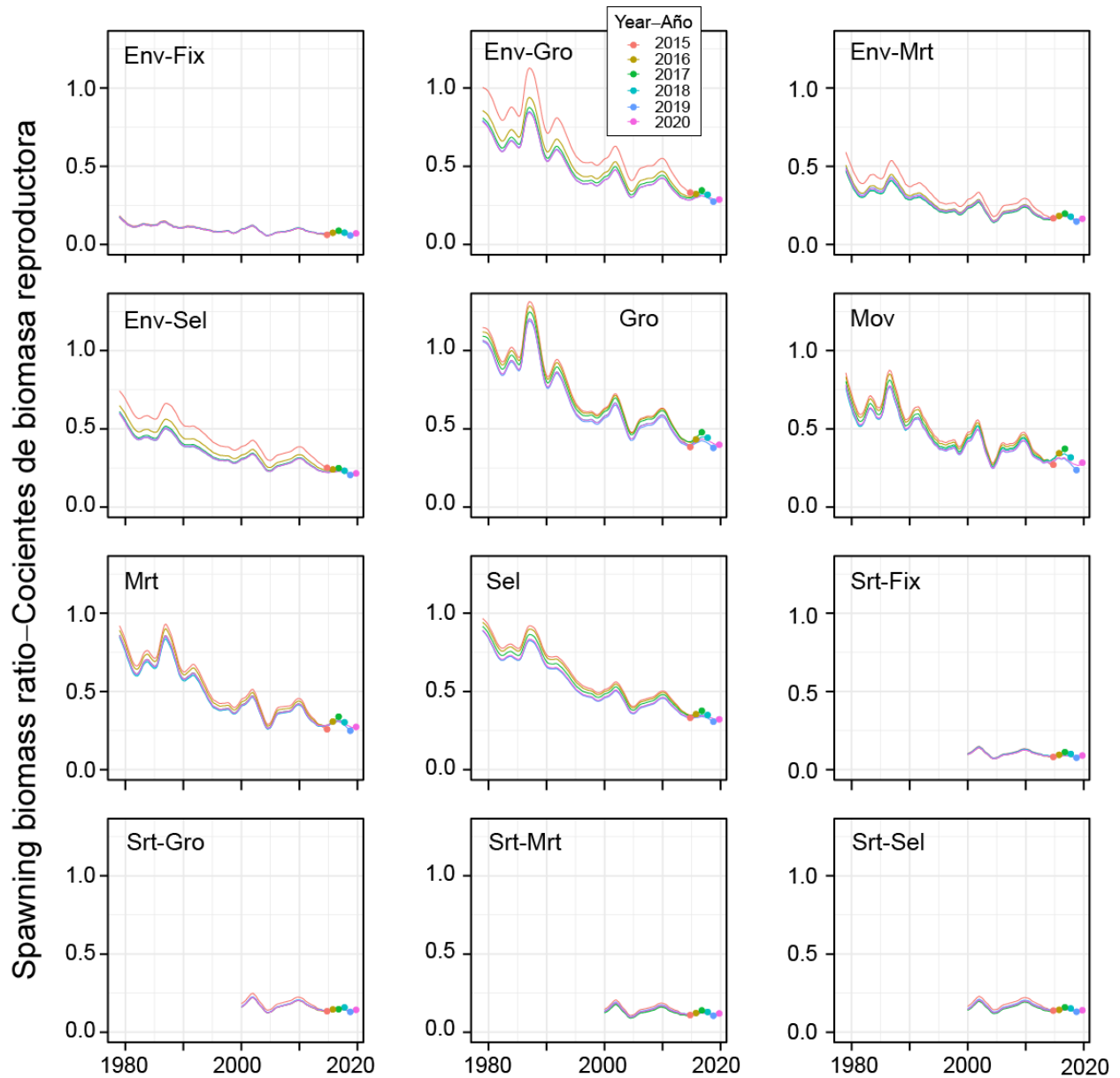
FIGURE A6. Retrospective patterns of the spawning biomass ratio (SBR) of bigeye tuna in the EPO. Year on the x-axis is the year in which the assessment was conducted. See Table 2 for explanation of model names.

FIGURA A6. Patrones retrospectivos del cociente de biomasa reproductora (SBR) del patudo en el OPO. El año en el eje x es el año en el que se realizó la evaluación. En la Tabla 2 se explican los nombres de los modelos.
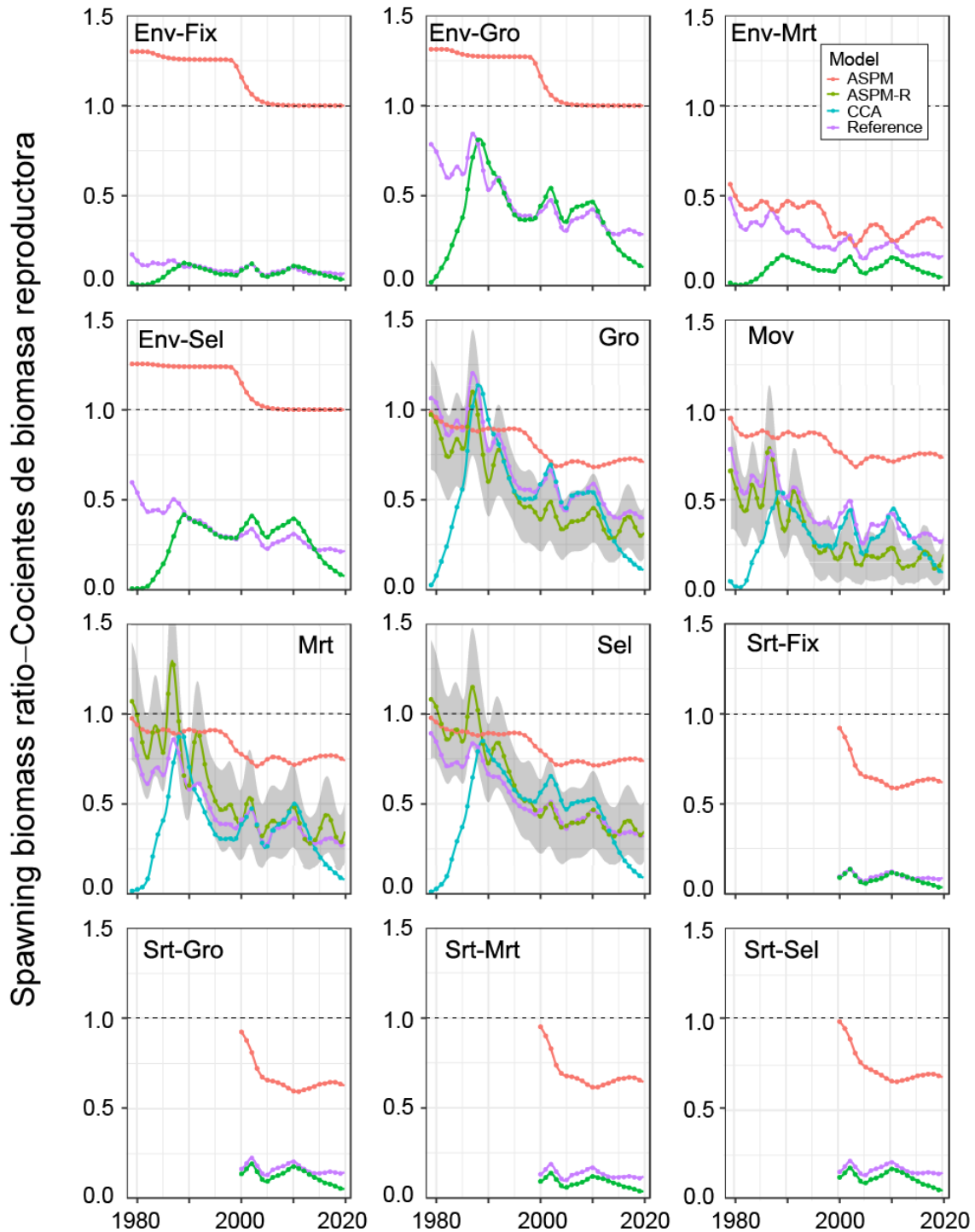
FIGURE A7. Comparison of estimated spawning biomass ratio (SBR) of bigeye tuna in the EPO among reference models and the corresponding ASPM, ASPM-R, and CCA models. The shaded area represents the 95% confidence interval of SBR estimated by ASPM-R. See Table 2 for explanation of model names.

FIGURA A7. Comparación del cociente de biomasa reproductora (SBR) estimado del patudo en el OPO entre los modelos de referencia y los modelos ASPM, ASPM-R, y ACC correspondientes. El área sombreada representa el intervalo de confianza de 95% del SBR estimado por ASPM-R. En la Tabla 2 se explican los nombres de los modelos.
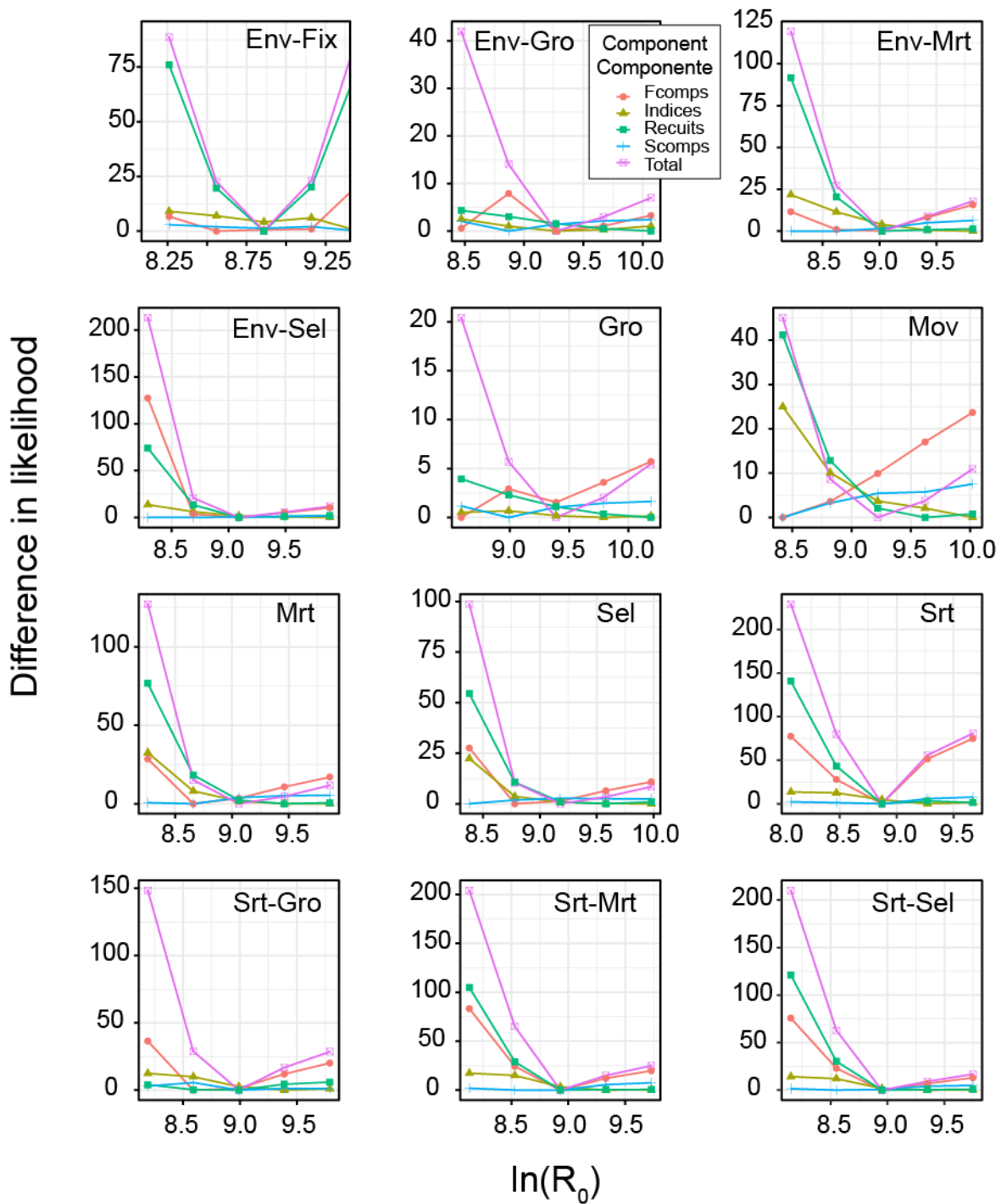
FIGURE A8. Comparison of R_0 likelihood profiles for the 12 reference models for bigeye tuna in the EPO. See Table 2 for explanation of model names.

FIGURA A8. Comparación de los perfiles de verosimilitud de R_0 para los 12 modelos de referencia para el patudo en el OPO. En la Tabla 2 se explican los nombres de los modelos.
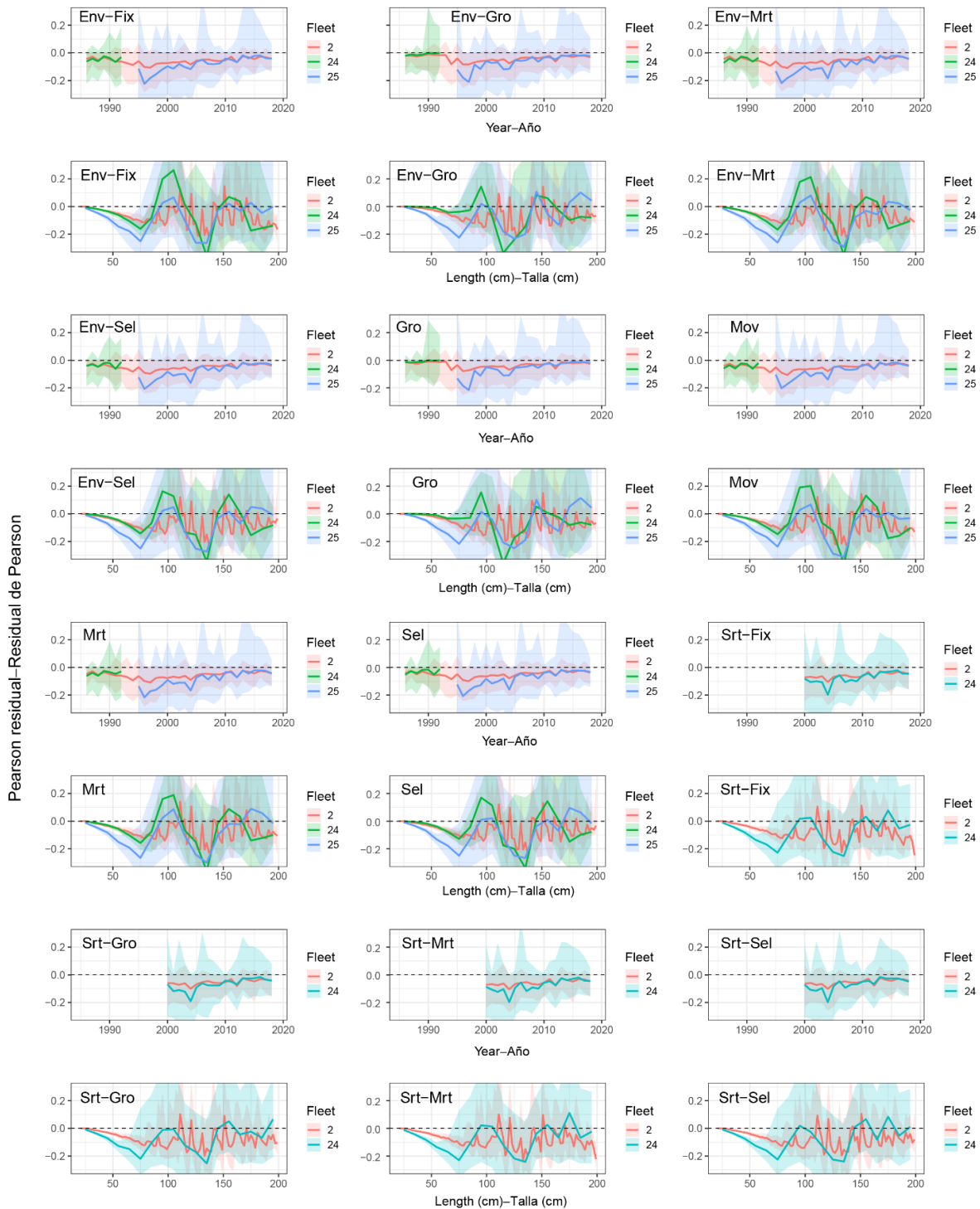
FIGURE A9. Pearson residual plot for the length compositions of Fishery 2 and Surveys 24 and 25 against time and length. The solid lines represent the median and the shaded areas cover the 25-75 percentile range. See Table 2 for explanation of model names.

FIGURA A9. Gráfica del residual de Pearson para las composiciones por talla de la Pesquería 2 y los estudios 24 y 25 contra tiempo y talla. Las líneas sólidas representan la mediana y las áreas sombreadas cubren el rango del percentil 25-75. En la Tabla 2 se explican los nombres de los modelos.

TABLE A1. The AICs of reference models under different steepness (*h*) assumptions. The three AICs compared in this table are the AIC of the model run without including age-at-length (AIC), longline indices of abundance (No LL indices), and longline compositions (No LL comps). The meaning of model name can be found in Table 2. NA: run does not converge/Hessian is not positive definite.

	1	2	3	4	5	6	7	8	9	10	11	12
	Env-Fix	Env-Gro	Env-Mor	Env-Sel	Gro	Mov	Mrt	Sel	Srt-Fix	Srt-Gro	Srt-Mor	Srt-Sel
AIC												
h=1	5245.92	5145.80	5226.02	4748.77	5149.12	5251.30	5238.62	5185.30	3890.20	3873.46	3889.82	3876.74
h=0.9	NA	5145.28	5226.96	4749.01	5149.86	5253.16	5240.76	5186.64	3891.98	3874.30	3890.94	3877.42
h=0.8	NA	5146.10	5228.36	4749.41	5150.82	5255.14	5243.40	5188.32	3894.98	3875.52	3892.82	3878.42
h=0.7	NA	5145.48	5230.50	4750.09	5152.10	5258.70	5246.70	5190.44	NA	3877.56	3895.28	3879.84
No LL indices												
h=1	5633.48	5603.85	5631.36	5595.18	5604.71	5662.87	5649.14	5604.02	4079.49	4072.22	4082.31	4074.51
h=0.9	NA	5603.60	5633.40	5596.12	5605.52	5665.47	5651.91	5605.57	4082.16	4073.54	4084.19	4075.69
h=0.8	NA	5604.78	5636.18	5597.37	5606.49	5668.64	5655.18	5607.49	4086.40	4075.42	4087.01	4077.30
h=0.7	NA	5604.56	5639.73	5599.04	5607.78	5670.85	5659.15	5609.88	NA	4078.22	4090.60	4079.42
No LL comps												
h=1	1681.70	1654.35	1661.04	1654.84	1662.38	1677.85	1665.14	1661.84	1315.63	1314.66	1314.50	1311.28
h=0.9	NA	1653.69	1660.89	1655.06	1663.40	1679.47	1666.37	1663.28	1316.18	1314.88	1314.70	1311.41
h=0.8	NA	1654.27	1660.89	1655.52	1664.72	1680.36	1668.03	1665.10	1317.30	1315.23	1315.18	1311.69
h=0.7	NA	1653.45	1661.47	1656.33	1666.43	1686.90	1670.25	1667.41	NA	1316.09	1316.06	1312.19