INTER-AMERICAN TROPICAL TUNA COMMISSION

SCIENTIFIC ADVISORY COMMITTEE

TENTH MEETING

San Diego, California (USA)
13-17 May 2019

DOCUMENT SAC-10 INF-F

EVALUATING INCONSISTENCIES IN THE YELLOWFIN ABUNDANCE INDICES

Carolina Minte-Vera, Mark N. Maunder, Alexandre Aires-da-Silva, and Haikun Xu

CONTENTS

SUMMARY

The recent fishing mortality (*F*) for yellowfin tuna is estimated to be above the maximum sustainable yield (MSY) level (F_{MSY}; *F* multiplier[1] = 0.89), which is a substantial change from the previous assessment. Investigations showed that it was due to the addition of new data for the longline-CPUE[2]-based index of abundance. Similar concerns were raised about the 2018 bigeye tuna assessment, which had an unprecedented change in the *F* multiplier, caused mainly by new longline index data (SAC-09 INF-B). Research under Project H.1.a to improve the bigeye assessment has identified several issues with the longline index that need to be addressed, and these issues are also valid for yellowfin.

There is an inconsistency between the index based on longline CPUE and the indices based on dolphin-associated purse-seine CPUE, and the stock assessment model cannot adequately fit both types of index (IATTC Stock Assessment Report 17). In addition, there have been significant changes in the length composition of the catches by the longline fishery (presentation, SAC-08) indicating possible changes in the operation of that fleet. We consider four hypotheses (change in targeting, mis-specified growth, not adequately accounting for spatial structure in the indices of abundance, and spatial stock structure) that might explain the discrepancies among the indices, and tested the first three, but none of them adequately explained the inconsistencies. More research is required to construct a stock assessment model for yellowfin in the eastern Pacific Ocean (EPO) that provides reliable management advice. A work plan has been developed to achieve this before the benchmark assessment in 2020.

---

[1] *F* multiplier: the value by which current fishing mortality (*F*) has to be multiplied to equal F_{MSY} (the fishing mortality corresponding to the maximum sustainable yield)

[2] Catch per unit of effort

1. INTRODUCTION

Benchmark assessments for bigeye and yellowfin tuna are scheduled for 2020, to provide management advice for 2021.kThe IATTC staff has been focusing on implementing the workplan to improve the bigeye tuna assessment (Project H.1.a), because in 2018 it was concluded that the assessment was not reliable enough to provide management advice (SAC-09 INF-B).

In planning the improvement work for the yellowfin assessment, the staff is considering several matters that have been identified in previous assessments (*e.g.* IATTC Stock Assessment Reports 17, 19) and the current update assessment (SAC-10-07). These include:

a. Changes in spatial distribution of effort for the Japanese fleet, whose data are used to represent the southern longline fishery, and potential changes in targeting, which may invalidate the use of the CPUE of this fishery as the main abundance index in the assessment model. This may require a temporal change in selectivity and catchability.
b. Implementation of a large-scale tagging program to address hypotheses about stock structure and regional differences in life-history parameters and depletion.
c. Improved estimates of growth, particularly for older fish.
d. Weighting of the different data sets that are fitted in the assessment model.
e. Refinement of fisheries definitions within the assessment model.
f. Implementation of time-variant selectivity, mainly for the purse-seine fisheries on floating objects.
g. Exploration of alternative assumptions about stock structure within the assessment model.

In addition to these planned improvements, the results of the bigeye research are shedding light on similar issues that have now also been identified in the current yellowfin assessment. This report discusses the following issues, and identifies the research needed to improve the 2020 benchmark assessment:

1. The management quantities are sensitive to the inclusion of the 2018 data for the longline index of abundance.
2. Inconsistencies between the main index of abundance, which is based on CPUE data for the Japanese longline fishery south of 15°N, and the indices based on the dolphin-associated purse-seine fisheries;
3. Significant changes in the length composition for the southern longline fishery, which are based on samples from the Japanese fleet, indicating possible operational changes.

2. SENSITIVITY OF THE STOCK ASSESSMENT TO THE LONGLINE INDEX OF ABUNDANCE

The yellowfin tuna assessment is fitted to five indices of abundance, but assumes that the southern longline CPUE index ("LL-S index") is the most reliable, and thus gives it the most weight. It is therefore not surprising that the model is sensitive to its inclusion. The model was run with and without the updated longline CPUE data, and the estimated management quantities were found to be sensitive to its inclusion (Table 1). In general, the longline data for the fourth quarter of the year are incomplete, and are not used in the model, so the comparison was made between model runs with the LL-S index data up to the third quarter of 2017 and with the LL-S index data up to the third quarter of 2018. With the updated data, the *F* multiplier is 0.89; without them, it is 1.00. Similarly, the SBR[3]/SBR_{MSY} with the updated data is 0.76, while without them it is 0.99. A similar sensitivity to the longline CPUE was found in the bigeye tuna assessment (SAC-09 INF-B). These differences imply substantial changes to the management advice that the staff formulates for the Commission. Therefore, ensuring that the LL-S index is a good measure of relative abundance is critical; if it is not, it might be appropriate to assign more weight to the dolphin-associated indices.

[3] SBR: spawning biomass ratio, the ratio of the current spawning biomass to that of the unfished population

3. INCONSISTENCY AMONG INDICES OF RELATIVE ABUNDANCE

The inconsistency between the main index of abundance, based on the CPUE of the Japanese longline fishery south of 15°N (LL-S) (Figure 1), and the indices based on the CPUEs of the Northern (DEL-N) and Inshore (DEL-I) dolphin-associated purse-seine fisheries ("DEL indices"), is of particular concern for the yellowfin assessment. The model estimates peaks in the DEL indices before they occur, and under-estimates the indices after 2010, but does the opposite for the LL-S index (IATTC Stock Assessment Report 17: Figure 2). The length composition for the LL-S fishery has also changed to larger fish recently, indicating possible changes in the operation of the longline fishery (presentation, SAC-08: Figure 3).

A correctly-specified model should be able to adequately fit all the indices of abundance within their sampling error. Because no model is perfectly correct, it is expected that some unmodeled process variation or model misspecification will cause the fit to the indices to be worse than indicated by the sampling error. This additional misfit should not be substantial, and should be minimized as much as possible. Data are assumed to be correct, but the staff's understanding of some data may be wrong or limited, thus they may be incorrectly modeled within the assessment. Such misunderstood data may not be useful, and may even be detrimental, because once translated through the observation model, they provide misleading information. It is also possible that some well-understood data are not informative, because the observation model is so complex that all the information contained in the data is used to estimate the parameters of the observation model itself. Therefore, the inconsistency between the longline and dolphin indices suggests that the population dynamics or observation models used in the yellowfin assessment are mis-specified.

3.1. Indices of abundance used in the yellowfin stock assessment

The yellowfin assessment includes several fisheries that can potentially be used to create CPUE-based indices of relative abundance: the purse-seine fisheries on dolphins, floating objects, and unassociated schools, and the longline fisheries. Indices based on purse-seine sets on floating objects (OBJ) are not considered reliable because they are likely to be hyperstable: most sets are made on fish-aggregating devices (FADs), and vessels use electronic devices to track FADs and approximate the abundance of the associated tuna, and only set on FADs that have tuna. Indices based on the southern dolphin-associated fishery (DEL-S) and the longline fishery north of 15°N (LL-N) are likewise not used because catches are low and/or the CPUE is variable. The CPUE of the longline fishery south of 15°N (LL-S) is assumed to be the most reliable, and is fitted in the model using a standard deviation of the lognormal-based likelihood function (approximately equal to the coefficient of variation (CV)) of 0.2. The standard deviation for the likelihood functions for the northern and inshore dolphin associated fisheries (DEL-N, DEL-I) and the northern and southern unassociated fisheries (NOA-N, NOA-S) are estimated inside the model, and are all 0.4 or greater. Therefore, the index based on the southern longline fishery is given substantially more weight than other indices of abundance included in the stock assessment.

3.2. Hypotheses

Given the importance of the LL-S index for estimating management quantities, it is important to investigate the inconsistencies with the DEL indices to make sure that they are not due to a misunderstanding of the longline CPUE data. Several hypotheses might explain the inconsistencies, including: 1) change in fishing behavior (*e.g*. targeting) by the longline fishery; 2) mis-specified growth; 3) inadequate consideration of spatial structure in the indices of abundance; and 4) spatial structure in the population. These four hypotheses are discussed further below, and the management consequences are presented in Table 1.

3.2.1. Hypothesis 1: Change in fishing behavior by the longline fishery

The increase in the average size of the fish in the catch in the southern longline fishery since about 2010 (Figure 3) indicates that the fishery may have changed its behavior, specifically the species or sizes targeted, and could also be related to a change in the spatial distribution of the fishery. The Japanese longline fishery now covers a smaller area than before, and CPUE and size of yellowfin have been observed to vary spatially (WSLL-01).

The consistency and longevity of this increase in average size indicates that it is not due to a few strong cohorts growing larger but to some other persistent characteristic. Possible explanations for increase include: a) changes in operational characteristics that cause larger fish to be selected; b) changes in fishing location; c) increased growth rates or maximum length; d) reduced natural mortality; e) reduced fishing mortality; or f) discarding of smaller yellowfin.

The hypothesis of a change in fishery behavior was investigated by estimating changes in the catchability for the index of abundance and the selectivity of the fishery; the latter is also used to predict the index of abundance in the fitting procedure. These changes improved the fit to both the longline (Figure 2a, b, 58 log-likelihood units) and dolphin (Figure 2a, 9 and 11 log-likelihood units for the DEL-N and DEL-I fisheries, respectively) indices of abundance as well as the longline length-composition data (Figure 3, 57 log-likelihood units). However, they did not correct the misfit in the timing of the peaks in CPUE, and also had important implications for management advice (Table 1).

3.2.2. Hypothesis 2: Mis-specified growth

The temporal difference in the peaks between the indices could be caused by differences in the size of the fish caught in each fishery. Mis-specified growth could assign the wrong age to yellowfin caught in each fishery, biasing the timing of strong cohorts in a way that would cause the peaks to occur in the wrong time periods. The growth curve currently used in the yellowfin assessment is fixed, based on parameters previously estimated within the stock assessment model (Maunder and Aires-da-Silva 2009) while fitting to conditional age-at-length data from counts of daily growth rings on otoliths (Wild 1986). However, otolith data are available only for fish up to about 5 years old and 160 cm long. Tagging data could be used to estimate the age-at-length for older individuals, but their spatial distribution is limited, and it is not clear how representative they are of the growth rates of yellowfin in the EPO.

To investigate the hypothesis of mis-specified growth causing the inconsistencies in the indices, the stock assessment model was run estimating all the growth parameters except those for the variation of length at age. Estimating growth improved the fit to the LL-S index (44 log-likelihood units), but degraded the fit to the DEL indices (10 and 17 log-likelihood units for the DEL-N and DEL-I indices, respectively) (Figure 4). Estimating growth did not fix the mismatch among the indices in the peaks in CPUE, or fit the increased average size in the longline fishery.

3.2.3. Hypothesis 3: Inadequate consideration of spatial structure in the indices of abundance

Indices of abundance should take account of spatial differences in fish density. Neither the LL-S nor the DEL indices deal with spatial structure adequately. The DEL indices are simply raw catch per day fished (CPDF), with no spatial component except the definitions of the fisheries (Figure 1). The LL-S index is based on a GLM that includes latitude and longitude, but these are invariant over time, and the index is weighted by data rather than area. New approaches are available that model the spatiotemporal variation and its correlations appropriately, and calculate indices of abundance weighted by area; similar approaches should be used to develop indices by length class as well. Preliminary analyses show that the impact of accounting for spatiotemporal variation in the CPUE and composition data is small for the longline index, but large for the dolphin indices. The indices estimated by Xu *et al*. (2019), using the VAST spatiotemporal model, show lower abundance since 2005 than those used in the current assessment (Figure 5), which makes the indices more consistent with each other. When the VAST indices were used in the stock assessment, they improved the model fit to both the LL-S index (30 log-likelihood units) and the DEL indices (30 and 43 log-likelihood units for DEL-N and DEL-I, respectively), particularly in recent years for the DEL indices (Figure 6). However, it did not correct the misfit of the timing of the peaks. The method could be extended to produce indices by length class, which may address changes in the length frequencies associated with the indices.

3.2.4. Hypothesis 4: Spatial structure in the population

The staff has not yet investigated this hypothesis, which posits that spatial structure in the population caused by restricted movement or ontogenetic movement may cause local depletion or differences in sizes among areas. Substantial sections of the areas defined for the dolphin and longline indices do not overlap (Figure 1); therefore, spatial patterns in the fishing effort, and differences in these patterns among gears, could cause inconsistencies in the indices of abundance. There is some evidence of restricted movement from tagging data (Schaefer *et al*. 2011, 2014), but there is no clear stock structure in the EPO, and separation by distance may be as likely an explanation as spatially-defined stocks (*i.e*. characteristics of fish differ among areas because exchange rates are low due to the distances involved, rather than because of some physical, chemical, or biological barrier). The tagging data are limited, and insufficient for constructing a spatial stock assessment model with movement. Analysis of CPUE and length-composition data for both the purse-seine and longline fleets indicates some spatial differences, which could be used to develop new areas for an areas-as-fleets approach.

4. DISCUSSION

The EPO yellowfin tuna assessment is highly dependent on the LL-S index of relative abundance. Several issues have been identified with the index, which uses data from the Japanese longline fleet only: in particular, the spatial extent of the fleet has changed over time, and recently the average size of yellowfin caught has increased. Indices based on the combined data of several fleets have been recommended (WSLL-01). In addition, there is an inconsistency between the LL-S index and the DEL indices: the peaks in abundance are out of step, and recent CPUE is higher for the DEL indices. The model cannot adequately fit both types of indices.

Three of the four hypotheses that might explain the inconsistencies among the indices were tested, but none of them did so adequately. It is possible that more than one of these hypotheses needs to be introduced simultaneously (*e.g*. time block for longline catchability and selectivity + spatiotemporal modelling of the DEL indices).

The length-composition data used in the stock assessment model are used to both a) determine the size of fish removed from the population and b) provide information for estimating other model parameters. Generally, it is assumed that the size composition data contain information for estimating the selectivity of the gear and cohort strength. However, the length-composition data also provide information on mortality rates, and consequently absolute biomass, through a catch-curve type process, since the catch is assumed to be known with little error. Therefore, it is important to correctly model the processes controlling the composition data (selectivity, growth, recruitment, *etc*.) and how they change over time.

More research is required to construct a stock assessment model for yellowfin tuna in the EPO that provides reliable management advice before the benchmark assessment in 2020. As part of the work plan to prepare for that assessment, a workshop is proposed to finalize the work on the longline indices of abundance and composition data, with the goal of developing indices by length class. This will require access to the operational-level data and corresponding size-composition data.

5. WORK PLAN

1. Complete the work on yellowfin stock and fishery structure.
2. Hold a workshop to finalize the work on the longline indices of abundance and composition data.
3. Develop purse-seine based indices of abundance using spatiotemporal methods by length class.
4. Develop an improved model for yellowfin growth.
5. Reevaluate yellowfin natural mortality.
6. Adequately weight the different data sets that are fitted in the assessment model using recommended methods and best practices.
7. Implement time-variant selectivity, mainly for the purse-seine fisheries on floating objects.

REFERENCES

Maunder, M.N., and A. Aires-da-Silva. 2009. Status of yellowfin tuna in the eastern Pacific Ocean in 2007 and outlook for the future. Inter-Amer. Trop. Tuna Comm., Stock Asses. Rep.9: 3-94.

Schaefer, K.M., D.W. Fuller, and B.A. Block. 2011. Movements, behavior, and habitat utilization of yellowfin tuna (*Thunnus albacares*) in the Pacific Ocean off Baja California, Mexico, determined from archival tag data analyses, including Kalman filtering. Fish. Res. 112, 22-37.

Schaefer, K.M., D.W. Fuller, and G. Aldana. 2014. Movements, behavior, and habitat utilization of yellowfin tuna (*Thunnus albacares*) in waters surrounding the Revillagigedo Islands Archipelago Biosphere Reserve, Mexico. Fish. Ocean. 23: 65-82.

Wild, A. 1986. Growth of yellowfin tuna, *Thunnus albacares*, in the eastern Pacific Ocean based on otolith increments. IATTC Bull. 18(6): 421-482.

Xu, H., Lennert-Cody, C. E., Maunder, M. N., 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.

TABLE 1. MSY and related quantities for the base case and sensitivity analyses (Hypotheses 1-3), based on average fishing mortality (*F*) for 2016-2018. B_{recent} and B_{MSY} are defined as the biomass, in metric tons, of fish 3+ quarters old at the start of the first quarter of 2019 and at MSY, respectively, and S_{recent} and S_{MSY} are defined as indices of spawning biomass (therefore, they are not in metric tons). C_{recent} is the estimated total catch for 2018.

TABLA 1. RMS y cantidades relacionadas para el caso base y los análisis de sensibilidad (Hipótesis 1-3), basados en la mortalidad por pesca (*F*) media de 2016-2018. Se definen $B_{reciente}$ y B_{RMS} como la biomasa, en toneladas, de peces de 3+ trimestres de edad al principio del primer trimestre de 2019 y en RMS, respectivamente, y $S_{reciente}$ y S_{RMS} como índices de biomasa reproductora (por lo tanto, no se expresan en toneladas). $C_{reciente}$ es la captura total estimada de 2018.

YFT	Base case	Sensitivity analyses - Análisis de sensibilidad			
		No longline CPUE update	H1: Change in selectivity and catchability	H2: Estimate growth	H3: Spatio-temporal DEL indices
	Caso base	Sin CPUE palangrera actualizada	H1: Cambio en selectividad y capturabilidad	H2: Crecimiento estimado	H3: Indices DEL espaciotem-porales
MSY-RMS (t)	254,974	254,872	248,890	306,849	254,960
B_{MSY}- B_{RMS} (t)	371,787	372,247	371,206	477,413	371,460
S_{MSY}- S_{RMS} (t)	3,638	3,642	3,660	7,752	3,638
B_{MSY}/B_0- B_{RMS}/B_0	0.31	0.31	0.31	0.31	0.31
S_{MSY}/S_0- S_{RMS}/S_0	0.27	0.27	0.27	0.29	0.27
C_{recent}/MSY- $C_{reciente}$/RMS	1	1	1.04	0.83	1
B_{recent}/B_{MSY}- $B_{reciente}/B_{RMS}$	0.84	1.03	1.12	1.33	0.73
S_{recent}/S_{MSY}-$S_{reciente}/S_{RMS}$	0.76	0.99	1.08	1.36	0.64
F multiplier-Multiplicador de *F*	0.89	1	1.14	1.63	0.83

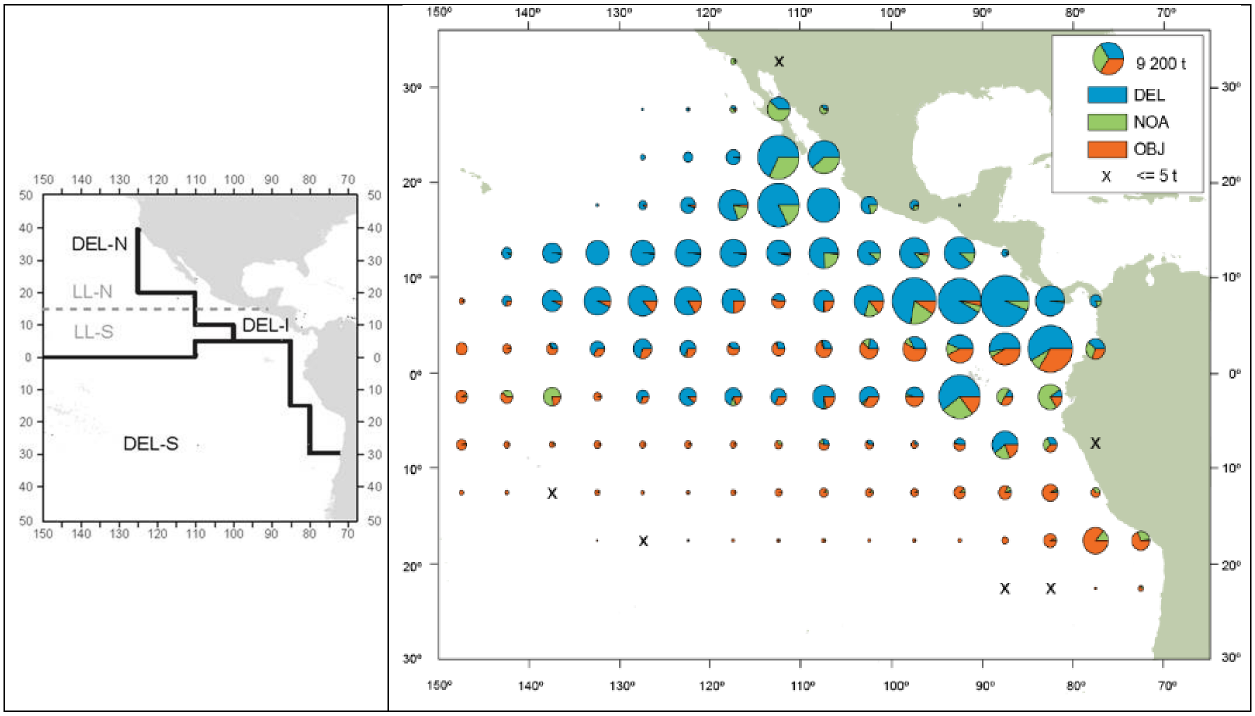

FIGURE 1. Longline (LL) and dolphin-associated (DEL) fisheries used as indices of abundance in the stock assessment (left), and average annual distributions of the purse-seine catches of yellowfin, by set type, 2012-2016 (right). The size of the circles is proportional to the amount of yellowfin caught in that 5° by 5° area (IATTC Fisheries Status Report, Figure A-1a).

FIGURA 1. Pesquerías palangreras (LL) y asociadas a delfines (DEL) usadas como índices de abundancia en la evaluación de la población (izquierda), y distribución media anual de las capturas cerqueras de aleta amarilla, por tipo de lance, 2012-2016 (derecha). El tamaño de cada círculo es proporcional a la cantidad de aleta amarilla capturado en la cuadrícula de 5° x 5° correspondiente (Informe de la Situación de la Pesquería de la CIAT, Figura A-1a).

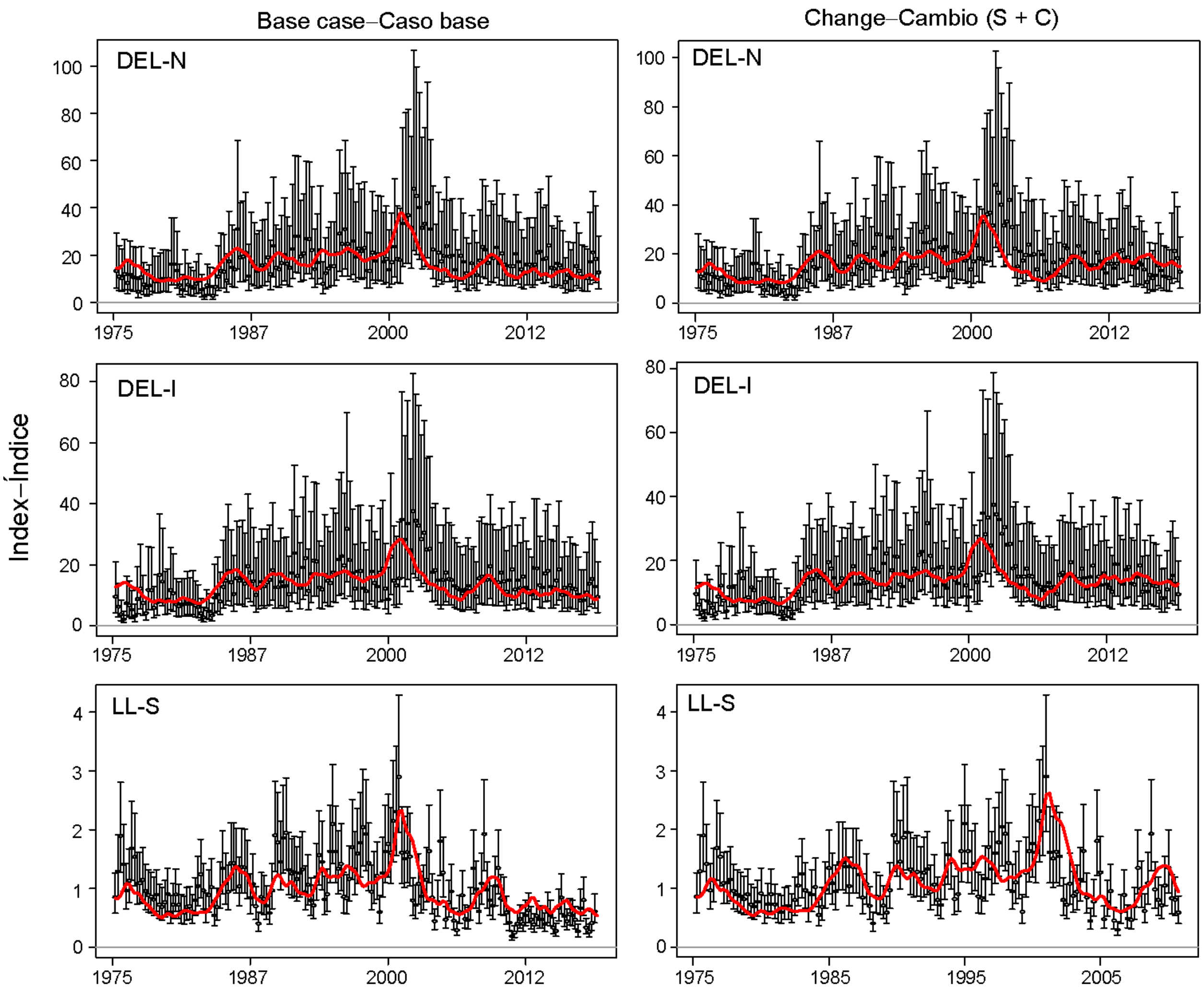

FIGURE 2a. Fits to the CPUE data from the base case (left) and from the model with a change in selectivity and catchability after 2010 (right), for the northern and inshore dolphin indices (DEL-N, DEL-I) and the southern longline index (LL-S). Note that the bottom right figure stops in 2010, when the change in selectivity and catchability occurs; later years are shown in Figure 2b.

FIGURA 2a. Ajustes a los datos de CPUE del caso base (izquierda) y del modelo con un cambio en la selectividad y capturabilidad después de 2010 (derecha), para los índices de las pesquerías asociadas a delfines del norte y costera (DEL-N, DEL-I) y el índice de la pesquería palangrera del sur (LL-S). Nótese que la figura inferior derecha se detiene en 2010, cuando ocurre el cambio en la selectividad y capturabilidad; en la Figura 2b se muestran los años posteriores .

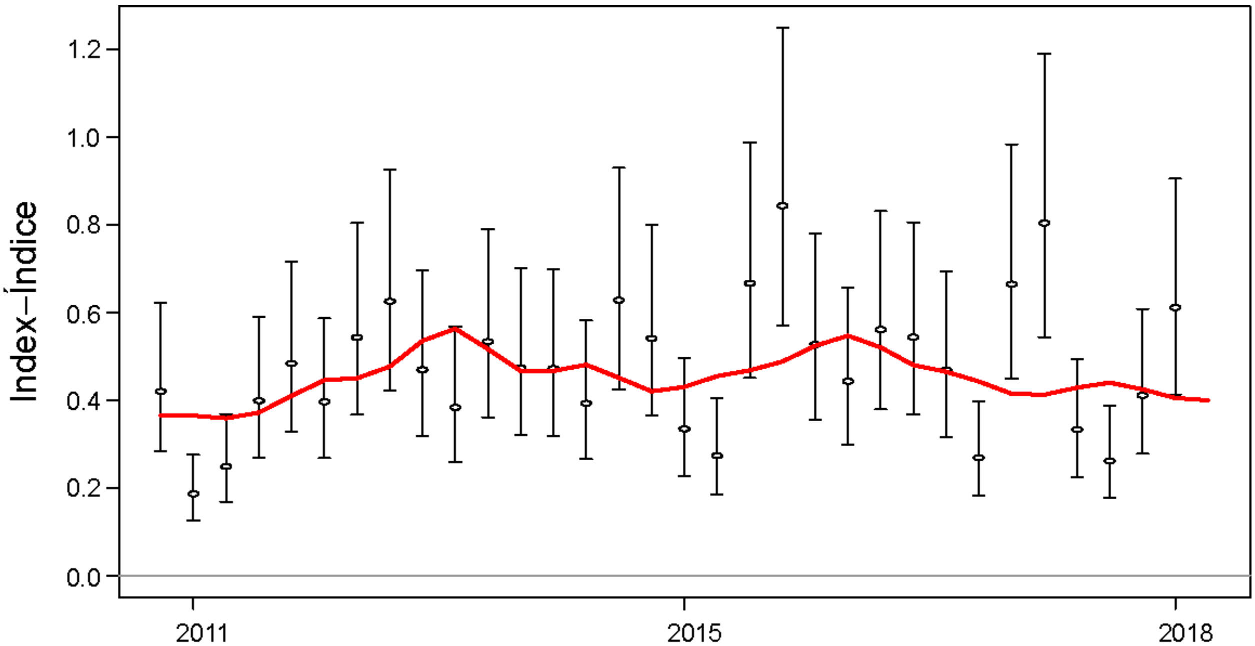

FIGURE 2b. Fits to the southern longline (LL-S) index of relative abundance after the change in selectivity and catchability in 2010.

FIGURA 2b. Ajustes al índice de abundancia relativa palangrera del sur (LL-S) después del cambio en la selectividad y capturabilidad en 2010.

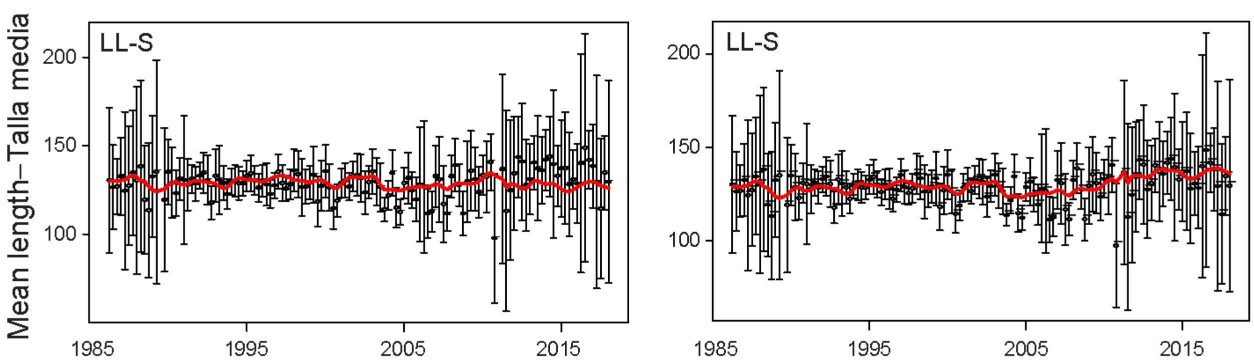

FIGURE 3. Observed and predicted average lengths of the fish in the catch for the southern longline fishery (LL-S), from the base case (left) and from the model with a change in selectivity and catchability after 2010 (right).

FIGURA 3. Tallas promedio observadas y predichas de los peces en la captura de la pesquería palangrera del sur (LL-S), del caso base (izquierda) y del modelo con un cambio en la selectividad y capturabilidad después de 2010 (derecha).

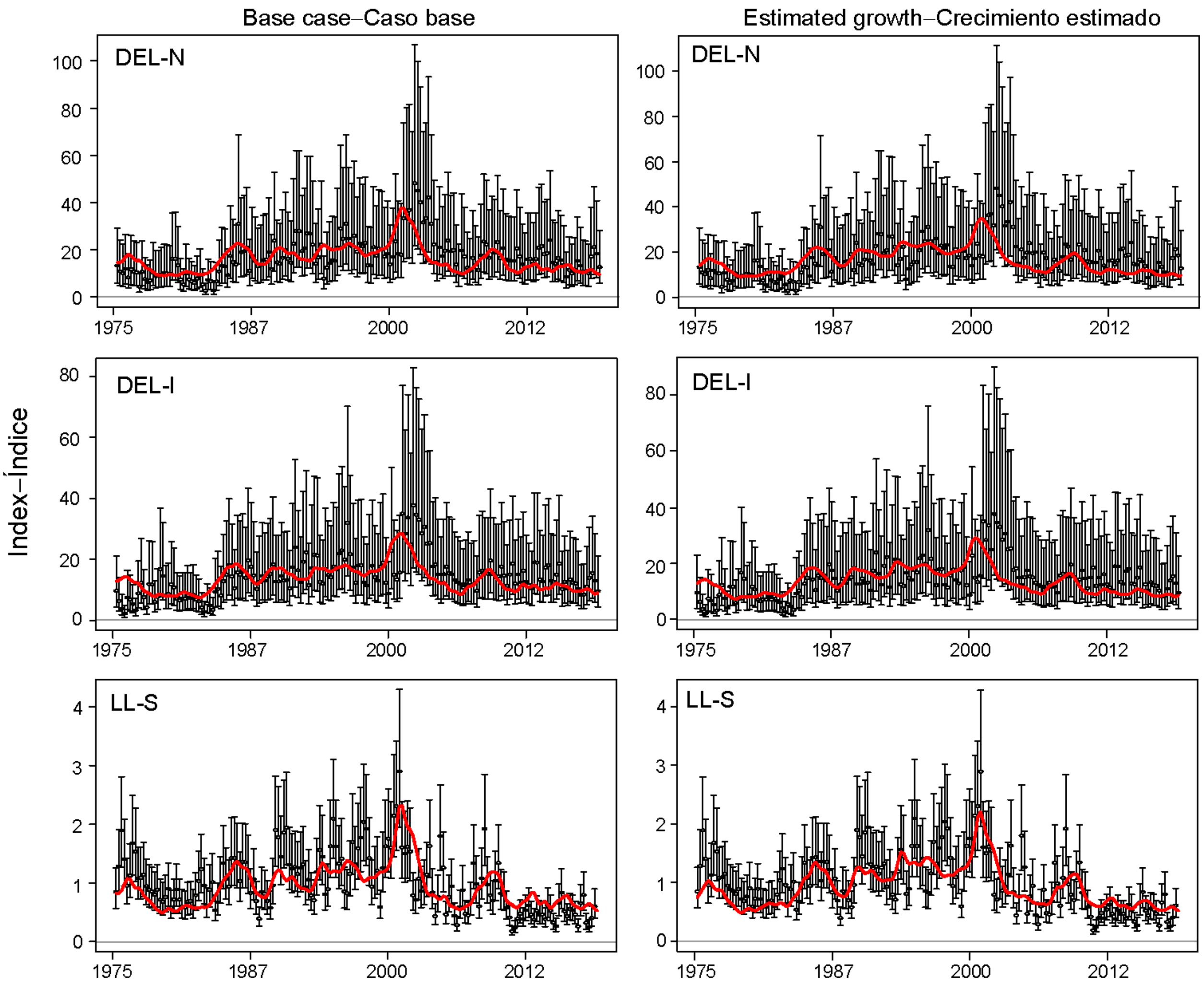

FIGURE 4. Fits to the CPUE data from the base case (left) and from the model that estimates growth (right), for the northern and inshore dolphin indices (DEL-N, DEL-I) and the southern longline index (LL-S).

FIGURA 4. Ajustes a los datos de CPUE del caso base (izquierda) y del modelo que estima el crecimiento (derecha), para los índices de las pesquerías asociadas a delfines del norte y costera (DEL-N, DEL-I) y el índice de la pesquería palangrera del sur (LL-S).

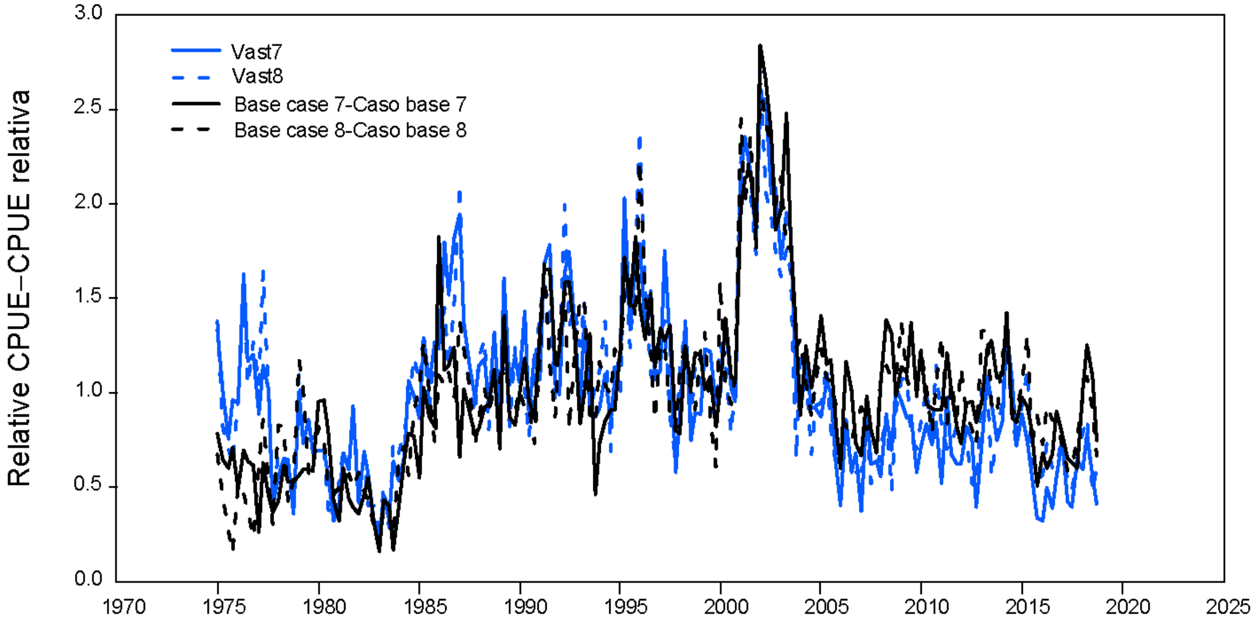

FIGURE 5. Nominal CPUE used in the current stock assessment (black lines), and generated using a spatiotemporal model (blue lines, Xu *et al*. 2019). All indices are scaled by their average.

FIGURA 5. CPUE nominal usada en la evaluación de poblaciones actual (líneas negras) y generada con un modelo espaciotemporal (líneas azules, Xu *et al.* 2019). La escala de todos los índices se ajusta a su promedio.

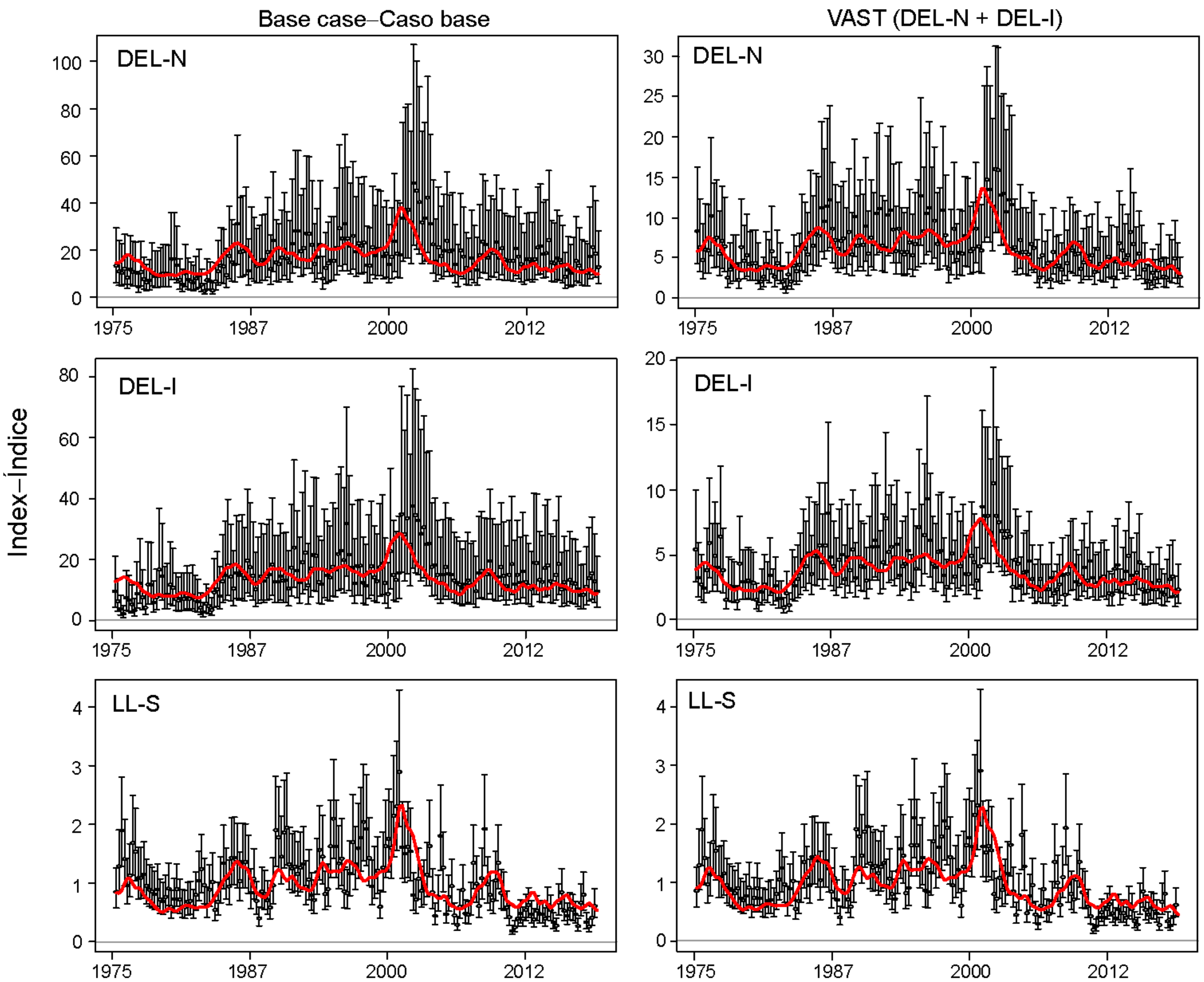

FIGURE 6. Fits to the CPUE data from the base case (left) and from the model that uses the dolphin associated indices of abundance based on the spatiotemporal model (VAST; from Xu *et al*. 2019, right), for the northern and inshore dolphin indices (DEL-N, DEL-I) and the southern longline index (LL-S).

FIGURA 6. Ajustes a los datos de CPUE del caso base (izquierda) y del modelo que usa los índices de abundancia de la pesquería asociada a delfines basados en el modelo espaciotemporal (VAST; de XU *et al.* 2019, derecha), para los índices de las pesquerías asociadas a delfines del norte y costera (DEL-N, DEL-I) y el índice de la pesquería palangrera del sur (LL-S).