INTER-AMERICAN TROPICAL TUNA COMMISSION

SCIENTIFIC ADVISORY COMMITTEE

15TH MEETING

La Jolla, California (USA)
10 - 14 June 2024

DOCUMENT SAC-15-02 REVISED

STOCK ASSESSMENT OF BIGEYE TUNA IN THE EASTERN PACIFIC OCEAN: 2024 BENCHMARK ASSESSMENT

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

Contents

EXECUTIVE SUMMARY

1. The 2024 benchmark assessment of bigeye tuna in the eastern Pacific Ocean continues to use a risk analysis approach to provide management advice. The risk analysis encompasses three levels of hypotheses structured hierarchically to address the main uncertainties in the assessment.
2. The reference models in this benchmark assessment show minor degrees of regime shift in recruitment. The degree of the regime shift in recruitment has significantly decreased from 140% (the base reference model in the last benchmark assessment) to only 20% (the base reference model in this benchmark assessment). Therefore, the regime shift hypothesis is no longer included as the overarching hypothesis in this benchmark assessment.
3. The significant decrease in the degree of the regime shift in recruitment results from the combination of changes made to the assessment model. Among these changes the most influential in reducing the degree of regime shift are adding one more time block to the selectivity of longline fishery fleets in 2011, improving the CPUE standardization model, and using the Lorenzen natural mortality curve for juvenile bigeye.
4. The three levels of hypotheses are structured to address (1) the misfit to the length composition data for the longline fishery that is assumed to have an asymptotic selectivity; (2) the degree of effort creep in the longline fishery; and (3) the steepness of the stock-recruitment relationship.
5. Four models (the initial reference model (Fix), estimating growth (Gro), dome shape selectivity for all fisheries (Sel), and estimating natural mortality (Mrt)) are considered for the first level hypothesis, three rates of annual increase in longline catchability (0%, 1%, 2%) are considered for the second level hypothesis, and three values of steepness (1.0, 0.9, 0.8) are considered for the third level hypothesis. The combination of the three levels of hypotheses results in 36 reference models, of which thirty-three are included in the risk analysis due to convergence issues with three models.
6. The four models considered for the first level hypothesis are equally weighted, the three rates of annual increase in longline catchability for the second level hypothesis are equally weighted, and the three values of steepness considered for the third level hypothesis are weighted based on expert judgement from the risk analysis for the last benchmark assessment (SAC-11-INF-F).
7. The overall results of the risk analysis, based on the thirty-three converged reference models, show unimodal probability distributions for management quantities. The shift from a bimodal to unimodal pattern in the distributions likely results from resolving the regime shift in recruitment in this benchmark assessment. The risk analysis indicates:
    a. 46.6% probability that the spawning biomass at the beginning of 2024 is below the target reference point (S_{MSY_d})
    b. 24.7% probability that the fishing mortality in 2021-2023 is above the target reference point (F_{MSY})
    c. 58.5% probability that the fishing mortality in 2017-2019 (the *status quo* period) was above the target reference point (F_{MSY})
    d. 0.2% probability that the spawning biomass at the beginning of 2024 is below the limit reference point (S_{Limit})
    e. 0.1% probability that the fishing mortality in 2021-2023 is above the limit reference point (F_{Limit})
8. The weighted 10-year projection under the current fishing mortality suggests there is a 50% probability that the spawning biomass ratio at the beginning of 2034 will be above 0.27.

1. INTRODUCTION

Bigeye tuna (*Thunnus obesus*) is a tropical tuna species inhabiting tropical and temperate waters of the Pacific, Atlantic, and Indian Oceans (Collette et al. 2001). They are fished by various methods in the eastern Pacific Ocean (EPO). Bigeye tuna has been the main target species of the longline fishery in the EPO since the 1970s, owing to its high commercial value in the global sashimi market (Matsumoto 2008). Before 1993, the distant water longline fishery was the primary method of harvesting bigeye tuna in the EPO, with an average annual catch of 88,000 metric tons from 1985 to 1992 (IATTC 2021). In contrast to longline fisheries that catch primarily large and mature bigeye, purse-seine fisheries catch mostly small and immature bigeye (Okamoto and Bayliff 2003, Xu et al. 2020). The three main types of purse-seine fisheries in the EPO include sets made on free-swimming tuna schools (NOA), on tunas associated with dolphin herds (DEL), and on tunas associated with floating objects (OBJ) (Lennert-Cody and Hall 2000, Maunder and Harley 2006). Of the three types, bigeye tuna in the EPO is most vulnerable to the purse-seine sets on floating objects, which before 1993 were a coastal fishery based mostly on natural objects such as tree trunks and kelp paddies (Lennert-Cody and Hall 2000). The purse-seine fishery on floating objects during this period caught about 5,000 metric tons of bigeye annually, which is much lower than the level of longline catches (IATTC 2021).

With the rapid development of Fish Aggregation Devices since 1993, the OBJ fishery gradually replaced the longline fishery as the dominant fishery catching bigeye tuna in the EPO (IATTC 2021, Xu et al. 2020). Fish Aggregation Devices are man-made floating objects placed in the water to attract tunas. They are commonly equipped with an echo-sounder to measure fish abundance and a GPS to report their geographic locations (Hall and Roman 2013). The OBJ fishery, which catches small bigeye tuna, has expanded substantially and rapidly in the tropical EPO from the coastal waters of the American continent to beyond the west management boundary (150°W) of the IATTC (Lennert-Cody and Hall 2000). This expansion of the OBJ fishery has a strong impact on the longline fishery that catches large bigeye tuna of the same stock (Matsumoto 2008, Okamoto and Bayliff 2003, Sun et al. 2019). Specifically, the longline catch of bigeye tuna in the EPO has declined significantly from 88% in 1993 to a historically low level of 23% in 2020 (IATTC 2021).

The last benchmark assessment for bigeye tuna in the EPO was conducted in 2020 (SAC-11-06). This benchmark assessment introduced a new approach to providing management advice at IATTC as it provided the basis for a risk analysis (SAC-11-08). The new risk analysis methodology uses several reference models that represent various plausible states of nature (assumptions) about the biology of the fish, the productivity of the stocks, or the operation of the fisheries, effectively incorporating uncertainty into the management advice. Forty-eight reference models were developed for the last benchmark assessment within a hierarchical framework to address three major uncertainties from the previous assessment. These uncertainties included the apparent regime shift in recruitment, the misfit to the length composition data for the longline fishery that was assumed to have an asymptotic selectivity, and the steepness of the stock-recruitment relationship. Under this risk analysis approach, the staff can explicitly evaluate the probability of breaching the reference points defined in the IATTC's harvest control rule for tropical tunas (C-16-02).

The last benchmark assessment for bigeye tuna in the EPO highlighted a concern regarding a bimodal pattern observed in the management quantities. In particular, two distinct groups of reference models used for bigeye tuna were identified based on the management quantities relative to maximum sustainable yield (MSY): pessimistic and optimistic models. The large difference between the MSY-related management quantities of these two groups resulted in model-combined joint probability distributions of management quantities, such as F/F_{MSY}, to show two distinct modes, a pattern that has generally been referred to as the "bimodal pattern" in the bigeye assessment. Although the overall combined results of the risk analysis were used for management advice, this bimodal pattern allowed for two distinct interpretations about stock status depending on the group of reference models interpreted (i.e. optimistic or pessimistic), either that fishing mortality for bigeye could be greatly increased or greatly decreased from the recent level to achieve the target reference point. Moreover, the risk analysis indicated that neither of the two scenarios is significantly more likely than the other, making it challenging to provide effective management advice.

This report presents the outcomes of the 2024 benchmark stock assessment for bigeye tuna in the EPO. Since the last benchmark assessment, several changes related to data (index of relative abundance and longline length compositions), biology (natural mortality and growth), and model specifications (selectivity and data weighting) have been made to the assessment model for bigeye tuna in the EPO. These changes incorporate elements from the panel recommendations of the two recent external reviews of the stock assessments (RVMTT-01-RPT and RVDTT-01-RPT). The combination of these changes effectively removed the apparent regime shift in recruitment estimates and the bimodal pattern in management quantities.

2. DATA

2.1. Fisheries and 'survey': overview

The 2nd external review of the bigeye assessment did not consider developing a spatial model for the EPO a high priority in the short term. Accordingly, the assessment models considered in this benchmark assessment are not spatially structured and use the 'areas-as-fleets' approach, which treats geographic areas as separate fleets with different selectivity curves in a single-stock assessment model. This approach implicitly assumes that the stock is homogenously distributed throughout its range and that any differences in composition data arise due to different contact selectivities (Hurtado-Ferro *et al.* 2014). However, it recognizes that fishing in different areas usually leads to different ages/sizes of fish being removed from the population due to spatial variation in age/size structure. Consequently, fisheries need to be defined spatially to achieve a relatively homogeneous fish distribution across each area. This approach ensures that each fishery's length composition is not influenced by the location of fishing activities (Punt 2019).

Since the last benchmark assessment, survey fleets have been disconnected from the fisheries structure, total catch, and catch composition. In the EPO, there were no fishery-independent surveys of tuna abundance and size composition, with the term "survey" in this context referring to a fleet that has data (e.g., abundance index and size composition) but takes no catch. For the "areas-as-fleets" approach on which the assessment is based, the abundance index and the associated composition data should reflect the conditions of the EPO-wide bigeye population (Maunder et al. 2020a). Therefore, the abundance index for a survey fleet should be computed using an area-weighting approach for the EPO rather than an area defined for the fishery. The composition data associated with the survey abundance index should be spatially weighted by fish abundance and aggregated across the entire spatial domain.

2.2. Fishery definitions

2.2.1. Method

A regression tree approach for analyzing length frequency data is used to provide gear and set type-specific fishery definitions. The regression tree algorithm (Lennert-Cody *et al.* 2013, Lennert-Cody *et al.* 2010) uses recursive partitioning to search for hierarchical binary decision rules that divide the data into more homogeneous subgroups. The binary decision rules are selected to provide the greatest decrease in the heterogeneity of length composition data, which is measured based on the Kullback–Leibler divergence. The regression tree algorithm has been recently included in an R package *FishFreqTree*, where fisheries length-frequency data, separated by gear (longline/purse-seine) and purse-seine set type (OBJ/NOA/DEL), are grouped by latitude, longitude, quarter, and cyclical-quarter. This R package is open-source and can be accessed at: https://github.com/HaikunXu/FishFreqTree.

There are three main differences between the regression tree analysis conducted for the last benchmark assessment and this benchmark assessment. The previous analysis is based on both catch-per-unit-effort (CPUE) and length frequency to find compromised spatial boundaries across gear and set type. In contrast, this analysis is based solely on length frequency and is conducted for each gear and set type to provide uncompromised gear and set type-specific fishery definitions. The habitat preference of bigeye tuna is size-specific, so fish caught by different gear types are likely to have distinct spatial patterns of age/size composition. As such, independent fishery definitions are more appropriate for this assessment model that utilizes the "areas-as-fleets" approach.

The second difference is in the source of longline composition data. The previous analysis is based partially on the longline length composition dataset that Japan submitted to the IATTC's public domain. This dataset is coarse and pre-aggregated by 5° latitude, 10° longitude, and 1 quarter (WSBET-02-02). This analysis uses new longline length composition data that Japan submitted through a Memorandum of Understanding with the IATTC. This dataset has a much finer spatial and temporal resolution (1° latitude, 1° longitude, and month) and includes additional useful information, such as the bin size associated with each length measurement.

The third difference is whether processing purse-seine length composition data for the regression tree analysis. In this regression tree analysis, the length measurements of bigeye tuna taken in the first and third months of a quarter are adjusted based on the growth curve to reflect the value they would represent if fish were measured in the middle of the quarter. Furthermore, to remove the influence of recruitment variation on observed length frequency, each length frequency observation is divided by the EPO-wide average length frequency for the corresponding quarter. These two data processing steps were not included in the previous regression tree analysis.

2.2.2. Longline fisheries

Longline fisheries are defined in this benchmark assessment using Japanese longline length composition data, which covers the period between 1986 and 2020. Before being analyzed by the regression tree algorithm, the data is filtered to include only commercial vessels' data collected at a spatial resolution of 1° x 1° and a bin size of 1 or 2 cm. Poorly sampled grids with less than four years of data between 1986 and 2020 are excluded from the dataset. The remaining data is aggregated by 5° x 5° spatially and quarter temporally into fifteen length bins (<70 cm, 70-80 cm, 80-90 cm, …, >190 cm).

The regression tree algorithm is specified to find five splits or define six longline fisheries for the EPO excluding the Hawaii corner (north of 10°N and west of 105°W), where a separate longline fishery is defined as in the previous benchmark assessment. The regression tree is hierarchical and may exhibit a certain degree of instability. Instead of selecting only the best candidate for each split, we consider the top four and two competing candidates for the first and second splits, respectively, and rank the eight (4 x 2) 5-split combinations according to the proportion of variance in the length-frequency data explained. Among the eight 5-split combinations, the best one selected for the longline fishery in the EPO explains 15.22% of the variance in the length-frequency data (Table 1). It is worth noting that the algorithm is originally specified to find four splits or define five fisheries, but there is one fishery (15°S - 5°S and 150°W - 105°W) having a bimodal pattern in aggregated length frequency profile (Figure 1). The best candidate for the fifth split (130°W) explains an additional 1.10% of the variance in the length-frequency data and significantly reduces the bimodal pattern by splitting the fishery spatially into two fisheries. Consequently, the best candidate for the fifth split is kept and the EPO including the Hawaiian corner is divided into seven areas for longline fisheries.

Given that longline catches are reported to the IATTC in numbers by some fleets and in weight by others, two longline fleets, the catch units of which are 1,000s of fish and metric tons, respectively, are defined for each longline area. In total, the benchmark assessment model includes fourteen longline fishery fleets.

2.2.3. Purse-seine fisheries on floating-objects

The definition of OBJ fisheries is based on length composition data collected by port samplers from the OBJ sets made by Class-6 vessels (Suter 2010). Port samplers collect data only from wells with catch from the same set type, sampling area, and year-month. Data before 2000 are removed from this analysis because the sampling protocol used by the IATTC port-sampling program changed in that year and the OBJ fishery was not fully expanded across the EPO during the 1990s. The raw data has a 5° x 5° spatial resolution and a 1 cm bin size from 1 cm to 201 cm. Poorly sampled grids with less than 4 years of data available since 2000 are removed from the dataset. The remaining length frequency data is then aggregated by quarter into fifteen 10 cm length bins (<30 cm, 30-40 cm, 40-50 cm, …, >170 cm) (Figure 2) and divided by the EPO-wide average length frequency for the corresponding quarter to remove the influence of recruitment variation.

Same as in the last benchmark assessment, this benchmark assessment includes five OBJ fishery fleets and therefore we specify the regression tree algorithm to find four splits. We consider the top four and two competing candidates for the first and second splits, respectively, and rank the eight (4 x 2) split combinations according to the proportion of variance in the length-frequency data explained. Among the eight 4-split combinations, the best one selected for the OBJ fishery in the EPO explains 10.46% of the variance in the length-frequency data (Table 2).

2.2.4. Purse-seine fisheries on free schools

The definition of NOA fisheries is based on length composition data collected by port samplers from the NOA sets made by Class-6 vessels (Suter 2010). Port samplers collect data only from wells with catch from the same set type, sampling area, and year-month, Data before 2000 are removed from this analysis because the sampling protocol used by the IATTC port-sampling program changed in that year. The raw length frequency data is aggregated by quarter into fifteen 10 cm length bins (<30 cm, 30-40 cm, 40-50 cm, …, >170 cm) and divided by the EPO-wide average length frequency for the corresponding quarter to remove the influence of recruitment variation.

The length frequency data for NOA sets are sparse both spatially and temporally (Figure 3). Moreover, NOA sets contribute to only a small percentage of bigeye catch in the EPO. We therefore include only two NOA fishery fleets in this benchmark assessment model. The best split selected for the NOA fishery in the EPO (i.e., 130°W) explains 9.95% of the variance in the length-frequency data (Table 3).

2.2.5. Summary

Twenty-two fishery fleets are defined for bigeye tuna in this benchmark assessment, classified by gear (purse-seine/longline), purse-seine set type (OBJ/NOA), area of operation (Figure 4), and unit of longline catch (numbers/weight) (Table 4). Due to a lack of length composition data and a negligible contribution to total bigeye catch, we pool both pole-and-line and DEL sets into the NOA sets in this benchmark assessment model. The twenty-two fishery fleets comprise fourteen longline fishery fleets, five OBJ fishery fleets, one OBJ discard fleet (further details in section 2.4.4), and two NOA fishery fleets. The aggregated length frequencies of bigeye tuna show a single mode in most fisheries (Figure 5), suggesting that most fisheries defined by the regression tree analysis do not include more than one cohort and the double-normal selectivity curve parameterization can be used.

2.3. Survey definition

In the last benchmark assessment, two longline survey fleets were defined based on the time of operation: 'early' (1979-1992) and 'late' (1995-2019). Catchability and selectivity were estimated separately for the two survey fleets and the coefficient of variation (CV) of the late index of abundance was fixed while that of the early index was estimated. The main reason for splitting the longline abundance index into two time periods was that gear configurations of Japanese longline vessels changed abruptly in 1993 and 1994. Specifically, both hooks-between-floats (HBF) and mainline material, two key indicators of hooks' depth distribution in the water column, changed rapidly in 1993-1994. As the depth distribution of bigeye tuna in the EPO is influenced by body size (Schaefer and Fuller 2010), these notable changes in gear configurations may have led to a change in catchability and selectivity for the survey fleet.

Unless there is evidence against constant survey catchability and selectivity, the current good practices for CPUE modeling advise against splitting the abundance index by time into separate non-overlapping time blocks (Hoyle et al. 2024). Splitting the abundance index by time wastes a large amount of information in the CPUE data, particularly the continuous trend of population abundance over a long period. Hoyle et al. (2024) argue that if the assessment model is misspecified, splitting the abundance index can introduce bias as the model may not be able to reliably scale abundance indices. Thus, analysts should at least consider whether the estimated change in catchability at the split makes sense. Regarding this point, we revisit the survey definition in this benchmark assessment by checking the estimated change in catchability and selectivity at the split.

Indeed, the last benchmark assessment model estimated similar catchability and selectivity for the early and late survey fleets. The estimated catchability for the early period (1.58 $\pm$ 0.39) is slightly higher than that for the late period (1.34 $\pm$ 0.13). The selectivity curves estimated for the two time periods are also closely aligned (Figure 7 in SAC-14-05). This result is contrary to expectations, as the catchability of the main target species (Japanese longline fishery in the EPO persistently targets bigeye tuna) tends to increase over time due to continuous improvements in fishing technology and knowledge. This counterintuitive result suggests that the assessment model is likely mis-specified and unable to reliably scale the two abundance indices. Consequently, one longline survey fleet that covers the entire model period (1979-2023) is defined for the model in this benchmark assessment.

2.4. Catch

2.4.1. Catch definition

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

- Retained: catch retained aboard the vessel
- Discarded: catch not retained aboard the vessel
- Total: retained catch + discard
- Unloading: retained catch unloaded from the vessel

2.4.2. Purse-seine

The information used to estimate the total catch by species comes from four main sources. Those sources are canneries, on-board observers, vessel logbooks, and in-port sampling by IATTC staff. The observer and logbook databases also contain other information about the catches, such as the location, date, and set type. Year is the only ancillary information available for the cannery data. Additionally, the port-sampling program for collecting length composition data has also provided information on species composition since 2000.

For this assessment, total catches 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 has changed over time. Estimates before 2000 are based on the recorded species totals in the cannery 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 ancillary information in the observer and logbook databases. Since 2000, port-sampling data have been used to determine the species composition of the total catch. The total catch of all three species combined (from cannery, observer, and logbook data)[1] 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 estimators can be found in Tomlinson (2002; 2004), Suter (2010) and in WSBET-02-06. Details of the port-sampling protocol in use since 2000 can be found in the appendix of Suter (2010). This catch estimation methodology, which is a design-based approach, is used to obtain the fleet-level Best Scientific Estimates (BSEs) of species composition of the catches for each purse-seine fishery fleet. The methodology has been integrated into a R package *BSE* that can be accessed at: https://github.com/HaikunXu/BSE.

Bias-adjustment was made for the BSE-estimated OBJ catches for the two years affected by the COVID-19 pandemic (2020 and 2021). The pandemic disrupted the collection of species and size composition data by IATTC port-samplers, leading to a systematic loss of port-sampling data from ports where much of the EPO BET catch is unloaded (SAC-13 INF-L). Given that the BSE algorithm relies heavily on the port-sampling data to predict the species composition of purse-seine catches, it is likely that the purse-seine catches estimated for the two COVID-19 years by the BSE algorithm are biased. Recent research conducted by Majumdar et al. (2023) suggests that the BSE algorithm overestimates bigeye catches in the OBJ fishery by 12.0% and 18.2% for 2020 and 2021, respectively. Consequently, adjustments were made to reduce each BSE-estimated quarterly OBJ catch for 2020 and 2021 by 12.0% and 18.2%, respectively.

2.4.3. Longline

The IATTC staff does not collect data on longline catches directly. Instead, they are reported annually to the IATTC by individual Members and Cooperating Non-Members (CPCs), according 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 longline fleets report catch and effort data 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, Korea, and Vanuatu) and coastal CPCs (principally Mexico and the United States).

For this assessment, longline catch data are aggregated in line with the new fishery definitions based on the area of operation (Figure 1). Because two longline fishery fleets are defined for each area, the catches are entered in their original units (1,000s of fish and metric tons), and the conversion between numbers and weight is done internally in the assessment model. Updated and new catch data for the longline fishery fleets are incorporated into the current assessment. The catch data for 2023 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 that covers the CPC's Exclusive Economic Zone (EEZ). The algorithm to calculate the catch by longline fishery fleet is described in WSBET-02-03, and the associated R code is available at https://github.com/HaikunXu/IATTCassessment/blob/master/R/ll_catch.R.

[1] If landing information from canneries is unavailable, catch information in the observer or vessel logbook databases, in that order, is used instead.

2.4.4. Discards

Two types of discards are considered in this benchmark assessment: those resulting from inefficiencies in the fishing process and those related to catch sorting. Examples of inefficiency are catches from a set exceeding the remaining storage capacity of the fishing vessel or dumping unwanted bycatch species, and 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 under the on-board observer program of the Agreement on the International Dolphin Conservation Program (AIDCP), using the methods in Maunder and Watters (2003). No observer data is available to estimate discards before 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 OBJ fisheries (fleets 15-19) represents retained catch plus discards resulting from inefficiencies in the fishing process. Fishery fleet 20 represents discards resulting from catch sorting in OBJ fisheries. They are treated separately, following the rationale of Watters and Maunder (2001), and are assumed to be composed of 2-4 quarters old bigeye. In fishery fleets 21 and 22 (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 are not available so the retained catch is assumed to represents the total catch (Table 4).

2.4.5. Summary

To facilitate the comparison of purse-seine and longline catches, the portion of longline catches that are recorded in numbers of fish is converted to weight inside the assessment model (Figure 6). The longline and purse-seine catches for bigeye tuna have several important features:

- Longline fisheries dominated bigeye catches before 1993, the OBJ fishery has become the main fishery for bigeye since then.
- The total annual catch has been relatively stable since 2005.
- Both longline and purse-seine catches dropped dramatically since 2021. The total bigeye catches in 2021-2023 reached the historically low level since 1979.

2.5. Index of relative abundance

Indices of relative abundance are a crucial input to stock assessment models as they directly inform the changes in population abundance over time (Francis 2011). Ideally, indices of abundance should be calculated using fishery-independent survey data, collected using the same fishing gear and operation across time to assure constant catchability and selectivity, and have a random or fixed sampling design in space. However, for most tuna species worldwide, including bigeye tuna in the EPO, survey data are not available. Therefore, indices of abundance are derived solely from fishery-dependent CPUE data. These data need to be standardized so that the abundance index is approximately proportional to population abundance (Maunder and Punt 2004). To achieve this, the standardization model needs to remove the part of the variation in the CPUE data that is not driven by changes in population abundance. Furthermore, the standardization model should impute fish abundance for unfished locations and use an area-weighting approach to compute the abundance index for the population for the entire spatial domain of the fishery (Thorson et al. 2015).

While both purse-seine and longline indices of abundance are available for bigeye in the EPO, this assessment includes only the longline index, which primarily informs the abundance trend of large bigeye.

Standardizing purse-seine indices of abundance is notoriously challenging as the relationship between abundance and fishing effort, and how it evolves with technological advancements over time, is unclear. Furthermore, 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.

2.5.1. Data source

In the assessment of bigeye tuna in the EPO, the survey fleet is based on fishery-dependent CPUE and length composition data collected by Japanese longline vessels that persistently target bigeye tuna. Among all distant-water longline vessels operated in the EPO, Japanese longline vessels have the highest spatial coverage within the EPO and the longest history of high-quality logbook data, providing the information needed for the standardization of a reliable abundance index with a large contrast across time. For the first time in the history of the benchmark assessment for bigeye tuna in the EPO, the longline index of relative abundance is derived from operational CPUE data collected by Japanese commercial longline vessels. This dataset records catch (in number of fish) and effort (in number of hooks), as well as useful gear or vessel information such as HBF and vessel ID, at the set level with a spatial resolution of 1° x 1°.

Several filters are applied to the Japanese operational longline CPUE dataset before it is fit to the spatio-temporal model for CPUE standardization. First, data before 1979 are excluded due to unavailable vessel ID. Previous analyses suggest that including vessel ID is necessary for standardizing Japanese longline CPUE data as different vessels have different efficiencies in catching bigeye tuna in the EPO (usually referred to as vessel effects). Newly introduced vessels tend to be more efficient in catching fish than newly retired vessels, so vessel turnover most likely results in an increase in average fishing efficiency over time. In other words, the index of abundance standardized by a model in which vessel effects are ignored tends to overestimate the temporal trend in the index.

Second, poorly sampled vessels and spatial cells are excluded from the dataset to which the CPUE standardization model is fit. Estimating vessel effects can be difficult if a vessel does not have enough years of data. Spatial imputation of fish density is unreliable and can lead to a biased index of abundance if a spatial cell is poorly sampled, especially on the edge of the spatial domain of the core habitat. Specifically, all vessels and spatial cells with less than 40 and 20 quarters of data, respectively, between 1979 and 2023 are removed for CPUE standardization. Also, the data in the Hawaii corner (north of 10°N and west of 105°W) are removed due to an assumption that bigeye caught in that area does not belong to the core EPO population.

Third, several minor filters are applied to remove outliers from the dataset. Specifically, we remove the sets with 1) an extreme (less than 5 or more than 25) or missing HBF; 2) more than 5,000 hooks; 3) a ratio of the number of bigeye to the number of hooks larger than 0.1; and 4) a latitude south of 25°S.

Both fishing effort and scale of the Japanese longline fleet operating in the EPO have declined almost linearly since around 1993 (Figure 7). The spatial coverage of the dataset decreased increasingly faster in the last decade, especially on the eastern side of the EPO (Figure 8). HBF increased abruptly from an annual average of 12 to that of 16 during 1993-1995 and remained relatively stable thereafter (Figure 9).

2.5.2. Standardization procedure

There is a need to standardize fishery-dependent CPUE data in the process of computing the index of relative abundance (Hoyle et al. 2024, Maunder and Punt 2004). The term "standardize" denotes the process of removing the impact of other factors on CPUE, ensuring that the standardized index of relative abundance is proportional to population abundance. It is well known that indices of relative abundance derived from fishery-dependent CPUE data are susceptible to biases stemming from factors such as preferential sampling (Conn et al. 2017, Diggle et al. 2010, Pennino et al. 2019), gear characteristics (Campbell

2015), and targeting (Chang et al. 2011, Winker et al. 2013). In contrast to statistically designed surveys, which collect data intending to ensure that the selection of sample locations is independent of fish abundance (with exceptions, such as in cases involving the establishment of a marine protected area; see Yalcin et al. 2023 for an example), fishery fleets operate non-randomly in space and seldom cover the entire spatial domain of the population during the period of interest.

Most CPUE standardization models implicitly assume that the choice of fishing locations is independent of fish abundance. This assumption is often violated in fishery operations due to financial and practical incentives. For example, fishers tend to fish in areas where the abundance of the target species is expected to be high. This phenomenon is referred to as preferential sampling and can lead to biased predictions of fish abundance. Unfished "holes" associated with fishery-dependent data can emerge due to preferential sampling, fishery restrictions, or economic considerations. The presence of these unfished holes can introduce significant bias to the trend of the abundance index, particularly if they cover a larger portion of the population spatial domain or exhibit a pronounced trend in their spatiotemporal distribution. Given the assumption of linear proportionality between population abundance and the abundance index, any bias in the abundance index may lead to biased estimates of stock status and, consequently, misguided management advice. Therefore, a CPUE standardization process is essential to mitigate these biases and ensure the reliability of assessments and subsequent management recommendations.

The principal challenge encountered in the standardization of longline CPUE data for bigeye tuna in the EPO arises from a notable contraction of the Japanese fleet fishing ground, particularly on the eastern side of the EPO since 2010 (Xu et al. 2020). The standardization model deals with the existence of numerous zero-count observations by using the delta approach (Lo et al. 1992). This approach separately models encounter probability (the probability of positive catch) and positive catch rate. Traditionally, the delta-generalized linear model (GLM) was employed for standardizing CPUE data for bigeye tuna in the EPO (Hoyle and Maunder 2006). Both encounter probability and positive catch rate in the GLM include a density temporal term, a spatial density term, and a catchability term (HBF). The spatial term is estimated for each 5° by 5° spatial cell without accounting for spatial autocorrelation. HBF, a gear characteristic widely acknowledged to influence tuna catch rates, informs the depth distribution of hooks in the water column. This GLM lacks a spatiotemporal density term for both encounter probability and positive catch rate, assuming implicitly that spatial effects on the abundance of bigeye tuna remain constant over time. This assumption contradicts a recent finding from a study by Satoh et al. (2021) that the spatial distribution of bigeye tuna in the EPO changed from year to year in response to environmental fluctuations. Moreover, simulation studies focusing on high-migratory pelagic species have shown that accounting for time-area interaction in longline CPUE standardization yields a less biased index of relative abundance (Grüss et al. 2019, Zhou et al. 2019). These considerations underscore the necessity of refining the CPUE standardization model to account for the dynamic spatiotemporal distribution of bigeye tuna.

Our current approach to CPUE standardization for bigeye tuna in the EPO involves the utilization of a spatiotemporal delta-generalized linear mixed model (GLMM). This type of model has gained prominence in recent years for standardizing fishery-dependent CPUE data, including for highly migratory species (Ducharme-Barth et al. 2022, Xu et al. 2019). Spatiotemporal GLMMs can account for time-area interaction by including a spatiotemporal term to both encounter probability and positive catch rate. In contrast to the traditional GLM, the spatiotemporal GLMMs explicitly consider spatial and temporal autocorrelation in spatial and spatiotemporal terms. An additional advantage of the spatiotemporal GLMM is its capacity to impute fish abundance in unfished areas based on spatial and temporal autocorrelation. Moreover, it can compute an area-weighted index of relative abundance over the entire spatial domain of the population of interest.

Despite the application of a more advanced spatiotemporal GLMM, the standardization of longline CPUE data for bigeye tuna in the EPO remains notably challenging. In the tropical EPO, bigeye tuna has been the main target species of the Japanese longline fishery since the 1970s, driven by its high commercial value in the global sashimi market (Matsumoto 2008). The Japanese longline fishery, upon which the CPUE standardization relies, historically operated extensively across the tropical EPO until about 2000. Since then, it has gradually withdrawn from the eastern part of the tropical EPO, presenting a systematic large-scale contraction of the fishing ground. This significant contraction necessitates the imputation of fish abundance by the spatiotemporal model for a large portion of the tropical EPO. Although the spatiotemporal model can perform imputation by using estimated spatial autocorrelation patterns, this process is susceptible to substantial bias due to a lack of neighboring data to inform the imputation for the large unfished area in the east. Adding to the complexity, the contraction of the fishing ground may result from depletion-driven preferential sampling, a phenomenon the spatiotemporal GLMM cannot explicitly account for in the imputation of fish abundance. Both the CPUE trends from the longline and purse-seine fishery, catching respectively large and small bigeye, exhibit a more rapid decrease in the eastern than the western tropical EPO. The higher depletion rate of the target species (i.e., bigeye tuna) may be why the Japanese longline fishery gradually moved out of the eastern side of the tropical EPO since 2000. Ignoring this preferential sampling process could lead the spatiotemporal model to overestimate fish abundance in unfished areas (Conn et al. 2017, Pennino et al. 2019).

VAST (Thorson and Barnett 2017) is chosen as the platform to standardize Japanese longline CPUE, which is computed as the number of bigeye caught per 1,000 hooks. VAST is an open-source R package (https://github.com/James-Thorson-NOAA/VAST) and has recently gained increasing popularity in standardizing fishery-dependent CPUE data for tunas (Ducharme-Barth et al. 2022, Maunder et al. 2020b, Satoh et al. 2021, Xu et al. 2019). As a delta-generalized linear mixed model, VAST separately models encounter probability and positive catch rate to account for zero-inflated catch rate observations. We specify VAST to use the logit and log link functions for the linear predictors of encounter probability and positive catch rate, respectively. The gamma link replaces the log link, which was used in the CPUE standardization model for the last benchmark assessment (Xu et al. 2018a), due mainly to the fact that model diagnostics, the quantile-quantile plot, suggests a superior performance of the gamma link in fit to the CPUE data for bigeye in the EPO. The four quarters are treated equally in VAST.

Both the linear predictors of encounter probability and positive catch rate include an intercept (year-quarter) term, a time-invariant spatial term, a time-varying spatiotemporal term, a vessel effect term, and a catchability (HBF modeled by a 2-knot spline) term. By using Template Model Builder (Kristensen et al. 2016), the intercept term and the catchability term are estimated as fixed effects; the spatial term, the spatiotemporal term, and the vessel effect term are estimated as random effects. Given that values at nearby locations are usually more similar than those at remote sites, the spatial and spatiotemporal random effects are both assumed to be autocorrelated in space. Specifically, VAST applies the Matérn function to describe the rate at which the correlation between random effects declines over space.

VAST computes the index of abundance by using an area-weighting approach. It first predicts fish density for each spatial knot and time and then sums the product of fish density and area of the knot over the spatial domain to derive the abundance index. Choosing the number of spatial knots needs to consider the trade-off between model accuracy and model efficiency. A total of 200 spatial knots is used in this spatiotemporal model to balance the two components. Considering that the CV of predicted fish density increases over time due to reduced sample size and spatial coverage, a bias-correction algorithm (Thorson and Kristensen 2016) is applied to remove the re-transformation biases in VAST-derived quantities.

Given that the spatial domain of the CPUE standardization model extends beyond the core fishing ground to encompass locations with relatively sparse CPUE data, the abundance index for this benchmark assessment is subject to greater influence by imputed fish densities for unfished locations. As such, it is crucial to address potential biases associated with the imputation process, particularly in this case where fishery-dependent CPUE data is likely preferentially sampled. As the extent of unfished locations expands over time due to the depletion-induced contraction of the Japanese longline fishery, the positively biased imputation plays an increasingly more important role in the area-weighted abundance index, leading to a hyper-stable abundance index.

The spatiotemporal terms, which describe how the spatial pattern of fish density changes over time, need to be interpolated for each location and time. In the CPUE standardization model developed for the last benchmark assessment, the spatiotemporal terms are assumed to be temporally independent but spatially correlated according to the Matérn function. Thus, the spatiotemporal terms for the unfished eastern EPO are interpolated solely based on data collected from the fished western EPO during the same year-quarter. If there is a spatial pattern in population depletion, this assumption ignores preferential sampling and can lead to positively biased imputations of bigeye density in the eastern EPO.

In the CPUE standardization model developed for this benchmark assessment, spatiotemporal terms are assumed to be correlated in both space and time. Specifically, the spatiotemporal terms are assumed to be spatially correlated according to the Matérn function and to follow a random-walk process in time. Under this assumption, the spatiotemporal terms for the unfished eastern EPO are interpolated based on data collected not only from the fished western EPO in the same year-quarter but also from the eastern EPO in adjacent fished years.

2.5.3. Standardized index of relative abundance

The spatiotemporal model achieves convergence with a positive-definite Hessian and a maximum gradient of 0.00054. The quantile-quantile plot indicates that the CPUE standardization model for bigeye tuna in the EPO fits well to the Japanese longline CPUE dataset (Figure 10). The spatiotemporal model estimates a dome-shaped effect of the HBF on the catchability of bigeye tuna in the Japanese longline fishery (Figure 11). The catchability is estimated to be highest at a HBF of approximately 17 and decreases when the HBF falls below or above this value. The standardized longline index of abundance reveals a declining trend in the abundance of large bigeye from the start of the time series until 2010, followed by a relatively stable level of bigeye abundance since 2011 (Figure 12a). However, due to a reduction in sample size and a contraction in spatial coverage of the CPUE data (Figure 7), the coefficient of variation (CV) of the abundance index has increased rapidly since 2020 when the last benchmark assessment for bigeye tuna in the EPO was conducted (Figure 12b). Consequently, the new (2020-2023) information that the longline index of abundance provides regarding the abundance trend is subject to large uncertainty. The primary factor contributing to the large index CV since 2020 is the lack of CPUE data in the western and especially in the eastern equatorial EPO regions (Figure 13).

The CV of the standardized abundance index is originally estimated by VAST based on sample size and sample distribution. The stock assessment model cannot account for all sources of process error, so the input index CV for the stock assessment model usually needs to be rescaled to a higher level than that estimated by the CPUE standardization model. For this benchmark assessment, the scaler is estimated internally by an age-structured production model that estimates recruitment deviations. The mean CV of the standardized abundance index is estimated to be 0.124.

2.6. Size compositions

2.6.1. Purse-seine fishery fleets

The length frequency data for the purse-seine fisheries are collected 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 (Suter 2010, Tomlinson 2002), which specifies the strata from which samples are collected: fish-carrying capacity of the vessel, set type, month, and area of catch (area definition can be found 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 2010; Vogel, 2014). The sampling coverage in terms of the percentage of the catch is lower (SAC-02-10). The sampling areas were designed for yellowfin tuna before the development of the OBJ fishery. Since 2000, both the 5° cell and the sampling area have been recorded for most samples (Lennert-Cody et al. 2012); the 5° cell has been recovered for many samples before 2000. Ideally, fifty fish of each species in the sampled well were measured, and samplers have alternated between counting fish by species and measuring fish for length since 2000. 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 WSBET-02-06 and the Appendix of Suter (2010).

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 is then converted to the 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 estimates of proportions at length make use of both the species counts and the length-measurement data. Details of the estimators can be found in WSBET-02-06. The staff developed a design-based algorithm (Best Scientific Estimates or BSE) to calculate length compositions for each purse-seine fishery fleet. This algorithm has been integrated into a R package *BSE* that can be accessed at: https://github.com/HaikunXu/BSE. The input sample size of purse-seine length composition data is specified to be the number of wells sampled to indirectly account for over-dispersion in length composition data (Figure 14). The size compositions with an input sample size of less than five wells are removed from this benchmark assessment.

2.6.2. Longline fishery fleets

2.6.2.a Data source

In the last benchmark assessment, the computation of length composition data for longline fishery fleets relies solely on length composition data from Japanese commercial longline vessels. However, concerns have been raised about the representativeness of the Japanese longline length composition data collected in recent years. The contribution of Japanese longline catch to the total longline catch has continuously decreased over time from nearly 100% before 1985 to less than 25% since 2017. Furthermore, both the spatial coverage (Figure 15) and the sample size (Figure 16) of the longline length composition data from Japan have decreased notably since the 2010s. As the composition data for fishery fleets should be weighted spatially by catch amount, it is reasonable to expand the source of composition data for longline fishery fleets to other CPCs.

In this benchmark assessment, we also include longline length composition data collected by Korean observers to provide joint length frequencies for longline fishery fleets. There are several reasons for choosing Korean observers' data as the additional source of composition data for longline fishery fleets. Firstly, Korea has recently replaced Japan as the most important longliner for bigeye tuna in the EPO. Secondly, our comparison shows no noticeable difference between the length composition data collected, for bigeye tuna in the same spatiotemporal window, by Japanese and Korean observers. The comparison also shows a pronounced difference between the length compositions collected by Korean fishers and observers for bigeye tuna in the EPO. This finding, which is supported by a previous SAC information paper (SAC-11 INF-K), is the main reason for not including length composition data collected by Korean fishers in this benchmark assessment. Lastly, a large portion of the grids where Korean length composition data are available are not covered by Japanese length composition data (Figure 15). Expanding the data source to multiple fleets allows for a more complete spatial coverage of the longline fishing ground in the EPO.

The Japanese longline length composition data for bigeye tuna in the EPO covers the period between 1986 and 2023. All length compositions before 2011 and after 2015 were collected by fishers and on-board observers, respectively. Between 2010 and 2015, there was a rapid transition of the data source from 100% fishers to 100% on-board observers. Length measurements from the Japanese longline fleet were recorded at various spatial resolutions and bin sizes. This benchmark assessment includes only those collected at a spatial resolution of 1° x 1° and a bin size of 1, 2, or 5 cm. The longline length composition data, collected by Korean observers at a spatial resolution of 1° x 1° and a bin size of 1cm, covers the period between 2013 and 2023. Different from the Japanese length composition data that spread out across the EPO, the Korean length composition data covers only the offshore EPO (Figure 15). For the three longline fisheries located in the offshore EPO (Fisheries 2-4), the Korean data contributed to about half of the longline length composition data since 2016 (Figure 16). Due mainly to the negative impact of the COVID-19 pandemic on the longline observer program, the spatial distribution of longline length composition data was very restricted, and only Korean data was available for bigeye tuna in the EPO after 2020 (Figure 15).

2.6.2.b Standardization procedure

The methodology for computing length composition data for longline fishery fleets has been improved. In the last benchmark assessment, length composition data for longline fishery fleets are computed by spatially raising raw length compositions to catch amount. This methodology has a significant limitation, as a large proportion of longline catches do not contribute to the computation of length frequencies for fishery fleets. This is due to the sparse distribution of longline length composition data in space (Figure 15). As a result, length frequencies computed by raising raw length compositions spatially to catch may not adequately represent fishery removal.

To overcome this issue, we develop length-specific spatiotemporal models to impute length frequency for the catches without corresponding length compositions. This new approach allows the computation of length frequencies for longline fishery fleets based on all, rather than a small percentage, of longline catches. The joint longline length frequencies are based on data collected by Japan and Korea, and the length-specific spatiotemporal model is fitted to Japanese and Korean length composition data simultaneously.

VAST is also chosen as the platform to standardize longline length frequency, which is aggregated across vessel flags (Japan and Korea) by year, month, 1° latitude, and 1° longitude. We specify VAST to use the logit and log link functions for the linear predictors of encounter probability and positive catch rate, respectively, for each length bin. Both linear predictors include an intercept (year-quarter) term, a time-invariant spatial term, and a time-varying spatiotemporal term. All three terms are assumed to be independent and identically distributed among length bins. Of the three terms, the intercept term is estimated as fixed effects and the other two terms are estimated as random effects. The spatial and spatiotemporal random effects are both assumed to be autocorrelated in space according to the Matérn function. Neither the catchability covariate term nor vessel effects term is included in this model because they are not available in this dataset. This VAST model also treats the four quarters equally.

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

The length compositions of a fishery fleet are catch raised within the spatial domain of the fishery. Specifically, the length frequency for a fishery fleet ($LF(F)$) in time t and length l is computed as:

$$LF(F)_{t,l} = \frac{\sum_s (c_{s,t} \times lf_{s,t,l})}{\sum_l \sum_s (c_{s,t} \times lf_{s,t,l})} \quad (Equation\ 1)$$

where c_s is the fleet-specific total catch in grid s and time t, and $lf_{s,t,l}$ is the length frequency in grid s, time t, and length l predicted by the length-specific spatiotemporal model. The fleet-specific total catch, reported in the number of fish, is extracted from the IATTC's database and has a spatial resolution of 5° x 5°. To match with this spatial resolution, we aggregate the predicted length frequencies from the length-specific spatiotemporal model from 1° x 1° to 5° x 5°. The longline length composition data are spatiotemporal model-based, to be consistent we also use model-based input sample size for the longline length composition data. Specifically, the input sample size is calculated by the length-specific spatiotemporal model to approximate the estimated imprecision for predicted length frequency (Thorson and Haltuch 2018). The size compositions with an input sample size of less than 100 fish are removed from this benchmark assessment.

2.6.3. Longline survey fleet

In addition to CPUE data, there is a need to standardize the composition data associated with the abundance index (Maunder et al. 2020a). The composition data for the survey fleet should represent the condition of the entire population. However, only a small portion of CPUE data has the corresponding composition data, indicating that composition data is distributed more sparsely in space than CPUE data. Survey length compositions should be spatially weighted by CPUE, so the spatiotemporal fields of both length frequency and fish abundance are necessary. The spatiotemporal field of fish abundance can be extracted from the spatiotemporal model that has been developed to provide the index of relative abundance. The spatiotemporal field of length frequency can be extracted from the spatiotemporal model that has been developed to provide size compositions for longline fishery fleets.

The length compositions of the survey fleet are CPUE-raised and area-weighted across the EPO. Specifically, the length frequency for the survey fleet ($LF(S)$) in time t and length l is computed as:

$$LF(S)_{t,l} = \frac{\sum_s (a_s \times d_{s,t} \times lf_{s,t,l})}{\sum_l \sum_s (a_s \times d_{s,t} \times lf_{s,t,l})} \quad (Equation\ 2)$$

where a_s is the area of grid s, and $d_{s,t}$ is the fish density in grid s and time t predicted by the spatiotemporal model for CPUE standardization, and $lf_{s,t,l}$ is the length frequency in grid s, time t, and length l predicted by the spatiotemporal model for length-frequency standardization. The input sample size of this composition data is estimated by the length-specific spatiotemporal model based on the method described in Thorson and Haltuch (2018). The size compositions with an input sample size of less than 500 fish are removed from this benchmark assessment.

2.7. 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 measurements of 254 bigeye tuna caught by purse-seine vessels in the EPO between 2000 and 2004 (Figure 17). 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 because otolith daily increments for large/older fish are very difficult to read and have not been validated by OTC-marking experiments. Schaefer and Fuller (2006) found no statistical difference between the growth models fitted to male and female otolith readings.

3. ASSUMPTIONS AND PARAMETERS

3.1. Biological and demographic information

3.1.1. Growth

Specifying the growth curve in the stock assessment of bigeye tuna remains challenging. Age-at-length data derived from otolith readings are available for fish up to four years of age (Schaefer and Fuller 2006). This is a narrow spectrum of ages of longevity of at least 15-16 years estimated from tagging studies (Langley *et al*. 2008). Otolith daily increments for large (old) fish are 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.

This benchmark assessment uses the growth cessation model for bigeye tuna, which is different from the Richards growth curve used in the last benchmark assessment (Figure 18). In the last benchmark assessment, the Richards growth curve (Schnute 1981) was estimated from an integrated model developed by Aires-da-Silva et al. (2015). This integrated model incorporates both otolith age-at-length data and length-increment tagging data into the estimation of growth parameters, so it improves the estimates of growth parameters than the model based on otolith age-at-length only. The main reason for updating the growth curve in this benchmark assessment is that the growth cessation model fits better than the Richards model to the tagging data for large bigeye (Maunder et al. 2018). In comparison to the old growth curve, the new growth curve suggests a larger and smaller length for bigeye below and above 30 quarters, respectively (Figure 18).

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. 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 estimated externally. The reason for estimating the standard deviation of length at age 0 internally by the assessment model is that the externally estimated value appears to be too small, as suggested by the poor fit of the length-composition data from the floating-object fisheries at small sizes.

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

$$w = 3.661 \times 10^{-5} \times l^{2.90182} \quad (Equation\ 3)$$

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

3.1.2. Natural mortality

Age-specific vectors of natural mortality (*M*) are assumed for bigeye in the EPO. The last benchmark assessment used sex-specific models and a natural mortality vector was provided for each sex. For both sexes, *M* was assumed to be 0.25 at age 0 and to decrease to 0.1 at 5 quarters of age (Figure 19). Female *M* was assumed to increase to 0.143 after the fish reach maturity. 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.

In this benchmark assessment, the *M* for juvenile bigeye was changed to the Lorenzen curve based on the good practice recommendation from some recent publications (Lorenzen 2022, Lorenzen et al. 2022). A new Lorenzen *M* curve was planned to be estimated using a cohort analysis approach with the new tagging data for bigeye. However, the cohort analysis model did not converge mainly because the reporting rates for both longline fisheries and tagging data before 2020 are unknown. Alternatively, the *M* for juvenile bigeye was calculated as a function of length (l) based on the curve that Lorenzen et al. (2022) provided:

$$M_l = M_{l50} \times \left(\frac{l}{l_{50}}\right)^{-1} \quad (Equation\ 4)$$

where $l50$ is the length at 50% maturity, and M_{l50} is fixed at the value used in the last benchmark assessment (i.e., 0.1). The *M* for adult bigeye remains the same as that used in the last benchmark assessment. The *M* for juvenile bigeye in this benchmark assessment is considered to be better than that in the last benchmark assessment because it follows the current good practice recommendation derived from scientific research and fits better to the *M* of Hampton (2000) (Figure 19).

3.1.3. Recruitment

It is assumed that bigeye can be recruited to the fishable population 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 (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 steepness, defined as the fraction of virgin recruitment that is produced if spawning biomass is reduced to 20% of its unexploited level. Steepness 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, estimating steepness in the assessment model is challenging due to a lack of contrast in spawning biomass, and due to other factors, like environmental influences, can cause recruitment to be extremely variable. If steepness is estimated as a free parameter in the model, it is estimated to be 1. However, simulation analyses have shown that steepness is frequently estimated to be 1 even when the true value is lower (Lee et al. 2012). In this assessment, three values of steepness (0.8, 0.9, and 1.0) are considered to account for this axis of uncertainty. The quarterly recruitment deviates are specified to have a standard deviation of 0.6. It is also important to note that the method proposed by Methot and Taylor (2011) is used to provide bias adjustment for recruitment.

3.1.4. Selectivity and data weighting

In this benchmark assessment, the decision regarding how to specify selectivity (including the form of the curve, whether to estimate the curve, and whether to add time blocks to the curve) and how to weight composition data is guided by a decision tree developed by the staff (Figure 20). Simulation studies in Privitera-Johnson et al. (2022) found that the double-normal selectivity is most robust to uncertainty in selectivity form, so all fishery and survey fleets in this benchmark assessment have double-normal selectivities. For each fleet, there are three options regarding the combination of selectivity and data weighting. Fleets with high catch amounts, can fit a double-normal selectivity curve to their composition data closely, and having rich composition data should use time-varying selectivity along with the Francis weighting (TA1.8 in Francis (2011)). Fleets with low catch amounts, unable to fit a double-normal selectivity curve to fit to its composition data closely, or having poor composition data, should not estimate selectivity (e.g., fix or mirror selectivity) and should not fit to its composition data. Fleets not falling into either category above should use constant selectivity and 20% of the Francis weight.

The decision regarding selectivity and data weighting for each fleet included in the assessment model is outlined in Table 5. Size compositions for fishery fleets are spatially weighted by catch within the respective area of operation. Here, fishery selectivity is defined as the combination of gear selectivity and availability, so any variation in fish availability or fleet distribution over time can result in time-varying fishery selectivity. In contrast, size compositions for the survey fleet are spatially weighted by fish abundance across the EPO, allowing survey selectivity to be treated as gear selectivity and approximately constant over time. In this assessment model, the trend of population abundance is primarily dictated by the index of relative abundance, whereas the scale of population abundance is heavily influenced by composition data. Mis-specifying selectivity can thus lead to a biased estimation of the population scale. The key philosophy behind the design of the decision tree is that the estimation of the population scale should rely mainly on the composition data of the survey fleet rather than that of fishery fleets, as the degree of mis-specification tends to be higher for time-varying fishery selectivity than time-invariant survey selectivity.

In theory, all data-rich fishery fleets should use time-varying selectivity to minimize the extent of selectivity mis-specification and consequently improve estimation accuracy (Martell and Stewart 2014, Xu et al. 2018b). This, however, will lead to estimating a substantial amount of additional selectivity parameters in this assessment model, which includes twenty-two fishery fleets. Considering the trade-off between estimation accuracy and model efficiency/stability, time-varying selectivity with consecutive decadal time blocks is only applied to fishery fleets with high catch amounts, rich composition data, and can fit a double-normal selectivity curve to its composition data closely (Table 5). Of those fleets, purse-seine fleets have three selectivity time blocks (<2000; 2000-2010; and >2010) and longline fleets have three selectivity time blocks (<1993; 1993-2010; and >2010). All other fishery fleets for which selectivity is estimated use time-invariant selectivity, with their composition data down-weighted by 80% to greatly reduce their influence on the scale of estimated population abundance (Table 5). The two NOA fishery fleets have low catch amounts and poor composition data, so their selectivities mirror those of two OBJ fisheries with similar observed length frequencies and their composition data are excluded from the assessment model.

4. BRIDGING ANALYSIS

A bridging analysis is conducted to illustrate the impacts of new changes on model results, using the "base" reference model (Env-Fix; see Table 2 in SAC-11-06 for model definition) from the last benchmark assessment as the platform. This reference model assumes that only one longline fishery fleet during 2011-2023 has an asymptotic selectivity and that both growth and natural mortality are known. Given that a bridging analysis was previously conducted in the exploratory analysis (SAC-14-05) to evaluate the effects of prior changes on model results, the current bridging analysis focuses on three key changes implemented after the exploratory analysis in 2023. To evaluate the impact of each change on model results, the three additional changes are introduced progressively into the assessment model in a stepwise manner. Consequently, a total of four models, including the "base" reference model from the last exploratory assessment (M0), are compared in this bridging analysis.

The first change (M1) is about selectivity and data weighting. In the exploratory analysis, the selecitivities of longline and floating-object fisheries are time-varying and constant, respectively. Additionally, the length compositions of all fishery fleets are weighted using the Francis method. In this benchmark assessment, the specifications of selectivity and data weighting of fishery fleets are determined by a decision tree developed by the staff (see Figure 20 and Table 5).

The second change (M2) is about the growth curve. In the exploratory analysis, the growth curve is assumed to be the Richards growth curve (Schnute 1981) that was estimated by Aires-da-Silva et al. (2015). In this benchmark assessment, the Richards growth curve is replaced by the growth cessation model (Maunder et al. 2018).

The third change (M3) is about *M*. In the exploratory analysis, the M vector is from a broken-stick model estimated by Aires-da-Silva and Maunder (2008) to approximate the M estimates from Hampton (2000). In this benchmark assessment, the M vector for juveniles (smaller than the length at 50% maturity) is calculated according to the Lorenzen natural mortality curve with a shape parameter of -1 (Lorenzen et al. 2022). Namely, *M* for juveniles is proportional to the inverse of body length.

Trajectories of estimated spawning biomass, spawning biomass ratio, and relative recruitment are used to assess the impact of the new changes on model results (Figure 21). The new selectivity and data weighting approach results in a slightly reduced scale of spawning biomass and spawning biomass ratio (M1 vs. M0). Updating the growth curve has a negligible impact on both spawning biomass and spawning biomass ratio (M2 vs. M1). Updating the *M* vectors leads to notably lower spawning biomass while spawning biomass ratio remains almost identical (M3 vs. M2). Additional, Updating the *M* vectors leads to a notable reduction in the degree of the regime shift in recruitment.

5. REFERENCE MODELS

5.1. Hypotheses for risk analysis

In this report, the latest version (3.30.22.beta) 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 is capable of accommodating models of varying complexity and fitting to diverse types of data. Similar to the previous benchmark assessment, a risk analysis framework is developed to incorporate uncertainties regarding several assumptions in the assessment model, explicitly considering them in the evaluation of stock status and formulation of management advice.

The first step in implementing the risk analysis is establishing plausible hypotheses that define the states of nature associated with the main sources of uncertainty. In the last benchmark assessment, the overarching hypothesis for bigeye tuna aimed to explain the apparent regime shift in recruitment estimates that coincided with the expansion of the floating-object fishery in the EPO. However, the degree of the regime shift in recruitment estimates from the new assessment models is reduced greatly, from 140% to 20% for the base reference model. As such, this overarching hypothesis is not included in this benchmark assessment. This significant decrease of the regime shift in recruitment results from the combination of changes made to the assessment model. Among these changes the most influential in reducing the regime shift are adding one more time block to the selectivity of longline fishery fleets in 2011, improving the CPUE standardization model, and using the Lorenzen natural mortality curve for juvenile bigeye. The reference models in this benchmark assessment address three major uncertainties within a hierarchical framework: (1) the misfit to the length-composition data for the longline fishery that is assumed to have asymptotic selectivity; (2) the degree of effort creep in the longline fishery; and (3) the steepness of the stock-recruitment relationship.

Level 1 hypothesis: Four models are included to address the misfit to the composition data for the longline fishery that is assumed to have asymptotic selectivity: (1) ignore the issue (Fix); (2) estimate the growth curve with a prior on L_{inf} (Gro); (3) estimate a dome-shape selectivity curve for the longline fishery that is assumed to have asymptotic selectivity (Sel); and (4) estimate the scaler of the natural mortality vector (Mrt). Additionally, a model where growth rather than natural mortality is sex-specific is explored to address the misfit, but it is not included for this hypothesis due to its worse misfit compared to the model where the issue is ignored (Fix). The four reference models are equally weighted. The decision to equally weight the four models is made based on the outcome of the two risk analysis workshops organized by the IATTC (WSRSK-01-RPT and WSRSK-02-RPT).

Level 2 hypothesis: Three levels of annual increasing rate of the longline catchability for bigeye are included to address the uncertainty in effort creep. Considering that bigeye is the main target species of the Japanese longline fishery in the EPO, its catchability in this fishery is expected to increase owing to advancements in fishing skill and technology. The review panel suggests considering a 1% annual increase in the catchability of bigeye in the longline fishery (RVMTT-01-RPT). Based on this recommendation, three annual increases in longline catchability (0%, 1%, and 2%) are considered to address this uncertainty, each equally weighted.

Level 3 hypothesis: Three steepness values (1.0, 0.9, and 0.8) are included to address the uncertainty in the shape of the stock-recruitment relationship. The three steepness values are weighted based on expert judgement from the risk analysis for the last benchmark assessment (SAC-11 INF-F).

In total, the combination of the three hypothesis yields $4 \times 3 \times 3 = 36$ reference models. All reference models are structured with quarterly time steps from the first quarter of 1979 to the last quarter of 2023. 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. Additionally, they are sex-structured models with sex-specific natural mortality. These reference models are fitted to indices of relative abundance and size compositions (and also age compositions for the reference models which estimate growth (Gro)) by maximizing 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 (without 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 deviation of 0.01.

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 through the last quarter of 2023.
3. Virgin recruitment.
4. Selectivity parameters for fisheries and surveys. In this assessment, the double-normal selectivity option is chosen for all asymptotic and dome-shaped selectivity curves. The number of

parameters estimated for the asymptotic and dome-shaped selectivity curves is 2 and 4, respectively.

5. 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 24 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. Natural mortality.

5.2. Assessment results

5.2.1. Model convergence

Of the thirty-six reference models for bigeye, thirty-three models converged with small maximum gradients, positive definite Hessians (Table 6). The results of the three reference models that did not converge with a positive definite Hessian (Model Fix with an effort creep of 0% and a steepness of 0.8 and Model Mrt with an effort creep of 1% and a steepness of 0.9 or 0.8) are not shown in this section.

5.2.2. Parameter estimates

The estimated growth curve, selectivity of Fishery 4, and natural mortality are shown in Figures 22-24. The difference in the growth curve is larger for old bigeye than young bigeye (Figure 22). The difference in the selectivity of Fishery 4 is negligible on the left branch of the dome but gets increasingly larger on the right branch of the dome (Figure 23). Model Mrt estimates a larger natural mortality than that fixed in other models (Figure 24). Specifically, the natural mortality for adult male is fixed at 0.1 while is estimated to be 0.13, 0.125, and 0.12 by Model Mrt that has a catchability assumption of 0%, 1%, and 2% annual increase, respectively.

5.2.3. Recruitment

The time series of annual recruitment estimates have several important features (Figure 25): (1) recruitment estimates are no sensitive to the assumptions regarding effort creep and, especially, regarding steepness; and (2) there is no pronounced regime shift in recruitment that is coincided with the expansion of the floating-object fishery. To quantify the extent of the regime shift in recruitment, the exponential difference between the mean recruitment deviation in the periods 1994-2023 and 1979-1993 was calculated (Table 6). Across all reference models in this benchmark assessment, minor degrees of regime shift in recruitment are observed, ranging from -20% to 27%. Consequently, these models are retained in the risk analysis based on this diagnostic.

5.2.4. Spawning biomass

The estimates of spawning biomass and spawning biomass ratio (the ratio of the spawning biomass of the current stock to that of the unfished stock) show considerable variability both within and across reference models. In general, these estimates demonstrate greater sensitivity to the degree of effort creep compared to that of steepness (Figure 26). This finding is anticipated, given that the longline index of abundance, which is directly influenced by the level of effort creep, is the most important indicator of spawning biomass.

5.2.5. Fishing mortality (*F*)

There have been notable fluctuations in the level of fishing mortality (*F*) for bigeye in the EPO. Across all reference models, it is evident that the *F* of bigeye aged less than 9 quarters experienced a substantial increase from close to zero before 1993 to historically high levels in 2020, followed by a paid decline thereafter (Figure 27). In comparison, the *F* of bigeye aged more than 12 quarters remained relatively stable since 1993, with a decreasing trend observed from around 2015 onwards. Additionally, all reference models indicate that the *F* of bigeye aged less than 9 quarters is notably lower compared to individuals 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 28). The discards of small bigeye in the floating-object fishery have a small, but detectable, impact on the depletion of the stock.

5.3. Diagnostics

5.3.1. Jitter analysis

A Jitter analysis is conducted for each reference model to evaluate whether the negative log-likelihood of the reference model has reached the global minimum. Due to time constraints, we compare only ten jitter runs with a jittering value of 0.02 for each reference model. This is the first diagnostic analysis conducted for each reference model to ensure that the reference model has converged at the global maximum likelihood estimation. All the thirty-three converged reference models pass the jitter diagnostics, confirming that the optimization process successfully identified the global minimum and that the results are reliable.

5.3.2. Fit to longline index of relative abundance

The longline index of relative abundance directly informs the population trend of large bigeye so it is critical to check whether each reference model fits closely to the longline index without an obvious residual pattern. Comparison of observed and predicted longline index suggests that the four models included for the first hypothesis fit well to the longline index under different assumptions of increasing rates in catchability (Figure 29).

5.3.3. Fit to longline composition data

To assess the fit of each reference model to the composition data for fishery 4, the predicted and empirical selectivity curves for fishery 4 are compared across all models included in the level 1 hypothesis. Empirical selectivity is calculated as the average observed catch at length divided by the average predicted population number at length from the assessment model. To facilitate the comparison, the empirical selectivity is scaled to have a maximum value of 1. When the assessment model fits well to a fishery's composition data, the two selectivity curves should closely align. This diagnostic analysis indicates that estimating a dome-shaped selectivity (Model Sel), growth (Model Gro), or natural mortality (Model Mrt) can improve the fit to the composition data for fishery 4. However, there remains a noticeable discrepancy between the empirical and predicted selectivity for fishery 4 in these three models (Figure 30).

5.3.4. Retrospective analysis

Retrospective analysis serves as a valuable tool for assessing the consistency of a stock assessment model is from one year to the next (Mohn 1999). Inconsistencies detected through the retrospective analysis can often indicate inadequacies in the model. Typically, this analysis is carried out by progressively eliminating the last year's data from the analysis while maintaining the same methodology and assumptions. It allows for an examination of how the inclusion of additional data impacts the resulting estimates of population attributes and management quantities. 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 benchmark assessment, a retrospective analysis is conducted by iteratively removing the data from the last year five times, and Mohn's rho for spawning biomass is calculated to quantify the extent of the retrospective pattern. Since the estimates of spawning biomass are found to be insensitive to the value of steepness explored in this benchmark assessment (Figure 26), this diagnostic analysis was only performed for the reference models with a steepness of 1. All the reference models with a steepness of 1 have positive Mohn's rhos, indicating a tendency to overestimate the spawning biomass in the terminal year. However, the values of Mohn's rho indicate that the degree of overestimation of the terminal year's spawning biomass is very small across all reference models (4-5%; Table 6). The retrospective patterns for spawning biomass, spawning biomass ratio, and recruitment can be found in Figures A1-A3.

5.3.5. 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 estimates of spawning biomass from the ASPM with and without recruitment deviates can be found in Figure A4.

5.3.6. R_0 likelihood profile

Virgin recruitment (R0), 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 R0 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 R0 is referred to as the R0 likelihood profile (Wang et al. 2009). The R0 likelihood profile is a diagnostic tool widely used to compare the influence of composition data and indices of relative abundance on absolute abundance. The R0 likelihood profile for each reference model with a steepness of 1 can be found in Figure A5.

6. 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.

6.1. Definition of reference points

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

6.1.1. Limit reference points

The spawning biomass limit reference point (S_{Limit}) is the threshold of S that should be avoided as 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, $S_{\text{current}}/S_{\text{Limit}}$, 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, $F_{\text{current}}/F_{\text{Limit}}$, 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.

6.1.2. 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 ($S_{\text{MSY\_d}}$) within five years or two generations. Here, $S_{\text{MSY\_d}}$ 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 $S_{\text{MSY\_d}}$ used to compute reference points for bigeye is the last quarter's S in the projection period. To provide management advice, $S_{\text{current}}/S_{\text{MSY\_d}}$, 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 $F_{\text{current}}/F_{\text{MSY}}$), are included in the management table.

The fishing mortality target reference point 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, $F_{\text{current}}/F_{\text{MSY}}$, 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 $F_{\text{current}}/F_{\text{MSY}}$ (F multiplier).

In the Kobe trajectory plot, the time series of $S_{\text{MSY\_d}}$ is computed based on two approximations: (1) $S_{\text{MSY\_d1}} = S_{0\_\text{d}}(S_{\text{MSY}}/S_0)$, where S_{0_d} is the dynamic spawning biomass in the absence of fishing and $S_{\text{MSY\_d}}/S_0$ is the depletion level that, under equilibrium, produces the MSY; (2) $S_{\text{MSY\_d2}}$, which is derived by projecting the population into the future under historical recruitment (bias adjusted) and $F = F_{\text{MSY}}$. The two approximations are weighted as follows to obtain the trajectory of $S_{\text{MSY\_d}}$ (t) in the Kobe plot:

$$S_{MSY_d}(t) = \big(1 - p(t)\big)S_{MSY_{d1}}(t) + p(t)S_{MSY_{d2}}(t) \quad (Equation\ 5)$$

where p increases linearly as a function of year (t) from 0 in the first year to 1 in the last 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 $S_{\text{MSY\_d}}$.

6.2. Estimates of stock status

According to the thirty-three converged reference models in this benchmark assessment, the spawning biomass of bigeye at the beginning of 2024 ranges from 45% to 292% of the spawning biomass at dynamic MSY and the fishing mortality of bigeye in 2021-2023 ranges from 42% to 136% of the fishing mortality at MSY (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 31).

According to the thirty-three converged reference models in this benchmark assessment, the spawning biomass of bigeye at the beginning of 2024 ranges from 134% to 346% of the spawning biomass at the limit level and the fishing mortality of bigeye in 2021-2023 ranges from 35% to 80% of the fishing mortality at the limit level (Table 7).

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 floating-object fishery expanded in 1993, the MSY was much greater than the current level (Figure 32). Since 1993, the ratio of floating-object catch to longline catch has increased persistently, causing the MSY to continue declining to the lowest level in 2022.

6.3. Joint probability and cumulative distribution functions for management quantities

Based on the estimates of management quantities related to the target reference points ($F_{current}/F_{MSY}$ and $S_{current}/S_{MSY\_d}$) and the associated CV from the thirty-three reference models (Figure 31), we calculate the joint probability and cumulative distribution functions for these two management quantities for bigeye. Each reference model's weight is equal to the product of the weights associated with the model, catchability assumption, and steepness assumption it represents. Since three of the thirty-six reference models are excluded from the final reference model pool due to poor convergence, the weights of remaining reference model need to be rescaled so that their sum across the thirty-three selected reference models equals 1. Specifically, the weight for steepness is rescaled scaled to sum to 1 for each combination of model and catchability assumption.
The joint distribution functions for both $F_{current}/F_{MSY}$ and $S_{current}/S_{MSY\_d}$ are unimodal (Figure 33). There is a 24.7% probability that $F_{current}$ is above F_{MSY} and a 46.6% probability that $S_{current}$ is below S_{MSY_d}. Among the four models included for the level 1 hypothesis, Models Fix and Gro are more pessimistic, estimating higher $F_{current}/F_{MSY}$ and lower $S_{current}/S_{MSY\_d}$ compared to Models Sel and Mrt (Figure 34). As expected, lower steepness values and higher rates of increase in longline catchability correspond to more pessimistic estimates of stock status: lower *S* and higher *F* relative to the reference points (Figure 34).

We also calculate the joint probability and cumulative distribution functions for the management quantities related to the limit reference points ($F_{current}/F_{Limit}$ and $S_{current}/S_{Limit}$). The joint distribution functions for both $F_{current}/F_{Limit}$ and $S_{current}/S_{Limit}$ are also unimodal (Figure 35), suggesting that there is a 0.1% probability that $F_{current}$ is above F_{Limit} and a 0.2% probability that $S_{current}$ is below S_{Limit}.

The trends of $F_{current}/F_{MSY}$ and $S_{current}/S_{MSY\_d}$ are similar among the four models included in the level 1 hypothesis (Figure 36). Models Fix and Gro estimate that $F_{current}/F_{MSY}$ surpassed 1 in around 2010, reached historically high levels in 2020, and decreased three years in a row to about 1 in 2023. Consequently, these two models estimate that $S_{current}/S_{MSY\_d}$ dropped below 1 since about 2010 and recovered slightly after 2020 due to the decrease in $F_{current}/F_{MSY}$ during the same time, whereas it is still below 1 in 2023. In comparison, Models Sel and Mrt estimate that $F_{current}/F_{MSY}$ reached historically high levels in 2020, which are slightly below 1, and decreased thereafter to below 0.7 in 2023. Consequently, these two models estimate that $S_{current}/S_{MSY\_d}$ decreased to historically low levels in 2019, which are above 1, and increase thereafter to above 1.3 in 2023.

In the last benchmark assessment, $F_{current}/F_{MSY}$ for the *status quo* period (2017-2019) was highly uncertain due to the apparent bimodal pattern in the joint distribution function. However, based on the new reference models in this benchmark assessment, the joint distribution function for $F_{current}/F_{MSY}$ in 2017-2019 is unimodal, suggesting that there is a 58.5% probability that the fishing mortality in 2017-2019 is higher than F_{MSY} (Figure 37).

6.4. 10-year projection under the current fishing mortality

There is a 24.7% probability that the fishing mortality in 2021-2023 was above the MSY level, so the spawning biomass is expected to increase to be above the MSY level if future fishing mortality remains at the current level. To assess the potential increase in future spawning biomass, each reference model conducts a 10-year projection under the current fishing mortality.

The projection ensemble indicates that the spawning biomass ratio will range from 0.133 to 0.395 at the beginning of 2034 (Figure 38). Among the four models included in the level 1 hypothesis, the weighted spawning biomass ratios at the beginning of 2034 for model Fix and Gro are relatively pessimistic (0.201 and 0.232, respectively) and those for Model Mrt and Sel are relatively optimistic (0.318 and 0.330, respectively). The weighted value across all reference models suggests that, under the current fishing mortality, there is a 50% probability that the spawning biomass ratio at the beginning of 2034 will be above 0.270.

7. FUTURE DIRECTIONS

7.1. Collection of new and updated information

The staff will continue the biological and tagging studies to improve the understanding of the biology of bigeye in the EPO, especially the growth and length-weight relationship of bigeye that can significantly impact assessment results. The staff also intends to continue its collaboration with Asian CPCs to improve the longline index of abundance for bigeye tuna. Due to the pronounced decrease in the spatial coverage of the Japanese longline fleet, the CV of the longline index of abundance for bigeye tuna has increased rapidly since 2020. Consequently, the longline index of abundance does not provide precise information on the temporal change of population abundance in the last three years. There is a strong need to combine CPUE data from multiple CPCs to increase the spatial coverage and sample size of longline data for CPUE standardization. To achieve this, the staff needs to get access to high-resolution longline CPUE data from the Asian CPCs simultaneously for at least three months.

7.2. Refinements to the assessment model and methods

The staff will continue improving the assessment model for bigeye tuna in the EPO. The following changes would be desirable for future assessments:

- Explore spatially-explicit stock assessment models.
- Explore purse-seine indices of abundance.
- Integrate the tagging growth-increment data into the stock assessment model.
- Incorporating estimates of total abundance from methods such as spatiotemporal tagging model and close-kin mark-recapture model into the stock assessment model.

ACKNOWLEDGEMENTS

Many IATTC and CPC staff provided data for the assessment. Japan provided the operational longline catch and effort data that was used to produce the longline index of abundance. IATTC staff members and CPC scientists provided advice on the stock assessment, fisheries, and biology of bigeye tuna.

REFERENCES

Aires-da-Silva, A.M., Maunder, M.N., Schaefer, K.M., and Fuller, D.W. 2015. 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.

Campbell, R.A. 2015. Constructing stock abundance indices from catch and effort data: Some nuts and bolts. Fisheries Research **161**: 109-130.

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.

Chang, S.-K., Hoyle, S., and Liu, H.-I. 2011. Catch rate standardization for yellowfin tuna (Thunnus albacares) in Taiwan's distant-water longline fishery in the Western and Central Pacific Ocean, with consideration of target change. Fisheries Research **107**(1-3): 210-220.

Collette, B.B., Reeb, C., and Block, B.A. 2001. Systematics of the tunas and mackerels (Scombridae). Fish physiology **19**: 1-33.

Conn, P.B., Thorson, J.T., and Johnson, D.S. 2017. Confronting preferential sampling when analysing population distributions: diagnosis and model ‐ based triage. Methods in Ecology and Evolution **8**(11): 1535-1546.

Diggle, P.J., Menezes, R., and Su, T.-l. 2010. Geostatistical inference under preferential sampling. Journal of the Royal Statistical Society Series C: Applied Statistics **59**(2): 191-232.

Ducharme-Barth, N.D., Grüss, A., Vincent, M.T., Kiyofuji, H., Aoki, Y., Pilling, G., Hampton, J., and Thorson, J.T. 2022. Impacts of fisheries-dependent spatial sampling patterns on catch-per-unit-effort standardization: A simulation study and fishery application. Fisheries Research **246**: 106169.

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.

Grüss, A., Walter III, J.F., Babcock, E.A., Forrestal, F.C., Thorson, J.T., Lauretta, M.V., and Schirripa, M.J. 2019. Evaluation of the impacts of different treatments of spatio-temporal variation in catch-per-unit-effort standardization models. Fisheries Research **213**: 75-93.

Hall, M., and Roman, M. 2013. Bycatch and non-tuna catch in the tropical tuna purse seine fisheries of the world. FAO fisheries and aquaculture technical paper(568): I.

Hampton, J. 2000. Natural mortality rates in tropical tunas: size really does matter. Canadian Journal of Fisheries and Aquatic Sciences **57**(5): 1002-1010.

Hoyle, S.D., Campbell, R.A., Ducharme-Barth, N.D., Grüss, A., Moore, B.R., Thorson, J.T., Tremblay-Boyer, L., Winker, H., Zhou, S., and Maunder, M.N. 2024. Catch per unit effort modelling for stock assessment: A summary of good practices. Fisheries Research **269**: 106860.

Hurtado-Ferro, F., Punt, A.E., and Hill, K.T. 2014. Use of multiple selectivity patterns as a proxy for spatial structure. Fisheries Research **158**: 102-115.

IATTC. 2021. The tuna fisheries in the eastern Pacific Ocean. Inter-Amer.Trop. Tuna Comm., 12th Scient. Adv. Com. Meeting: SAC-12-03.

Kristensen, K., Nielsen, A., Berg, C., Skaug, H., and Bell, B. 2016. Template model builder TMB. J. Stat. Softw **70**: 1-21.

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., and Hall, M. 2000. The development of the purse seine fishery on drifting Fish Aggregating Devices in the eastern Pacific Ocean: 1992-1998.

Lennert-Cody, C., and Tomlinson, P. 2010. Evaluation of aspects of the current IATTC port sampling design and estimation procedures for catches of tunas by purse-seine and pole-and-line vessels. Inter-Amer. Trop. Tuna Comm., Stock Assessment Report **10**: 279-309.

Lennert-Cody, C.E., Maunder, M.N., Aires-da-Silva, A., and Minami, M. 2013. Defining population spatial units: Simultaneous analysis of frequency distributions and time series. Fisheries Research **139**: 85-92.

Lennert-Cody, C.E., Minami, M., Tomlinson, P.K., and Maunder, M.N. 2010. Exploratory analysis of spatial–temporal patterns in length–frequency data: An example of distributional regression trees. Fisheries Research **102**(3): 323-326.

Lo, N.C.-h., Jacobson, L.D., and Squire, J.L. 1992. Indices of relative abundance from fish spotter data based on delta-lognornial models. Canadian Journal of Fisheries and Aquatic Sciences **49**(12): 2515-2526.

Lorenzen, K. 2022. Size-and age-dependent natural mortality in fish populations: Biology, models, implications, and a generalized length-inverse mortality paradigm. Fisheries Research **255**: 106454.

Lorenzen, K., Camp, E.V., and Garlock, T.M. 2022. Natural mortality and body size in fish populations. Fisheries Research **252**: 106327.

Majumdar, A., Lennert-Cody, C.E., Maunder, M.N., and Aires-da-Silva, A. 2023. spatio-temporal modeling for estimation of bigeye tuna catch in the presence of pandemic-related data loss using parametric adjacency structures. Fisheries Research **268**: 106813.

Martell, S., and Stewart, I. 2014. Towards defining good practices for modeling time-varying selectivity. Fisheries Research **158**: 84-95.

Matsumoto, T. 2008. A review of the Japanese longline fishery for tunas and billfishes in the eastern Pacific Ocean, 1998-2003. Bull IATTC **24**: 1-187.

Maunder, M.N., and Deriso, R.B. 2014. Proposal for biomass and fishing mortality limit reference points based on reduction in recruitment. IATTC Stock Assessment Report **15**: 193-206.

Maunder, M.N., Deriso, R.B., Schaefer, K.M., Fuller, D.W., Aires-da-Silva, A.M., Minte-Vera, C.V., and Campana, S.E. 2018. The growth cessation model: a growth model for species showing a near cessation in growth with application to bigeye tuna (Thunnus obesus). Marine Biology **165**(4): 76.

Maunder, M.N., and Harley, S.J. 2006. Evaluating tuna management in the eastern Pacific Ocean. Bulletin of Marine Science **78**(3): 593-606.

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., and Punt, A.E. 2004. Standardizing catch and effort data: a review of recent approaches. Fisheries research **70**(2-3): 141-159.

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. 2020a. 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.

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 tropical Pacific Ocean. Inter-Am. Trop. Tuna. Comm. Bull. **22**: 433-582.

Maunder, M.N., Xu, H., Lennert-Cody, C., Valero, J.L., Aires-da-Silva, A., and Minte-Vera, C.V. 2020b. Implementing reference point-based fishery harvest control rules within a probabilistic framework that considers multiple hypotheses. Inter-Amer.Trop. Tuna Comm., 11th Scient. Adv. Com. Meeting: SAC-11 INF-F.

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.

Okamoto, H., and Bayliff, W.H. 2003. A review of the Japanese longline fishery for tunas and billfishes in the eastern Pacific Ocean, 1993-1997. Inter-american tropical tuna commission bulletin **22**(4): 221-431.

Pennino, M.G., Paradinas, I., Illian, J.B., Muñoz, F., Bellido, J.M., López ‐ Quílez, A., and Conesa, D. 2019. Accounting for preferential sampling in species distribution models. Ecology and evolution **9**(1): 653-663.

Privitera-Johnson, K.M., Methot, R.D., and Punt, A.E. 2022. Towards best practice for specifying selectivity in age-structured integrated stock assessments. Fisheries Research **249**: 106247.

Satoh, K., Xu, H., Minte-Vera, C.V., Maunder, M.N., and Kitakado, T. 2021. Size-specific spatiotemporal dynamics of bigeye tuna (Thunnus obesus) caught by the longline fishery in the eastern Pacific Ocean. Fisheries Research **243**: 106065.

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.

Schaefer, K.M., and Fuller, D.W. 2010. Vertical movements, behavior, and habitat of bigeye tuna (Thunnus obesus) in the equatorial eastern Pacific Ocean, ascertained from archival tag data. Marine Biology **157**: 2625-2642.

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

Sun, C.-H., Maunder, M.N., Pan, M., Aires-da-Silva, A., Bayliff, W.H., and Compeán, G.A. 2019. Increasing the economic value of the eastern Pacific Ocean tropical tuna fishery: Tradeoffs between longline and purse-seine fishing. Deep Sea Research Part II: Topical Studies in Oceanography **169**: 104621.

Suter, J.M. 2010. An evaluation of the area stratification used for sampling tunas in the eastern Pacific Ocean and implications for estimating total annual catches.

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.

Thorson, J.T., and Kristensen, K. 2016. Implementing a generic method for bias correction in statistical models using random effects, with spatial and population dynamics examples. Fisheries research **175**: 66-74.

Thorson, J.T., Shelton, A.O., Ward, E.J., and Skaug, H.J. 2015. Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. ICES Journal of Marine Science **72**(5): 1297-1310.

Tomlinson, P. 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.

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.

Winker, H., Kerwath, S.E., and Attwood, C.G. 2013. Comparison of two approaches to standardize catch-per-unit-effort for targeting behaviour in a multispecies hand-line fishery. Fisheries Research **139**: 118-131.

Xu, H., Lennert-Cody, C.E., Maunder, M.N., and Minte-Vera, C.V. 2019. Spatiotemporal dynamics of the dolphin-associated purse-seine fishery for yellowfin tuna (Thunnus albacares) in the eastern Pacific Ocean. Fisheries research **213**: 121-131.

Xu, H., Maunder, M.N., Minte-Vera, C.V., Valero, J.L., Lennert-Cody, C., and Aires-da-Silva, A. 2020. Bigeye tuna in the eastern Pacific Ocean, 2019: benchmark assessment. Inter-Amer.Trop. Tuna Comm., 11th Scient. Adv. Com. Meeting: SAC-11-06.

Xu, H., Minte-Vera, C.V., Maunder, M.N., and Aires-da-Silva, A. 2018a. 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., Thorson, J.T., Methot, R.D., and Taylor, I.G. 2018b. A new semi-parametric method for autocorrelated age-and time-varying selectivity in age-structured assessment models. Canadian Journal of Fisheries and Aquatic Sciences(999): 1-18.

Yalcin, S., Anderson, S.C., Regular, P.M., and English, P.A. 2023. Exploring the limits of spatiotemporal and design-based index standardization under reduced survey coverage. ICES Journal of Marine Science: fsad155.

Zhou, S., Campbell, R.A., and Hoyle, S.D. 2019. Catch per unit effort standardization using spatio-temporal models for Australia's Eastern Tuna and Billfish Fishery. ICES Journal of Marine Science **76**(6): 1489-1504.

TABLES

TABLE 1. The best five-split combination selected by the regression tree algorithm for the longline fishery for bigeye tuna in the eastern Pacific Ocean. The last column shows the percentage of variance in the length-frequency data being explained.
TABLA 1. La mejor combinación de cinco divisiones seleccionada por el algoritmo de árbol de regresión para la pesquería palangrera de atún patudo en el Océano Pacífico oriental. La última columna muestra el porcentaje de varianza explicada en los datos de frecuencia de talla.

Split	Key	Value	Variance explained
Split1	Latitude	15°S	8.07%
Split2	Longitude	105°W	10.91%
Split3	Latitude	5°S	13.01%
Split4	Longitude	90°W	14.12%
Split4	Longitude	130°W	15.22%

TABLE 2. The best four-split combination selected by the regression tree algorithm for the OBJ fishery for bigeye tuna in the eastern Pacific Ocean. The last column shows the percentage of variance in the length-frequency data being explained.
TABLA 2. La mejor combinación de cuatro divisiones seleccionada por el algoritmo de árbol de regresión para la pesquería OBJ de atún patudo en el Océano Pacífico oriental. La última columna muestra el porcentaje de varianza explicada en los datos de frecuencia de talla.

Split	Key	Value	Variance explained
Split1	Longitude	115°W	7.54%
Split2	Latitude	0°	8.83%
Split3	Longitude	100°W	9.89%
Split4	Longitude	140°W	10.46%

TABLE 3. The best split selected by the regression tree algorithm for the NOA fishery for bigeye tuna in the eastern Pacific Ocean. The last column shows the percentage of variance in the length-frequency data being explained.
TABLA 3. La mejor combinación seleccionada por el algoritmo de árbol de regresión para la pesquería NOA de atún patudo en el Océano Pacífico oriental. La última columna muestra el porcentaje de varianza explicada en los datos de frecuencia de talla.

Split	Key	Value	Variance explained
Split1	Longitude	130°W	9.95%

TABLE 4. 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 area definition.

TABLA 4. Flotas pesqueras y de “estudio” definidas para la evaluación de referencia del atún patudo en el OPO. PS = cerco; LL = palangre; OBJ = lances sobre objetos flotantes; NOA = lances no asociados; DEL = lances sobre delfines. Ver la definición de las áreas en la Figura 1.

Fleet Number	Fleet type	Fleet name	Gear	Set type	Area	Catch data	Unit
1	Fishery	LL-n-A1	LL	-	1	Retained catch only	1,000s
2		LL-n-A2			2		
3		LL-n-A3			3		
4		LL-n-A4			4		
5		LL-n-A5			5		
6		LL-n-A6			6		
7		LL-n-A7			7		
8	Fishery	LL-w-A1	LL	-	1	Retained catch only	tons
9		LL-w-A2			2		
10		LL-w-A3			3		
11		LL-w-A4			4		
12		LL-w-A5			5		
13		LL-w-A6			6		
14		LL-w-A7			7		
15	Fishery	OBJ-A1	PS	OBJ	1	Retained catch + discards (inefficiency)	tons
16		OBJ-A2			2		
17		OBJ-A3			3		
18		OBJ-A4			4		
19		OBJ-A5			5		
20		OBJ-disc-EPO			1-5	Discards (size-sorting)	tons
21	Fishery	NOADEL-A1	PS	NOA+DEL	1	Retained catch + discards (all)	tons
22		NOADEL-A2			2		
23	Survey	LL-survey-EPO	LL	-	2-7	-	-

TABLE 5. The decisions for selectivity and composition data weighting according to each fishery's catch amount and composition data quality. The rules on which this decision table is based are illustrated as a flowchart in Figure 20. Column "Double-normal" indicates whether the length compostion data of the fleet can be fit well in the assessment model by using a double-normal selectivity curve. Column "Data quality" indicates the relative quality of the fleet's length composition data. Column "Time blocks" indicates whether and how the selectivity of the fleet is time-varying. Column "Weighting scaler" indicates how length composition data are weighted in comparison to the Francis weighting method.

TABLA 5. Decisiones de ponderación de los datos de selectividad y composición en función de la cantidad de captura de cada pesquería y de la calidad de los datos de composición. Las reglas en las que se basa esta tabla de decisiones se ilustran en forma de diagrama de flujo en la Figura 20. La columna "Normal doble" indica si los datos de composición por talla de la flota pueden ajustarse bien en el modelo de evaluación utilizando una curva de selectividad normal doble. La columna "Calidad de los datos" indica la calidad relativa de los datos de composición por talla de la flota. La columna "Bloques temporales" indica si la selectividad de la flota varía con el tiempo y de qué manera. La columna "Escalador de ponderación" indica cómo se ponderan los datos de composición por talla en comparación con el método de ponderación de Francis.

Fleet number	Fleet type	Fleet name	Catch amount	Double-normal	Data quality	Selectivity	Time blocks	Weighting scaler
1	Fishery	LL-n-A1	Low	No	High	Fixed	NA	0
2		LL-n-A2	High	Yes	High	Estimated	1993; 2010	1
3		LL-n-A3	High	Yes	High	Estimated	NA	0.2
4		LL-n-A4	High	Yes	High	Estimated	1993; 2010	1
5		LL-n-A5	High	Yes	High	Estimated	1993; 2010	1
6		LL-n-A6	Low	Yes	High	Estimated	NA	0.2
7		LL-n-A7	Low	Yes	High	Estimated	NA	0.2
8	Fishery	LL-w-A1	Low	NA	NA	Mirror F1	NA	NA
9		LL-w-A2	High	NA	NA	Mirror F2	NA	NA
10		LL-w-A3	High	NA	NA	Mirror F3	NA	NA
11		LL-w-A4	High	NA	NA	Mirror F4	NA	NA
12		LL-w-A5	High	NA	NA	Mirror F5	NA	NA
13		LL-w-A6	Low	NA	NA	Mirror F6	NA	NA
14		LL-w-A7	Low	NA	NA	Mirror F7	NA	NA
15	Fishery	OBJ-A1	Low	Yes	High	Estimated	NA	0.2
16		OBJ-A2	High	Yes	High	Estimated	2000; 2010	1
17		OBJ-A3	High	No	High	Estimated	NA	0.2
18		OBJ-A4	High	Yes	High	Estimated	2000; 2010	1
19		OBJ-A5	Low	Yes	High	Estimated	NA	0.2
20		OBJ-disc-EPO	Low	NA	NA	Fixed	NA	0
21	Fishery	NOADEL-A1	Low	Yes	Low	Mirror F15	NA	0
22		NOADEL-A2	Low	Yes	Low	Mirror F19	NA	0
23	Survey	LL-survey-EPO	NA	Yes	High	Estimated	NA	1

TABLE 6. The diagnostics metrics for all the reference models that have a positive definite Hessian. Gradient is the final gradient of the assessment model, R shift quantifies the degree of recruitment shift after the expansion of the floating-object fishery, and Mohn's rho quantifies the degree of retrospective bias in spawning biomass.

TABLA 6. Métricas de diagnóstico para todos los modelos de referencia que tienen una matriz hessiana positiva definida. El gradiente es el gradiente final del modelo de evaluación, Rshift cuantifica el grado de cambio en el reclutamiento tras la expansión de la pesquería sobre objetos flotantes y Rho de Mohn cuantifica el grado de sesgo retrospectivo en la biomasa reproductora.

Number	Model	Catchability	Steepness	Gradient	R shift	Mohn's rho
1	Fix	0%	1.0	0.002	21%	0.05
2	Fix	0%	0.9	0.039	24%	
3	Fix	0%	0.8	0.000	27%	
4	Fix	1%	1.0	0.000	8%	0.05
5	Fix	1%	0.9	0.004	13%	
6	Fix	1%	0.8	0.002	16%	
7	Fix	2%	1.0	0.000	-4%	0.04
8	Fix	2%	0.9	0.001	2%	
9	Gro	0%	1.0	0.000	12%	0.04
10	Gro	0%	0.9	0.048	15%	
11	Gro	0%	0.8	0.004	18%	
12	Gro	1%	1.0	0.000	-1%	0.04
13	Gro	1%	0.9	0.004	3%	
14	Gro	1%	0.8	0.001	7%	
15	Gro	2%	1.0	0.001	-12%	0.04
16	Gro	2%	0.9	0.003	12%	
17	Gro	2%	0.8	0.011	-3%	
18	Sel	0%	1.0	0.001	-2%	0.04
19	Sel	0%	0.9	0.015	0%	
20	Sel	0%	0.8	0.023	1%	
21	Sel	1%	1.0	0.000	-11%	0.04
22	Sel	1%	0.9	0.001	-9%	
23	Sel	1%	0.8	0.006	-6%	
24	Sel	2%	1.0	0.001	-20%	0.04
25	Sel	2%	0.9	0.005	-17%	
26	Sel	2%	0.8	0.000	-14%	
27	Mrt	0%	1.0	0.001	-6%	0.04
28	Mrt	0%	0.9	0.016	-3%	
29	Mrt	0%	0.8	0.001	1%	
30	Mrt	1%	1.0	0.000	-13%	0.05
31	Mrt	2%	1.0	0.002	-20%	0.05
32	Mrt	2%	0.9	0.001	-14%	
33	Mrt	2%	0.8	0.002	-9%	

TABLE 7. Management table for bigeye in the EPO. $S_{current}$, S_0, S_{MSY_d}: spawning biomass (metric tons) at the beginning of 2024, in a unfished equilibrium state, and at dynamic MSY, respectively; $F_{current}$ and F_{MSY}: fishing mortality between 2021-2023 and at MSY, respectively; S_{LIMIT} and F_{LIMIT}: limit reference points for spawning biomass and fishing mortality, respectively; $C_{current}$: total catch of bigeye in 2023 (metric tons); MSY_d: dynamic MSY; p(): probability.
TABLA 7. Tabla de ordenación para el patudo en el OPO. S_{actual}, S_0, S_{RMS_d}: biomasa reproductora (toneladas métricas) al principio de 2024, en estado de equilibrio en ausencia de pesca y en RMS dinámico, respectivamente; F_{actual} y F_{RMS}: mortalidad por pesca entre 2021-2023 y en RMS, respectivamente; $S_{LÍMITE}$ y $F_{LÍMITE}$: puntos de referencia límite para biomasa reproductora y mortalidad por pesca, respectivamente; C_{actual}: captura total de patudo en 2023 (toneladas métricas); RMS_d: RMS dinámico; p(): probabilidad.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
	0%-1.0	**0%-0.9**	**0%-0.8**	**1%-1.0**	**1%-0.9**	**1%-0.8**	**2%-1.0**	**2%-0.9**	**2%-0.8**
Fix									
MSY	87779	84598	82775	93262	89938	87776	99745	96355	NA
MSY_d	82360	82041	83401	83439	83917	85955	84889	86595	NA
$C_{current}$/MSY_d	0.79	0.79	0.78	0.78	0.77	0.75	0.76	0.75	NA
S_{MSY}/S_0	0.17	0.23	0.27	0.17	0.23	0.27	0.17	0.23	NA
$S_{current}/S_0$	0.17	0.17	0.18	0.14	0.14	0.14	0.11	0.11	NA
$S_{current}/S_{LIMIT}$	2.26	2.25	2.28	1.82	1.80	1.83	1.46	1.43	NA
p($S_{current}<S_{LIMIT}$)	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.02	NA
$F_{current}/F_{LIMIT}$	0.56	0.61	0.66	0.61	0.67	0.73	0.65	0.73	NA
p($F_{current}>F_{LIMIT}$)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	NA
$S_{current}/S_{MSY\_d}$	1.10	0.76	0.62	0.93	0.63	0.51	0.78	0.52	NA
p($S_{current}<S_{MSY\_d}$)	0.25	0.99	1.00	0.74	1.00	1.00	0.99	1.00	NA
$F_{current}/F_{MSY}$	0.81	1.01	1.15	0.88	1.11	1.27	0.96	1.22	NA
p($F_{current}>F_{MSY}$)	0.04	0.53	0.81	0.15	0.77	0.94	0.36	0.92	NA
Gro									
MSY	91802	86940	83853	97331	92194	88852	103674	89882	94844
MSY_d	85554	82617	82128	86497	83761	83680	87648	85005	86113
$C_{current}$/MSY_d	0.76	0.79	0.79	0.75	0.77	0.77	0.74	0.76	0.75
S_{MSY}/S_0	0.16	0.22	0.26	0.16	0.22	0.26	0.16	0.23	0.26
$S_{current}/S_0$	0.21	0.20	0.20	0.17	0.16	0.16	0.14	0.10	0.13
$S_{current}/S_{LIMIT}$	2.67	2.63	2.62	2.19	2.14	2.13	1.78	1.34	1.71
p($S_{current}<S_{LIMIT}$)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00
$F_{current}/F_{LIMIT}$	0.49	0.54	0.60	0.53	0.59	0.65	0.56	0.80	0.72
p($F_{current}>F_{LIMIT}$)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00
$S_{current}/S_{MSY\_d}$	1.42	0.94	0.76	1.23	0.81	0.64	1.05	0.45	0.54
p($S_{current}<S_{MSY\_d}$)	0.02	0.65	0.98	0.09	0.95	1.00	0.35	1.00	1.00
$F_{current}/F_{MSY}$	0.68	0.88	1.03	0.74	0.96	1.13	0.80	1.36	1.24
p($F_{current}>F_{MSY}$)	0.00	0.19	0.57	0.01	0.40	0.78	0.03	0.99	0.92
Sel									
MSY	107670	100765	96229	109706	103041	98752	113725	107229	103071
MSY_d	99041	91821	88179	95668	89277	86596	94006	88618	87082
$C_{current}$/MSY_d	0.65	0.71	0.74	0.68	0.73	0.75	0.69	0.73	0.74
S_{MSY}/S_0	0.15	0.22	0.26	0.16	0.22	0.26	0.16	0.22	0.26
$S_{current}/S_0$	0.27	0.26	0.26	0.21	0.20	0.20	0.16	0.16	0.15
$S_{current}/S_{LIMIT}$	3.46	3.44	3.43	2.67	2.64	2.63	2.09	2.03	2.01
p($S_{current}<S_{LIMIT}$)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$F_{current}/F_{LIMIT}$	0.36	0.40	0.43	0.42	0.47	0.51	0.48	0.54	0.60

	1	2	3	4	5	6	7	8	9
	0%-1.0	**0%-0.9**	**0%-0.8**	**1%-1.0**	**1%-0.9**	**1%-0.8**	**2%-1.0**	**2%-0.9**	**2%-0.8**
p(F_{current}>F_{LIMIT})	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$S_{\text{current}}/S_{MSY\_d}$	2.06	1.40	1.15	1.66	1.12	0.91	1.35	0.90	0.72
p(S_{current}<$S_{\text{MSY\_d}}$)	0.00	0.05	0.24	0.01	0.25	0.72	0.04	0.78	1.00
$F_{\text{current}}/F_{\text{MSY}}$	0.49	0.64	0.74	0.58	0.75	0.88	0.67	0.87	1.02
p(F_{current}>F_{MSY})	0.00	0.00	0.02	0.00	0.02	0.19	0.00	0.15	0.56
Mrt									
MSY	117855	107300	101302	115396	NA	NA	115795	108586	104466
MSY_d	106534	94892	90471	99015	NA	NA	94495	89245	89498
C_{current}/MSY_d	0.61	0.68	0.72	0.65	NA	NA	0.69	0.73	0.72
S_{MSY}/S_0	0.11	0.20	0.25	0.12	NA	NA	0.14	0.21	0.26
S_{current}/S_0	0.25	0.24	0.24	0.19	NA	NA	0.14	0.14	0.13
$S_{\text{current}}/S_{\text{LIMIT}}$	3.26	3.15	3.06	2.46	NA	NA	1.87	1.77	1.71
p(S_{current}<S_{LIMIT})	0.00	0.00	0.00	0.00	NA	NA	0.00	0.00	0.00
$F_{\text{current}}/F_{\text{LIMIT}}$	0.35	0.40	0.45	0.43	NA	NA	0.50	0.57	0.65
p(F_{current}>F_{LIMIT})	0.00	0.00	0.00	0.00	NA	NA	0.00	0.00	0.00
$S_{\text{current}}/S_{MSY\_d}$	2.92	1.52	1.15	2.05	NA	NA	1.47	0.84	0.62
p(S_{current}<$S_{\text{MSY\_d}}$)	0.00	0.01	0.19	0.00	NA	NA	0.00	0.94	1.00
$F_{\text{current}}/F_{\text{MSY}}$	0.42	0.61	0.74	0.53	NA	NA	0.65	0.91	1.10
p(F_{current}>F_{MSY})	0.00	0.00	0.01	0.00	NA	NA	0.00	0.22	0.75

TABLE 8. Management quantities for bigeye tuna in the EPO. Combined (E) is the expected value across models. Combined (p=0.5) is the median of the distribution across models.
TABLA 8. Cantidades de ordenación para el atún patudo en el OPO. Combinado (E) es el valor esperado en todos los modelos. Combinado (p=0.5) es la mediana de la distribución entre modelos.

	Fix	Gro	Sel	Mrt	Combined (E)	Combined (p=0.5)
$F_{\text{current}}/F\text{MSY}$	1.01	0.66	0.93	0.70	0.82	0.79
$p(F_{\text{current}} > F\text{MSY})$	0.49	0.10	0.34	0.08	0.25	
$F_{\text{current}}/F_{\text{LIMIT}}$	0.64	0.46	0.59	0.45	0.54	0.53
$p(F_{\text{current}} > F_{\text{LIMIT}})$	0.00	0.00	0.00	0.00	0.00	
$S_{\text{current}}/S_{MSY\_d}$	0.77	1.67	0.94	1.34	1.29	1.05
$p(S_{\text{current}} < S_{\text{MSY\_d}})$	0.83	0.23	0.57	0.27	0.47	
$S_{\text{current}}/S_{\text{LIMIT}}$	1.87	2.49	2.14	2.72	2.30	2.21
$p(S_{\text{current}} < S_{\text{LIMIT}})$	0.00	0.00	0.00	0.00	0.00	

FIGURES

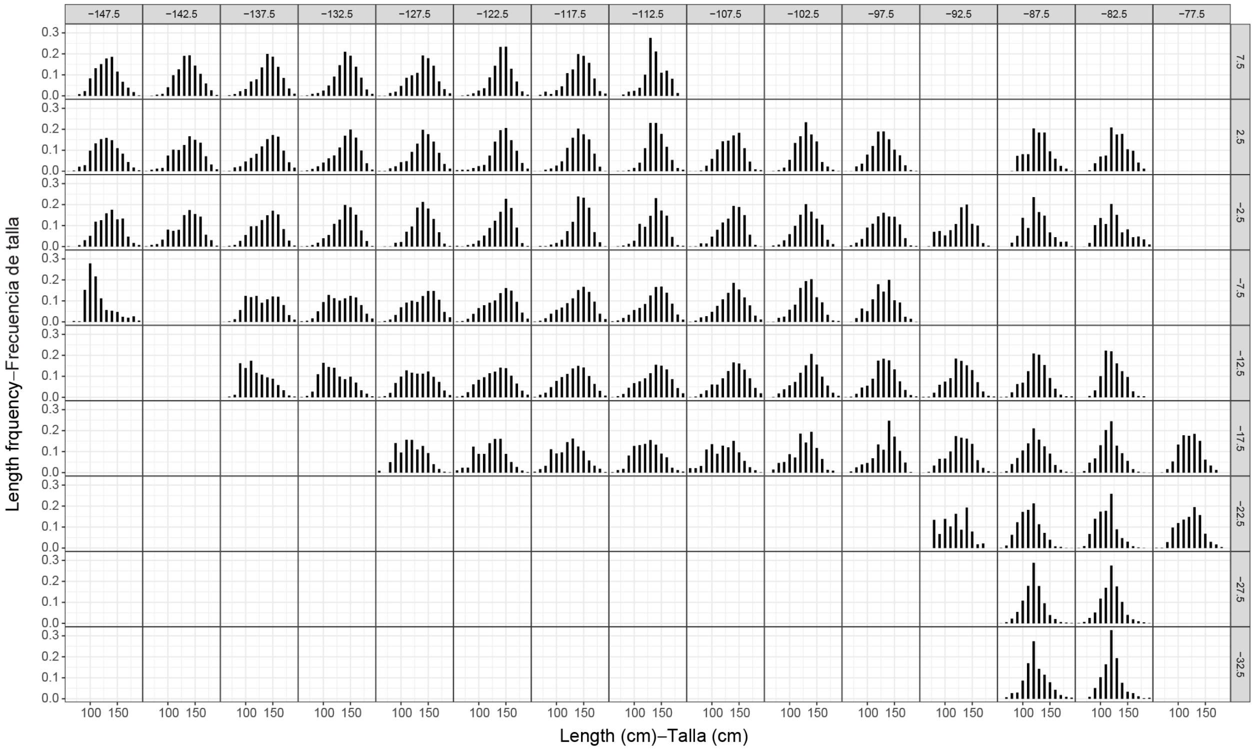

FIGURE 1. Average length frequency of bigeye by grid cell in the Japanese longline fishery.
FIGURA 1. Frecuencia de talla promedio del patudo por celda en la pesquería palangrera japonesa.

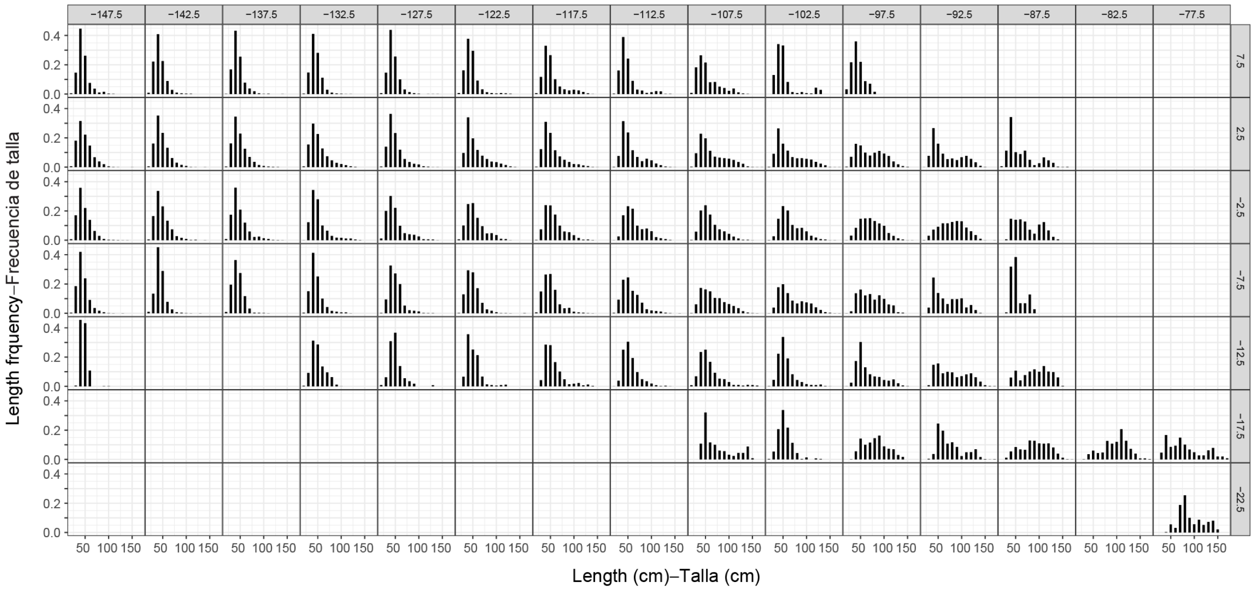

FIGURE 2. Average length frequency of bigeye by grid cell in the floating-object fishery.
FIGURA 2. Frecuencia de talla promedio del patudo por celda en la pesquería sobre objetos flotantes.

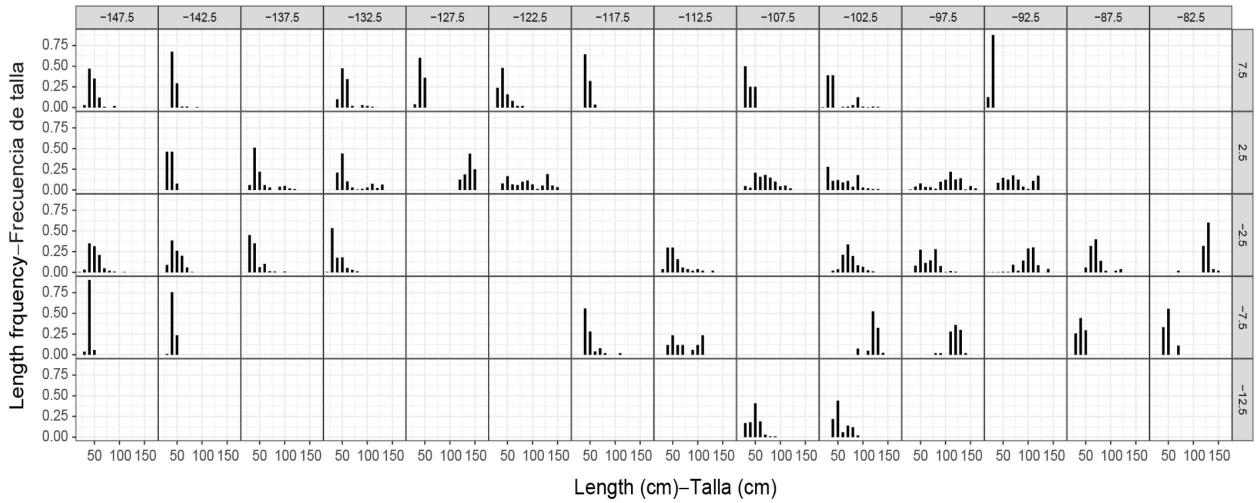

FIGURE 3. Average length frequency of bigeye by grid cell in the unassociated fishery.
FIGURA 3. Frecuencia de talla promedio del patudo por celda en la pesquería no asociada.

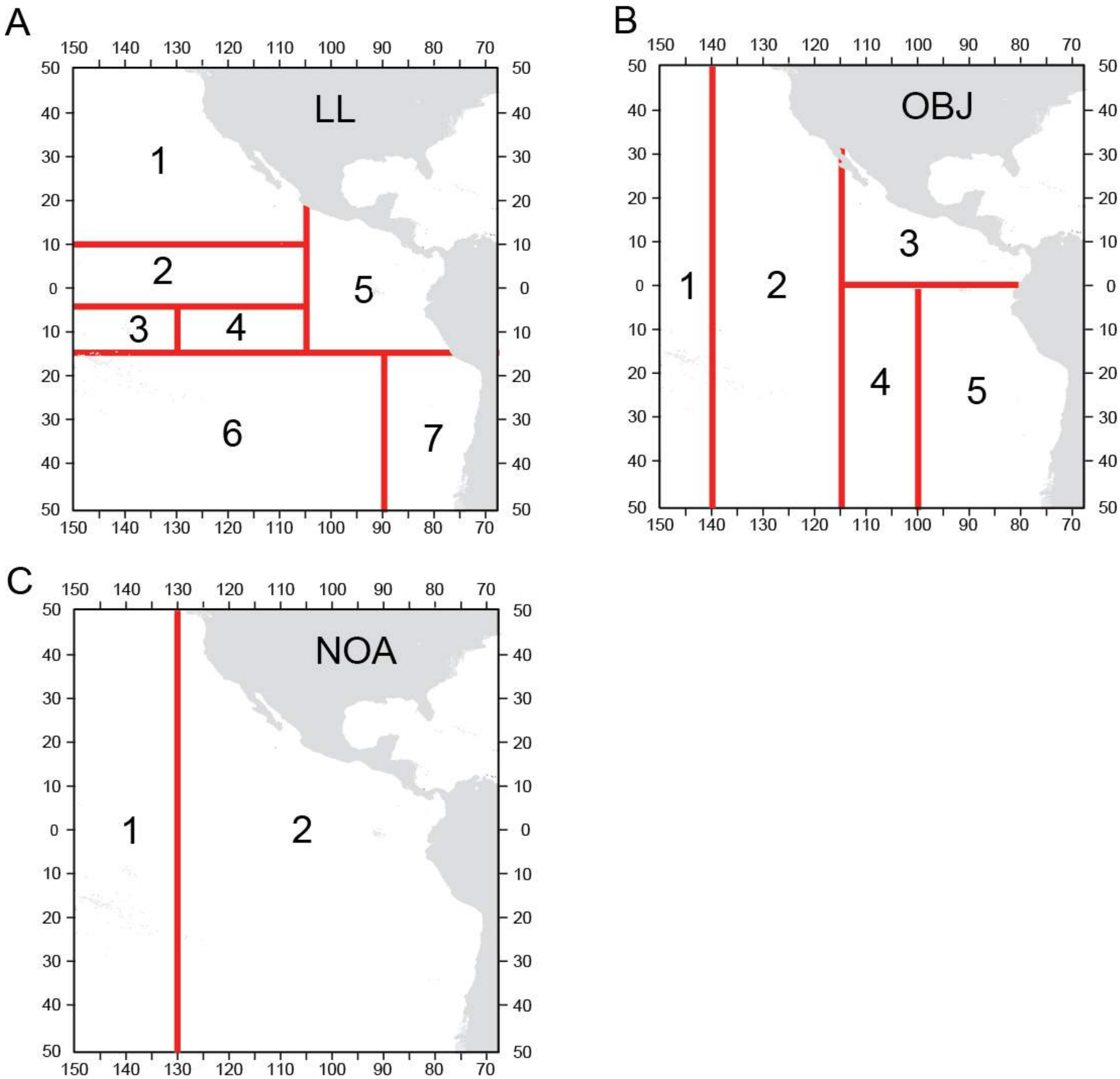

FIGURE 4. Summary of area definitions for the longline (LL), floating-object (OBJ), and unassociated (NOA) fishery fleets in the stock assessment of bigeye tuna in the EPO.

FIGURA 4. Resumen de las definiciones de áreas para las flotas de las pesquerías palangrera (LL), sobre objetos flotantes (OBJ) y no asociada (NOA) en la evaluación del atún patudo en el OPO.

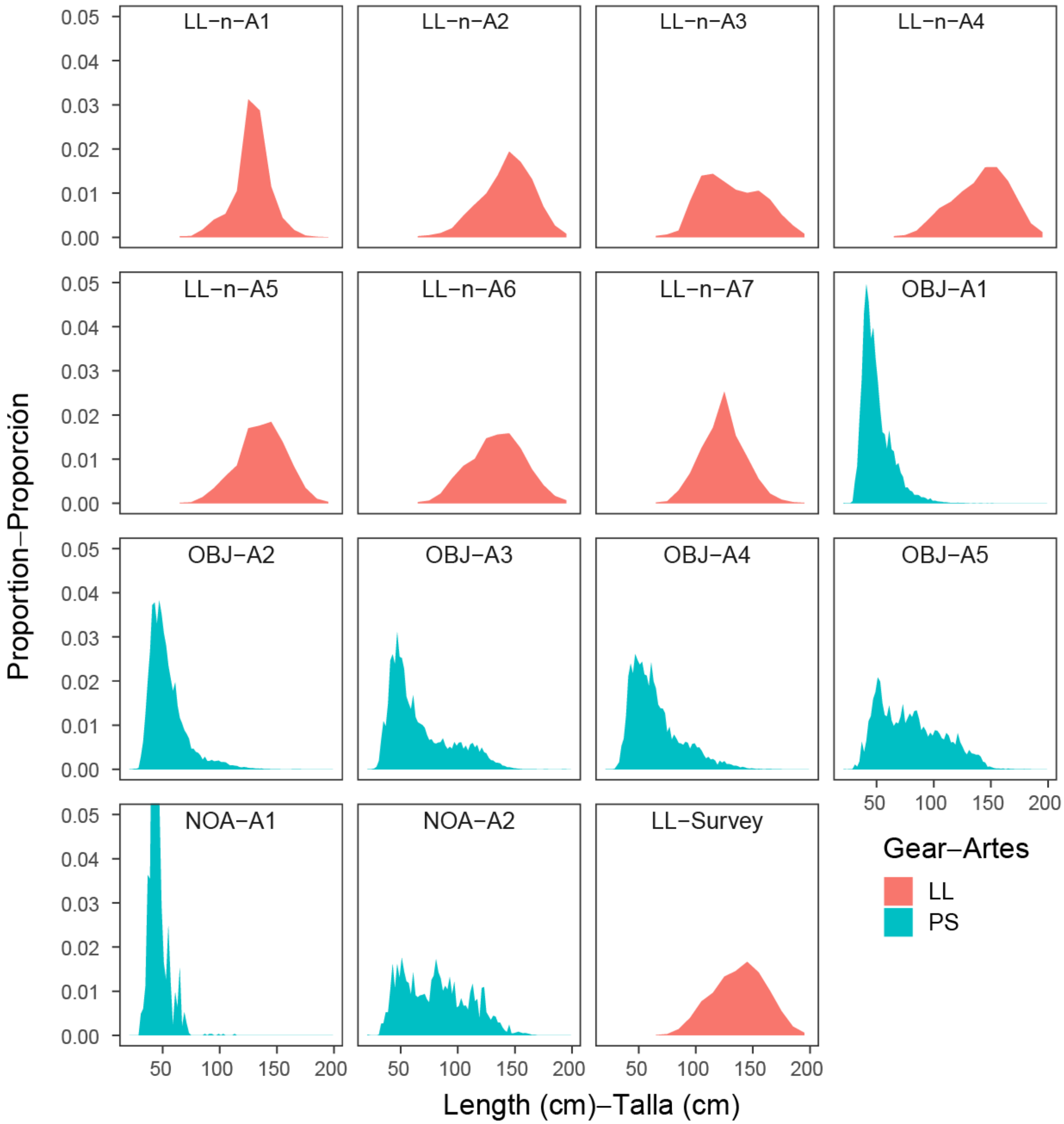

FIGURE 5. Sample-size weighted length frequency of bigeye tuna observed by each fishery and survey in the benchmark assessment model.

FIGURA 5. Frecuencia de talla ponderada por tamaño de muestra de atún patudo observada por cada pesquería y flota de estudio en el modelo de evaluación de referencia.

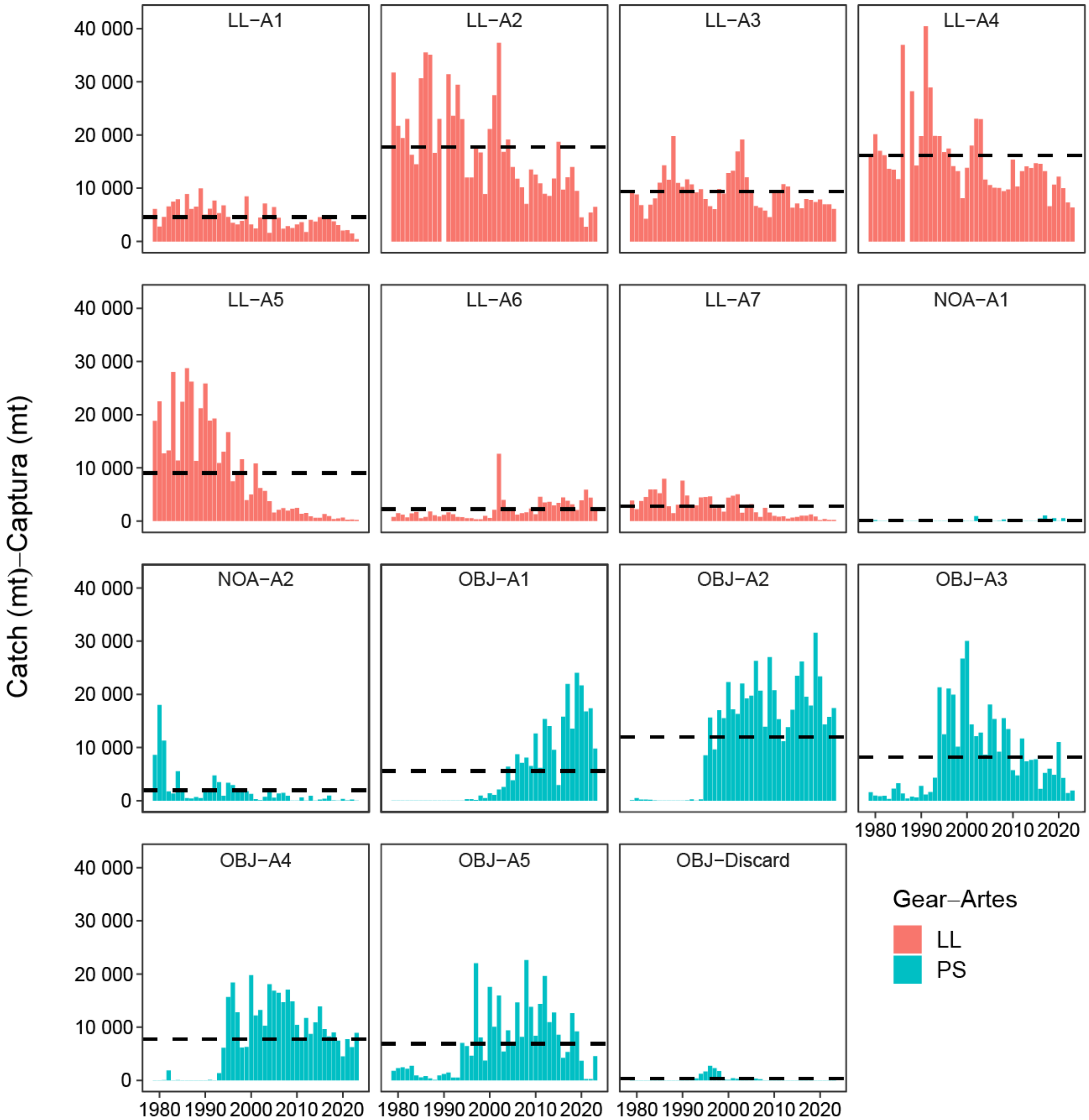

FIGURE 6a. Annual catches (metric tons) of bigeye tuna in the eastern Pacific Ocean by fishery in 1979-2023. To facilitate the comparison, the catches of fisheries 1-7 are converted by the stock assessment model from number to weight.

FIGURA 6a. Capturas anuales (toneladas métricas) de atún patudo en el Océano Pacífico oriental, por pesquería, en 1979-2023. Para facilitar la comparación, las capturas de las pesquerías 1-7 son convertidas por el modelo de evaluación de poblaciones de número a peso.

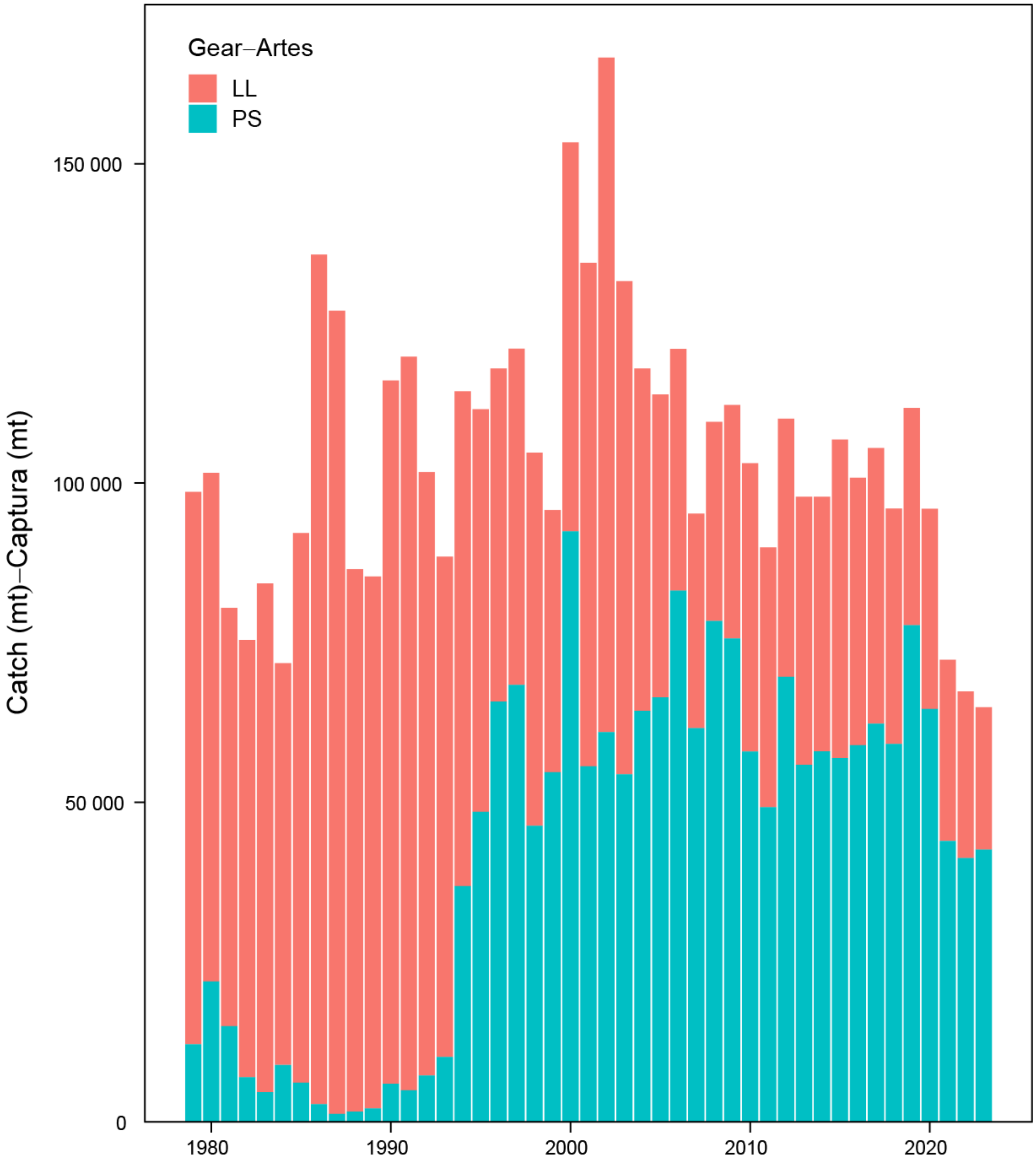

FIGURE 6b. Annual catches (metric tons) of bigeye tuna in the eastern Pacific Ocean by gear type in 1979-2023. To facilitate the comparison, the catches of fisheries 1-7 are converted by the stock assessment model from number to weight.

FIGURA 6b. Capturas anuales (toneladas métricas) de atún patudo en el Océano Pacífico oriental, por tipo de arte, en 1979-2023. Para facilitar la comparación, las capturas de las pesquerías 1-7 son convertidas por el modelo de evaluación de poblaciones de número a peso.

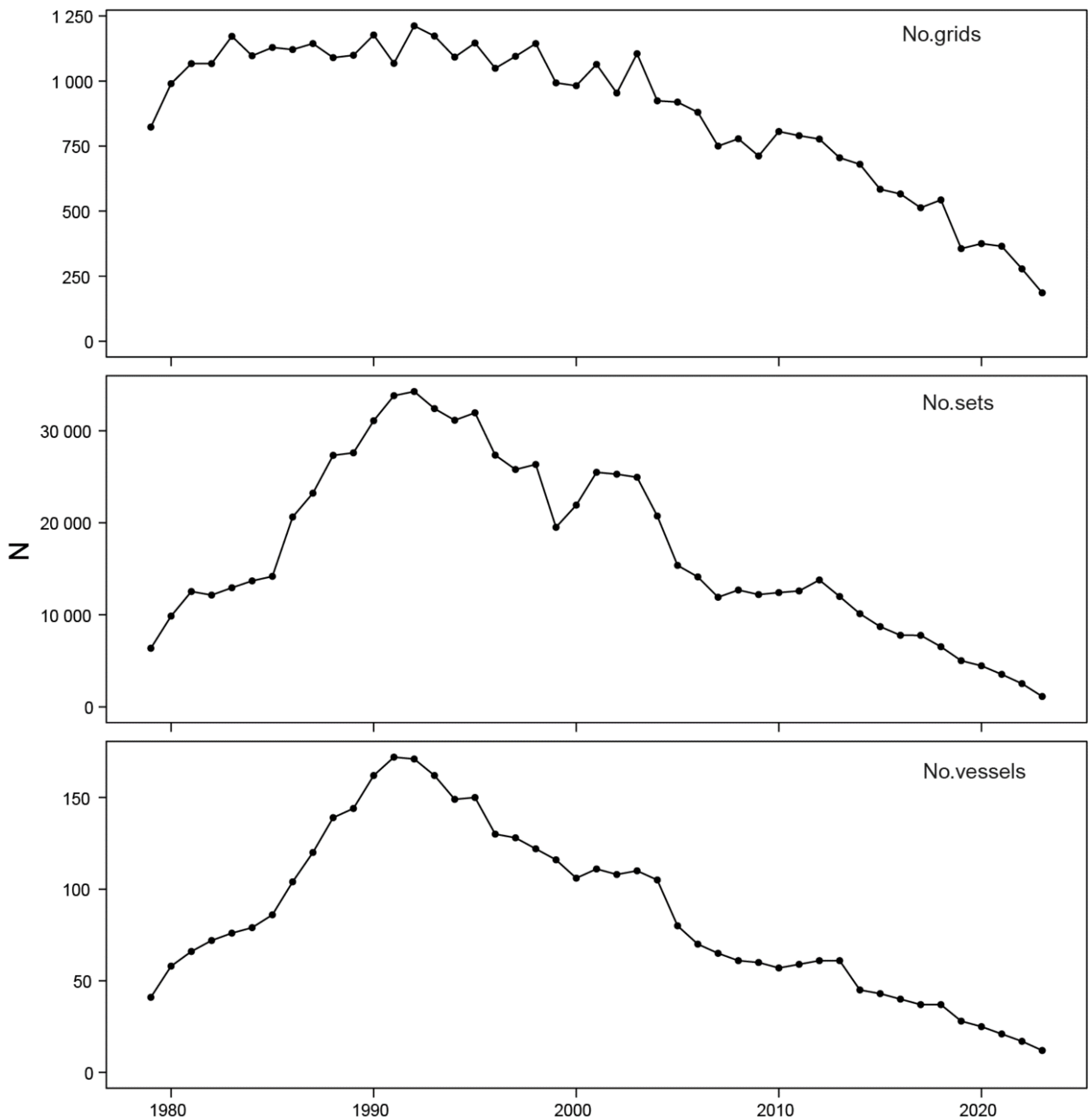

FIGURE 7. Time series of the number of 1° x 1° grid cells (top panel), sets (middle panel), and vessels (bottom panel) covered by the Japanese longline CPUE dataset between 1979 and 2023.
FIGURA 7. Series de tiempo del número de celdas de 1° x 1° (panel superior), lances (panel central) y buques (panel inferior) cubiertos por el conjunto de datos de CPUE de palangre de Japón entre 1979 y 2023.

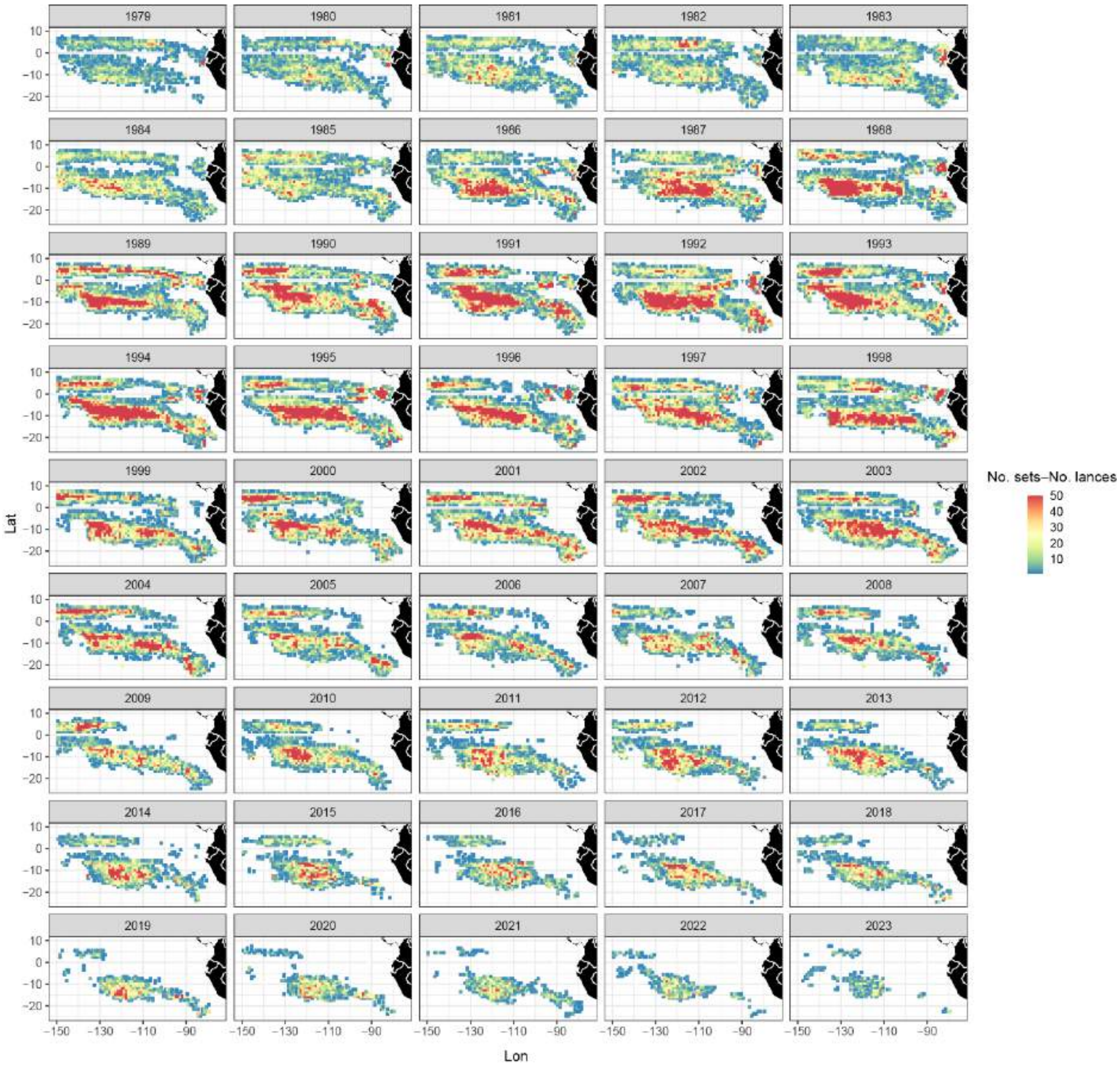

FIGURE 8. Spatial distribution of the annual number of sets made by the Japanese longline fleet operated in the eastern Pacific Ocean between 1979 and 2023.

FIGURA 8. Distribución espacial del número anual de lances de la flota palangrera de Japón en el Océano Pacífico oriental entre 1979 y 2023.

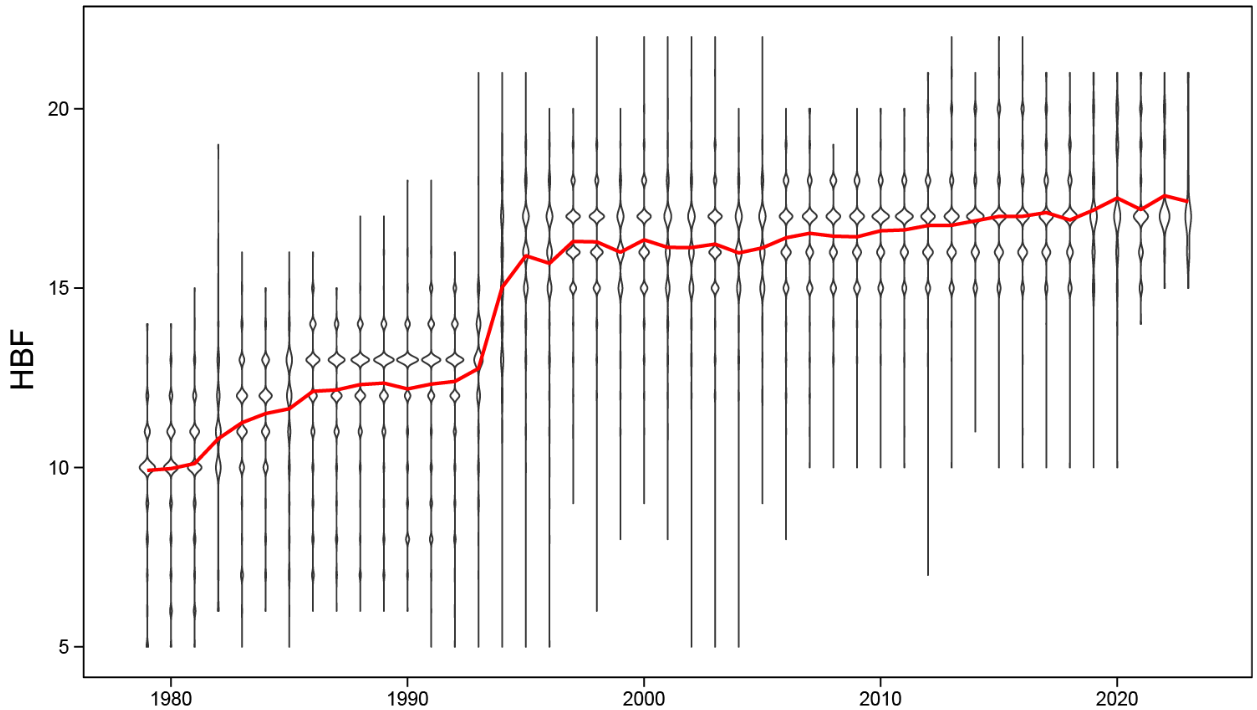

FIGURE 9. The violin plot of hooks-between-floats recorded in the Japanese longline CPUE dataset between 1979 and 2023. The red line represents effort-weighted values for each year.
FIGURA 9. Gráfica de violín de los anzuelos entre flotadores registrados en el conjunto de datos de CPUE de palangre de Japón entre 1979 y 2023. La línea roja representa los valores ponderados por esfuerzo para cada año.

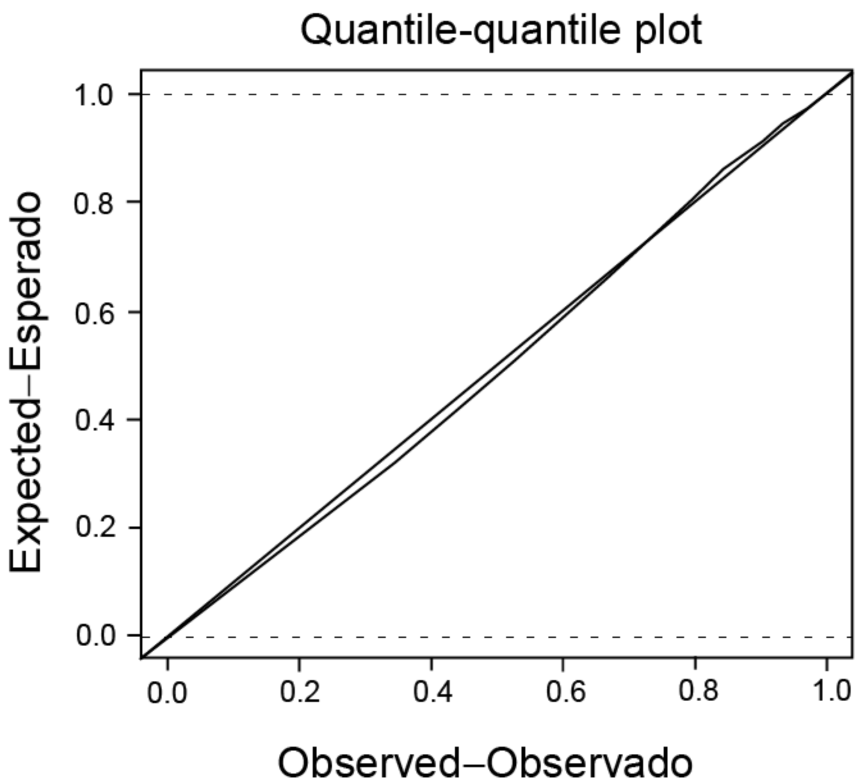

FIGURE 10. The quantile-quantile plot for the CPUE standardization model for bigeye tuna in the Japanese longline fishery.
FIGURA 10. Gráfica cuantil-cuantil del modelo de estandarización de la CPUE para el patudo en la pesquería palangrera de Japón.

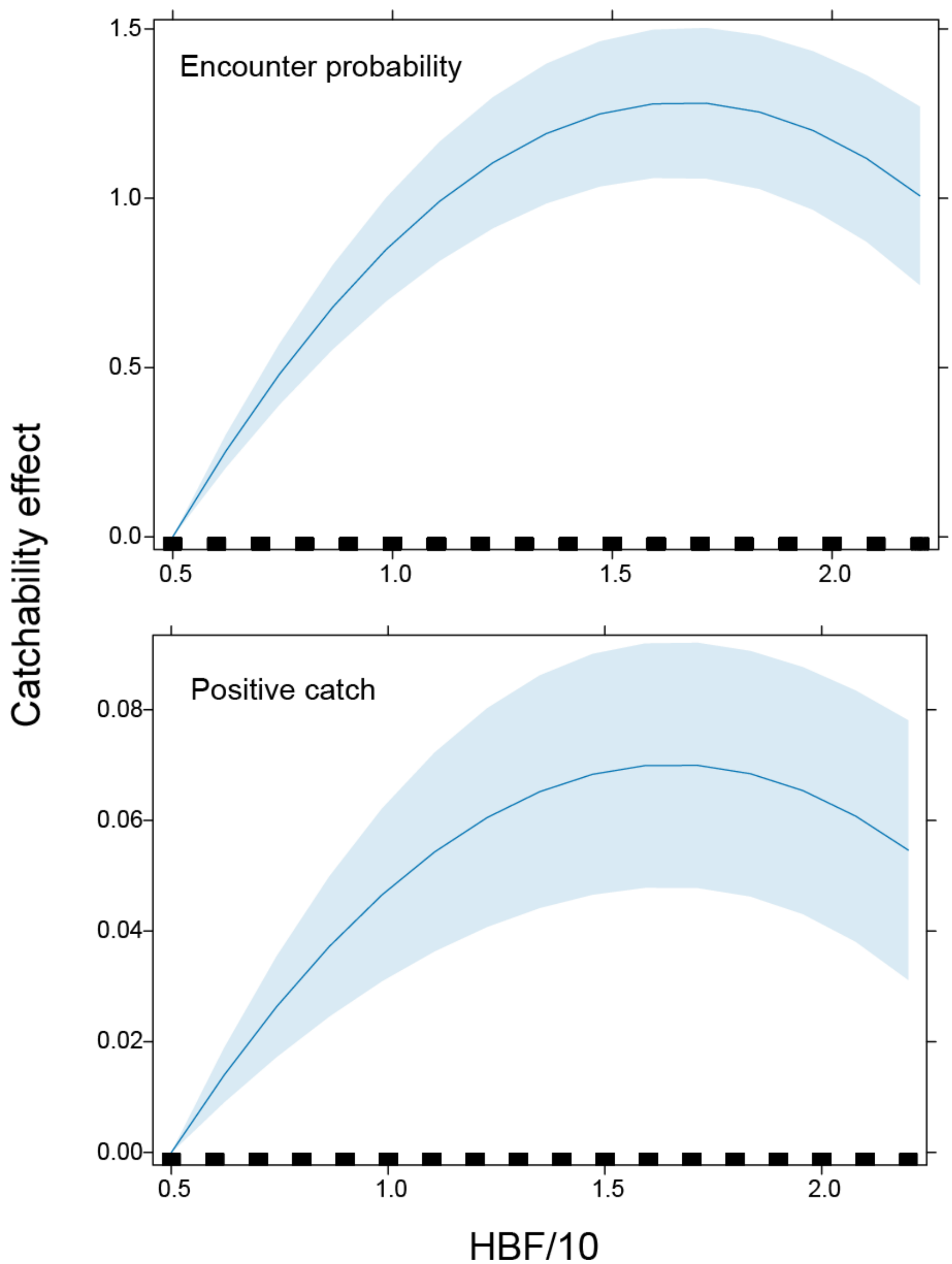

FIGURE 11. Estimated catchability effects of hooks-between-floats (HBF) on the encounter probability (top panel) and positive catch rate (bottom panel) of bigeye tuna in the Japanese longline fishery.

FIGURA 11. Efectos estimados de la capturabilidad de los anzuelos entre flotadores (AEF) sobre la probabilidad de encuentro (panel superior) y la tasa de captura positiva (panel inferior) del atún patudo en la pesquería palangrera de Japón.

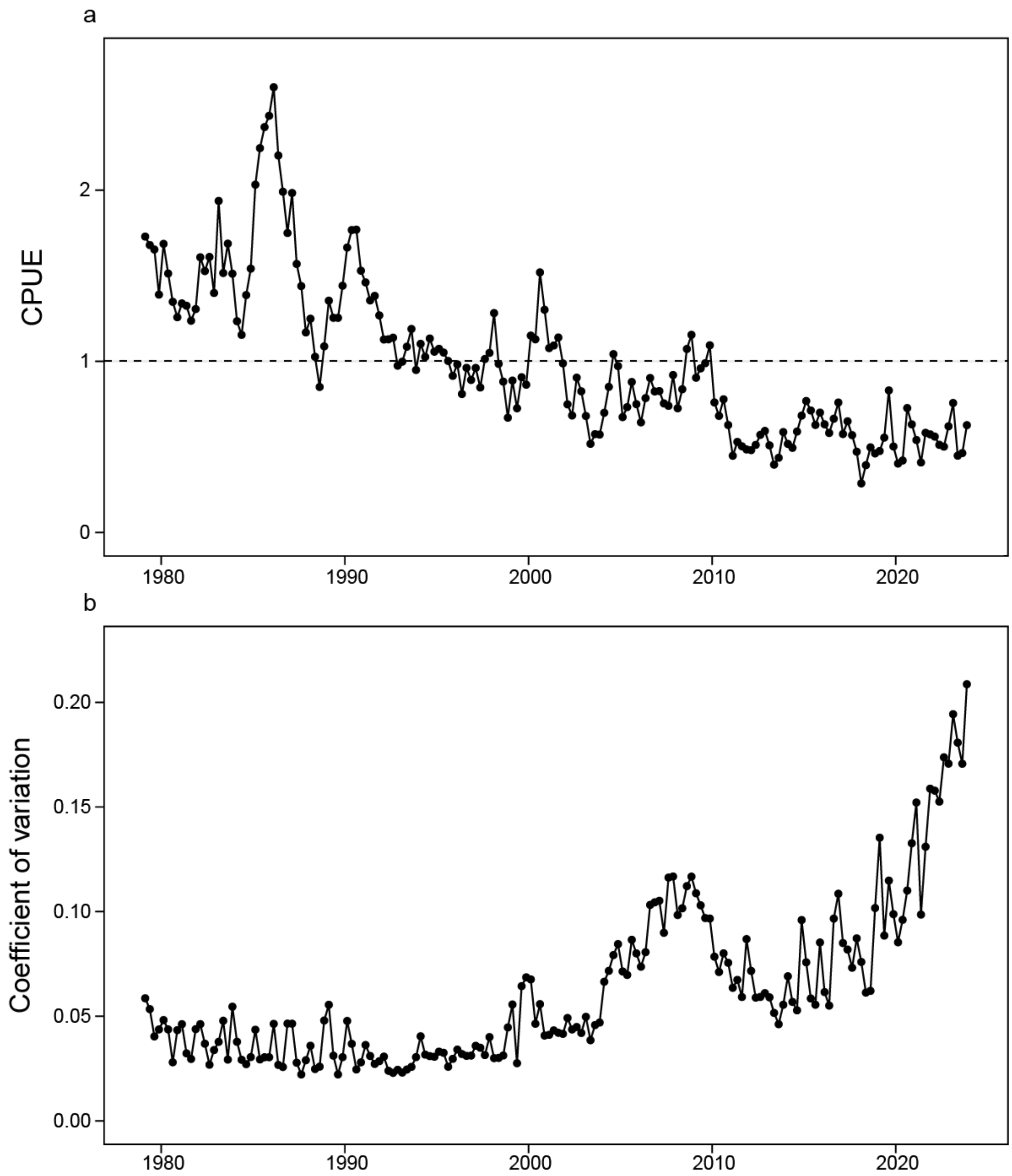

FIGURE 12. The standardized index of abundance and the associated coefficient of variation estimated by the spatiotemporal model for bigeye tuna.
FIGURA 12. Índice estandarizado de abundancia y coeficiente de variación asociado estimados por el modelo espaciotemporal para el atún patudo.

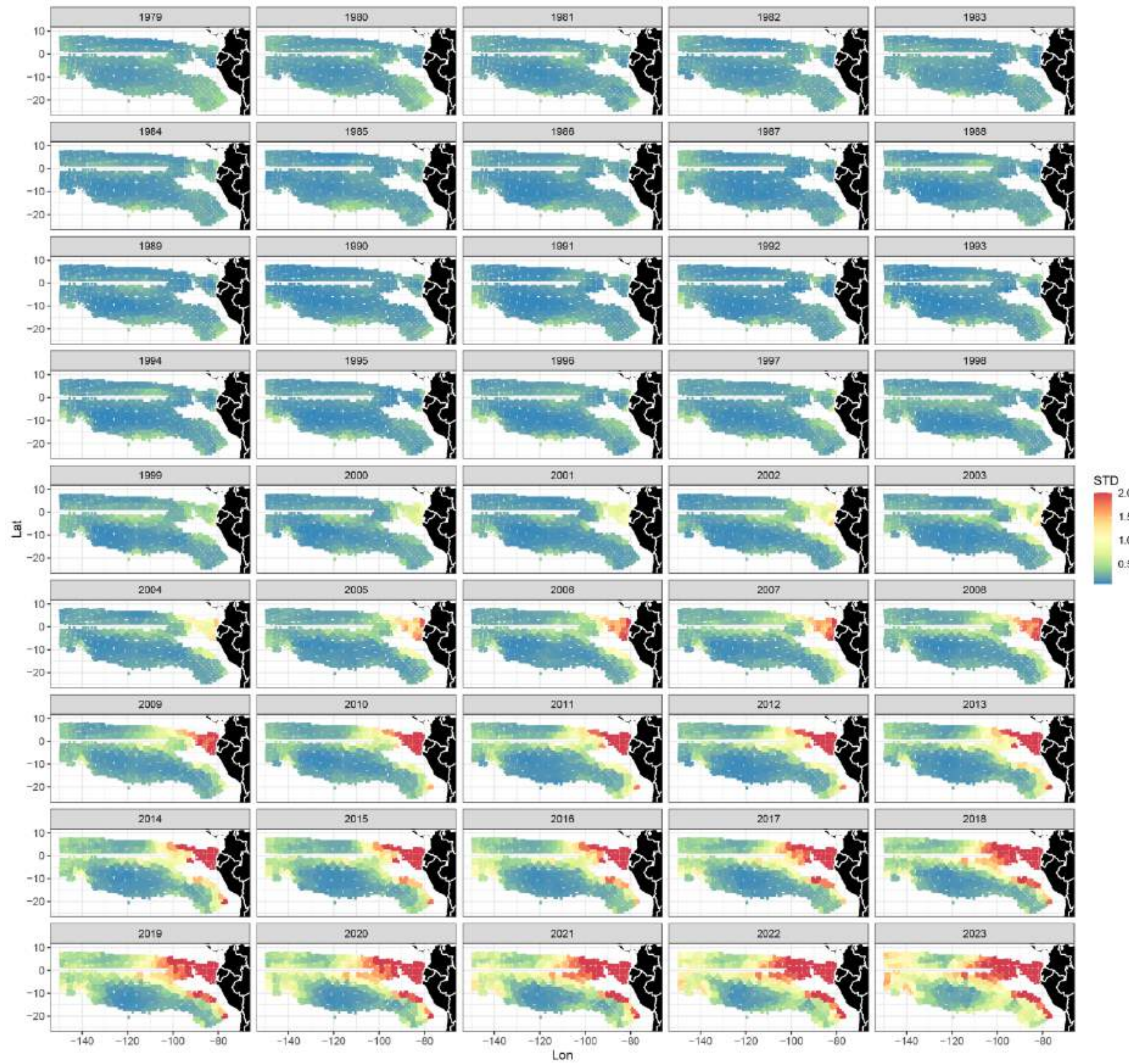

FIGURE 13. The standard deviation of predicted log density from the CPUE standardization model by year between 1979 and 2023.

FIGURA 13. Desviación estándar de la densidad logarítmica predicha a partir del modelo de estandarización de la CPUE, por año, entre 1979 y 2023.

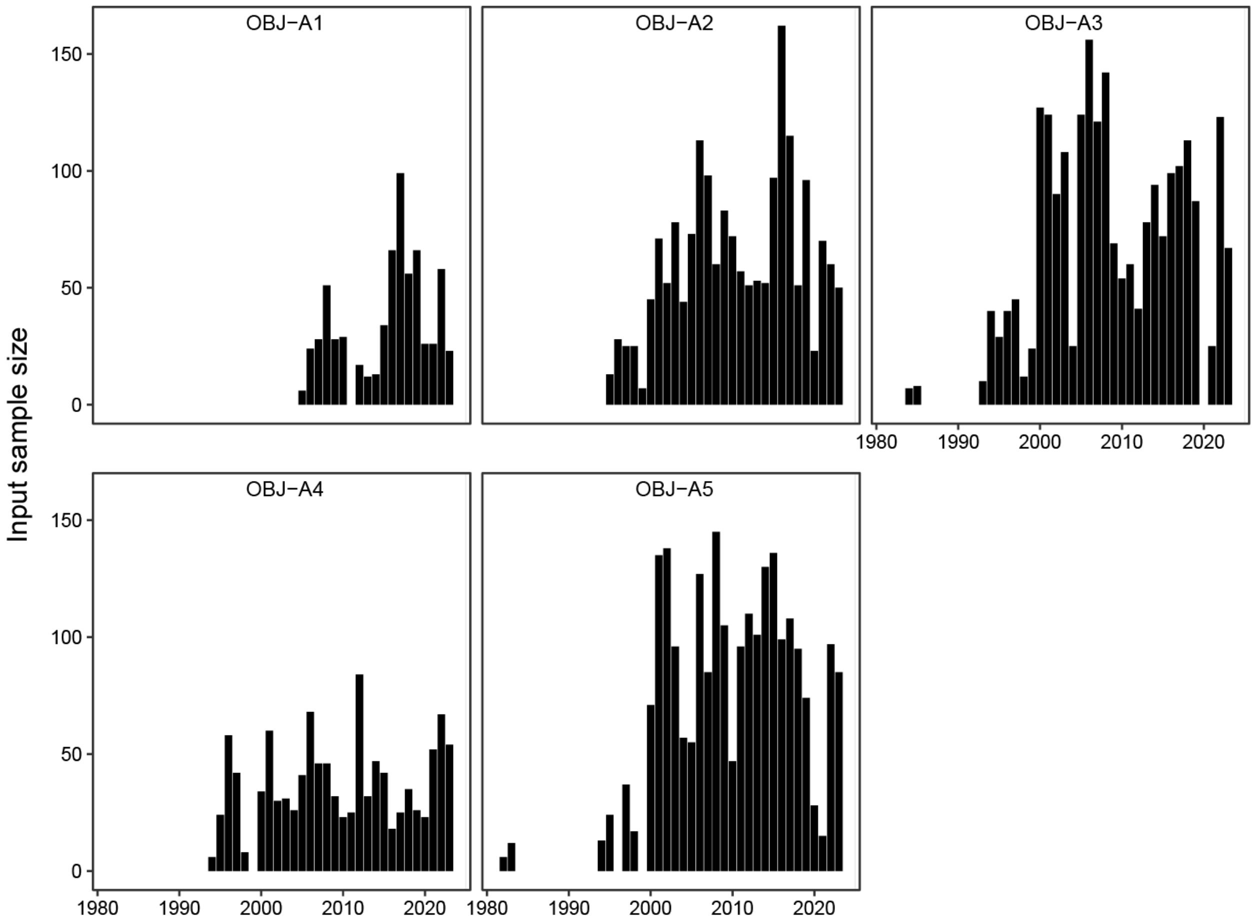

FIGURE 14. The number of wells sampled, by year and fishery, in the floating-object length composition data used in the benchmark assessment.

FIGURA 14. Número de bodegas muestreadas, por año y pesquería, en los datos de composición por talla de la pesquería sobre objetos flotantes utilizados en la evaluación de referencia.

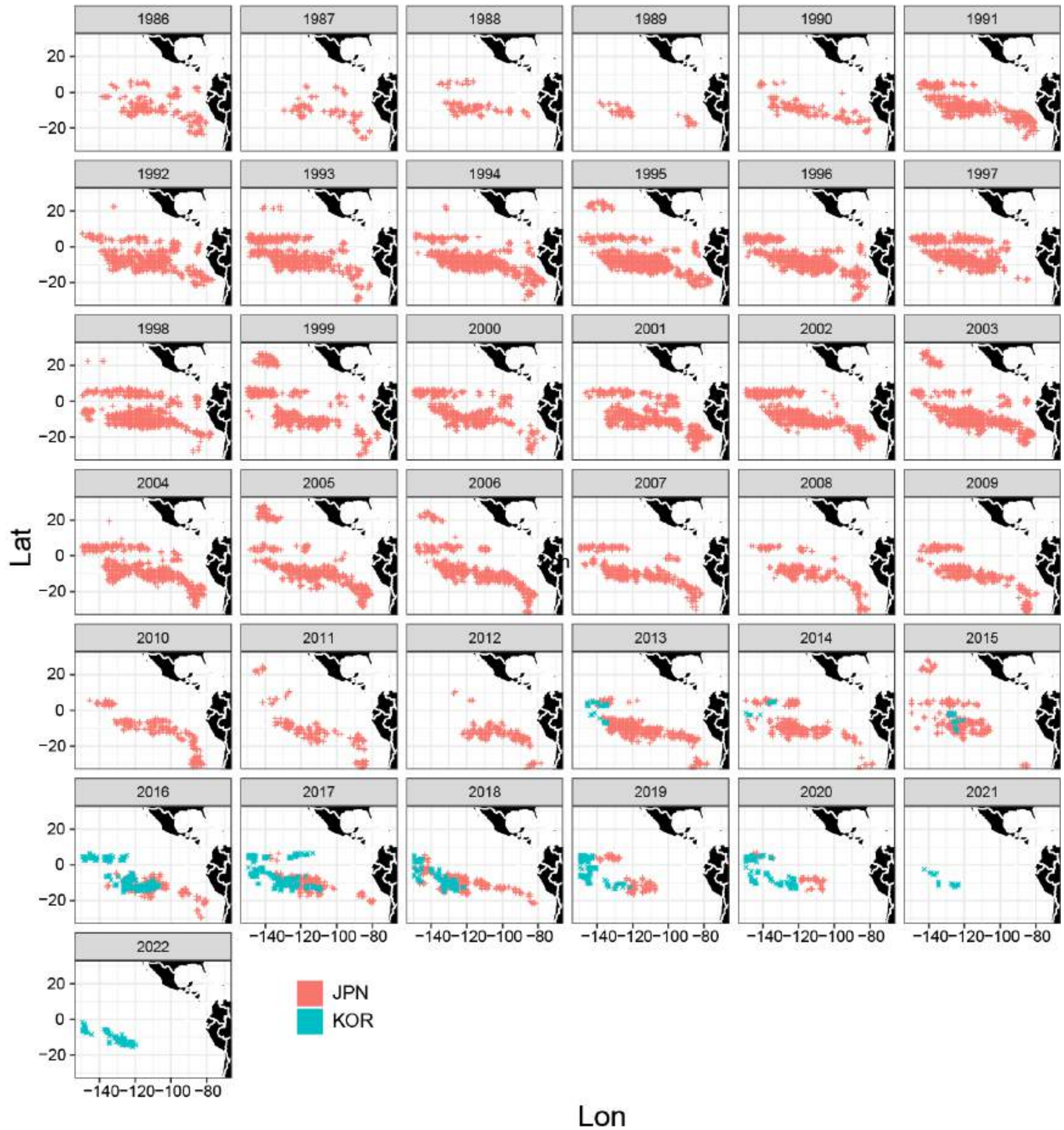

FIGURE 15. The spatiotemporal distribution of the Japanese and Korean longline length composition data used in the benchmark assessment.

FIGURA 15. Distribución espaciotemporal de los datos de composición por talla de palangre de Japón y Corea utilizados en la evaluación de referencia.

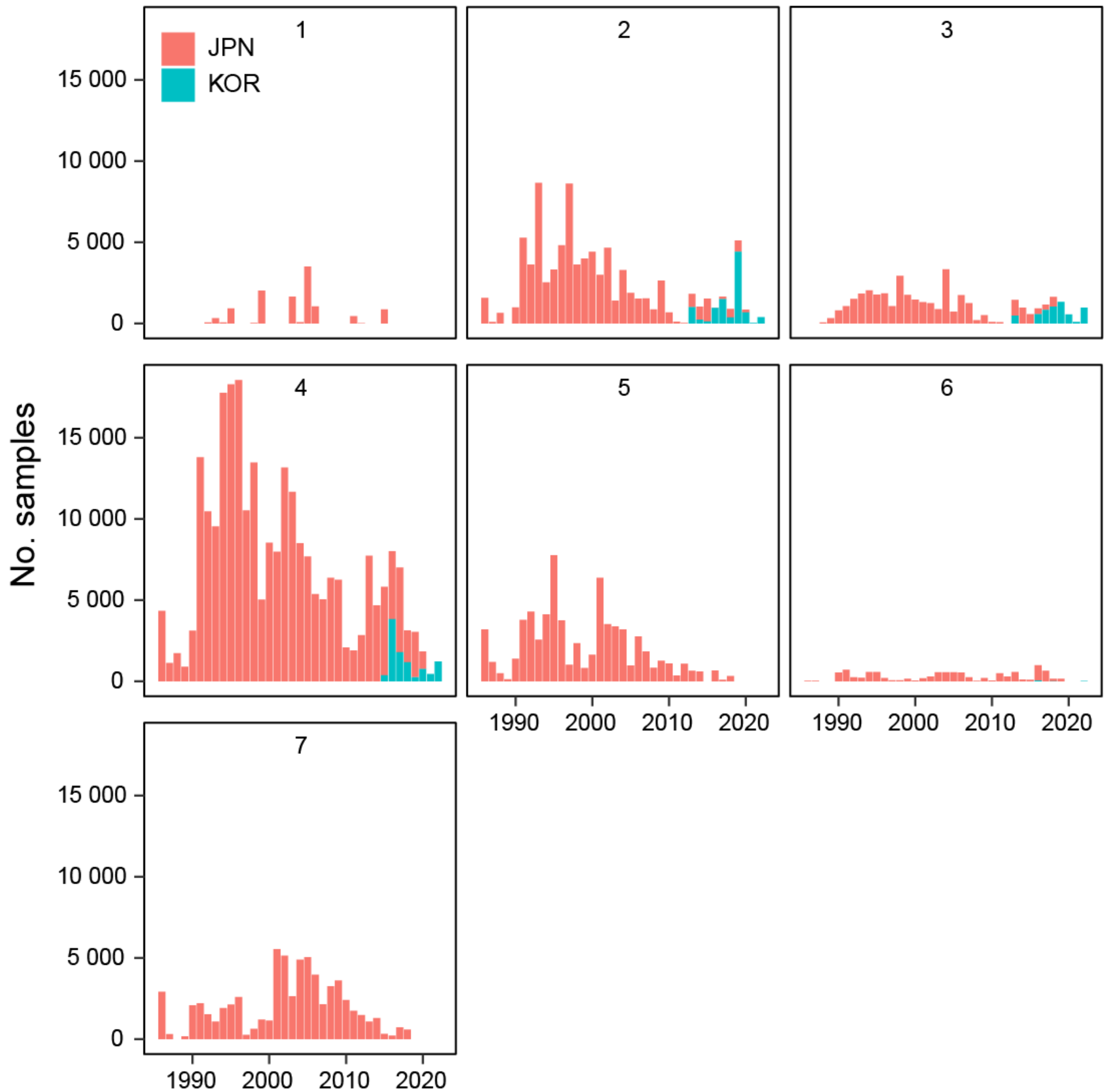

FIGURE 16. The number of fish sampled, by year and fishery, in the Japanese and Korean longline length composition data used in the benchmark assessment.

FIGURA 16. Número de peces muestreados, por año y pesquería, en los datos de composición por talla de palangre de Japón y Corea utilizados en la evaluación de referencia.

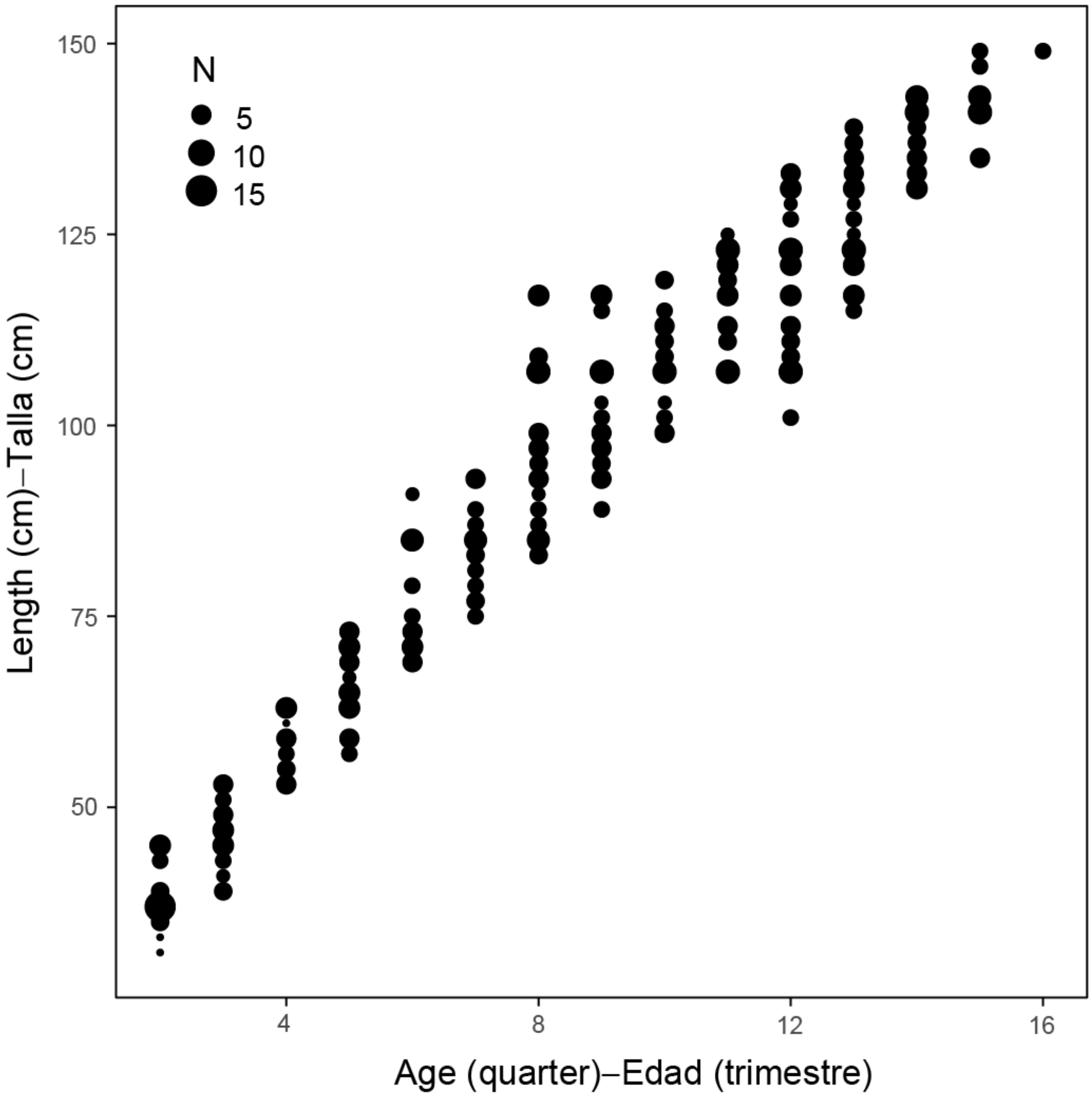

FIGURE 17. Age conditional on length for bigeye tuna in the EPO. The size of black dots represents the number of fish (N) for each age by 10 cm intervals.
FIGURA 17. Edad condicionada a la talla para el patudo en el OPO. El tamaño de los puntos negros representa el número de peces (N) para cada edad a intervalos de 10 cm.

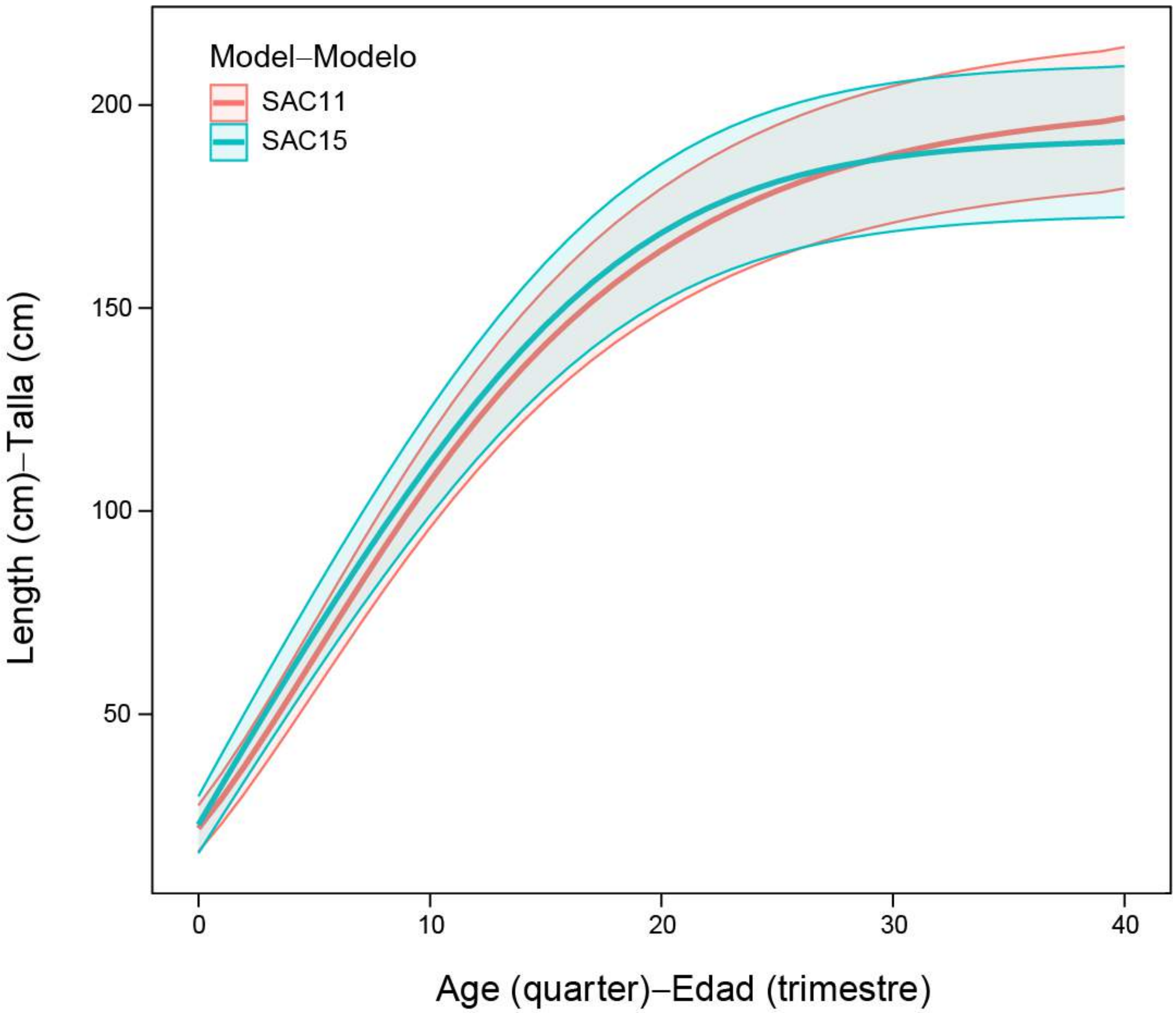

FIGURE 18. Comparison of the growth curve for bigeye tuna in the last benchmark assessment (SAC-11; Richards model) and in this benchmark assessment (SAC-15; growth cessation model). The shaded areas represent the 95% confidence interval of length at age.

FIGURA 18. Comparación de la curva de crecimiento del atún patudo en la última evaluación de referencia (SAC-11; modelo de Richards) y en esta evaluación de referencia (SAC-15; modelo de cese de crecimiento). Las áreas sombreadas representan el intervalo de confianza del 95% de la talla por edad.

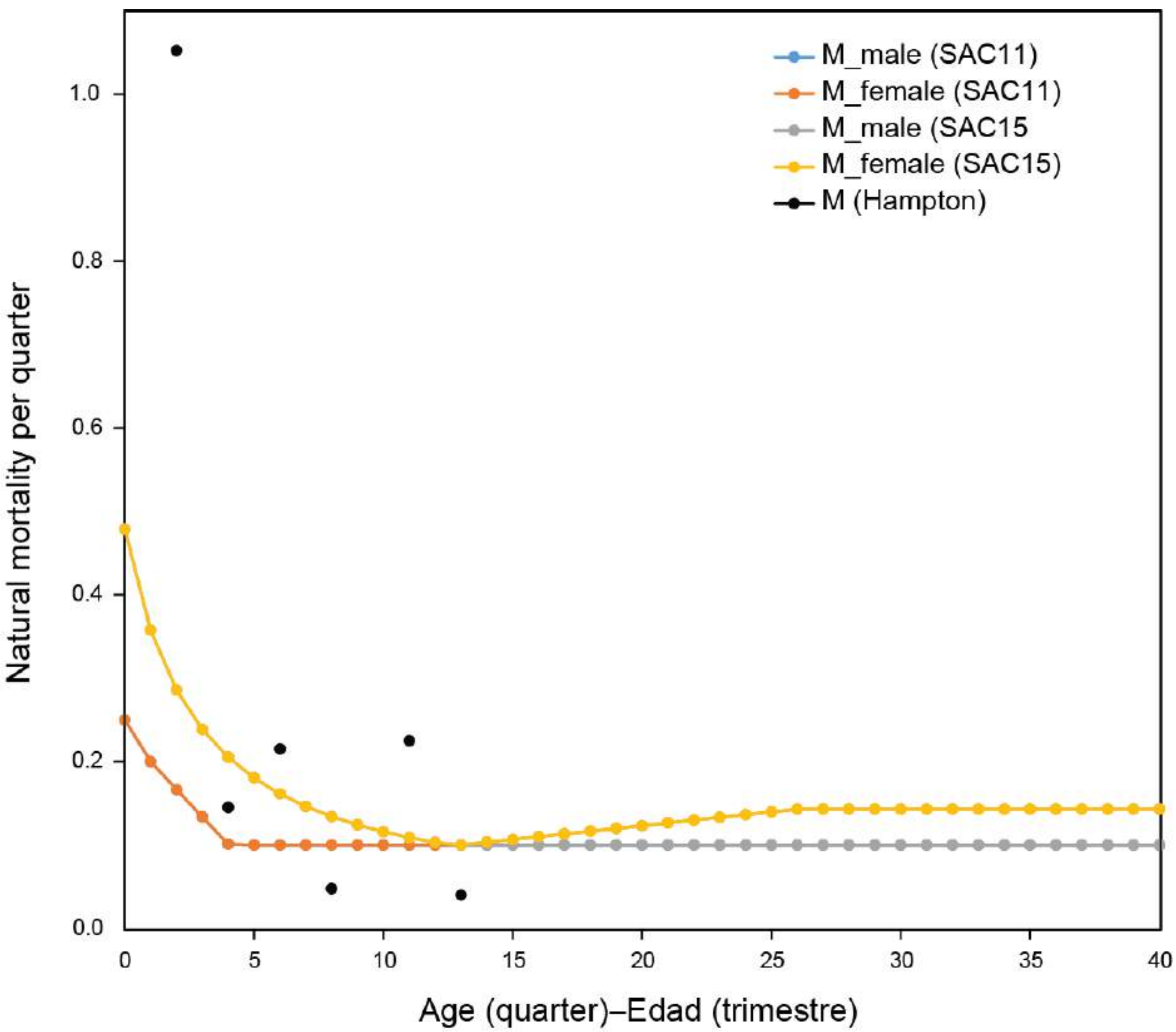

FIGURE 19. Comparison of the sex-specific natural mortality vectors for bigeye tuna in the last benchmark assessment (SAC-11) and this benchmark assessment (SAC-15). The natural mortality rates estimated by Hampton (2000) are also included for reference.

FIGURA 19. Comparación de los vectores de mortalidad natural por sexo del atún patudo en la última evaluación de referencia (SAC-11) y esta evaluación de referencia (SAC-15). También se incluyen como referencia las tasas de mortalidad natural estimadas por Hampton (2000).

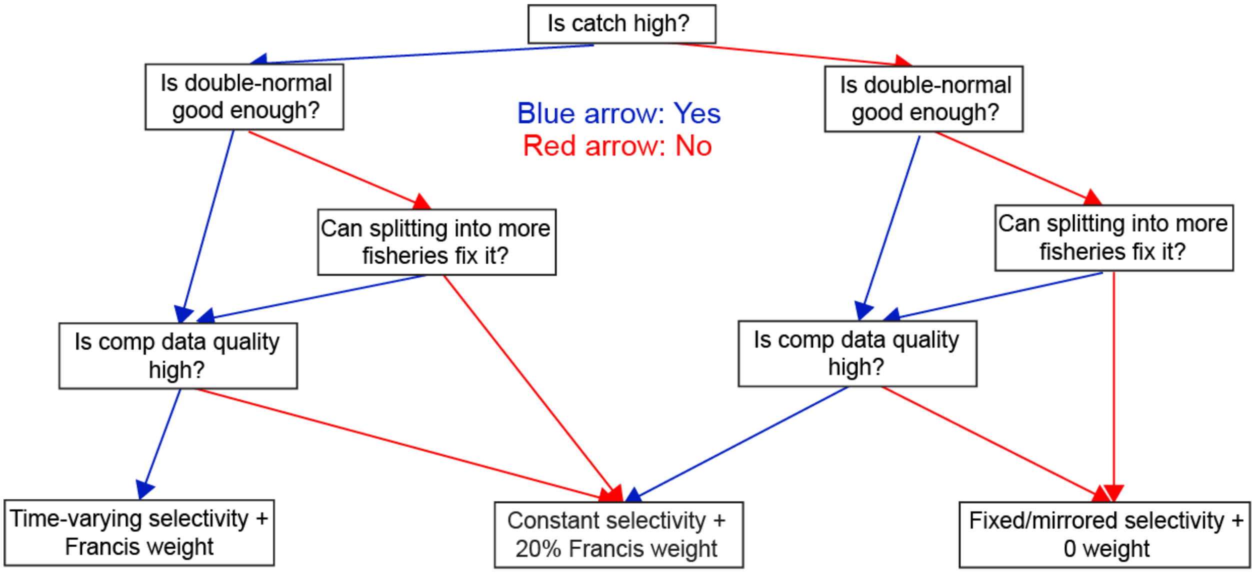

FIGURE 20. The decision tree on which the selectivity form and composition data weighting in this benchmark assessment are based.

FIGURA 20. Árbol de decisión en el que se basa la forma de la selectividad y la ponderación de los datos de composición en esta evaluación de referencia.

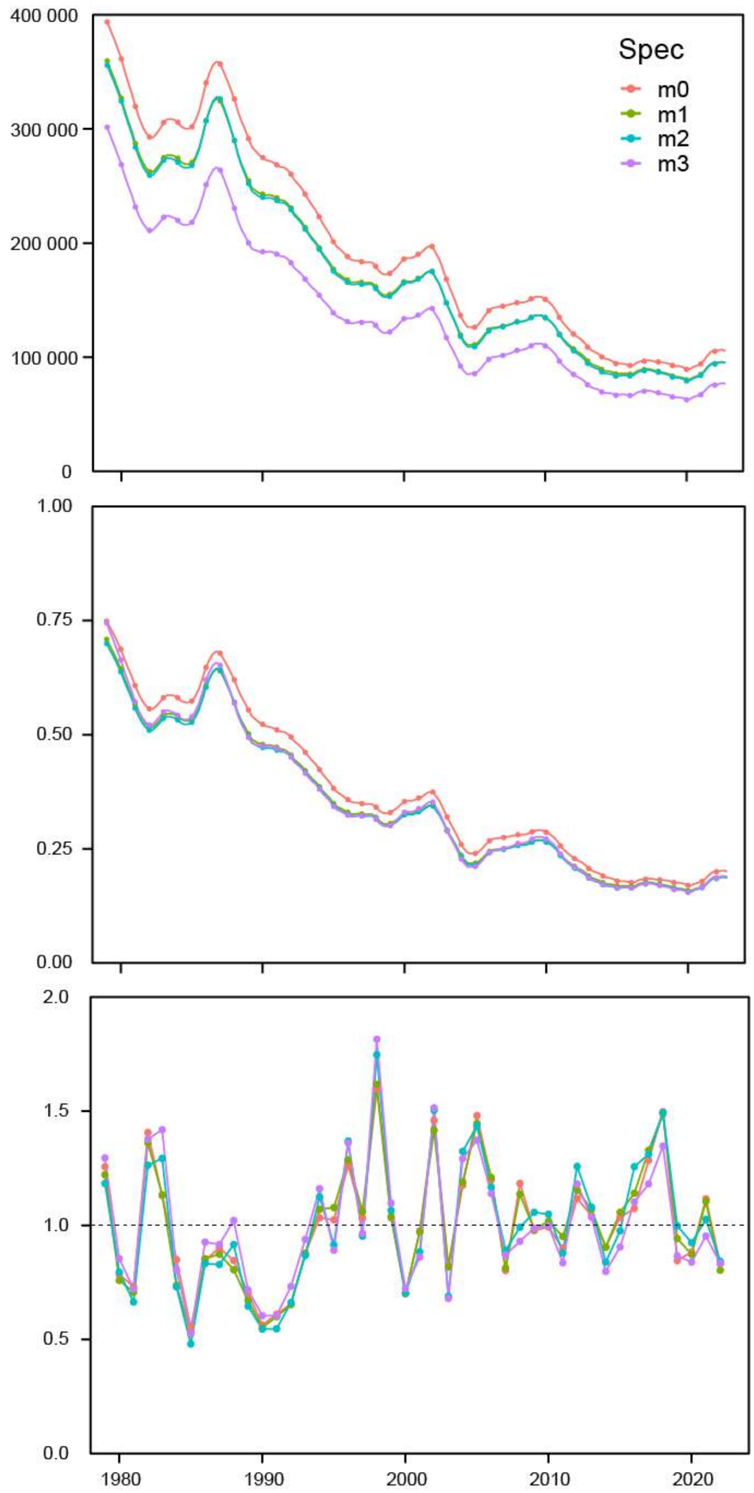

FIGURE 21. Comparison of estimated spawning biomass (top), spawning biomass ratio (middle), and relative recruitment (bottom) from the four models compared in the bridging analysis.

FIGURA 21. Comparación de las estimaciones de biomasa reproductora (arriba), cociente de biomasa reproductora (centro) y reclutamiento relativo (abajo) de los cuatro modelos comparados en el análisis de transición.

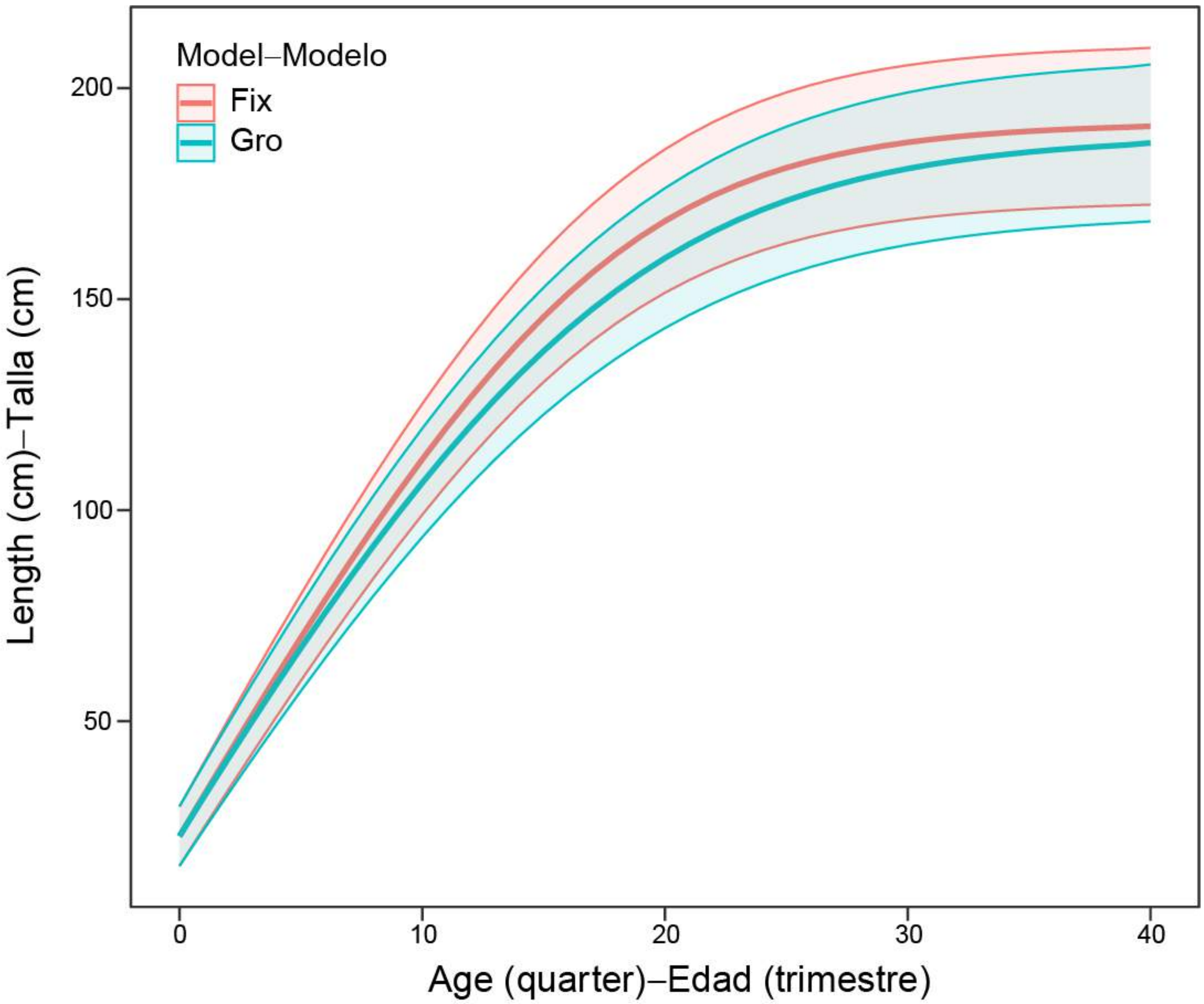

FIGURE 22. The fixed and estimated growth curves for bigeye tuna in this benchmark assessment. Only the estimated growth curve with constant catchability and steepness of 1.0 is shown because it is not sensitive to these two hypotheses.

FIGURA 22. Curvas de crecimiento fijo y estimado del atún patudo en esta evaluación de referencia. Solo se muestra la curva de crecimiento estimado con capturabilidad constante e inclinación de 1.0 porque no es sensible a estas dos hipótesis.

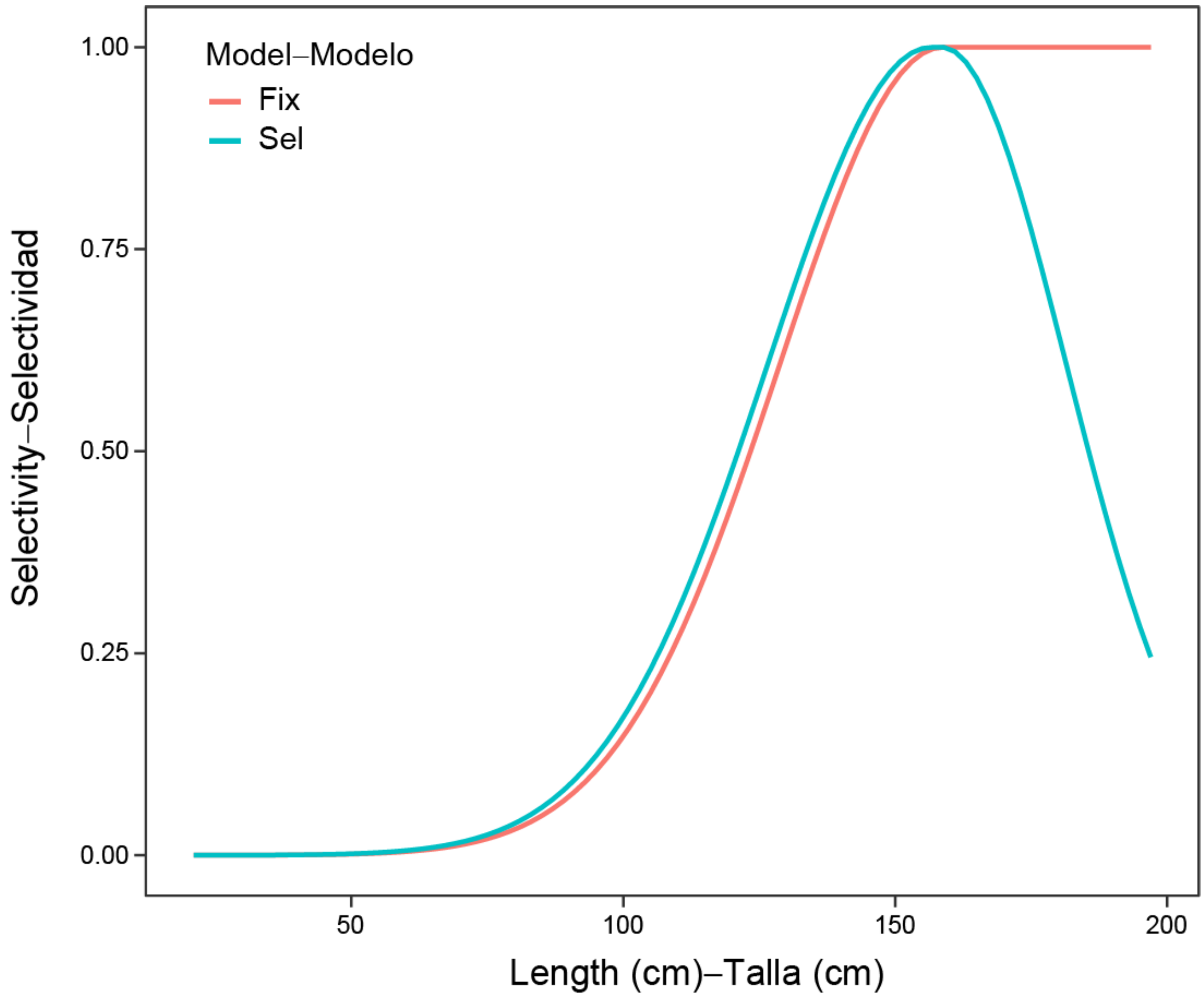

FIGURE 23. The fixed and estimated selectivity curves for the longline fishery in Area 4 (Fishery 4) between 2011 and 2023. Only the estimated selectivity curve with constant catchability and steepness of 1.0 is shown because it is not sensitive to these two hypotheses.

FIGURA 23. Curvas de selectividad fija y estimada para la pesquería de palangre en el Área 4 (Pesquería 4) entre 2011 y 2023. Solo se muestra la curva de selectividad estimada con capturabilidad constante e inclinación de 1.0 porque no es sensible a estas dos hipótesis.

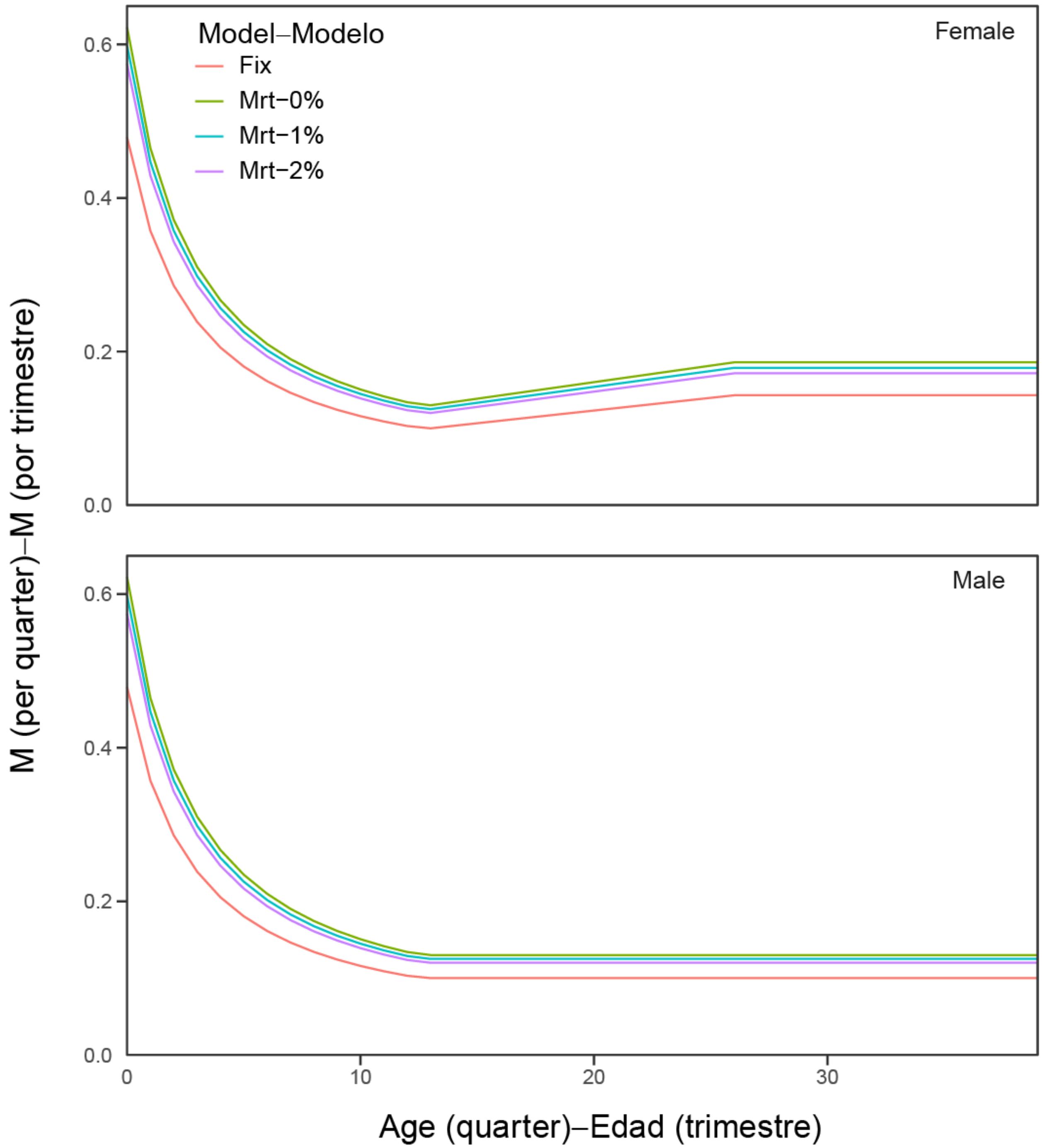

FIGURE 24. The fixed and estimated sex-specific natural mortality vectors used in this benchmark assessment. Natural mortality is estimated to be slightly different under different assumptions of the annual increasing rate of longline catchability.

FIGURA 24. Vectores de mortalidad natural por sexo fija y estimada utilizados en esta evaluación de referencia. Se estima que la mortalidad natural es ligeramente diferente bajo distintos supuestos de la tasa anual creciente de capturabilidad de palangre.

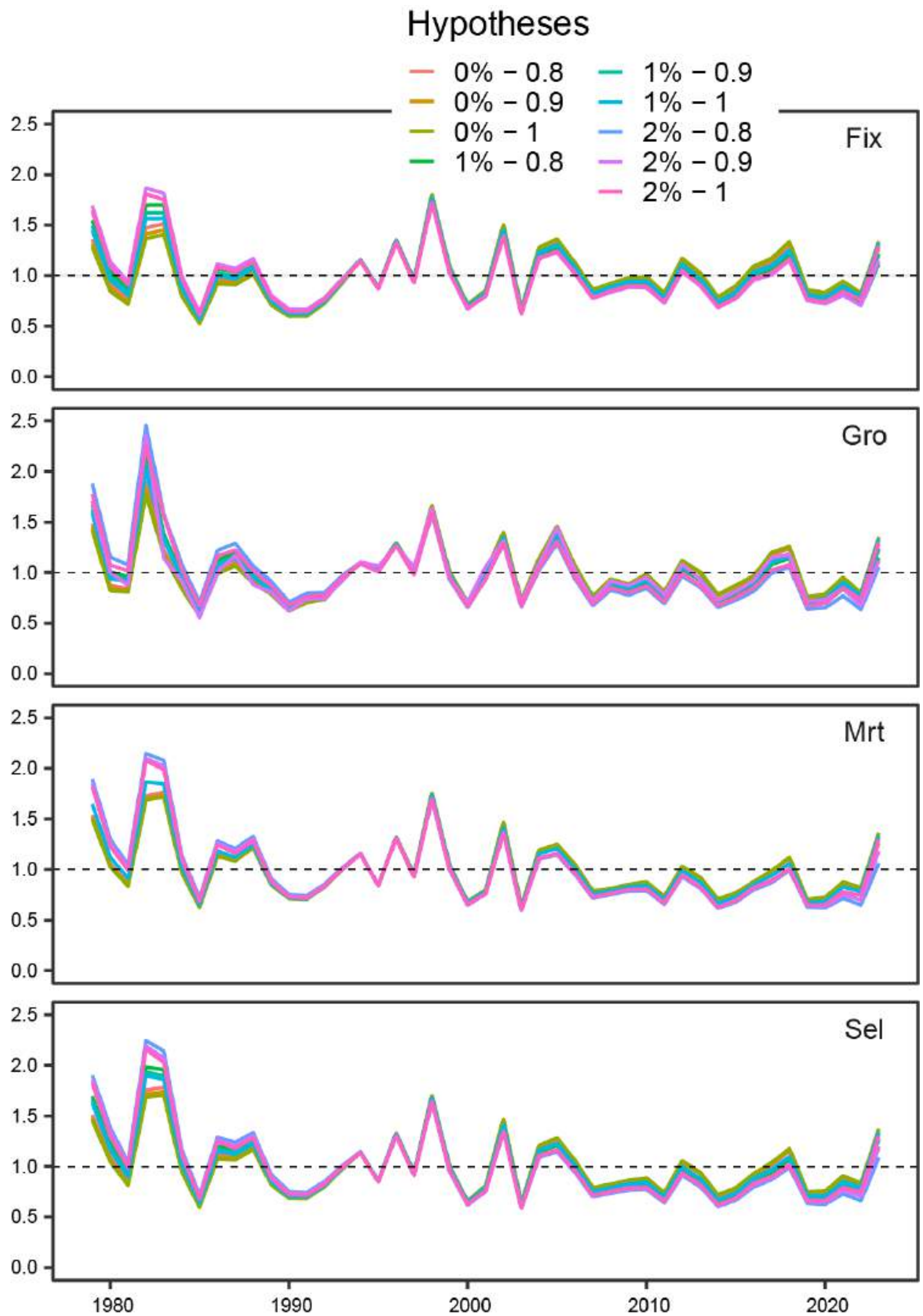

FIGURE 25. Comparison of estimated relative annual recruitment of bigeye tuna between 1979 and 2023.
FIGURA 25. Comparación del reclutamiento anual relativo estimado del atún patudo entre 1979 y 2023.

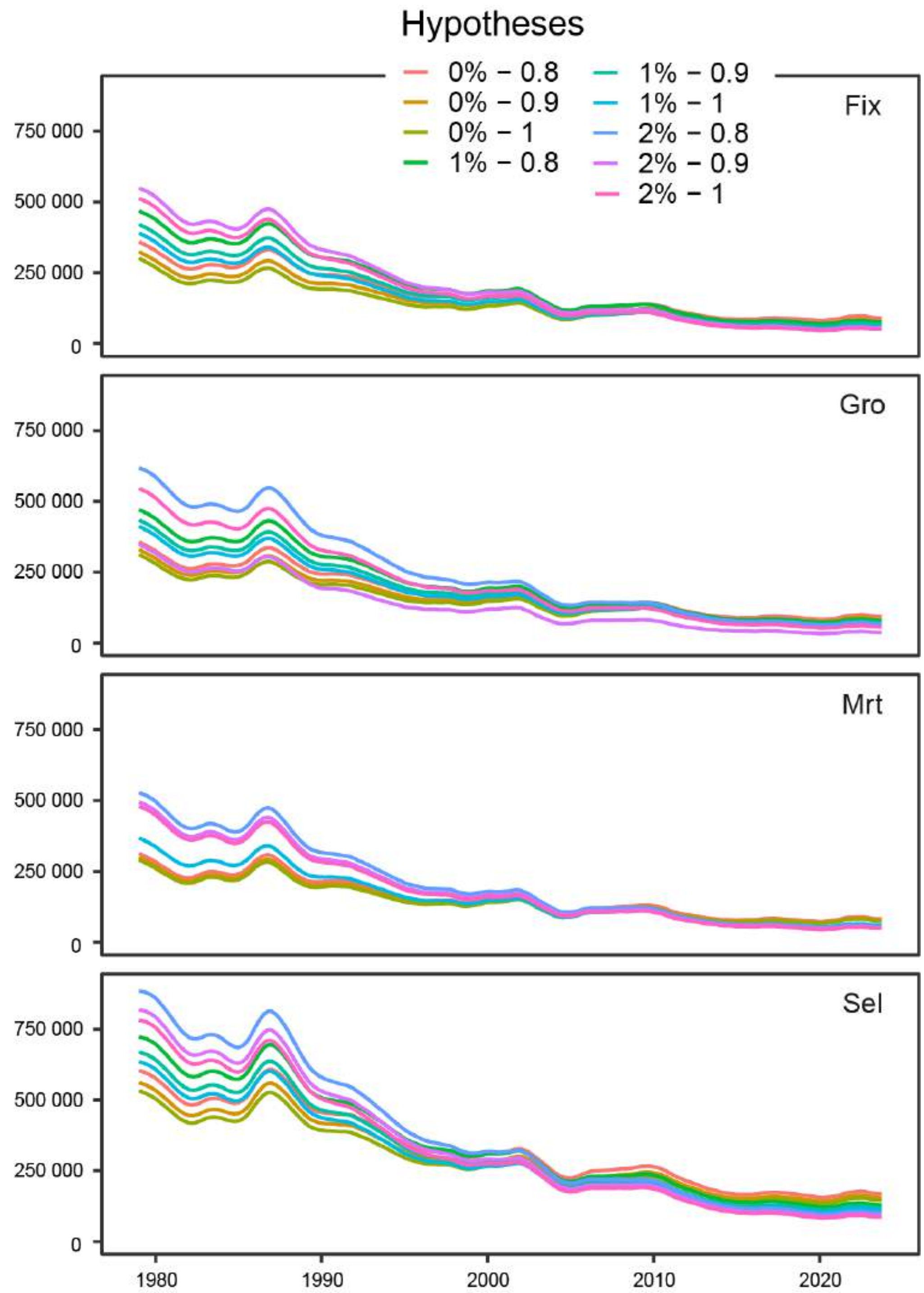

FIGURE 26a. Comparison of estimated spawning biomass of bigeye tuna between 1979 and 2023.
FIGURA 26a. Comparación de la biomasa reproductora estimada del atún patudo entre 1979 y 2023.

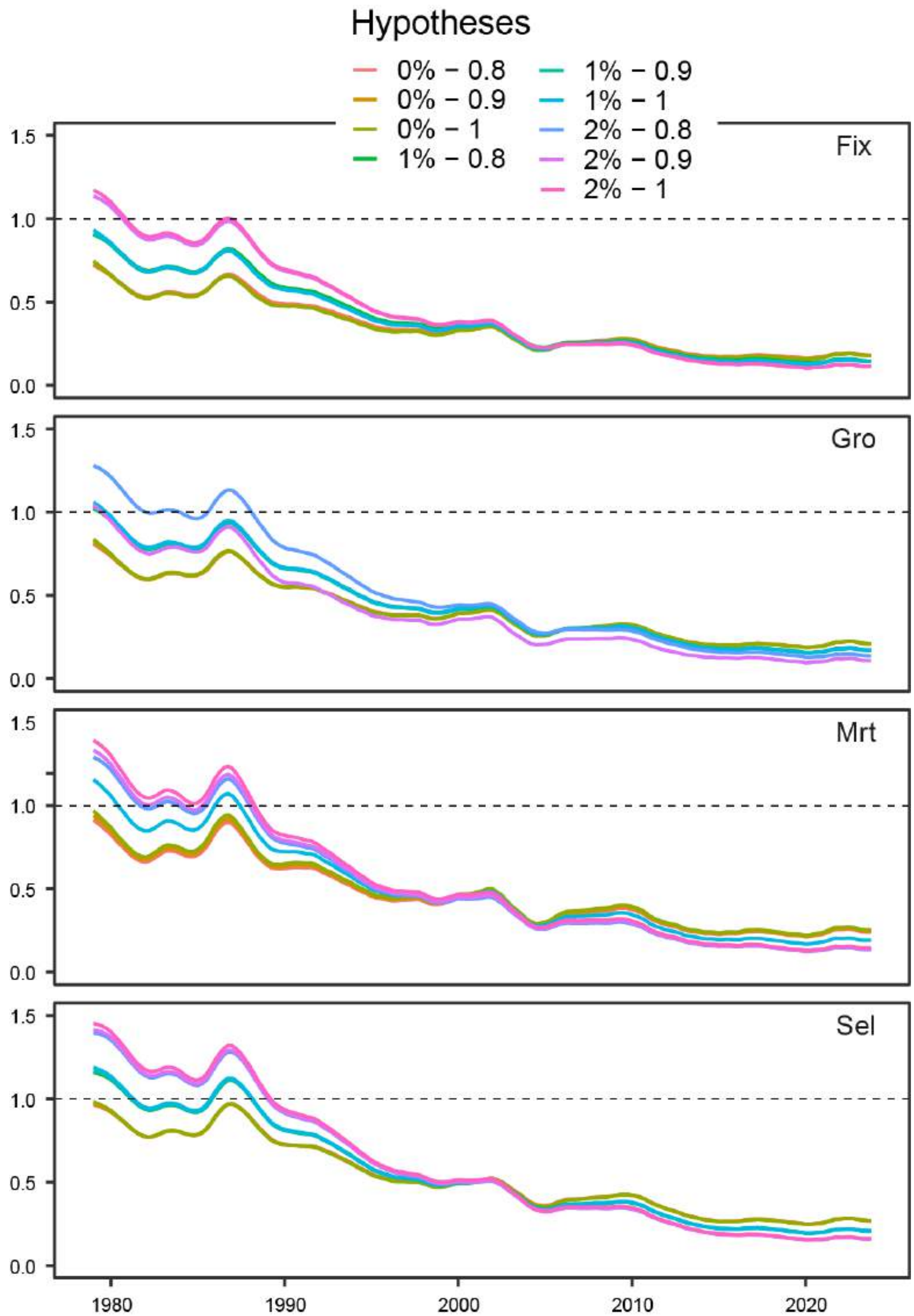

FIGURE 26b. Comparison of estimated spawning biomass ratio of bigeye tuna between 1979 and 2023.
FIGURA 26b. Comparación del cociente de biomasa reproductora estimado del atún patudo entre 1979 y 2023.

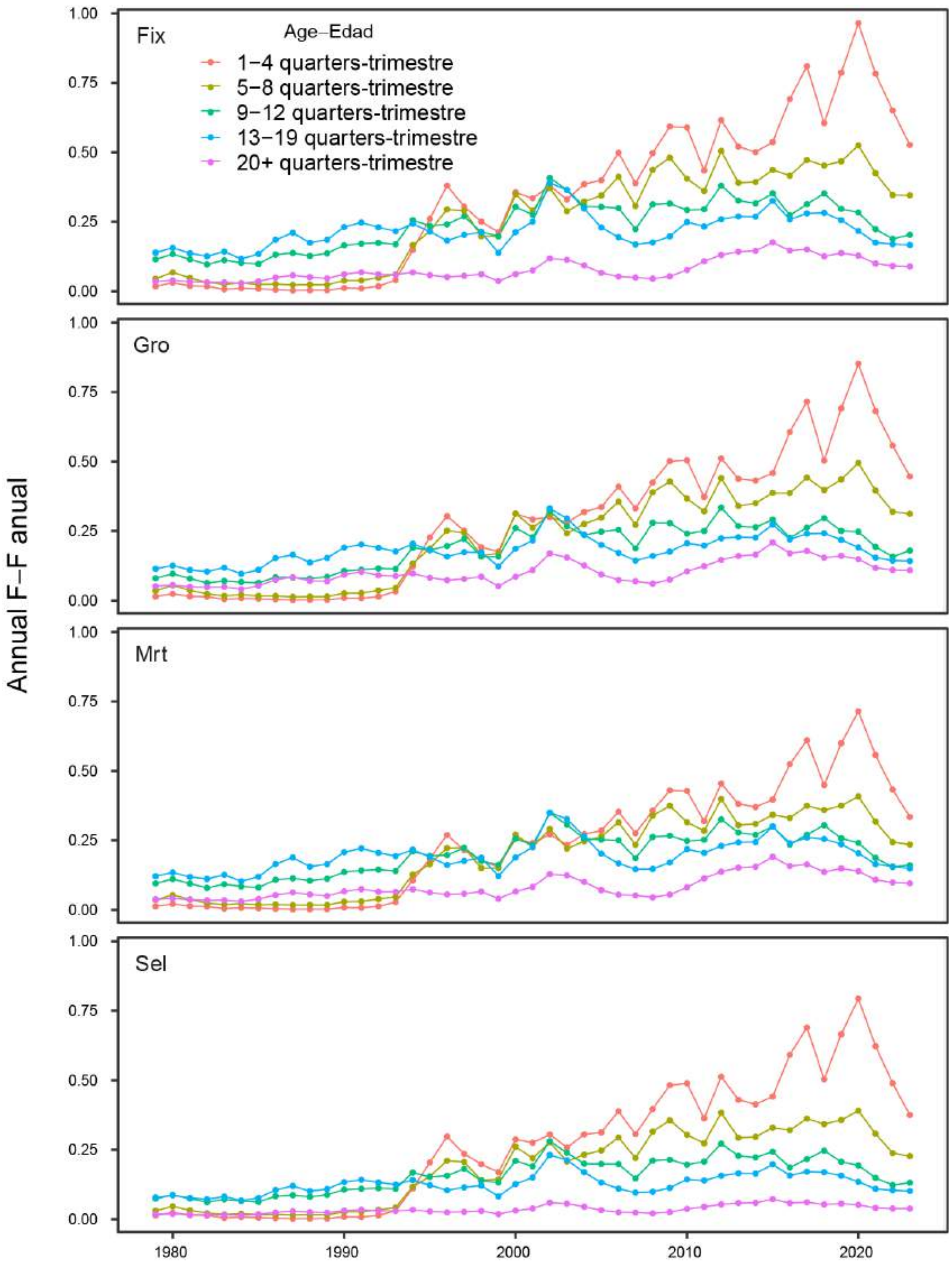

FIGURE 27. Comparison of average annual fishing mortality, by age groups, of bigeye tuna between 1979 and 2023. The values for each model and age group are weighted across the second- and third-level hypotheses.

FIGURA 27. Comparación de la mortalidad por pesca anual promedio, por grupos de edad, del atún patudo entre 1979 y 2023. Los valores para cada modelo y grupo de edad se ponderan en las hipótesis de segundo y tercer nivel.

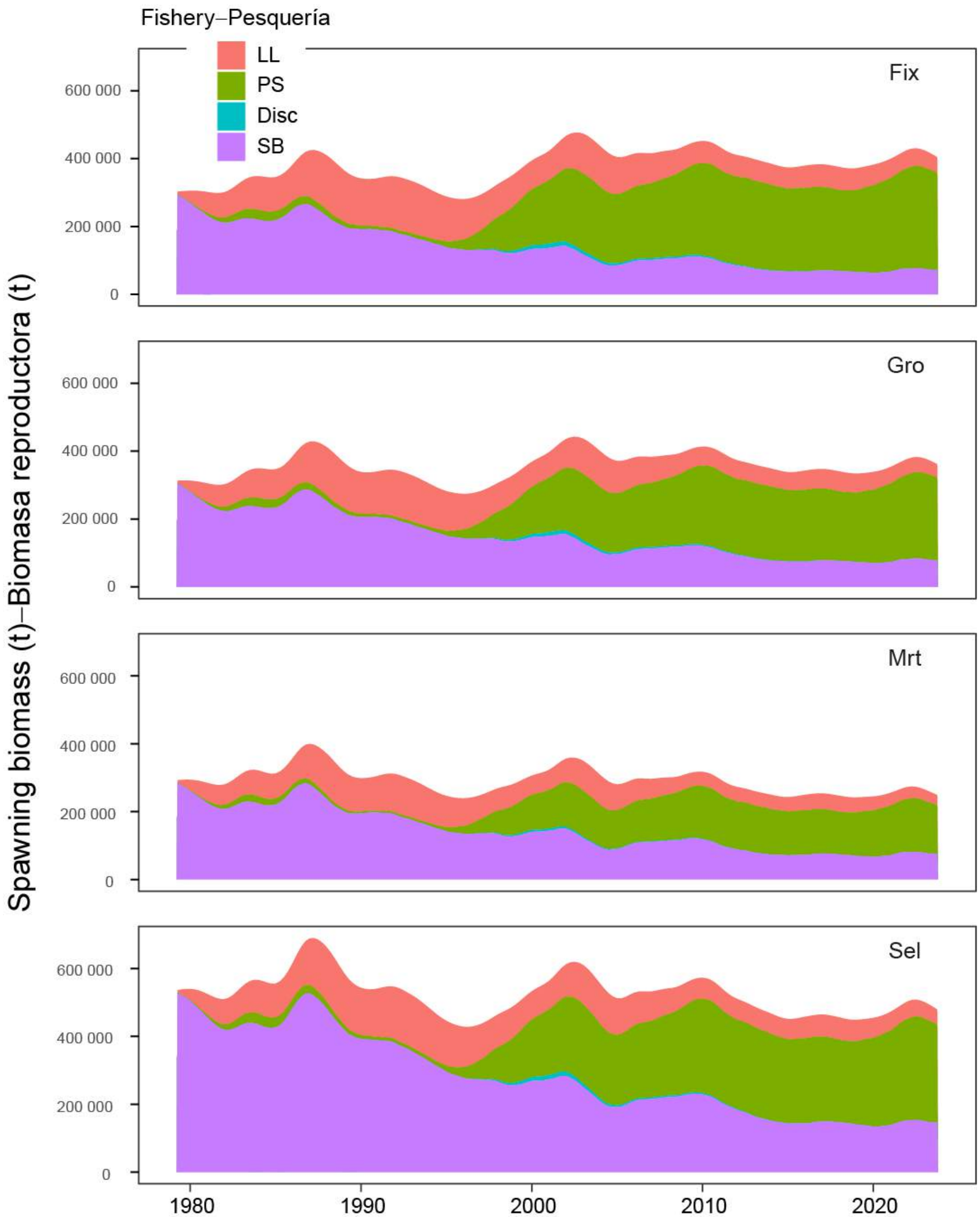

FIGURE 28. Comparison of spawning biomass trajectory of a simulated population of bigeye tuna that was never exploited (top line) and that predicted by the stock assessment model (bottom line). The shaded blue, green, and red areas show the proportional impact of the discard, purse-seine, and longline fishery, respectively.

FIGURA 28. Comparación de la trayectoria de la biomasa reproductora de una población simulada de atún patudo que nunca fue explotada (línea superior) y la predicha por el modelo de evaluación (línea inferior). Las áreas sombreadas en azul, verde y rojo muestran el impacto proporcional de las pesquerías de descarte, cerco y palangre, respectivamente.

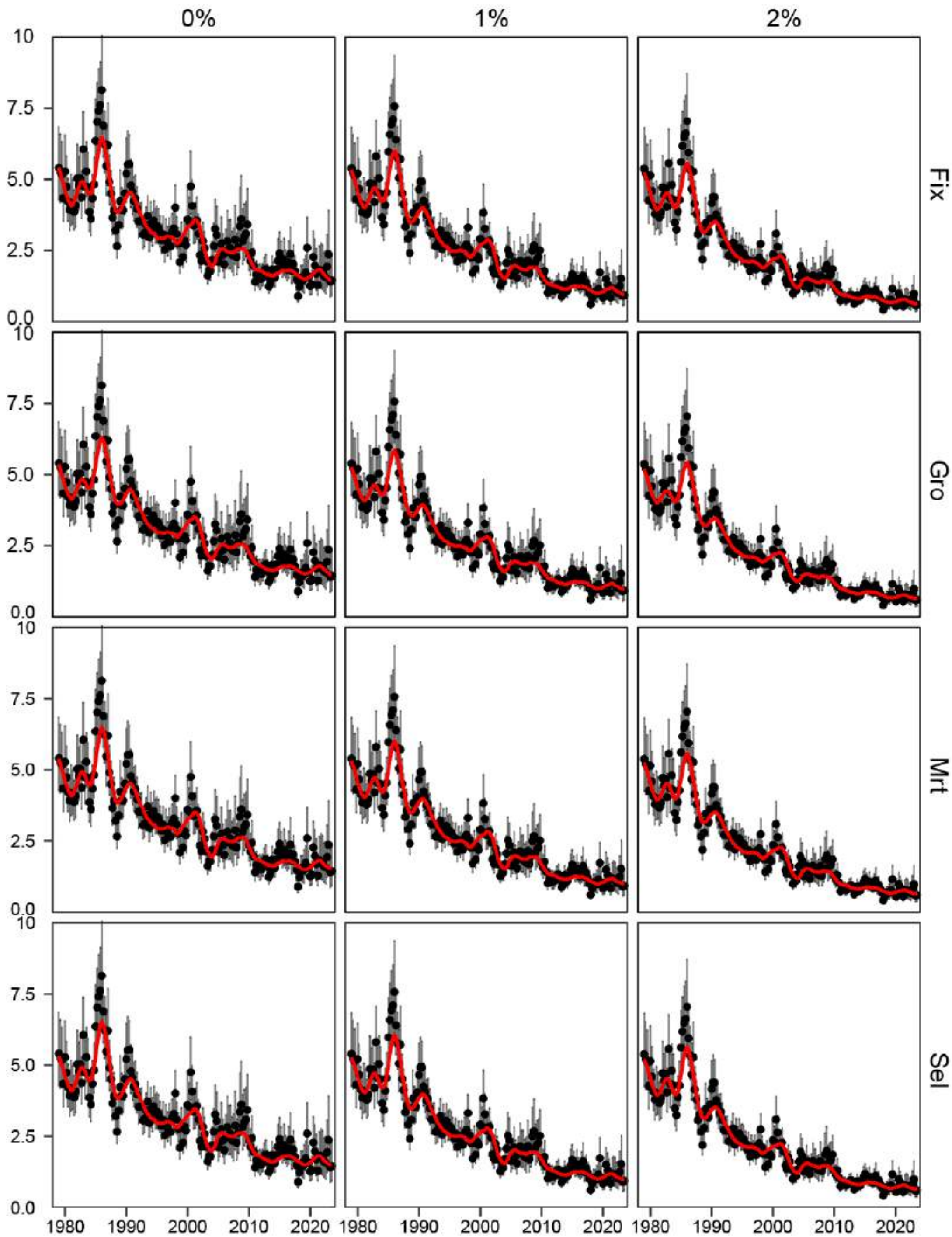

FIGURE 29. Fit to the longline index of relative abundance. The black dots and error bars represent the observed values and their 95% confidence interval. The solid red lines are predicted values from the stock assessment model.

FIGURA 29. Ajuste al índice de abundancia relativa de palangre. Los puntos negros y las barras de error representan los valores observados y su intervalo de confianza del 95%. Las líneas rojas son los valores predichos a partir del modelo de evaluación.

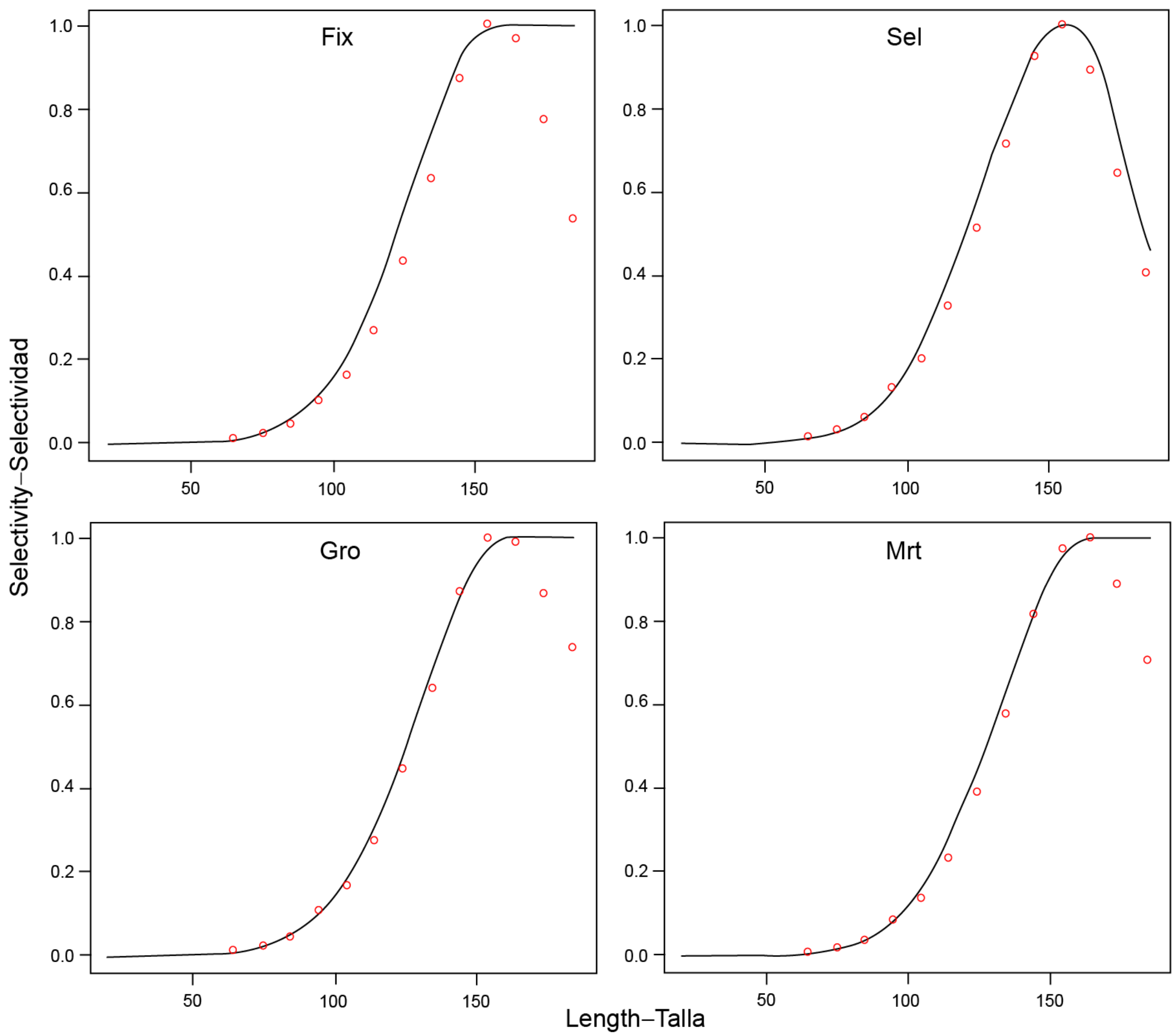

FIGURE 30. Comparison of estimated (black line) and empirical (red dots) selectivity for Fishery 4 between 2011 and 2023.

FIGURA 30. Comparación de la selectividad estimada (línea negra) y empírica (puntos rojos) para la Pesquería 4 entre 2011 y 2023.

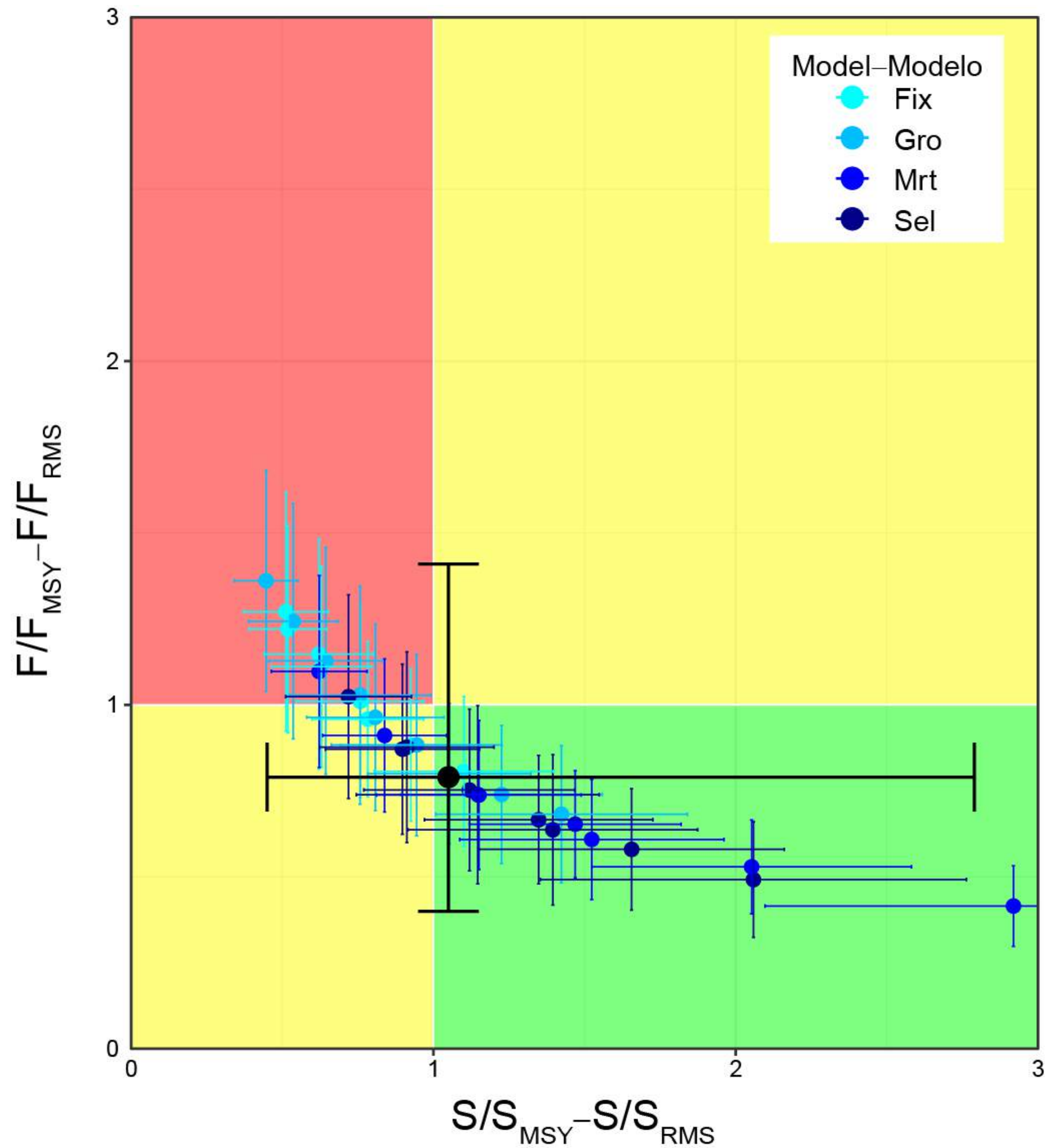

FIGURE 31. Kobe plot of the most recent estimates of spawning biomass (*S*) and fishing mortality (*F*) relative to their MSY reference points (S_{MSY_d} and F_{MSY}) from the thirty-three reference models. Each dot is based on the average *F* over the most recent three years, 2021-2023, and the error bars represent the 95% confidence interval of model estimates. The black dot and error bars represent the medium and 95% confidence interval of combined values, respectively.

FIGURA 31. 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} and F_{RMS}) de los 33 modelos de referencia. Cada punto se basa en la *F* promedio de los últimos tres años, 2021-2023, y las barras de error representan el intervalo de confianza del 95% de las estimaciones de los modelos. El punto negro y las barras de error representan el intervalo de confianza medio y del 95% de los valores combinados, respectivamente.

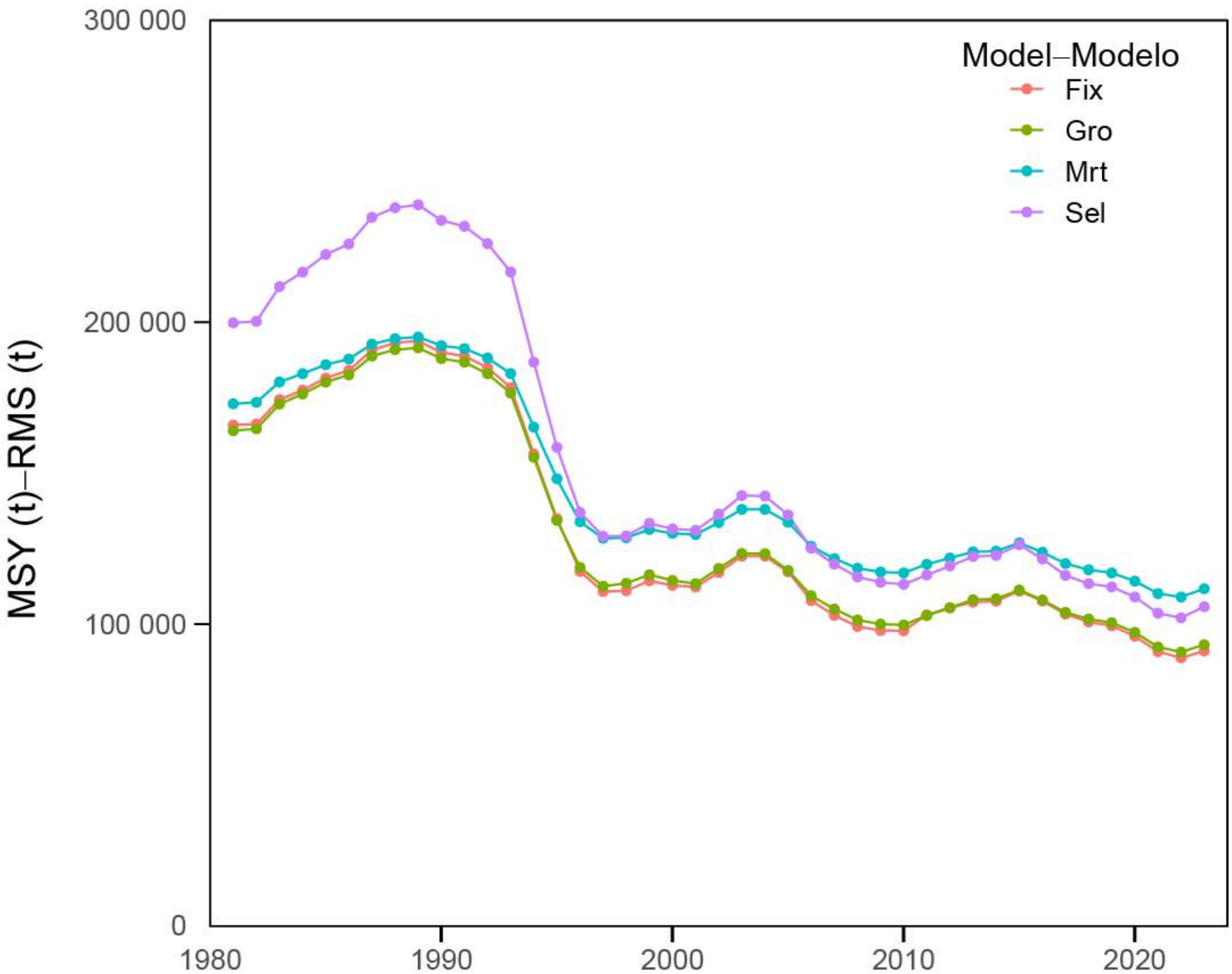

FIGURE 32. Estimates of maximum sustainable yield for the four models considered in the level 1 hypothesis using the average age-specific fishing mortality during each three years. The values for each model and age group are weighted across the second- and third-level hypotheses.

FIGURA 32. Estimaciones del rendimiento máximo sostenible para los cuatro modelos considerados en la hipótesis de nivel 1 utilizando la mortalidad por pesca promedio por edad durante cada tres años. Los valores de cada modelo y grupo de edad se ponderan en las hipótesis de segundo y tercer nivel.

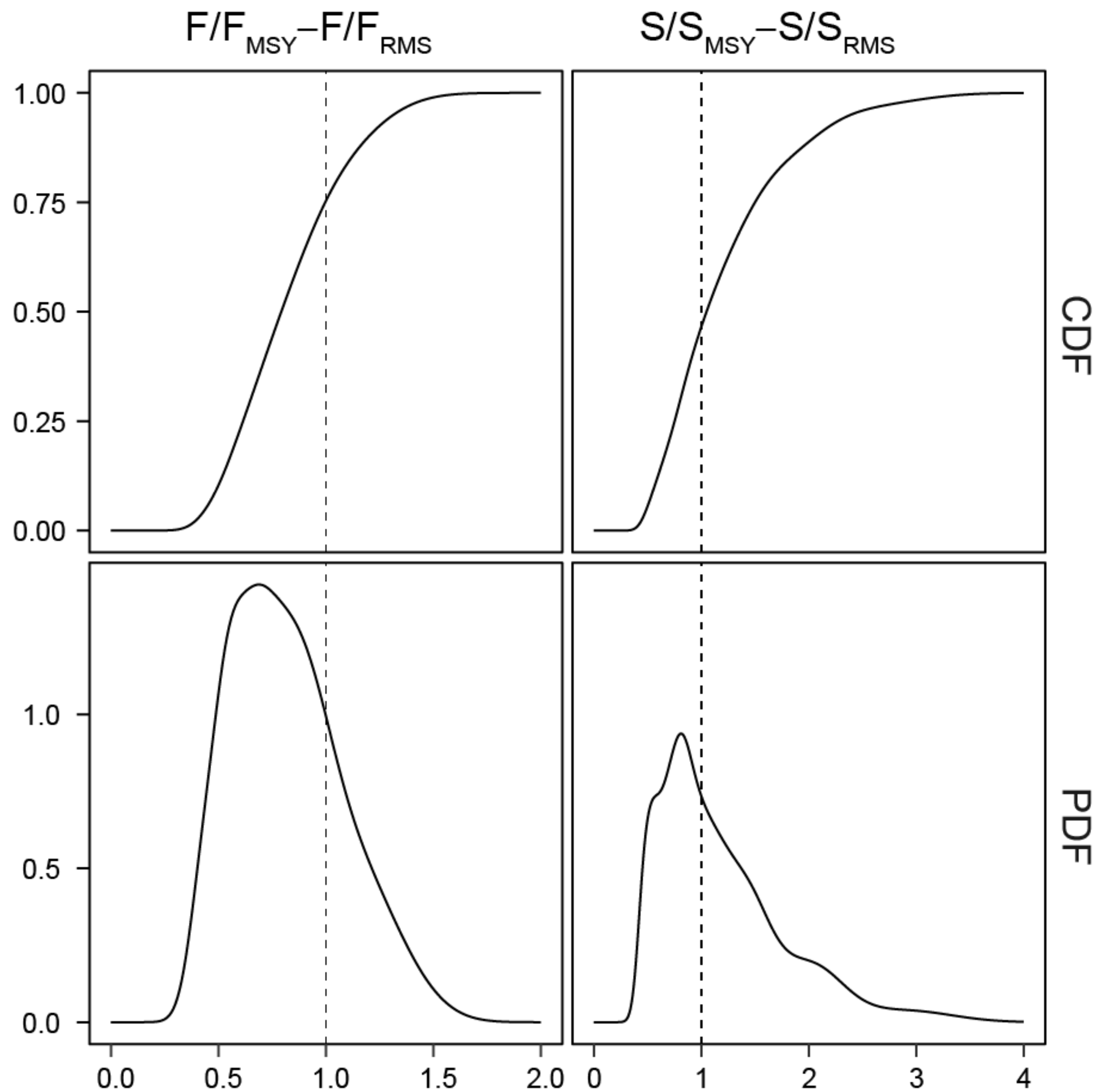

FIGURE 33. The joint probability and cumulative distribution functions for spawning biomass (S) in the first quarter of 2024 and fishing mortality (F) in 2021-2023 relative to their MSY reference points (S_{MSY_d} and F_{MSY}).

FIGURA 33. Funciones de distribución acumulada y de probabilidad conjunta para la biomasa reproductora (*S*) en el primer trimestre de 2024 y la mortalidad por pesca (*F*) en 2021-2023 en relación con sus puntos de referencia de RMS (S_{RMS_d} y F_{RMS}).

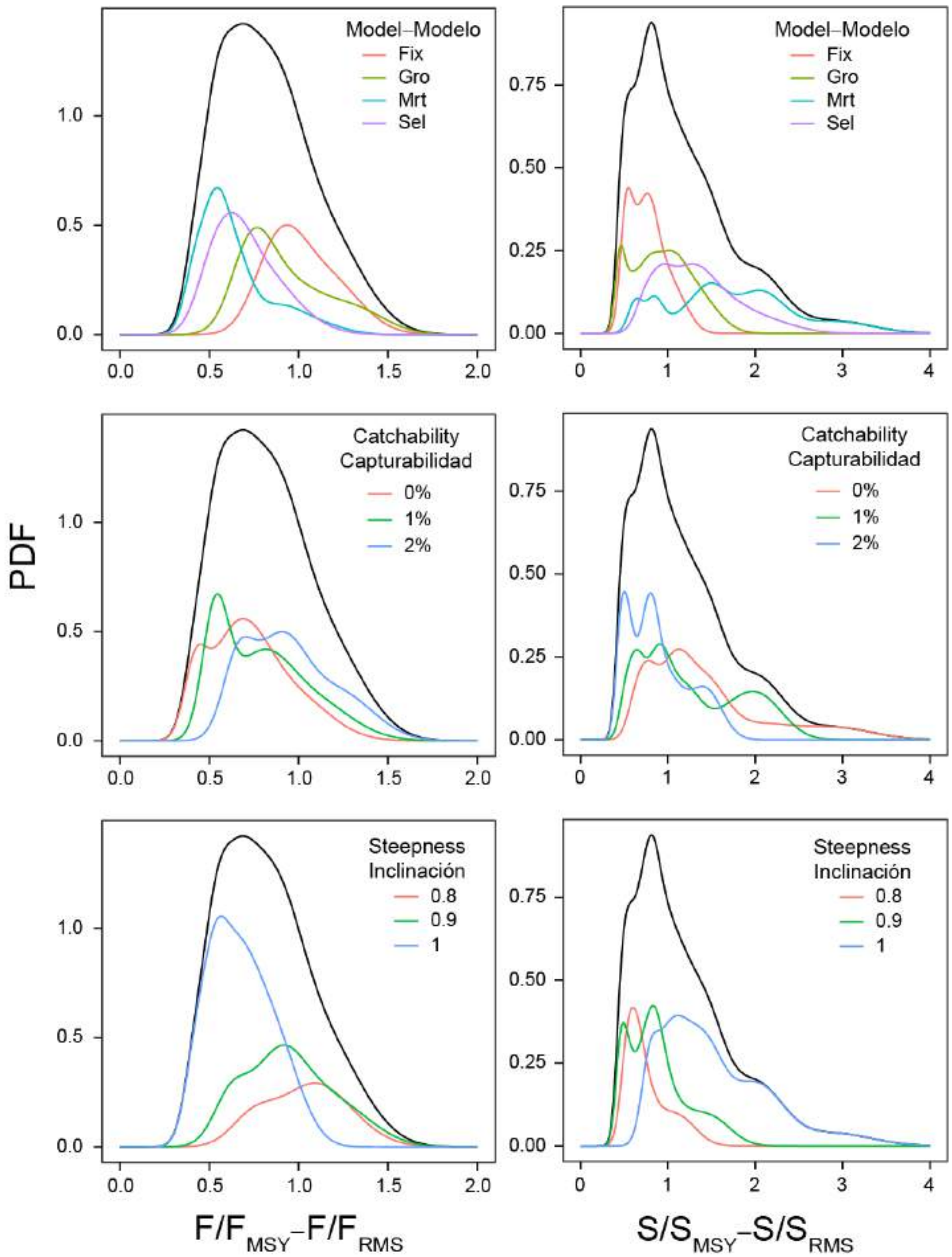

FIGURE 34. The joint probability distribution functions for F_{recent}/F_{MSY} and $S_{current}/S_{MSY}$ broken down into different components of the three hypotheses to address: (top) the misfit to the length-composition data for the longline fishery that is assumed to have asymptotic selectivity; (middle) the degree of effort creep in the longline index of relative abundance; and (bottom) the steepness of the stock-recruitment relationship.

FIGURA 34. Funciones de distribución de probabilidad conjunta para F_{actual}/F_{RMS} y S_{actual}/S_{RMS} desglosadas en diferentes componentes de las tres hipótesis a abordar: (arriba) el ajuste inadecuado a los datos de composición por talla para la pesquería de palangre con selectividad asintótica supuesta; (en medio) el grado de progresión del esfuerzo en el índice de abundancia relativa de palangre; y (abajo) la inclinación de la relación población-reclutamiento.

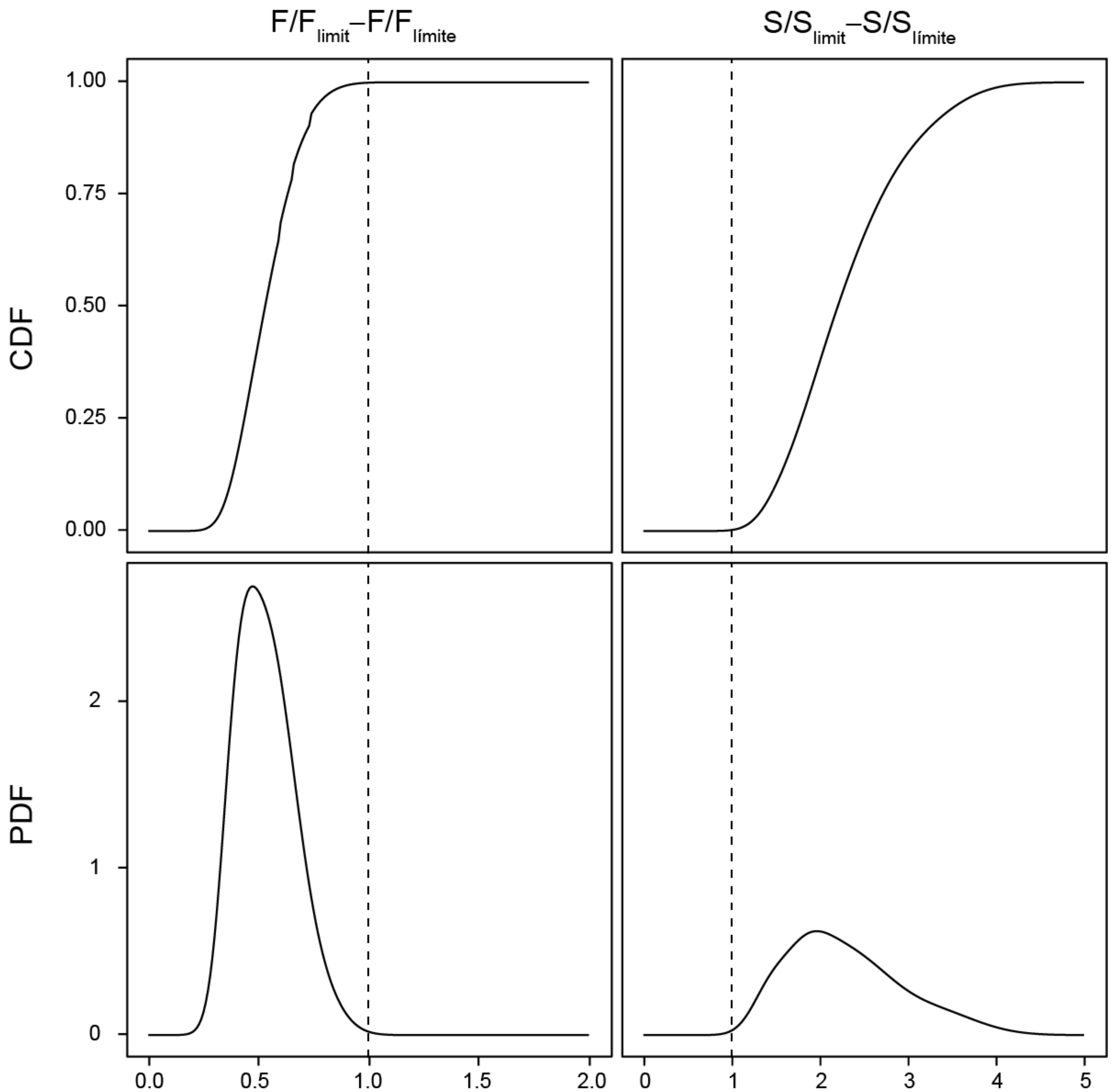

FIGURE 35. The joint probability and cumulative distribution functions for spawning biomass (S) in the first quarter of 2024 and fishing mortality (F) in 2021-2023 relative to their limit reference points (S_{Limit} and F_{Limit}).

FIGURA 35. Funciones de distribución acumulada y de probabilidad conjunta para la biomasa reproductora (*S*) en el primer trimestre de 2024 y la mortalidad por pesca (*F*) en 2021-2023 en relación con sus puntos de referencia límite ($S_{Límite}$ y $F_{Límite}$).

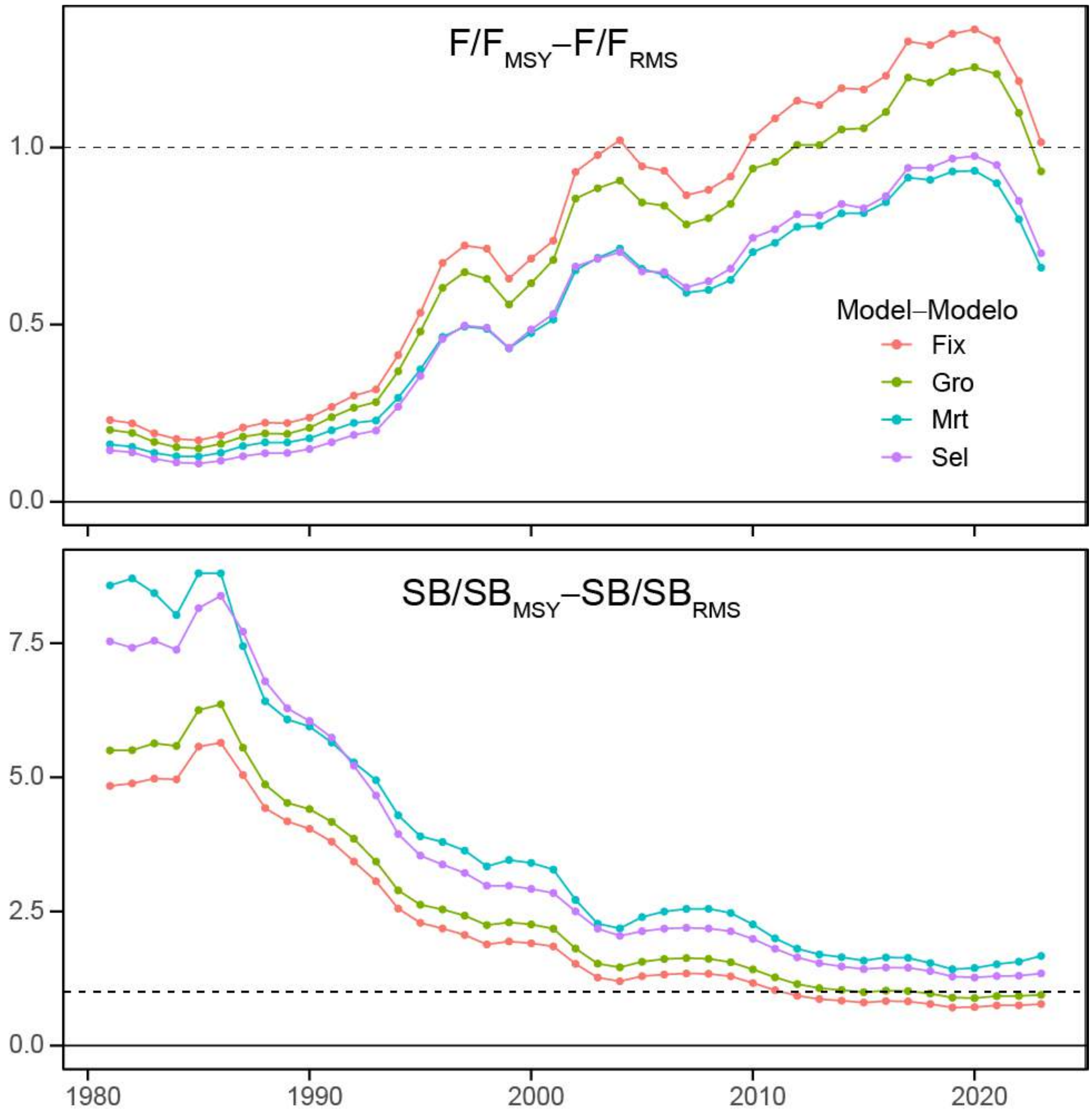

FIGURE 36. Time series of estimated spawning biomass (S) and fishing mortality (F) relative to their MSY reference points (S_{MSY_d} and F_{MSY}) for the four models considered in the level 1 hypothesis. The values for each model are weighted across the second- and third-level hypotheses and each dot for F is based on the average F over three years.

FIGURA 36. Series de tiempo de la biomasa reproductora (*S*) y la mortalidad por pesca (*F*) estimadas en relación con sus puntos de referencia de RMS (S_{RMS_d} y F_{RMS}) para los cuatro modelos considerados en la hipótesis de nivel 1. Los valores de cada modelo se ponderan en las hipótesis de segundo y tercer nivel y cada punto para *F* se basa en la *F* promedio a lo largo de tres años.

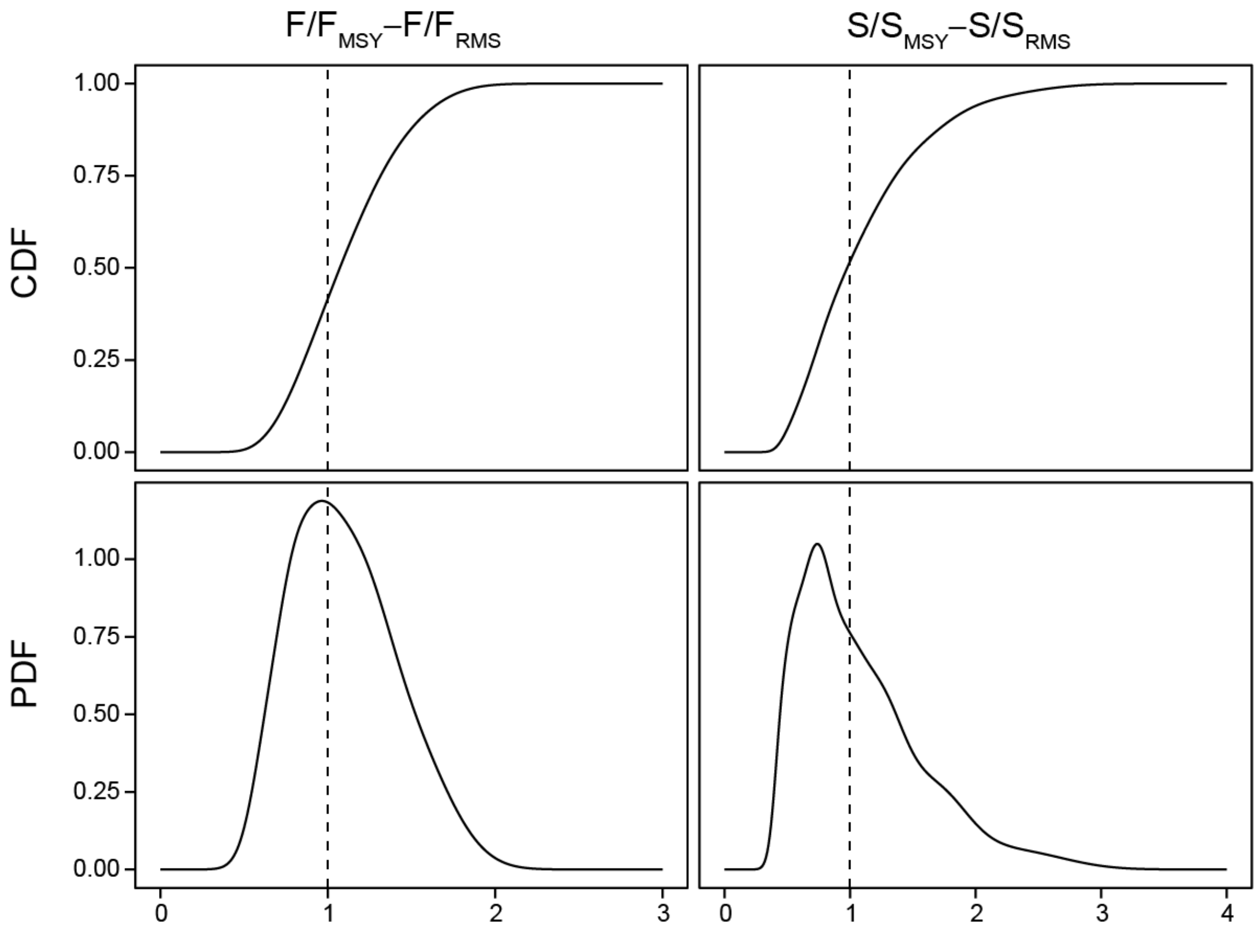

FIGURE 37. The joint probability and cumulative distribution functions for spawning biomass (*S*) in the first quarter of 2020 and fishing mortality (*F*) in 2017-2019 relative to their MSY reference points (S_{MSY_d} and F_{MSY}).

FIGURA 37. Funciones de distribución acumulada y de probabilidad conjunta para la biomasa reproductora (*S*) en el primer trimestre de 2020 y la mortalidad por pesca (*F*) en 2017-2019 en relación con sus puntos de referencia de RMS (S_{RMS_d} y F_{RMS}).

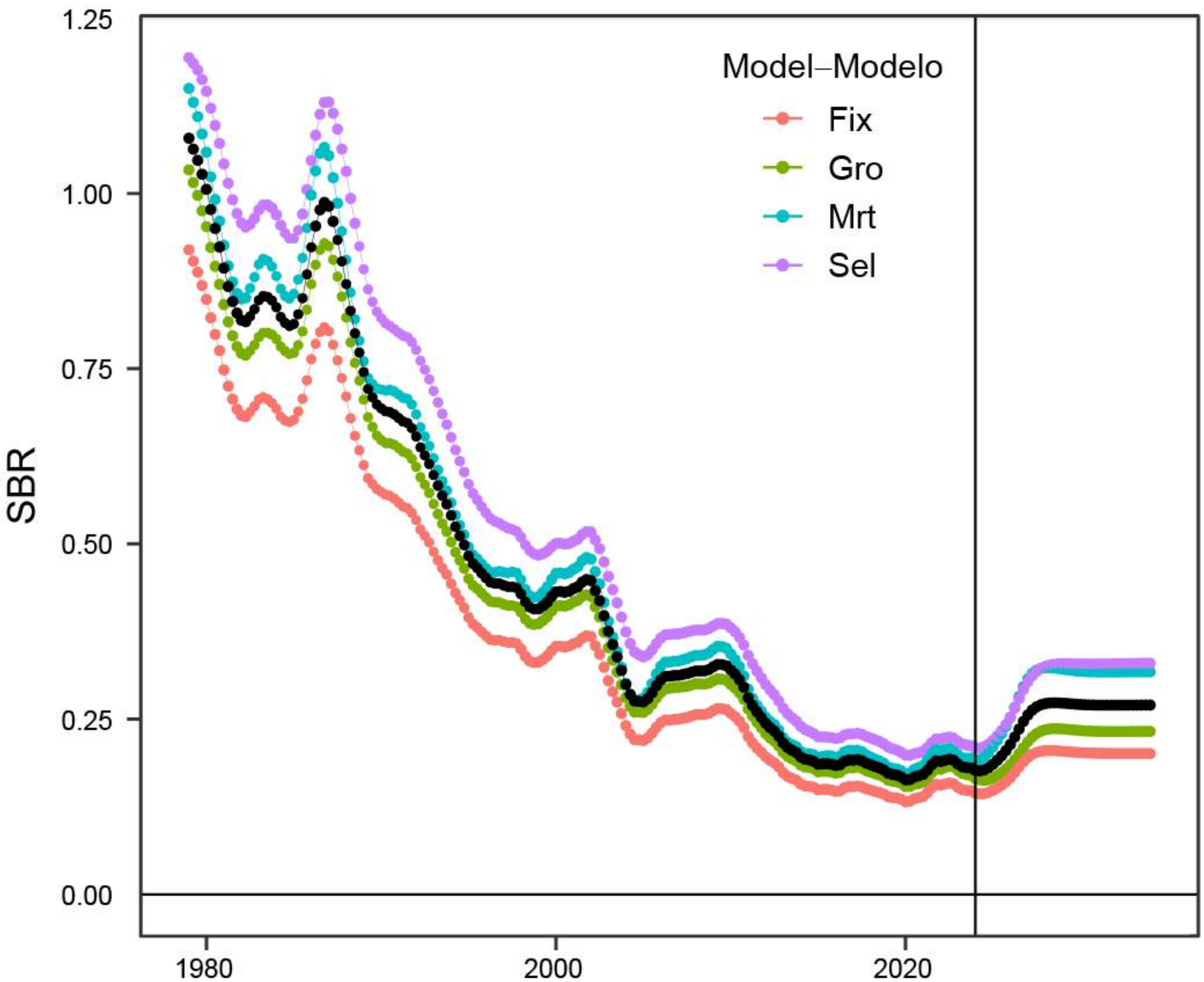

FIGURE 38. The 10-year (2024-2033) projection of spawning biomass ratio under average recruitment and current fishing mortality. The color dots represent weighted values across the second- and third-level hypotheses and the black dots represent the weighted values across all thirty-three reference models.

FIGURA 38. Proyección a 10 años (2024-2033) del cociente de biomasa reproductora bajo reclutamiento promedio y mortalidad por pesca actual. Los puntos de colores representan los valores ponderados en las hipótesis de segundo y tercer nivel y los puntos negros representan los valores ponderados en los 33 modelos de referencia.

APPENDIX

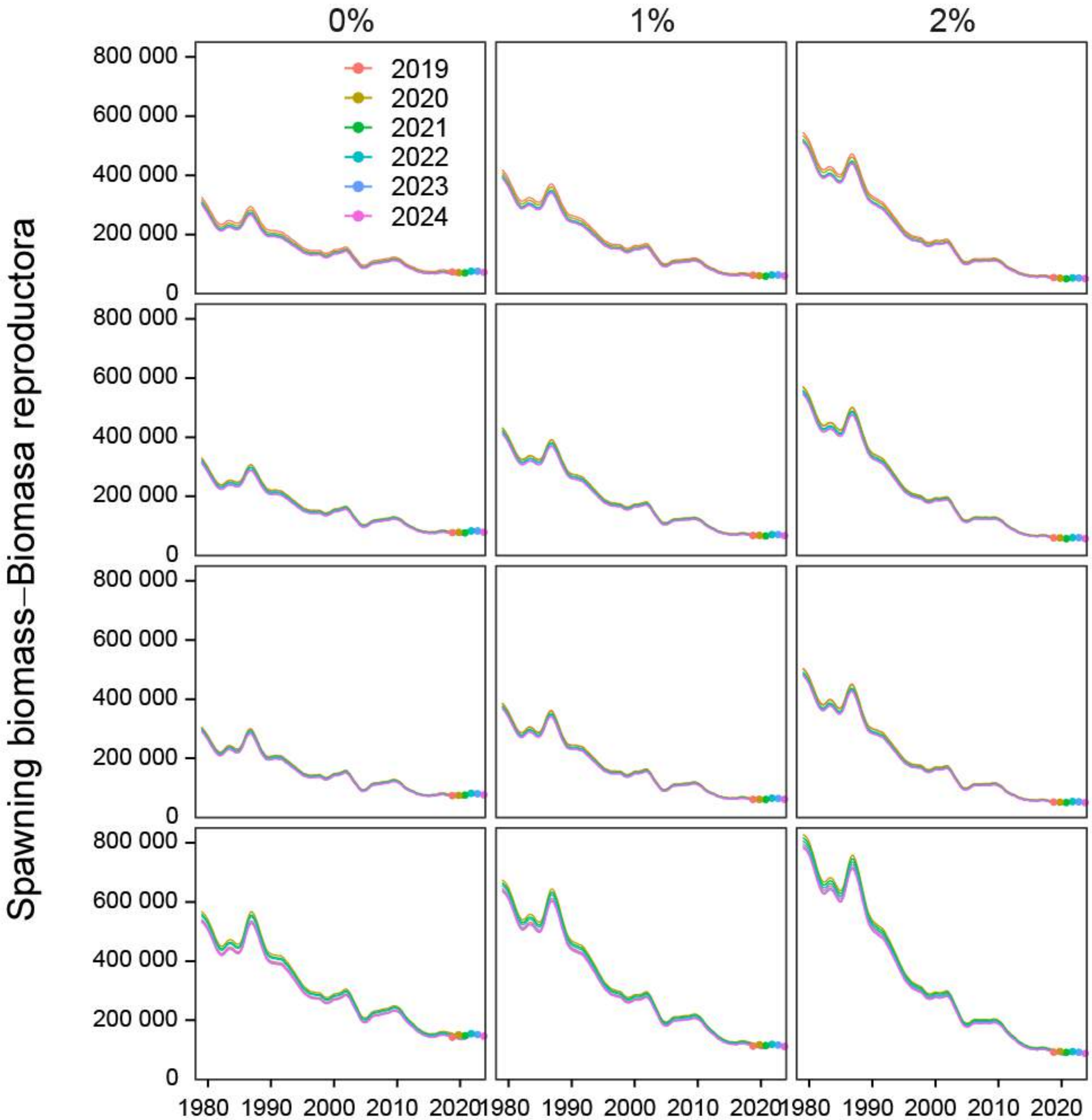

FIGURE A1. The retrospective pattern for the spawning biomass of bigeye tuna in the EPO.
FIGURA A1. Patrón retrospectivo de la biomasa reproductora del patudo en el OPO.

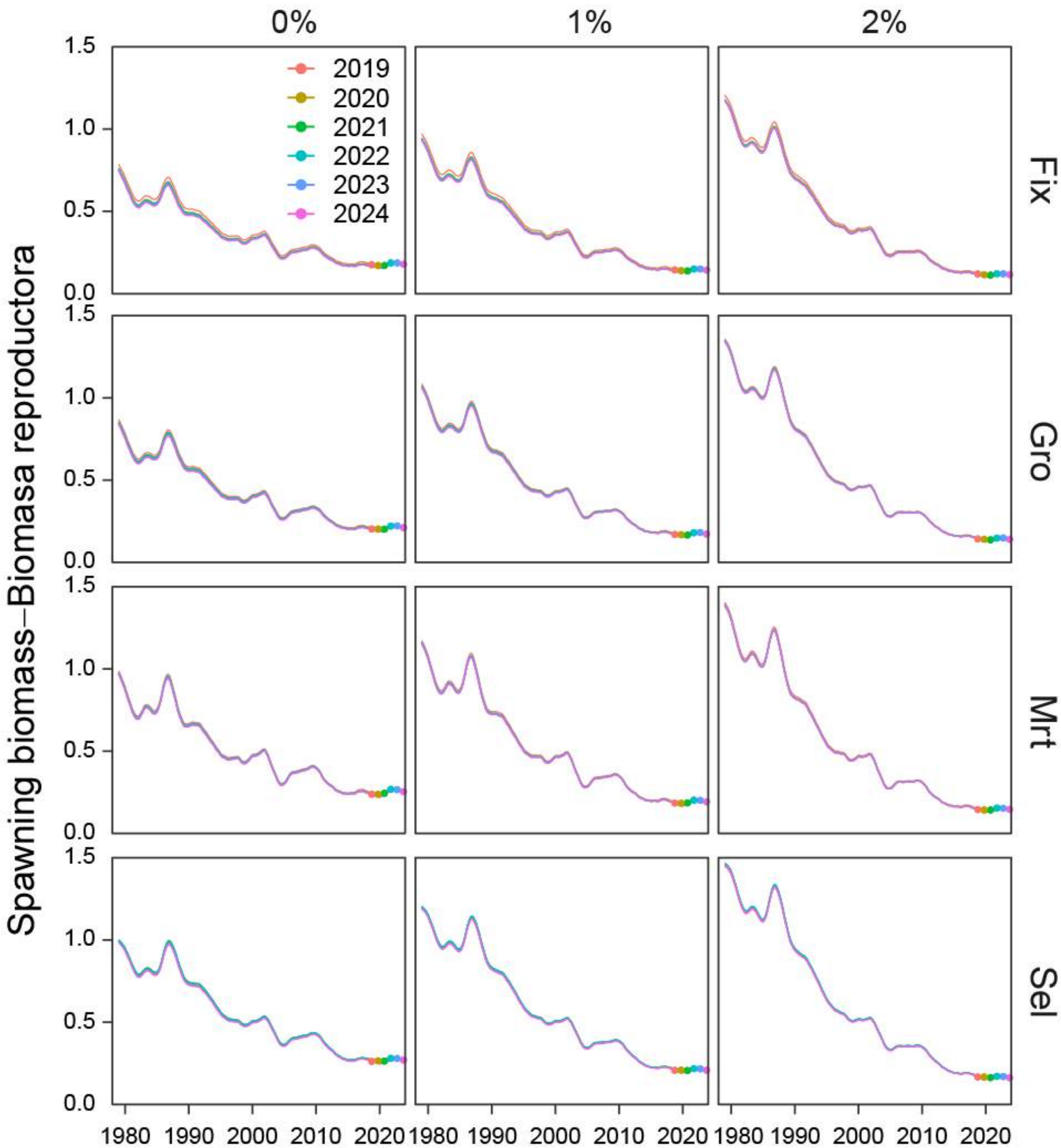

FIGURE A2. The retrospective pattern for the spawning biomass ratio of bigeye tuna in the EPO.
FIGURA A2. Patrón retrospectivo del cociente de biomasa reproductora del patudo en el OPO.

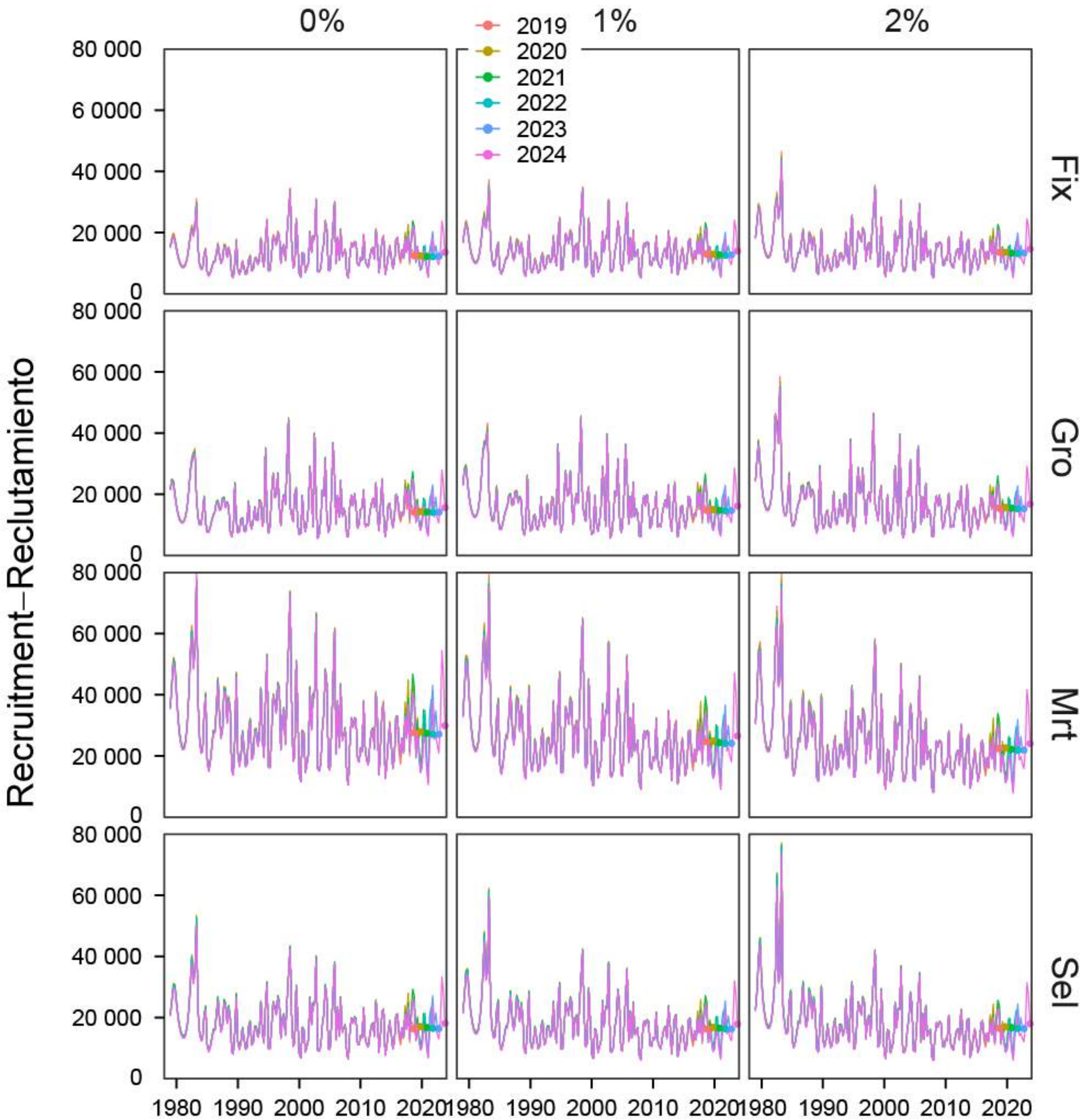

FIGURE A3. The retrospective pattern for the recruitment of bigeye tuna in the EPO.
FIGURA A3. Patrón retrospectivo del reclutamiento de patudo en el OPO.

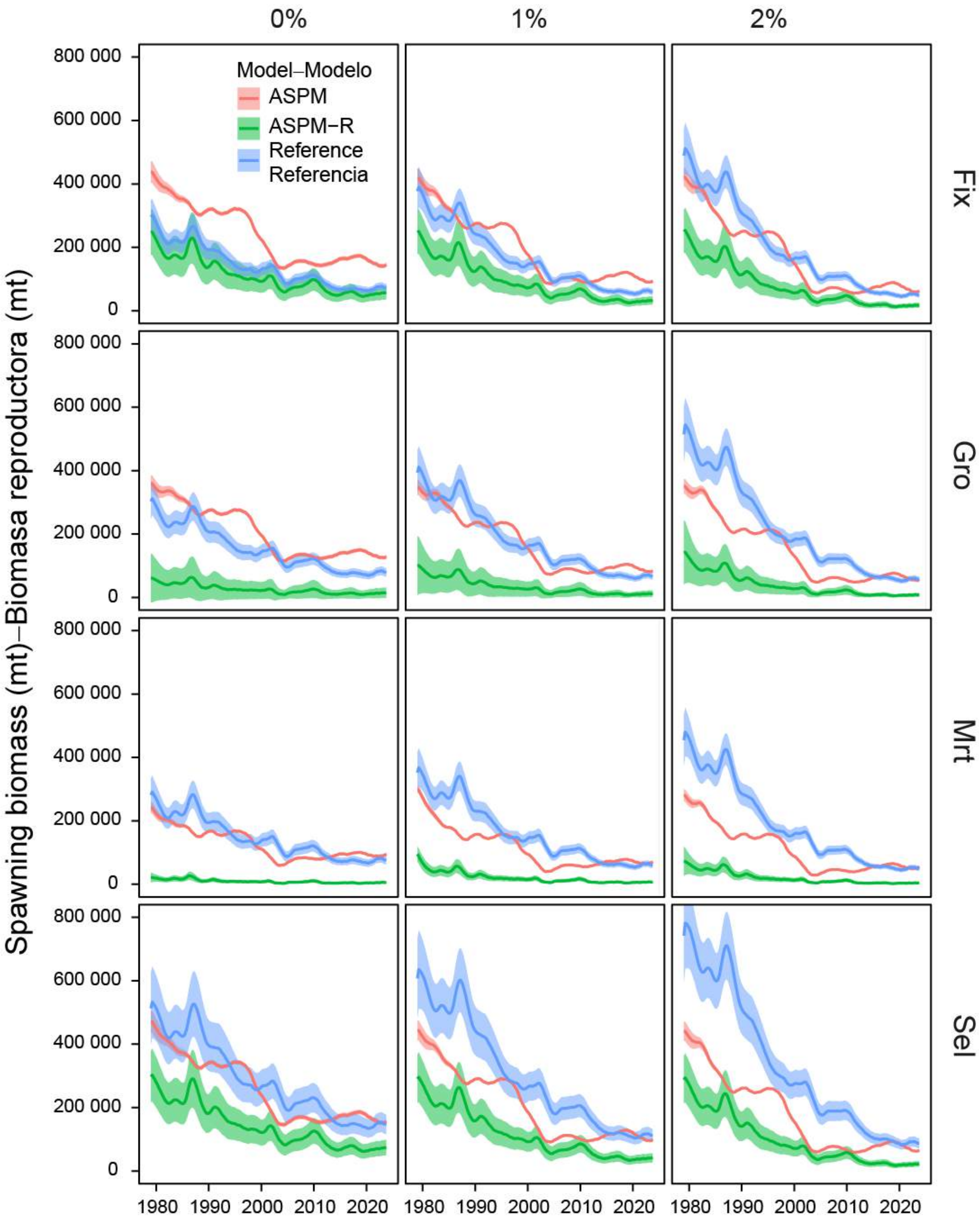

FIGURE A4. Comparison of estimated spawning biomass of bigeye tuna in the EPO from the reference model with a steepness of 1 and the corresponding age-structured production model with (ASPM-R) and without (ASPM) recruitment deviates. The shaded area represents the coefficient of variation of estimated spawning biomass.

FIGURA A4. Comparación de la biomasa reproductora estimada del atún patudo en el OPO del modelo de referencia con inclinación de 1 y el modelo de producción estructurado por edad correspondiente con (ASPM-R) y sin (ASPM) desviaciones del reclutamiento. El área sombreada representa el coeficiente de variación de la biomasa reproductora estimada.

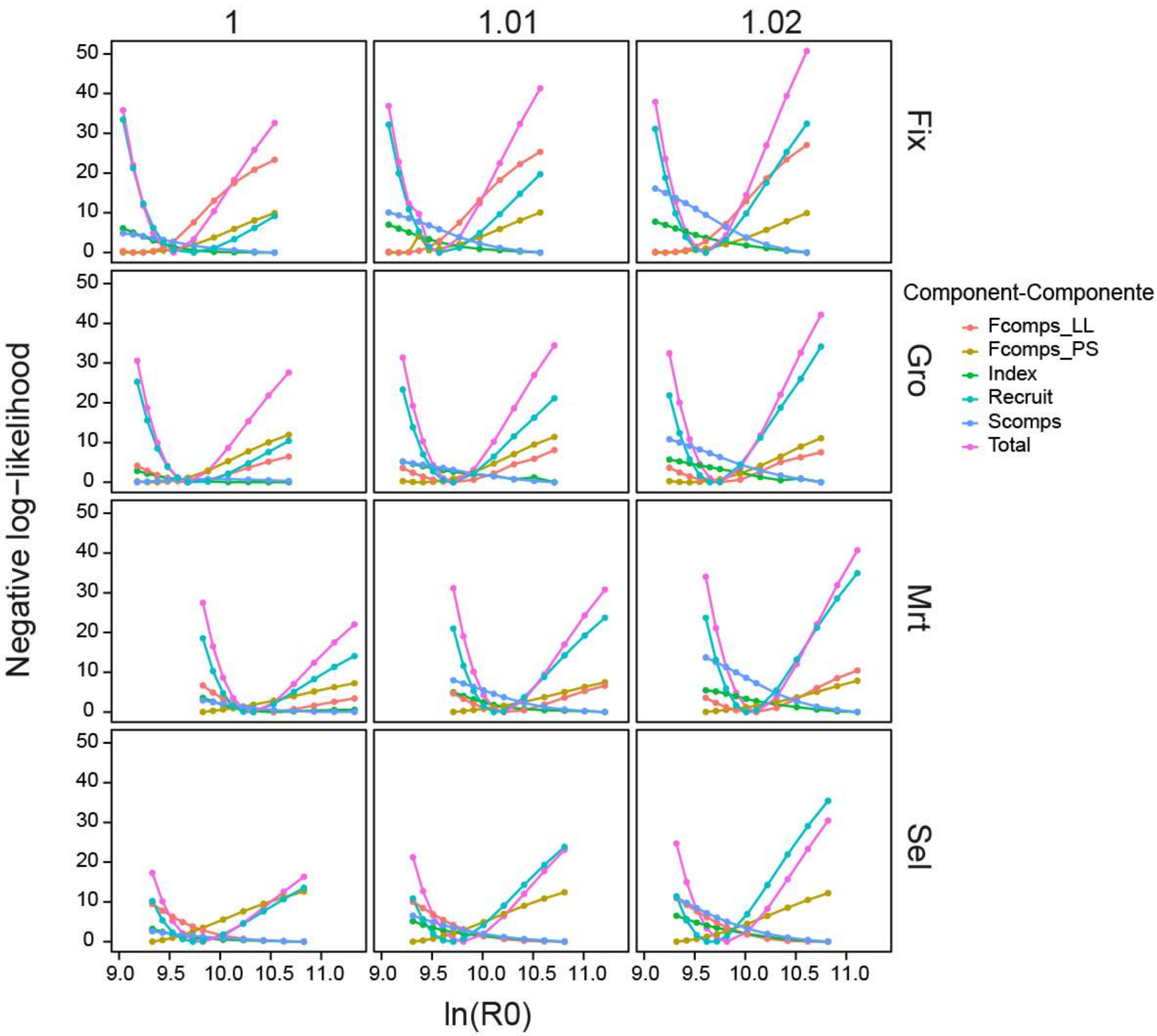

FIGURE A5. The R_0 likelihood profile for the reference models with a steepness of 1.
FIGURA A5. Perfil de verosimilitud de R_0 para los modelos de referencia con una inclinación de 1.