Report for the Center of Independent Experts on the stock assessment of bigeye tuna in the Western and Central Pacific Ocean

Stephen J. Smith
383 Portland Hills Drive
Dartmouth, Nova Scotia
Canada B2W 6R4

Table of Contents

Executive summary

The Pacific Islands Fisheries Science Center (PIFSC) requested an independent review by the Center of Independent Experts of the 2006 stock assessment of bigeye tuna in the Western and Central Pacific Ocean (WCPO). This assessment applied a similar modeling approach to that used in last year's assessment, however there were a number of changes with respect to data treatment that merited an external review. As for most stock assessments, the data are not perfect and assumptions must be made for their analysis. The fact that these assumptions and methods of processing the data have been changing for this stock assessment over time is not unusual relative to other stock assessments. A number of questions were raised in this report concerning some of the current procedures but these questions are concerned with detail and clarity. Research recommendations dealing with the objective determination of effective sample size, evaluation the value of the tagging data and options for the lack of effort for the Philippines and Indonesian fisheries have been provided here. The MULTIFAN-CL model developed for bigeye tuna in the western and central Pacific Ocean adequately captures the population dynamics for this species given the data on hand.

Background

The Oceanic Fisheries Programme (OFP) of the Secretariat of the Pacific Community, with collaboration from scientists participating in the Scientific Committee of the Western Central Pacific Fisheries Commission, is responsible for conducting the stock assessment for bigeye tuna in the western and central Pacific Ocean.

Results of the 2006 assessment indicate that overfishing of bigeye tuna is likely to be occurring in the Western Central Pacific Ocean (WCPO). While the stock is not yet in an overfished state ($B_{current}/B_{MSY} > 1$), further biomass decline is likely to occur at 2001−2004 levels of fishing mortality at long-term average levels of recruitment. The current level of biomass is 28% of the unexploited level ($B_{current}/B_{currentF=0} = 0.28$). The assessment provides the basis for scientific advice on the status of the stock that is provided regularly at both national and regional levels, and directly influences U.S. policy on resource utilization.

The Pacific Islands Fisheries Science Center (PIFSC) has requested an independent review by the Center of Independent Experts of the stock assessment of bigeye tuna in the Western and Central Pacific Ocean (WCPO). This assessment applied a similar modeling approach to that used in last year's assessment, however there were a number of changes with respect to data treatment that merited an external review.

Description of review activities

This review consisted of a desk review of one document entitled "Stock assessment of bigeye tuna in the western and central Pacific Ocean, including an analysis of management options" (Hampton et al. 2006). This assessment appears to be conducted on an annual basis and the two previous years' documents (Hampton et al. 2005, 2004) were obtained from the internet[1] for background on changes in the procedures over time.

The review was requested to address the following items.

1. Comments on the adequacy and appropriateness of data sources for stock assessment.

2. A review of the assessment methods: determine if they are reliable, properly applied, and adequate and appropriate for the species, fisheries, and available data.

3. An evaluation of the assessment model configuration, assumptions, and input data and parameters (fishery, life history, and spawner-recruit relationships): determine if data are properly used, input parameters seem reasonable, models are appropriately configured, assumptions are reasonably satisfied, and primary sources of uncertainty accounted for.

4. Comments on the proposed population benchmarks and management parameters (e.g., MSY, F_{msy}, B_{msy}, M_{SST}, M_{FMT}); if necessary, recommended values for alternative management benchmarks (or appropriate proxies) and clear statements of stock status.

5. Evaluate the adequacy, appropriateness, and application of the methods used to project future population status.

6. Suggested research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices.

Summary of the findings

The background material for the statement of work (Annex) for this project stated that while the assessment model was more or less the same as last year there were a number of important changes including differences in the relative weightings applied to the different model regions. The assessment document (Hampton 2006) does not actually give this change as such but does list the following as being important (page 25).

a. The weight frequency sample data were reprocessed to account for temporal and fishery specific changes in the conversion factors used to convert processed weights (usually gilled and gutted) to whole fish weights. The principal effect of this change was to increase the weight (in whole weight) of bigeye sampled by the Japanese longline fisheries subsequent to 1973 (see Langley et al. 2006) and, thereby, reduce the magnitude of the decline in fish size from the longline fishery over the model period.

[1] http://www.spc.int/oceanfish/Html/WCPFC/

b. A change in the bigeye length-weight relationship was included in the model, applying a relationship more consistent with established values for the species. The relationship predicts a marginally higher weight-at-length compared to the relationship applied in the 2005 assessment.

c. Selectivity was parameterized to allow declining selectivity of older fish for the principal (LL ALL 1–6) fisheries. In the previous assessment, all longline fisheries were constrained to be non-decreasing with increasing age and, thereby, have full selectivity for the oldest age classes.

d. The base-case assessment (LOWSAMP) applied a lower effective sample size to the length and weight-frequency data compared to the 2005 assessment. This gives greater influence to the effort data included in the model, resulting in trends in exploitable biomass for the principal longline fisheries being more consistent with the catch and effort series. The HIGHSAMP model applies effective sample sizes that are equivalent to those used in all of the 2005 bigeye tuna assessment runs.

e. Only the general linear modeling (GLM) approach was applied to the standardization of the longline effort series. The alternative statistical habitat based standardization (SHBS) approach used in the 2005 assessment was not used in the current assessment.

f. There was a change in the application of the regional scaling effects in the calculation of the standardized effort series for the principal longline fisheries. This resulted in an increased weighting to the LL ALL 2 longline CPUE index and, consequently, a higher total biomass estimated for this region compared to the 2005 assessment.

g. A sensitivity analysis was undertaken to investigate the effect of a substantial change in the regional structure of the model with the inclusion of an additional region in the western equatorial region encompassing the fisheries in Indonesian and Philippines waters.

I will address these changes where necessary under their respective review questions below.

1. Comments on the adequacy and appropriateness of data sources for stock assessment.

As the authors note the only additional data for this assessment relative to the previous year were the 2004 fishery data for longlines, the Philippines and Indonesia, and the 2005 fishery data for purse seines. The 2005 data for longline, the Philippines and Indonesia landings were not available for this assessment. As in previous years, the data used in the assessment were catch, effort, length and weight frequencies and tagging data. Overall, the data are adequate and appropriate for the stock assessment but I do have some questions about the regional stratification scheme and some of the data used.

While the kinds of data used are not new, the assessment does explore grouping the data using a seven region spatial stratification scheme instead of the currently used six region scheme. The new boundaries between region 2 and 4, and region 1 and 3 in the seven region spatial stratification scheme in Figure 3b do not make sense with respect to the distribution of longline catch for bigeye tuna. On page 5 of the current report, the authors report that that this change from a boundary at 20°N to one at 10°N was in response to the study by Langley (2006). However, Langley (2006) notes that the highest ranking options in the analysis of bigeye were a northern boundary at 20°N and the southern boundary at the equator. Apparently, the northern boundary at 10°N was optimal for yellowfin tuna but not so for bigeye. The combined analysis for the two species favored the 10°N boundary but this may simply reflect the dominance of the yellowfin data in the analysis. The southern boundary at 10°S was unchanged from the six region scheme (although the maps in Figure 3 do seem to indicate some change which just may reflect the resolution of the mapping tools) and based upon my understanding of Langley's results probably reflects more on yellowfin requirements than on bigeye requirements.

There may be administrative or other reasons why the spatial stratification scheme must use data for both species but details on these reasons were not available in the document being reviewed here. On page 5 of the current document, the 10°N was suggested as being useful for isolating the area of operation of the purse seines but the more important boundary for this fishery appears to be in an east/west direction in region 4 rather than a north/south direction (Figure 3b).

The lack of effort data for the Philippines and Indonesian fisheries is quite problematic and the substitution of proxy data proportional to the catch in combination with a low penalty weight was used in the model. Given the trends in catch given in Figure 5, I would expect to see fairly flat CPUE curves in Figure 6 for the gears used in the Philippines and Indonesian fisheries (PHID misc, PH HL). While a flat CPUE is presented for the PHID misc 3 fishery in Figure 6, the trend for the PH HL 3 fishery exhibits a great deal of temporal pattern. Where does the information behind this pattern come from? Does this CPUE trend have any influence on the final results?

The caption for Figure 6 also refers to the PH RN fishery but there is nothing in either Figure 5 or 6 on this fishery and I did not find any reference to it anywhere else in the document.

2. A review of the assessment methods: determine if they are reliable, properly applied, and adequate and appropriate for the species, fisheries, and available data.

As noted above (e) the longline effort series was standardized using only the general linear modeling (GLM) approach. The alternative statistical habitat based standardization (SHBS) approach used in previous assessments was not used and no justification or reference was given.

In the 2004 assessment (Hampton et al. 2004) the SHBS approach was claimed to have represented “an overall improvement in the fit to the various sources of data [relative to the GLM approach] and was designated as the base-case assessment.” Then in the 2005 document (page 12) the finding was that “The GLM-based analyses provide much better fits to the data and prior assumptions than the SHBS-based analyses. The main source of improvement was in the fits to the total catches and a lower penalty component for the effort deviations. This indicates that the GLM-based longline effort was more consistent with the observed catches than the SHBS effort.”

While there are probably background papers on what may have changed in the GLM and SHBS analysis of the effort data it would be helpful to have some insight from the current document into the reasons why these methods slip in and out of fashion. After all, the SHBS method was the subject of a primary scientific paper (Bigelow et al. 2002) and given that the resulting catch-per-unit effort indices are the only abundance indices for these fisheries, annual changes in the preferred indices used for the assessment seem capricious at best.

The MULTIFAN-CL population model was used in this report and many previous assessments for this stock. This model is appropriate for the data being used here and this model and associated software has received extensive testing for this and other tuna fisheries in the Pacific (e.g., Fournier et al. 1998, Hampton and Fournier 2001). Note that the url for the MULTIFAN user manual was given as http://www.multifan-cl.org/userguide.pdf which does not work and should be changed to http://www.multifan-cl.org/Downloads/MFCLdoc-1.pdf.)

3. An evaluation of the assessment model configuration, assumptions, and input data and parameters (fishery, life history, and spawner-recruit relationships): determine if data are properly used, input parameters seem reasonable, models are appropriately configured, assumptions are reasonably satisfied, and primary sources of uncertainty accounted for.

The configuration of the model has changed little from previous years with the exception of the introduction of a more severe correction for effective sample size than was used in the past. Overall, the model configuration appears to appropriate for the data on hand, however I have do have some comments and questions on the effective sample size correction, the implications of the model results and the usefulness of the tagging data.

The issue of estimating the effective sample size for cluster sampling has plagued many stock assessments that depend upon samples from fish catches or survey tows for length and weight measurements. While it is reasonable to investigate the impact of different effective sample sizes, I am unable to find the justification for the change from the HIGHSAMP approach previously used to the new and recommended LOWSAMP approach in this document. How were the factor of 0.02 and the maximum of 20 arrived at? Can the work of Miller and Skalski (2006) provide any guidance here on methods of calculating effective sample size?

The authors state that Figure 12 indicates the model has difficulty fitting the length frequencies from some of the fisheries, in particular those exhibiting multi-modal distributions such as LL All 1-2, LL HW 2 and LL PG 3. I assume that the observed histograms are based on the sample information as detailed in Figure 8 or Figure 13. Either there is in an error in Figure 8 or the sample sizes of the recent samples for LL ALL 1 are so small they do not appear after the 1960s.

What I don't understand here is what is actually plotted as the predicted length frequencies in Figure 12. Are these predicted frequencies aggregated over the time period corresponding to when observations were available or are they aggregated over the entire period for the fishery? If the latter why would one expect evidence of a good fit for cases such as LL HW 2 and LL PG 3 where samples were only available since the early 1990s in the former case and only since the late 1990s in the latter (looks like only one year sampled in Figure 8) when the predicted frequencies are aggregated since 1952? The time trend for median fish length for LL PG 3 indicates larger median lengths in the past than now even though there were no data from this fishery prior to the late 1990s to base this on.

How does one interpret what is happening to biomass in Figure 37? This figure indicates that biomass declines in the absence of fishing — region 1 and 2 seem to be particularly problematic. On page 13, the authors state that they made an adjustment for the impact of exploitation on recruitment but there does not seem to be any effect here. Comparing Figures 37 and 28 (recruitment series) suggests that the trends in Figure 37 for the unfished biomass were mainly a function of the observed recruitment series. This suggests that in the case of regions 1 and 2 increases in the biomass due to growth were not able to compensate for losses due to natural mortality — does this make sense? If the fishery is not responsible for the decline in some of these regions then what was?

What exactly is the value of the tagging data to the model? The associated log likelihood for these data is small relative to the length and weight frequency data. The tag release data was confined to regions 3, 4 and 5 and in the six region model the number of releases and returns are relatively small especially when compared to the seven region model. The tag return part of the population dynamics model seems only to fit purse seine in region 3 well and fit poorly to the data from other areas and gears (Figure 18).

The narrow confidence intervals for yield were attributed to the high precision associated with the stock recruitment relationship and the steepness estimate. What was the standard error of the steepness estimate and are there profile likelihood bounds available for this parameter? As noted fisheries data are not very informative about the parameters of the stock recruitment curve and a penalty/prior had to be used to constrain the steepness parameter for this model. What happens when steepness is estimated without the penalty?

4. Comments on the proposed population benchmarks and management parameters (*e.g., MSY, Fmsy, Bmsy, MSST, MFMT*); if necessary, recommended values for alternative management benchmarks (or appropriate proxies) and clear statements of stock status.

The authors have chosen to go with the commonly used reference points $B_{current}/B_{MSY}$ and $F_{current}/F_{MSY}$ for bigeye tuna, although they suggest that these reference points will be more variable than those based on ratios of equilibrium measures. Overall, the findings with respect to overfishing and overfished are consistent here within the context of the model results.

5. Evaluate the adequacy, appropriateness, and application of the methods used to project future population status.

The methods used to project future status with respect to the WCPFC-2 conservation and management arrangements appear to be appropriate given the uncertainties regarding future recruitment trends. However, the lack of knowledge about the effort expended by the Philippine and Indonesian fisheries combined with the fact that these fisheries tend to catch small fish suggests that this information should be used to frame another "axis" of uncertainty for the projections. That is, a scenario with lower recruitment trends for region 3 and the selectivity given by the seven region model for the PHID MISC.

The projections for evaluating the impact of temporal closures for the purse seine fisheries are comprehensive and well supported by the model and data.

6. Suggested research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices.

   i. Exploration of methods for objectively determining the effective sample size for length and weight frequency samples.
   ii. Evaluation of the value of the tagging data to the population model.
   iii. Explore other methods of determining effort for the Philippine and Indonesian fisheries, perhaps investing in an index fishermen program for future monitoring of these fisheries.

As a general comment on the document, note that in a number of the multi-paneled figures, the individual panels differed in scale for the y-axes which was confusing. The following figures should be redrawn to have the same scale for ease of comparison across regions, gear, etc., Figures 5, 6, 7, 11, 19, 25, 28 (except for WCPO panel), 31 (except for WCPO panel), 37 (except for WCPO panel), 40 and 41.

Conclusions/recommendations

The MULTIFAN-CL model developed for bigeye tuna in the western and central Pacific Ocean adequately captures the population dynamics for this species given the data on hand. However, as always improvements can be made to the data and model. Based upon my reading of previous stock assessments and background papers it is obvious that there is a continuous improvement program in place. The questions and comments presented in this review are intended to contribute to this improvement.

Comparing the previous two stock assessments (Hampton et al. 2004, 2005) with the current indicates that each year is an update with much of the text copied from the previous year. New procedures or modifications to methods are usually referenced to other papers presented at the same meeting. This boilerplate approach is not unique to the Scientific Committee of the Western Central Pacific Fisheries Commission as it speaks to the workload limitations of the authors having to do more than one major stock assessment a year. For those attending the scientific committee meetings this document format will work because the other documents and their authors are in attendance. However, this format does make it difficult to review when the document is presented on a standalone basis as was done here. Fortunately, the associated documents (see Bibliography) are available on the internet allowing me to delve a little deeper into the work behind the results presented in the review document. However, if the answers to the questions raised in this review are already documented, I did not see these in the documents that I accessed and did not have time to search every document posted on the SPC site.

Respectfully submitted on 1 October 2007,

Stephen J. Smith
383 Portland Hills Drive
Dartmouth, Nova Scotia
Canada, B2W 6R4
902-446-4404 (residence)
902-426-3317 (office)
smithsj@mar.dfo-mpo.gc.ca

Bibliography

Bigelow, K.A., J. Hampton, and N. Miyabe (2002) Application of a habitat-based model to estimate effective longline fishing effort and relative abundance of Pacific bigeye tuna (*Thunnus obesus*). Fisheries Oceanography. 11: 143–155.

Fournier, D.A., J. Hampton and J.R. Sibert. 1998. MULTIFAN-CL: a length-based, age-structured model for fisheries stock assessment, with application to South Pacific albacore *Thunnus alalunga*. Can. J. Fish. Aquat. Sci. 55: 2105–2116.

Hampton, J. and D.A. Fournier. 2001. A spatially disaggregated, length-based, age-structured population model of yellowfin tuna (*Thunnus albacares*) in the western and central Pacific Ocean. Mar. Freshwater Res. 52: 937–963.

Hampton, J., P. Kleiber, A. Langley, A., and K. Hiramats. K. 2004. Stock assessment of bigeye tuna in the western and central Pacific Ocean. Working Paper SA-2, SCTB 17, Majuro, Marshall Islands, 9−18 August, 2004.

Hampton, J., P. Kleiber, A. Langley, Y. Takeuchi and M. Ichinokawa. 2005. Stock assessment of bigeye tuna in the western and central Pacific Ocean. WCPFC-SC1 SA WP-2, Noumea, New Caledonia, 8−19 August 2005.

Hampton, J., A. Langley and P. Kleiber. 2006. Stock assessment of bigeye tuna in the western and central Pacific Ocean, including an analysis of management options. WCPFC-SC2-2006/SA WP-2, Manila, Philippines, 7–18 August 2006.

Langley, A. 2006a. Summary report from yellowfin and bigeye stock assessment workshop. ME WP-1, WCPFC-SC2, Manila, Philippines, 7−18 August 2006.

Langley, A., H. Okamoto, P. Williams, N. Miyabe and K. Bigelow 2006. A summary of the data available for the estimation of conversion factors (processed to whole fish weights) for yellowfin and bigeye tuna. ME IP-3, WCPFC-SC2, Manila, Philippines, 7–18 August 2006.

Miller, T.J. and J.R. Skalski. 2006. Integrating design- and model-based inference to estimate length and age composition in North Pacific longline catches. Can. J. Fish. Aquat. Sci. 63: 1092–1114.

Annex

Consulting Agreement between the University of Miami and Stephen Smith

Statement of Work

CIE Review of the stock assessment of bigeye tuna in the Western and Central Pacific Ocean

Background

The Pacific Islands Fisheries Science Center (PIFSC) requests an independent review of the stock assessment of bigeye tuna in the Western and Central Pacific Ocean (WCPO). The Oceanic Fisheries Programme (OFP) of the Secretariat of the Pacific Community, with collaboration from scientists participating in the Scientific Committee of the Western Central Pacific Fisheries Commission, is responsible for conducting the assessment. Results of the 2006 assessment indicate that overfishing of bigeye tuna is likely to be occurring in the WCPO. While the stock is not yet in an overfished state ($Bcurrent\ B_{MSY} > 1$), further biomass decline is likely to occur at 2001−2004 levels of fishing mortality at long-term average levels of recruitment. The current level of biomass is 28% of the unexploited level ($Bcurrent\ Bcurrent{,}F{=}0 = 0.28$). While this assessment applied a similar modeling approach to that used in last year's assessment, there were a number of important changes including differences in the relative weightings applied to the different model regions. The assessment provides the basis for scientific advice on the status of the stock that is provided regularly at both national and regional levels, and directly influences U.S. policy on resource utilization.

Review Requirements

The most recent stock assessment of bigeye tuna in the WCPO was completed by the OFP in 2006, with collaboration from U.S. scientists from NOAA Fisheries, PIFSC, and two reviewers are requested to review the assessment. The reviewers should be familiar with various subject areas involved in the review: tuna biology; analytical stock assessment, including population dynamics theory, integrated stock assessment models, and estimation of biological reference points; and MULTIFAN-CL and AD Model Builder. No travel is required and the reviewers will be provided with the necessary documentation, the current assessment of bigeye tuna in the WCPO.

Each reviewer's duties should not exceed 7 days.

A written report from each reviewer is required. The report generated by each reviewer shall include:

1.Comments on the adequacy and appropriateness of data sources for stock assessment.

2.A review of the assessment methods: determine if they are reliable, properly applied, and adequate and appropriate for the species, fisheries, and available data.
3.An evaluation of the assessment model configuration, assumptions, and input data and parameters (fishery, life history, and spawner recruit relationships): determine if data are properly used, input parameters seem reasonable, models are appropriately configured, assumptions are reasonably satisfied, and primary sources of uncertainty accounted for.
4.Comments on the proposed population benchmarks and management parameters (*e.g., MSY, Fmsy, Bmsy, MSST, MFMT*); if necessary, recommended values for alternative management benchmarks (or appropriate proxies) and clear statements of stock status.
5.Evaluate the adequacy, appropriateness, and application of the methods used to project future population status.
6.Suggested research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices.

The PIFSC will provide copies of the current assessment to the CIE for distribution to the reviewers.

Products

No later than October 1, 2007 each reviewer will submit a written report of the findings, analyses, and conclusions to the CIE.

Submission and Acceptance of Reviewer's Report

The CIE shall provide via e-mail the final reports of the consultants in pdf format to Mr. William Michaels (William.Michaels@noaa.gov) for review by NOAA Fisheries and approval by the COTR, Dr. Stephen K. Brown by October 15, 2007. The COTR shall notify the CIE via e-mail regarding acceptance of the report. Following the COTR's approval, the CIE shall provide the COTR with pdf versions of the final report.

ANNEX 1: Contents of Panelist Report

1. The report shall be prefaced with an executive summary of findings and/or recommendations.

2. The main body of the report shall consist of a background, description of review activities, summary of findings (including answers to the questions in this statement of work), and conclusions/recommendations.

3. The report shall also include as separate appendices the bibliography of all materials provided by the Center for Independent Experts and a copy of the statement of work.

Budget

1. Salary ($600/day for a maximum of 7 days)	$4,200