ISC/24/PBFWG-1/08

Assumptions and its alternatives for the assessment model in the 2024 Stock Assessment of Pacific Bluefin Tuna

Hiromu Fukuda, Yohei Tsukahara, Kirara Nishikawa

Highly Migratory Resources Division, Fisheries Resources Institute,
Japan Fisheries Research and Education Agency
2-12-4, Fukuura, Kanazawa-ku, Yokohama, Kanagawa 236-8648, JAPAN

Feb. 2024

Working document submitted to the ISC Pacific bluefin tuna Working Group, International Scientific Committee for Tuna and Tuna-Like Species in the North Pacific Ocean (ISC), from 29 February to 7 March 2024, Kaohsiung, Taiwan.

1 Introduction

The last full stock assessment for Pacific bluefin tuna (hereafter PBF) was carried out in February 2020 (ISC PBFWG, 2020), and the fishery data from 2017 to 2021 were updated in the 2022 stock assessment (ISC PBFWG, 2022). In both assessments, the WG recognized and listed several issues on those assessment reports as unsolved future issues. From the 2022 assessment onwards, the PBFWG reviewed their past assessments and summarized possible areas to improve the assessment model for the benchmark assessment scheduled in 2024. In keeping with above mentioned reviews and studies in the intersessional meetings, this document proposes; 1.) use of the short-term model; 2.) a minor modification to the 2022 short-term model; 3.) fitting to the newly available size composition data; 4.) reducing the residuals for the size composition data; 5.) improving the retrospective diagnostics; 6) fitting to the newly available index of abundance.

In this document, we started from the 2022 PBF stock assessment short-term model using the Stock Synthesis (SS) version 3.30.14.08 (Methot & Wetzel, 2013; Fukuda., 2021) to explore the possible model improvements with the order of the above-mentioned proposals.

2 Model and Data

2.1 Basic model configuration

The model assumes a single well-mixed stock for PBF and does not consider a spatially explicated structure. All the catch and size composition data are temporally stratified into the following 4 quarters of July-September, October-December, January-March, and April-June. Those quarters (Jul-Sept, Oct-Dec, Jan-Mar, and Apr-Jun) are assigned to 1st, 2nd, 3rd, and 4th seasons, respectively as the fishing year of PBF. The time period modeled in this assessment is 1952-2022 including the updated recent two years.

2.2 Data

The data for this model are based on the 2022 stock assessment and updated data submitted by the WG members. It should be noted that some data submitted after January 2024 were not reflected in this document because of the time constraint for the model explorations. The data submitted after January will be included in the model at the assessment meeting. The detailed information of those input data (the catch, discard, abundance index, and size composition) is described in Nishikawa et al. (2024). It also should be noted that the numbering of the fishery and abundance index (e.g. Fleet and Survey) were re-ordered to fix the inconsistent order of fishery definition, which has been changed in ad-hoc manner during the past assessments, into a simple sequence by fishery type and member.

3 Model Exploration

3.1 Short-term PBF assessment model

The 2022 PBF base case modeled data during 1952-2020 fishing year. Although it did not show any evidence of further improvements on the model convergence, it showed inflexibility to the changes of the productivity assumptions (e.g. lower steepness). The current base case allows for model convergence at slightly lower level of steepness only (h=0.99). Since the reason for this convergence issue might be that the population is observed at a very low relative stock size, and the model is fine-tuned to explain data under the current assumption, Fukuda (2021) re-constructed the PBF SS model without very high consecutive catches observed in 1981-1982 by starting the model from 1983 (short-term model). The model performance was evaluated by the residual analysis and model convergence for several alternative assumptions about the steepness and natural mortality. The biomass time series such as the SSB and recruitment estimated by the short-term model were basically identical with those from the long-term model. In that study, the short-term model brought some advantages such as higher flexibility to the alternative assumptions about the steepness, shorter run time, and keeping its high model performance in terms of the model fits to the data (Fukuda 2021).

The short-term model was utilized in the 2022 assessment only for the sensitivity analysis to the alternative assumptions about the productivity but for the assessment base case because 2022 assessment was the data update assessment. Since 2024 assessment is the bench-mark assessment, this could be a time to discuss changing the assessment time period. It should be noted that some of the empirical quantities used for the PBF management (e.g. historical median) could be brought from the 2022 base case in a form of relative value over the unfished biomass (%SSB0). The fishery impact (Wang et al., 2009), which was another metrics often mentioned during the discussion of the IATTC-WCPFC NC joint working group on PBF management, in historic period also could be brought from a previous estimate as a relative fishery impact ratio in each fishery group or could be estimate using an updated long-term model.

The authors recommend to use the short-term model in the 2024 assessment while maintaining the long-term model as one of the sensitivity analyses.

3.2 A minor modification to the 2022 short-term model

The observed length compositions for the Japanese longline fishery in fishing season 1 (Fleet 1) indicated a steep decrease in selection and a high uncertainty in parameter estimation in terms of the standard deviation (S.D.). Given the parametric length-based selectivity currently used, parameters associated with the descending limb of dome-shaped selectivity have little information on their values because selectivity is changing rapidly within a couple of size bins. The working group in November 2023 explored alternative data structure (e.g., smaller size bin), alternative selectivity shape (e.g., estimating the end selectivity of the 6-parameter double normal), and alternative time-block (e.g. eliminating the time-block from Fleet 1) to resolve this issue. Although a smaller length bin did not contribute to improve the parameter uncertainty, estimation of the end parameter of the double normal and time-block elimination resulted in a lower S.D. than that of the 2022 short-term model (Asai et al., 2023). Since the time block-block for Fleet 1 in the short-term model, which was a remnant of the long-term model, was a mis-specified time-block when any size data was not available (1983-1992), this needs to be eliminated. Estimation of the end-parameter of the double normal selectivity for Fleet 1 needs to be confirmed whether that parameter hit to its boundary during the stock assessment meeting using the finalized data set.

3.3 Fitting to Newly available size composition data

The size composition for the Japanese tuna purse seine operating in the Sea of Japan (JTPS_SOJ; Fleet 6) has been obtained by the port sampling (Fukuda et al., 2012). A part of this fleet initiated the farming operation in early 2010’s, and it has been increased a proportion of the farming operation in recent year. Nishikawa et al. (2023) estimated and submitted the size composition for the farming operation of JTPS_SOJ (assigned as Fleet 7)**.** The composition for the farming operation was mainly occupied by the same age-class with the Fleet 6 except age-3, which was not shown up in the farming operation. Nishikawa et al. (2023) also tested how to model this new fleet in the PBF assessment model, and they suggested two methods to estimate the length selectivity of this fleet. The first one was an independent estimation of the length-based time invariant selectivity and the age-based time varying selectivity for the farming operation fleet. The other one was a sharing selectivity with the Fleet 6 (TPS-SOJ), which assumed the length-based time invariant asymptotic selectivity and the age-based time varying selectivity for ages from 3 to 7, with fitting to the size composition data of the farming operation.

Although the former (independent estimates) showed slightly smaller residuals than the later one (sharing selectivity), it required 40 parameters to estimate more than the sharing selectivity method. Because the sharing selectivity method also could replicate both size composition data of Fleets 6 and 7 (market landing and farming), the authors recommend applying the sharing selectivity method.

3.4 Improve the residuals for the size composition

Fukuda (2023) conducted comprehensive retrospective analysis to identify the cause of the systematic retrospective error in the 2022 base case model. It compared the retrospective analysis for ASPM-R_{fix} type models with a size composition from each fleet one by one. It was indicated that the residuals for size composition data from Fleet 2 (Japanese longline in fishing season 1-3) and Fleet 5 (Japanese Tuna purse seine in the Pacific Ocean side) lead a part of the systematic retrospective pattern. A model which down-weighted size composition data for those fleets showed somewhat smaller degree of the retrospective pattern (ISC, 2023). In here, some examples to reduce the residuals for the size composition data of Fleet 2 and 5 were provided.

3.4.1 Japanese Longline in fishing season 1-3 (Fleet 2)

Fleet 2 (Japanese LL season 1-3) is a longline fishery fleet, which was not related to the JLL abundance index (S1), operated outside the main fishing season (fishing season 4). The separation of the catch and size composition data for JLL in season 1-3 logically sounds since there might be a seasonal difference in the selectivity and/or availability for this fishery. The average catch amount for Fleet 2 had been quite small (< 200 tons/year) until 2019, and the average input sample size for size composition data was also small (c.a. 3.8/year) accordingly.

The size composition data of Fleet 2 showed a multi-modal distribution with several spikes at the fish length of smaller than 150 cm. Although this kind of distribution was usually difficult to be replicated by a simple functional form of the selectivity, a time-invariant length based double normal selectivity was assumed for Fleet 2 because of the small amount of catch and input sample size. To replicate this kind of multi-modal distribution, a flexible form of the selectivity (e.g. time varying non-parametric age specific selectivity) was necessary, but it also needs a large number of parameters to estimate (more than 5 parameters in each year).

Given the small amount of catch observed for this fleet, the authors proposed to eliminate most of those spiky composition data. For this, the catch and size composition data by JLL for the fishing season 4 during FY2017-2022 when the large amount of small PBF were observed in the size composition data, were moved back to Fleet 1. The observation in the composition data at 152 cm bin and smaller bins were eliminated before 2020 to maintain the consistency between the size composition data and the data filtering process for the JLL abundance index (Survey 1). Two selectivity time blocks in 2020 FY and a period between 2021 and 2022 were specified for the Fleet 1 to separate the selectivity during the CPUE associated period (1983-2019) and later period (fishery period). The selectivity of JLL in the fishing season 1 to 3 (Fleet 2) for the period when the catch amounts were remained small (1983-2020 FY) could be assumed as the time invariant selectivity, which was estimated based on the composition data of one of the typical year (i.e. 2020). The difference between the 2022 assessment and this proposal was summarized in Table 1.

Table 1. Catch, Size composition and time block for the selectivity of Fleet 1 and Fleet 2

F1JPN_LL(S4)		
	Simple update	**Modified model**
Catch	1983-2016 (S4)	1983-2022 (S4)
Size	1993-2016 (S4)	1993-2022 (S4)
Time block	No	2020-2020 2021-2022
F2JPN_LL(S1-3)		
	Simple update	**Modified model**
Catch	1983-2022 (S1-3) 2017-2022 (S4)	1983-2022 (S1-3)
Size	1993-2022 (S3) 2017-2022 (S4)	2020-2022 (S3)
Time block	No	2021-2022

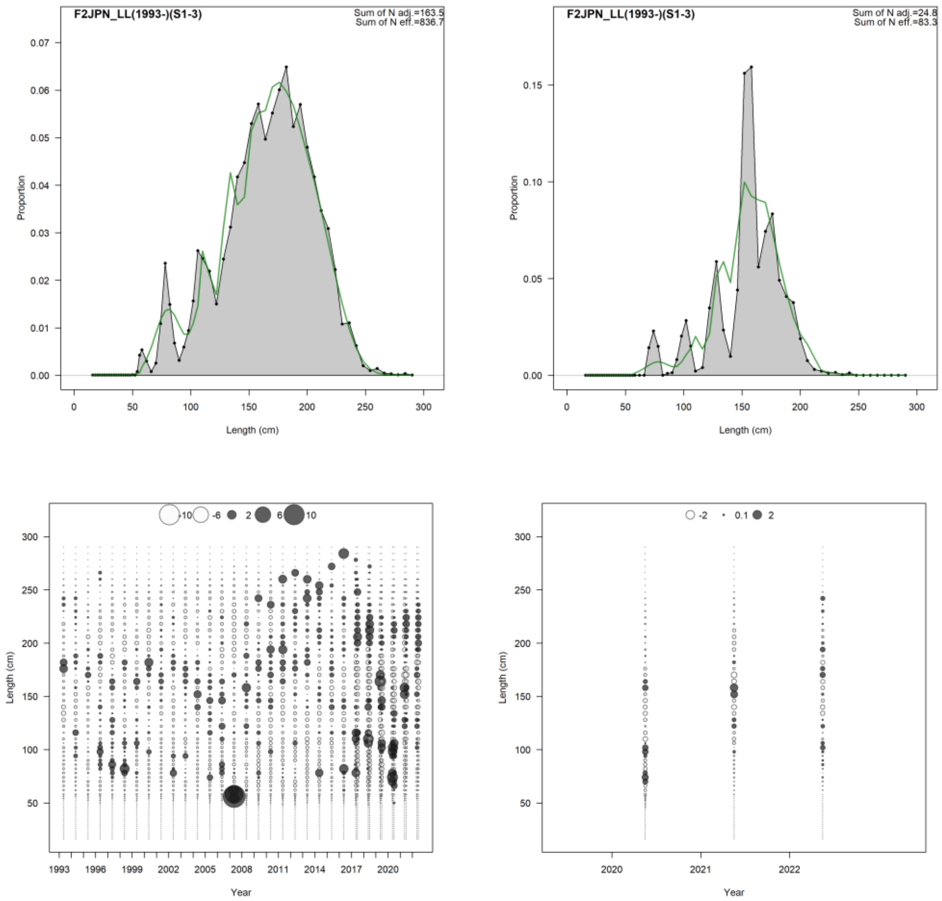

Fig. 1 Aggregated size composition data for Fleet 2 (upper panels) and residual plots (lower panels) from the simple update model (left) and the modified model (right).

3.4.2 Japanese Tuna Purse seine operating in the Pacific side (Fleet 5)

As shown in Fukuda et al. (2019), there was a big shift of selectivity for this Fleet 5 from “small & large PBF fishery” to “only large PBF fishery” due to the management change. To reflect this phenomenon in the model, two time-blocks has been introduced as the age specific non-parametric selectivity at 2011, which was the initial year of the new management and after 2014 when larger PBF than previous years were observed in the data.

Although this specification of the selectivity worked well in the 2020 assessment, some obvious residuals were observed in recent year (e.g. 2019 and 2021 fishing year). This indicated a possible change in the fishing practice and/or availability due to the migration of fish. A time invariant selectivity within the terminal time-block (from 2015 to 2022) could not allow the model to depict the observed length composition data.

To reduce the residuals for this size composition data, we added two more time-blocks within the period of 2015-2022, namely 2015-2018, 2019-2019, and 2020-2022. This flexibility in selectivity enabled the model to depict the recent size composition data more closely (Fig. 2). The authors recommend to have additional selectivity parameters for this fleet.

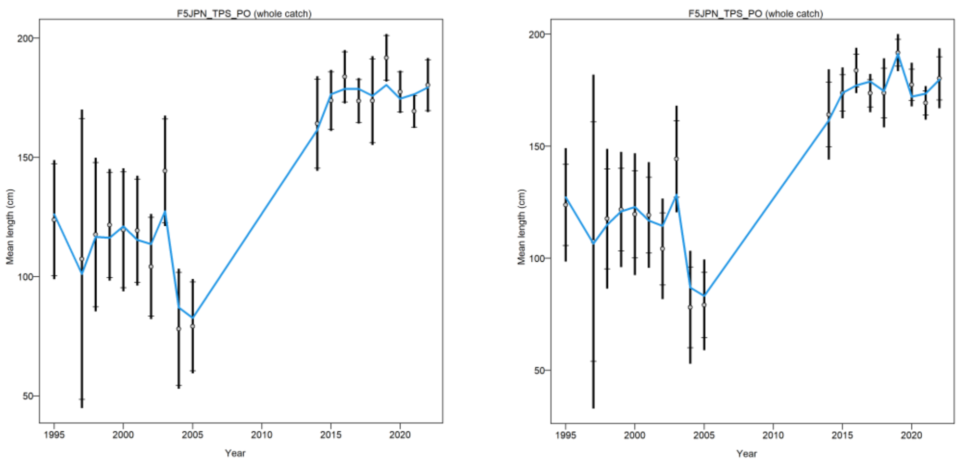

Fig. 2 Mean size for Fleet 5 (Jpn tuna PS in PO) with 95% confidence intervals based on current sample sizes from the simple update model (left) and the modified model (right).

3.5 Improving the retrospective diagnostics

Since the 2020 PBF stock assessment, a small but persistent retrospective underestimation of terminal SSB in recent several years was recognized and it was carried to the current assessment (Fig. xx), which was simply updated from the most recent assessment conducted in 2022. This issue was also pointed out at the 18th meeting of the scientific committee of the Western and Central Pacific Fishery Commission (WCOFC SC18, 2022) as one of the uncertainties of the assessment. This error was likely indicated that the model could not anticipate the rapid recovery of the SSB when the observations are peeled back from the terminal year. The recovery of the SSB was basically informed by two consistent standardized abundance indices based on the catch per unit effort (CPUE) of longline fleets from two different fishing nations. And the recovery trend was also observed in the result using very simple stock assessment model (e.g. age structured production model (ASPM)) that fitted to the catch alone.

Fukuda (2023) conducted comprehensive retrospective diagnostics using the PBF dynamics model based on a simple ASPM-R with alternative assumptions for the data (adult index, recruitment index, and composition data by fleet) and productivity (steepness, natural mortality, recruitment). This analysis suggested that the input data, in particular Japanese troll recruitment index (S3) in recent year and the residuals for the Fleet 2 and 5 (as mentioned in the above sub-section), were possible contributing factors.

In this study, a simply updated model without the recruitment index after 2010 as well as that after 2013 were diagnosed (Fig. 3b &3c). The results suggested that both of the models eliminating the recent recruitment index contributed to improve the systematic error, although the model without that index after 2013 showed clearer negative error than that of the model without the index after 2010. The Mohn's rho values for those two models were almost identical (c.a. -0.08). It should be noted that the all the models tested in this sub-section showed a lower Mohn's rho values (in absolute value) than 0.2, which is a criteria commonly used (Mohn, 1999). A smaller Mohn's rho value in a simple updated model (figure 3a) than the 2022 assessment base case model does not mean the improvement of the assessment model but just because of a rapid recovery of the SSB in recent year, and it made a relative retrospective error over the overall biomass change in the full-data model smaller.

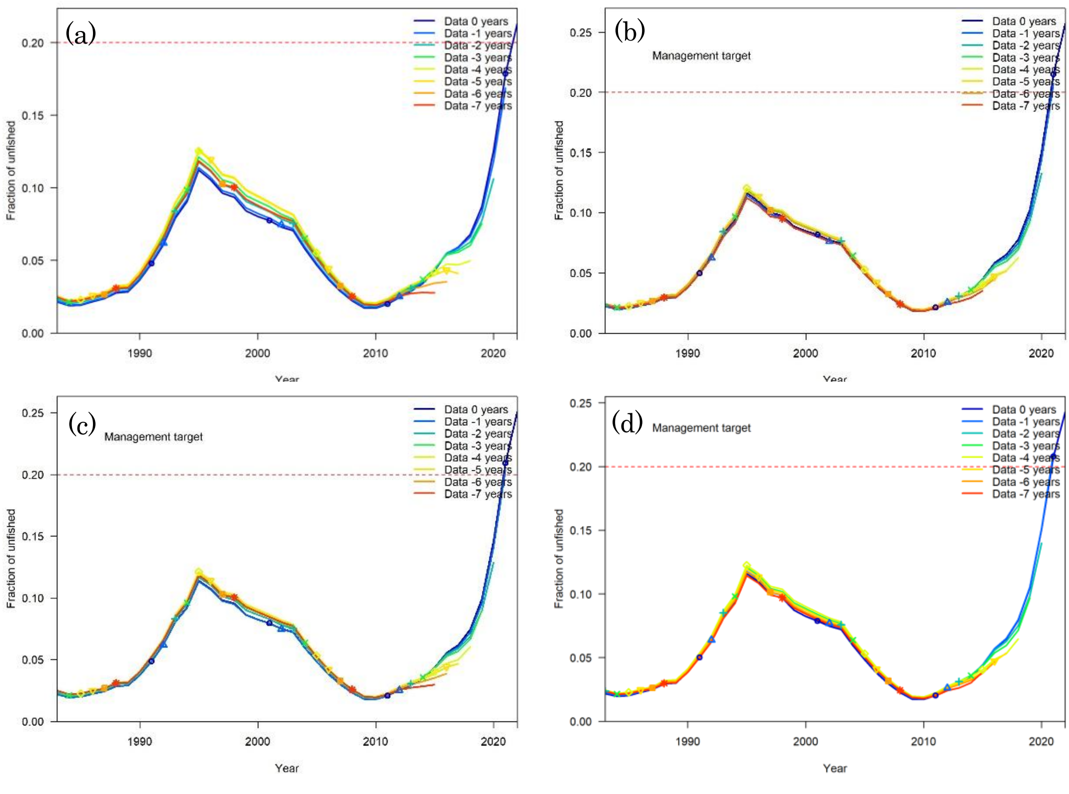

Fig. 3 SSB calculated for each peeled model for (a) the simple update model, (b) simple update model without Jpn troll index after 2011; (c) simple update model without Jpn troll index after 2013; and (d) simple update model without Jpn troll index after 2010 and reduced residuals for Fleet 1, 2, 5.

I also diagnosed the model eliminating the recruitment index after 2010 with additional change in the size composition data and selectivity pattern written in the sub-sections 3.4.1 and 3.4.2 (Fig. 3d). Main contributing factors for a systematic negative retrospective pattern were somehow fixed in this model, and the model showed most consistent SSB estimates among those 4 models (Mohn's rho = c.a. -0.02). As a conclusion of this sub-section, the authors recommend to exclude recent recruitment index data points (e.g. 2011-2016) for the sake of the internal consistency of the model.

3.6 Fitting to newly available index of abundance

3.6.1 Taiwanese longline CPUE based index

In the current assessment model, Taiwanese longline (TLL) CPUE based index is a single index of abundance, which was maintained its continuity up to the terminal year (e.g. 2022 fishing year). TLL index currently used was standardized using a delta-generalized linear mix model (delta-GLMM) without consideration of the spatial effect (Chang et al., 2020). This was a traditional method and the PBFWG chose this method to prioritize the length of time series, where a CPUE standardized by the vector auto regressive spatiotemporal model (VAST) which considered spatial effect, was also available.

Additionally, (Yuan et al., 2023) further improved CPUE standardization method using the VAST model incorporating SST and age group data. Although this new method has a shorter time series than the traditional GLMM index due to the data availability, it has an advantage which incorporated the spatial and size effect on the catchability of PBF. Given the recent situation of TLL fleet, whose operation dynamics might have changed, this new method may have an advantage.

Fukuda and Chang (2023) then, compared the performance of the candidate TLL indices in the 2022 assessment model, and they found that the age 6-8 specific TLL index standardized by VAST with age and SST data showed a high consistency with the Japanese longline (S1) index of abundance, which was another longline CPUE based index of abundance.

Given a short-time series of the age-specific VAST CPUE, the same tests were performed for all the TLL indices submitted by Yuan et al. (2024) (Table 2) with additionally updated two years data. The method to compare the performance within the PBF population model was same with the Fukuda and Chang (2023). We used ASPM-R_{fix} whose recruitment deviation and size selectivity were specified at the values estimated by the same preparation run without any TLL index (Lee et al., 2020). Then, ASPM-R_{fix} was re-run with a candidate TLL index one by one. The fully integrated assessment models with alternative TLL index were also compared.

ASPM-R_{fix} Models 18, 19, and 22, which were fitted to the TLL spatial index for age 6-8 (S10), age 9-11 (S11), and age 18+ (S14), respectively, showed relatively low RMSE values for JLL index (S1) (Table 2). Those results indicated that those TLL indices might bring more consistent information about the population scale with that of the JLL index. However, the RMSE value for those candidate TLL indices themselves by each ASPMR_$_{fix}$ was generally high (RMSE $\geq$ 0.4). This might suggest a possible inconsistency of those age specific indices with the estimated CAA and recruitment deviations based on the other data sources in the PBF assessment model.

On the other hand, the TLL south GLMM index (S5) could be replicated well in both of the fully integrated model and the ASPM_R_{fix} (models 2 and 13). It was not quite sure whether this indicates a good consistency between the TLL S5 index and other data sources in the model, but its notable that the ASPM_R_{fix} with S5 also showed a fairly good fit to the JLL S1 index (RMSE =0.27).

From those results, the authors suggested two TLL indices as a candidate for the base case index. One is the traditional TLL index which showed a fair consistency with the JLL index while being replicated well by the ASPM_Rfix and fully integrated models. The other one was the TLL age 6-8 index standardized by the VAST (S10) model using length data and SST. This index continuously showed a good consistency with the JLL index by ASPM-R_{fix} diagnostics although there was an obvious misfit for the S10 index itself.

3.6.2 Japanese Recruitment Monitoring Survey index

As shown in this study (section 3.5) as well as the Fukuda (2023), Japanese troll CPUE based abundance index in recent year caused a systematic retrospective pattern. If the WG decided to eliminate the troll CPUE based index for some recent years, then, how to deal with the alternative recruitment index would be subject to decide.

Table 3 showed a comparison of the two data structure (use of the recruitment monitoring index or not) by three model structures (Fully integrated model, ASPM-R_{fix}, and ASPM-R_{est}). In ASPM-R_{fix}, the model 4 (w/ recruitment monitoring index after 2010) showed a better fit to the S1 index in terms of the RMSE value. ASPM-R_{fix} with the recruitment monitoring index (Model 4) showed slightly better fit to the JLL index in the last 2 years. This indicated that the recruitment information brought by the S4 index was consistent with the JLL S1 index (Fig. x). It was also notable that the Model 4 (ASPM-R_{fix} w/ monitoring index) showed a better data-matching with the S10 (TLL age 6-8 index). Since the S10 index is the abundance index of the youngest adult fish continuing until the terminal year, this index would theoretically have information about 2016 recruitment year class in youngest. A lower RMSE value in ASPM-R_{fix} w/ the recruitment monitoring index would be a sign of consistency between the recruitment information derived by the recruitment monitoring index and those brought by the TLL age 6-8 index, and this relationship would validate the reliability of recruitment monitoring index during 2011-2016 data points.

Although it was difficult to validate the monitoring index for most recent year (if we had that information in our assessment data, we will not search an alternative recruitment index), a lower size likelihood after 2010 in the ASPM-R_{est} with recruitment monitoring index (model 6) than that without monitoring index (model 5) might suggested the consistency between the recruitment monitoring index and size composition data for young cohorts in a recent year. From those results, now there is a choice for the PBFWG about the Recruitment monitoring index;

1). Continuing not to use the recruitment monitoring index;
2). Use recruitment monitoring index during 2011-2022;
3). Use recruitment monitoring index during 2011-2016 (use a part of the index which was validated through the analysis);
4). Use recruitment monitoring index during 2018-2022 (use a part of the index with avoiding a low data coverage period (Fujioka et al., 2023).

The authors recommended to use the monitoring index for entire period or during 2018-2022 (case 2 or 4).

Table 2 Model structure, index fitted in the model, model parameter setting, and root-mean-square-error (RMSE) for the short-term model with alternative of the Taiwanese longline CPUE based abundance index. Shaded cell indicated that that index was not included in the likelihood function of the model.

Fully integrated model		Model setting				RMSE													
Model No	TLL index tested	selex	Rdev	JLL index	Jtroll index	S1JPN_LL	S2	S3_Troll	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14
1	NoTLL	estimated	estimated			0.255	0.140	0.174	0.606	0.401	0.445	0.607	0.574	0.617	0.554	0.859	0.921	0.877	0.424
2	S5TWLL_GLMM_South	estimated	estimated			0.290	0.137	0.170	0.589	0.235	0.292	0.424	0.374	0.417	0.537	0.700	0.616	0.542	0.403
3	S6TWLL_GLMM_Whole	estimated	estimated			0.298	0.137	0.170	0.587	0.223	0.274	0.435	0.378	0.426	0.554	0.730	0.590	0.492	0.398
4	S7TWLL_GeoSt_South	estimated	estimated			0.290	0.138	0.171	0.577	0.295	0.365	0.286	0.281	0.279	0.422	0.553	0.650	0.658	0.516
5	S8TWLL_GeoSt_Whole	estimated	estimated			0.292	0.138	0.170	0.579	0.253	0.321	0.312	0.283	0.299	0.456	0.593	0.593	0.573	0.458
6	S9TWLL_GeoSt_All_age	estimated	estimated	fitted	fitted during 1983-2010	0.287	0.139	0.171	0.578	0.278	0.349	0.298	0.284	0.288	0.435	0.568	0.643	0.643	0.496
7	S10TWLL_GeoSt_age6-8	estimated	estimated			0.247	0.142	0.173	0.582	0.400	0.449	0.528	0.514	0.536	0.241	0.865	1.072	1.051	0.442
8	S11TWLL_GeoSt_age9-11	estimated	estimated			0.293	0.139	0.179	0.584	0.343	0.415	0.376	0.376	0.370	0.543	0.247	0.872	1.074	0.479
9	S12TWLL_GeoSt_age12-14	estimated	estimated			0.333	0.135	0.165	0.597	0.217	0.276	0.381	0.319	0.365	0.699	0.675	0.265	0.341	0.471
10	S13TWLL_GeoSt_age15-17	estimated	estimated			0.308	0.137	0.170	0.605	0.236	0.285	0.459	0.400	0.455	0.720	0.830	0.460	0.195	0.488
11	S14TWLL_GeoSt_age18+	estimated	estimated			0.274	0.138	0.167	0.593	0.358	0.402	0.558	0.522	0.570	0.538	0.840	0.817	0.766	0.285
ASPM-R_fix		Model setting				RMSE													
Model No	TLL index tested	selex	Rdev	JLL index	Jtroll index	S1JPN_LL	S2	S3_Troll	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14
12	NoTLL					0.233	0.149	0.170	0.605	0.448	0.490	0.645	0.615	0.654	0.535	0.879	0.982	0.960	0.490
13	S5TWLL_GLMM_South					0.270	0.134	0.172	0.607	0.382	0.422	0.615	0.574	0.627	0.607	0.861	0.845	0.770	0.390
14	S6TWLL_GLMM_Whole					0.275	0.133	0.171	0.607	0.379	0.416	0.622	0.578	0.636	0.633	0.867	0.815	0.727	0.382
15	S7TWLL_GeoSt_South					0.284	0.215	0.181	0.604	0.387	0.440	0.552	0.529	0.557	0.500	0.839	0.981	0.957	0.436
16	S8TWLL_GeoSt_Whole					0.271	0.170	0.178	0.605	0.389	0.438	0.579	0.549	0.587	0.529	0.847	0.944	0.907	0.421
17	S9TWLL_GeoSt_All_age	fixed at Model1's values		fitted	-	0.272	0.191	0.179	0.604	0.400	0.451	0.568	0.545	0.573	0.501	0.847	0.992	0.972	0.453
18	S10TWLL_GeoSt_age6-8					0.244	0.136	0.174	0.603	0.501	0.548	0.647	0.631	0.647	0.460	0.902	1.143	1.186	0.638
19	S11TWLL_GeoSt_age9-11					0.255	0.137	0.173	0.606	0.401	0.445	0.609	0.575	0.619	0.555	0.859	0.920	0.875	0.424
20	S12TWLL_GeoSt_age12-14					0.350	0.221	0.175	0.613	0.388	0.405	0.690	0.627	0.710	0.868	0.941	0.611	0.436	0.395
21	S13TWLL_GeoSt_age15-17					0.353	0.237	0.182	0.613	0.393	0.409	0.699	0.634	0.719	0.886	0.951	0.603	0.422	0.398
22	S14TWLL_GeoSt_age18+					0.262	0.139	0.170	0.607	0.391	0.428	0.632	0.589	0.646	0.627	0.871	0.832	0.750	0.395

Table 3 Model structure, index fitted in the model, model parameter setting, and root-mean-square-error (RMSE) for the short-term model with alternative of the recruitment monitoring index (S4). Shaded cell indicated that that index was not included in the likelihood function of the model.

Fully integrated model		Model setting				RMSE														All_size_Likelihood
Model No	Recruitment_index_tested	selex	Rdev	Jtroll_index (S3)	Rmoni(S4)	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	_after_2010
1	Rind2010_No_Rmoni	estimated	estimated	to 2010	-	0.255	0.140	0.174	0.606	0.401	0.445	0.607	0.574	0.617	0.554	0.859	0.921	0.877	0.424	639.5
2	Rind2010_Rmoni_2011				2011-	0.255	0.138	0.174	0.302	0.379	0.425	0.569	0.538	0.577	0.487	0.830	0.912	0.870	0.416	646.5
ASPM-R_fix		Model setting				RMSE														Likelihood_2011-
Model No	Recruitment_index_tested	selex	Rdev	Jtroll_index (S3)	Rmoni(S4)	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	All_size
3	Rind2010_No_Rmoni	Fixed(model 1)	Fixed(model 1)	-	-	0.233	0.149	0.170	0.605	0.448	0.490	0.645	0.615	0.654	0.535	0.879	0.982	0.960	0.490	639.8
4	Rind2010_Rmoni_2011	Fixed(model 2)	Fixed(model 2)			0.229	0.155	0.170	0.302	0.424	0.467	0.612	0.582	0.619	0.480	0.851	0.955	0.930	0.475	646.4
ASPM-R_est		Model setting				RMSE														Likelihood_2011-
Model No	Recruitment_index_tested	selex	Rdev	Jtroll_index (S3)	Rmoni(S4)	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	All_size
5	Rind2010_No_Rmoni	Fixed(model 1)	estimated	to 2010	-	0.163	0.096	0.128	0.631	0.422	0.462	0.620	0.593	0.629	0.403	0.936	1.118	1.070	0.437	729.7
6	Rind2010_Rmoni_2011				2011-	0.167	0.097	0.125	0.151	0.422	0.463	0.614	0.588	0.621	0.447	0.910	1.130	1.086	0.439	719.3

4 General conclusion

This document summarized a discussion point for the 2024 benchmark assessment with some additional analysis to provide a rationale for decision make to change any assessment assumption. Although this document treated really wide range of the proposals to modify the assessment time period, newly available size data, the selectivity parameterization methods, and the abundance index used for the assessment, the estimated population scale was robust among all of tested models as well as the previous assessment because of the strong relationship between the catch and abundance index in the historical period.

The slope of the recovery after 2011 might be different by the choice of the assumptions listed in this document, however, all of the tested runs showed a very rapid recovery of SSB.

It also should be noted that all of the model runs and diagnostics were conducted under the implicit assumption of "catch is relatively reliable than the other observations".

5 Reference paper

Asai, S., Tsukahara, Y. and Fukuda, H. 2023. Alternative input data bin format of length-composition from Japanese longline fishery for robust estimation of its selectivity. ISC/23/PBFWG-01/07

Fujioka K., Asai, S., Tsukahara, Y., Fukuda, H., and Nakatsuka, S. 2023. Recruitment abundance index of immature Pacific bluefin tuna, derived from real-time monitoring survey data of troll fisheries. ISC/PBFWG-01/01

Fukuda, H. 2019. Estimation of the PBF length-composition for the Japanese purse seine operating in the Pacific side. ISC/19/PBFWG-02/05

Fukuda H. 2021. How to improve a flexibility of the Stock Synthesis model for Pacific bluefin stock to the alternative assumptions with keeping its performance. ISC/21/PBFWG-02/12

Fukuda, H., Kanaiwa, M., Tsuruoka, I. and Takeuchi, Y. 2012. A review of the fishery and size data for the purse seine fleet operating in the Japan Sea (Fleet 3). Working paper submitted to the ISC PBF Working Group Meeting, 10-17 November 2012, Honolulu, Hawaii, USA. ISC/12/PBFWG-03/03: 12p. Available at: http://isc.fra.go.jp/pdf/PBF/ISC12_PBF_3/ISC12_PBFWG-3_03_Fukuda.pdf

Fukuda H. 2023. Comprehensive model diagnostics to investigate the cause of a systematic retrospective pattern of SSB in Pacific Bluefin tuna Stock Synthesis model used for the 2022 assessment. ISC/21/PBFWG 02/12

Fukuda, H. and Chang, S.K. 2023. Selection of an abundance index and its selectivity for the 2024 PBF assessment. ISC/23/PBFWG-02/07

ISC. 2020. Stock assessment of Pacific Bluefin Tuna in the Pacific Ocean in 2020. https://isc.fra.go.jp/pdf/ISC20/ISC20_ANNEX11_Stock_Assessment_Report_for_Pacific_Bluefin_Tuna.pdf

ISC. 2022. Stock assessment of Pacific Bluefin Tuna in the Pacific Ocean in 2022. https://isc.fra.go.jp/pdf/ISC22/ISC22_ANNEX13_Stock_Assessment_for_Pacific_Bluefin_Tuna.pdf

ISC. 2023. 24th Meeting of the International Scientific Committee for Tuna and Tuna-Like Species in the North Pacific Ocean, Annex 4, Report of the Pacific Bluefin tuna working group intersessional meeting. https://isc.fra.go.jp/pdf/ISC24/ISC24_ANNEX04-Report_of_the_PBFWG_Workshop_November_27-December_1_2024-FINAL.pdf

Lee, H.H., Piner, K.R., Maunder, M.N., Fukuda, H. 2020. New data tests age-structured production model (ASPM) for Pacific bluefin tuna. ISC/20/PBFWG-01/10

Methot Jr., 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 J. Mar. Sci., 56, pp. 473-488

Nishikawa K., Chang, S.-K., Kwon, Y. Lee, J., Dreyfus, M., Lee, H.-H., Fukuda, H., and Nakatsuka, S. 2024. Input data of Pacific bluefin tuna fisheries for stock assessment model, Stock Synthesis 3; Update for 2024 assessment. ISC/24/PBFWG-01/07

Nishikawa, K. and Fukuda, H. 2023. estimation of the length composition for the Japanese tuna purse seine with new data collected at PBF farming operation using stereoscopic camera. ISC/23/PBFWG-02/02

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

Yuan, T. L., Chang, S.K. Xu, H. Developing abundance indices for Taiwanese PBF longline fishery using GLMM and VAST, incorporating SST and size data. ISC/24/PBFWG-01/05