Report to the Center for Independent Experts

Review of the 2005 stock assessment of Yellowfin tuna in the western and central Pacific Ocean

Michel Bertignac

October 2006

Executive summary

- Overall, data and assessment are considered adequate for evaluating stock status of yellowfin in the western and central Pacific Ocean, relative to reference points and to provide management advice; the quantity and quality of work going into the assessment is impressive. The very large variety of data used in the assessment are probably the best available to the analysis at the time. The Multifan-CL methodology provides a rigorous and adequate basis for assessing the stock biomass and fishing mortality rates. However, several potential sources of bias and uncertainty in the data and assumptions made for the assessment leave some ground for further improvement of the analysis. Furthermore, some issues may need some clarification. Some suggestions for improvement are made in the report.

- There seems to be an important source of uncertainty in the historical trend in the catch from the Philippines and Indonesia fishery. Owing to the key importance of this fleet in the recent trends in abundance and fishing mortality levels, this may raise some questions on the results, and more specifically on the amplitude of the recent trends in abundance and fishing mortality, and as a consequence on the conclusion that over-fishing is taking place. Thus, improving knowledge about the development of this fishery should be given a high priority for future assessments.

- Some tagging data useful as fishery independent information is currently used in the assessment but is limited in time. If possible, some new tagging experiment should be carried out. No abundance information is currently provided by fishery independent surveys, which means that stock assessment relies heavily on fisheries-dependent indices of abundance. Even if this is not easy to carry out for tuna species, the possibility of collecting such data should be investigated. The same applies for environmental data which could provide useful further information on abundance and recruitment.

- Although several sensitivity analyses have been carried out (standardisation methods of some longline fishing effort and natural mortality), the impact of the recent substantial changes in some key structural assumptions of the model (spatial structure, methods for estimating the relative weights of areas for abundance) has not been investigated (or at least is not presented in the report). A comparison of an update of the 2004 assessment with current assessment could have been carried out. Furthermore, some retrospective analyses are also missing.

- Some lack of fit on the length and weight distributions has been observed. Further to proposals made in the report for improving the fit, a time variable selectivity for some fishery could be tested.

- Currently the assessment indicates that over-fishing is likely to occur but that the stock does not appear to be over-fished. That conclusion is based on certain assumptions, notably those relating to the value selected in the models for M (natural mortality

coefficient) and *h* (recruitment steepness parameter). Given the information currently available, I believe that the choice decided on in the current assessment was well founded. Further work could however be carried out on those issues, either by investigating other type of S/R relationship or looking at a range of "plausible" values to be used in sensitivity analysis for the steepness parameter. As they are more robust to S/R uncertainties, spawner per recruit (SPR) benchmark levels could be applied. The use of management strategy evaluation would be helpful to identify alternative robust population benchmarks.

Background

The Center for Independent Experts (CIE) at the University of Miami contracted the author to review the stock assessment of yellowfin tuna in the Western and Central Pacific Ocean (WCPO). This independent review was requested by the Pacific Islands Fisheries Science Center (PIFSC). 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 2005 assessment indicate that overfishing of yellowfin tuna is likely to be occurring in the WCPO. The current assessment is more pessimistic than previous yellowfin assessments for the WCPO. The most influential change in the current assessment was due to differences in the relative weightings applied to the different model regions, essentially down-weighting the proportion of the total longline exploitable biomass in the non-equatorial 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 activities

The assessment document was provided on 13 September, 2006, or one month ahead of the deadline to send back the review to CIE; thus, enough time was provided to carry out the review. I was afforded the opportunity to raise questions to the assessment scientists/authors. Instead of a conference call which was initially planned between reviewers and authors, it was decided to send questions via email. This was done on the 6 October 2006. Replies from the authors were forwarded to the reviewers on the 3 November 2006 and provided satisfactory answers to the clarifications that were asked.

Summary of comments

This review is structured on the basis of the term of references provided by the CIE. The points of the ToR are addressed one by one. Thus, a discussion on the data used for the assessment is carried out, followed by some comments on the assessment methodology and structural assumption and parameterisation of the model. Then, the proposed population benchmark and

management parameters are discussed. From my understanding, the 5th point of the ToR is not valid as there was no projection carried out in the report provided. Finally, some suggestions for research priorities are made.

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

The stock assessment of yellowfin tuna in the Western and Central Pacific Ocean uses, in an integrated way, a very large variety of data sources. This includes catch, effort, length-frequency and weight frequency data for the fisheries defined in the analysis together with tag release-recapture data. Such data cover a large part of the available information about the stock and its dynamics and allow the assessment model to estimate key parameters of those dynamics such as abundance, growth, mortality and movement. Furthermore, the fact that it is used in an integrated way helps in the identification of problems, such as when different sources of data deliver conflicting signals (which is the case with length and weight distribution data in the present assessment).

The sampling program for lengths and weight from the commercial fishery is quite extensive with enviable sampling levels, covers a very long period of time (some longline fisheries have been sampled since the early 50s), and is carried out both at port and from observers onboard vessels. This needs to be noted as sampling is an extremely time consuming and logistically difficult activity, especially when dealing with such a large area encompassing so many different countries and exploited by so many different fisheries. However, one area of concern is what seems to be the "poor" knowledge of the level of historical catch in the Philippines and Indonesian fisheries. Those fisheries account for a substantial proportion of the catch in area 3 of the model and, according to available information, is said to have increased steadily over the past decade. As area 3 is of primary importance in the current stock assessment, the strength of this catch increase and the relative importance of this fishery may need to be further investigated.

A considerable amount of tagging data is available for the analysis with about 40,000 yellowfin released during a large scale program conducted from 1989 to 1992 and close to 5,000 recovered. This data set provides valuable information on fish movement, mortality rates, and growth. It is however limited in time (the early 1990s) and current assessments indicate a sharp increase in fishing mortality since that period when tagging was carried out. Although tagging experiments are expensive, they could still be repeated (both at a large scale and using archival tagging) and used to validate the more recent level of fishing mortality estimated (which are only estimated by fishery data) and provide additional information on movement, natural mortality, and growth.

Data on growth and age are available from both tagging and otolith sampling, and this permits a useful comparison of estimated growth rates produced by the assessment model using length distribution from commercial samples with those obtained from other sources of data.

Reliable indices of stock size for recruits, juveniles, and adults are very important in stock assessment. At present, the abundance indices are strongly driven by time series of catches and GLM standardised effort from the main longline fisheries. To get reliable indices, it is better to collect data independently from the fishery, which is not the case in the present assessment. I

know that methods for collecting such data for tunas are not yet available; however, this could/should be envisaged in the future.

No environmental data are currently incorporated into the base-case analysis. Such data are used however for standardizing the longline cpue data in 2 sensitivity analysis (using a statistical habitat-based model for effort standardization). In the report, a reference is made to the SEAPODYM simulations (Lehodey, 2005) which explicitly use such data. As this could provide some extra information, for instance on recruitment level, catchability and/or population distribution, the use of environmental data into the assessment model could also be envisaged in the future, a possibility which has been confirmed by the authors when answering specific questions from reviewers.

In conclusion and despite these minor reservations, data sources currently used can be considered adequate and appropriate for the stock assessment presented in the report and may be the best available to the analysis at the time. A question mark remains on the Philippines and Indonesia data quality. This may have consequences on the current diagnosis on recent trends in key population parameters (biomass and fishing mortality). This uncertainty on the assessment is maybe not stressed enough in the report.

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.

The assessment of yellowfin tuna in the Western and Central Pacific Ocean is carried out using MULTIFAN-CL (Fournier et al., 1998; Hampton and Fournier, 2001; http://www.multifan-cl.org/) which is apparently becoming one of the standard methods for tuna stock assessment in the Pacific Ocean. This is a statistical, size-based, age-structured and spatially structured stock assessment model. It has variable, region specific recruitment, fleet specific selectivity, and time varying catchability. Many types of data are used to provide estimates of the unknown parameters of the model. The methods also provide measures of uncertainty about estimates and predictions of the model.

Historically, this model has been first developed to assess the South Pacific albacore stock (Fournier et al., 1998) and then has been updated and adapted to tropical tuna stocks such as yellowfin and bigeye. This process of adapting a model to the available data and to the stock and fishery specificities, without guaranteeing an adequate model, is probably the best way to proceed. As a consequence, it appears that the model incorporates some of the key features which are important for the species and fisheries assessed here, namely:

i) a spatial structure which is very important for a stock (and a fishery) with such a large distribution and which accounts for the spatial heterogeneities of stocks and fisheries and for the age-dependent movements of fish among areas, the latter of which is quite important for all large tuna stocks,
ii) the possibility of using the information contained in available tagging data which helps in estimating fish movements, growth and mortality rates, and,

iii) the possibility of using both length-frequency and weight frequency data which allows to make the best use of available fishery data.

Both methods used to standardize the longline catch and effort data are appropriate; some, based on habitat are even innovative and should be encouraged (Langley et al, 2005). However, the GLM approach is carried out independently from the population model. A better approach may be to integrate the GLM into the population model.

Overall, the assessment methods can thus be considered as appropriate for the species**,** fisheries, and available data. They provide an adequate and rigorous scientific basis for assessing stock biomass and fishing mortality rates and conducting projections.

On the aspect of reliability, to the best of my knowledge, this seems to be the case for the estimation of the main parameters. The current model has undergone extensive testing, in some cases using an operational model to generate data (Labelle, 2004) and it was found that MF-CL provides estimates that can be considered reliable for management purposes with some reservations on the estimates of M at age. This has been confirmed by the authors. It should be noted however that, as a consequence, in the current base-case assessment, M is fixed. There are limits to such an exercise as "when discrepancies are noted, it is difficult to determine whether the fault lies with the operational model or with the assessment model" (Sibert, 2004).

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.

Overall, the assessment model configuration, for such a complex model, is adequate with reasonable parameter values and structural assumptions. Furthermore, when some data sources or methods are considered uncertain, sensitivity analyses have been conducted to see the impact this may have on the assessment results. This is the case for the effort standardisation of the main longline fleets. A large number of diagnostics of model fits are produced and as the authors write, "the model diagnostics did not indicate any serious failure of model assumptions, although some departures from the model's assumptions were identified in several areas". The comments made in the case of a lack of fit are in general appropriate and suggestions made to improve the model are convincing.

However, despite the clarification provided by the author, there remain some issues in the current report that have not been sufficiently investigated and thus may need further consideration. There were substantial changes in the structural assumptions of the model used for the 2005 assessment as compared to previous years. These include a different spatial stratification with consequences on the definition of the fisheries, the use of modified methods for longline effort standardisation resulting in an increase in the proportion of the biomass in the equatorial region, and some modification in the parameterisation of the selectivity. Overall, the rationale given by the authors to justify the changes is satisfactory. The larger number of areas should permit to better account for spatial heterogeneity of data, as the use of a selectivity function is less demanding in number of parameters than the separate age specific coefficients used in the

previous version of the stock assessment model. However, from the current analysis, it is still difficult to know if changes in the perception of the stock status obtained in 2005 are simply due to the incorporation of new data (one more year) or a result of mainly to changes in the structure of the model itself. It would have been useful, for comparisons with previous assessments, to conduct some kind of sensitivity analysis to those changes by updating the 2004 assessment with new data from 2005 (keeping the same structural assumptions) and compare that with the assessment presented in the 2005 report. Retrospective analysis would have been useful in determining the sensitivity of the model vai the addition of more recent data and to understand how the precision of the assessment for the more recent period. Furthermore, it seems that area 3, which accounts for the majority of the total catch of yellowfin, is still quite heterogeneous in terms of data. Splitting this area in sub-areas might prove useful. The fact that now, due to a re-weighting of the relative biomass between areas, the model results are mainly driven by this area should also be looked into more carefully. This is another argument in favour of a "zoom in" into this important area which could be investigated by the authors for further development of the model structure.

On the specific issue of selectivity, it may be advisable to consider time variable parameters at least for some fisheries as it seems that having constant selectivity over such a long time period (for some fisheries) constitutes a very strong assumption. This could be one of further reasons why the model poorly fits several length distributions. If carried out, this modification to the model should be tested to see if adding more complexity to the model is worthwhile (i.e., the trade-off between better fit and model complexity).

I have no basis to judge if the value of natural mortality rate in the fixed M runs is sound. However, further investigations on the reasons why there is a large difference between the fixed values chosen by the authors and the estimated one (when M is estimated by the model) need to be carried out.

At the aggregate level, the stock recruitment relationship is assumed to follow a Beverton and Holt model. The main reason given by the authors for the use of an S/R model is the need to undertake yield analysis. To avoid constraining the recruitment too much, only a weak relationship between recruitment and spawning biomass is allowed. The more generic question of model validation needs consideration when, as it seems to be the case here, a stock-recruitment model (B&H) is apparently poorly supported by the data and steepness is often difficult to estimate. Testing alternative (e.g. non-parametric) models should be attempted. This is quite important as the yield analysis is based on extrapolations beyond the range of observations which depend on the assumed model. In this respect, in the current assessment, the left part of the S/R relationship is obviously poorly determined as no data are available to fit the ascending part of the relationship. This should lead to high uncertainty on the level of equilibrium yield at high fishing mortality rates.

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

The methods to estimate population benchmarks and management parameters are appropriate and scientifically sound. The model produces a large series of quantities useful for management purposes. The statistical nature of the model allows for the estimation of uncertainties associated with all estimated parameters and management quantities. Probability distributions for the ratio of Bcurrent to Bmsy and Fcurrent to Fmsy are estimated using the likelihood profile which is appropriate. This provides useful information on the current status of the stock relative to the biomass and fishing thresholds. The yellowfin assessment is more pessimistic than the previous assessment. This is, according to the authors, due mainly to an increase in the relative biomass in the tropical areas. It is likely that overfishing is occurring (Fcurrent/FMSY $\geq$ 1), but it might not yet be in an overfished state (Bcurrent/BMSY > 1 for most of the models explored). Further biomass decline is likely to occur if fishing mortality is kept at current levels (2001-2003 average) which will probably move the yellowfin stock to an overfished state ($\widetilde{B}_{current}/\widetilde{B}_{msy}$=0.69–1.00). Several useful fishery impact indicators are also produced. These indicate that the yellowfin fishery has substantially reduced the biomass of the stock in the equatorial region while the temperate regions are more lightly exploited. This reduction in biomass is attributable mainly to the Indonesian and Philippines fishery. During the same time fishing mortality has increased both on juveniles and adults. Several additional comments can be made:

As already stated above (question 3), the MSY based approach is sensitive to the stock recruitment relationship. Further work needs thus to be carried-out on that issue, either by investigating another type of relationship or by looking at a range of "plausible" values to be used in sensitivity analysis for the steepness parameter. Spawner per recruit (SPR) benchmark levels are more robust to such uncertainties, so their use should be considered The use of a management strategy evaluation would be useful to identify alternative robust population benchmarks.

Despite these reservations which address the uncertainty surrounding the estimates of key reference points based on MSY, other, more robust fundamental indicators are useful for providing management guidance. For instance, fishing impact has clearly increased in recent years (F is continuously increasing and biomass is decreasing). Such trends argue against the long-term sustainability of current harvest levels.

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

As no projection is presented in the report, I have no comment on this particular issue.

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

A few suggestions for future research are listed below:

On data

- Improve the quality of the Philippines and Indonesia catch data.

- Carry out a new tagging experiment to validate the more recent level of fishing mortality estimated (which is only estimated by fishery data) and provide additional information on movement, natural mortality, and growth.

- Investigate the possibility of incorporating environmental data into the assessment model.

On the model and its configuration

- Include the possibility of using time-variable selectivity for some fisheries.

- Consider a "zoom-in" into the tropical areas by splitting areas 3 and 4 in sub-areas.

- Investigate the reason for the bias in natural mortality as shown by the model-testing which has been carried out.

- Incorporate the GLM into the population model.

- Test the use of alternative S/R relationship and/or, if using a Beverton and Holt model, carry out sensitivity analysis on the impact of steepness parameter value on management benchmark values.

- Carry out a management strategy evaluation to identify alternative robust population benchmarks.

Appendix 1

Document provided for the review

Hampton, J., P. Kleiber, A. Langley, Y. Takeuchi, and M. Ichinokawa. 2005. Stock assessment of yellowfin tuna in the western and central Pacific Ocean. WCPFC SC1 SA WP-1, Noumea, New Caledonia, 8−19 August 2005.

Further bibliography consulted for and cited in the review

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

Labelle, M. 2004. Testing the accuracy of MULTIFAN-CL assessments of the western and central Pacific Ocean yellowfin tuna (*Thunnus albacares*) fisheries. *Fish. Res.* Submitted.

Langley, A., Bigelow, K., Maunder, M. and Miyabe, N. 2005. Longline CPUE indices for bigeye and yellowfin in the Pacific Ocean using GLM and statistical habitat standardisation methods. WP SA-8, WCPFC-SC1, Noumea, New Caledonia, 8−19 August 2005.

Lehodey, P. 2005. Application of SEAPODYM to the Pacific pelagic ecosystem: recent results and perspectives. WP EB-8, WCPFC-SC1, Noumea, New Caledonia, 8−19 August 2005.

Sibert, J. 2004. Comparison of stock assessment methods using an operational model. SCTB17 Working paper. MWG-4 21p.

Appendix 2

Statement of work.

Consulting Agreement between the University of Miami and Dr. Michel Bertignac

--

September 13, 2006

Background

The Pacific Islands Fisheries Science Center (PIFSC) requests an independent review of the stock assessment of yellowfin 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 2005 assessment indicate that overfishing of yellowfin tuna is likely to be occurring in the WCPO. The current assessment is more pessimistic than previous yellowfin assessments for the WCPO. The most influential change in the current assessment was due to differences in the relative weightings applied to the different model regions, essentially down-weighting the proportion of the total longline exploitable biomass in the non-equatorial 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 yellowfin tuna in the WCPO was completed by the OFP in 2005, with collaboration from Japanese and U.S. scientists, 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, consisting of the current assessment of yellowfin tuna in the WCPO. The reviewers' duties should not exceed 7 days each, and a written report from each reviewer is required. At a mutually acceptable point mid-way through the review, the CIE shall arrange a conference call between the reviewers and the NMFS scientists who participated in developing the assessment. The purpose of this call is to provide the reviewers an opportunity to ask questions and to discuss the assessment. The report generated by each reviewer shall include the following.

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, recommend values for alternative management benchmarks (or appropriate proxies) and provide 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.

Schedule and Deliverables

No later than October 13, 2006, each reviewer shall submit their individual writtenreport1 that addresses points 1-6 above. See Annex I for additional details on the report outline and contents. Each report shall be sent to Dr. David Die, via email at ddie@rsmas.miami.edu , and to Mr. Manoj Shivlani, via email at mshivlani@rsmas.miami.edu.

Submission and Acceptance of CIE Reports

The CIE shall provide the final individual reviewer reports in pdf format for review for compliance with this Statement of Work and approval by NOAA Fisheries to the COTR, Dr. Stephen K. Brown (Stephen.K.Brown@noaa.gov), no later than October 27, 2006. The COTR shall notify the CIE via e-mail regarding acceptance of the reviewers' reports. Following the COTR's approval, the CIE shall provide pdf format copies of the reviewers' reports to the COTR.