UM Independent System for Peer Reviews
Consultant Report on:

Yellowfin Tuna in South-West Pacific

6-20 October 2006

Paul A. Medley
Sunny View
Jack Hole
Alne, YO61 1RT
UK
paul.medley@virgin.net

Contents

Executive Summary

This report is based upon a review of the 2005 stock assessment of yellowfin in the western and central Pacific Ocean. The review was carried out using documents and through questions answered by email.

The data appear to be of reasonable quality and adequate for stock assessment and scientific advice. There are a number of problems with the data which the scientists are addressing and the various research recommendations, which have been made at the WCPO Scientific Committee meetings, are not repeated in this review.

The assessment model appears to be reliable, is properly applied, incorporates the available data (with a few minor exceptions) and adequately describes the changes in the fisheries. The 2005 sensitivity analyses are well chosen, and suggest a focus of research on the longline fishing power change.

The management advice provided in the 2005 assessment was adequate for a broad assessment of the status of the stock. The 2006 assessment gives a much improved assessment for the purposes of management advice, showing a time series of the development of the fishery in terms of indicators and reference points and a fishery projection, among other things.

The most important recommendation of this review is to improve the standardisation of the Japanese longline effort data. The Japanese longline catch and effort data is the primary index of abundance for the stock assessment. The standardisation does not account for increasing fishing power and is generally poorly structured. Overall, it is not clear that the resulting standardised indices are much better than the nominal CPUE.

Other recommendations are:

- The purse seine and pole and line catches from region 1 should be included in future even if only best-guess size compositions are used, such as borrowing the selectivity from other fisheries.
- Given the uncertainty surrounding the Indonesian and Philippine catches, it would be useful to construct sensitivity analyses based on likely catch reporting scenarios.
- A risk analysis, such as a decision table for example, would be a better way to present uncertainty to management currently done through sensitivity analyses.
- Retrospective analyses should be developed as part of the diagnostics of the assessment. Some other diagnostics are also suggested.
- The assessment documentation should perhaps be improved if it is planned that external reviews will be carried out regularly.
- Future research in the short term should focus on improving CPUE standardisation and the catch data. In the longer term, oceanographic research should try to focus on explaining catch rate changes due to aggregation, and changes in productivity which is most likely to affect condition factors of fish and recruitment. Improving models of recruitment is difficult, but would have the greatest impact on the assessment and management advice.

Background

The Pacific Islands Fisheries Science Center (PIFSC) requested 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, was responsible for conducting the assessment. 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.

Before this review was completed, the 2006 assessment was submitted. Although this review primarily considers the 2005 assessment, the 2006 assessment is taken into account where appropriate. Both assessments have essentially the same base model, although the sensitivity analyses have changed.

Review Activities

The main report was supplied by email and other supporting articles were provided on request or loaded down from the Secretariat of the South Pacific Community website. The review covered a number of articles describing the background and development of various aspects of the model. This review focused on those aspects which seem to have the greatest influence on the assessment.

Completion of the review was delayed due to circumstances which prevented the assessment scientists being unable to answer questions rapidly.

Data Sources

Summary of Findings

The data are total catch, tagging data, catch and effort by fleet, length compositions and weight compositions. The primary driving forces behind the assessment results in terms of data are the overall catch and longline catch and effort data. Other fleets have low influence probably primarily because their catchability is allowed to change, whereas longline catchability is fixed. The longline effort data have been standardised. Catches show an almost continuous increase since 1950, and the longline CPUE shows a corresponding decline. There is no evidence of any increase in stock size during this period. This lack of contrast will make some of the advice from the model more uncertain.

The primary uncertainties in the model are related to the main driving forces. Reported catches of the Indonesian and Philippine surface fisheries are thought to be less reliable than other fleets, but have influence on the overall depletion. There were also inconsistencies reported between the size compositions and catch-effort indices. As the main indicator of abundance, it is important that longline catch effort data used in the model is as reliable as possible.

The Indonesian and Philippine catches contribute significantly to the overall fishing mortality of yellowfin. The data are not considered reliable, but it is difficult to quantify any error. The chances are that a constant bias, or raising the time-independent error of the annual catches, would have a limited affect on results as the model could adjust to these errors. Perhaps of greater concern might be changes in trends, due to changes in reporting, for example. There have been several projects in

Indonesia, which should in theory have improved data collection and reporting, and may have biased the estimate of the change in catches over the years around Indonesia. Decreasing the trend in catch may decrease the estimate of fishing mortality. The assessment scientists are aware of these problems and are clearly doing what they can.

The model is fully age structured, so all data are interpreted in terms of age composition. This involves fitting selectivity as well as catchability for all fleets, and therefore size composition data is required for all catches. For this reason some catches which did not have size data, namely purse seine and pole and line of region 1, were not included in the assessment. The overall catch of these fisheries is low, and it is reasonable to assume that would not have much impact on the assessment.

Much of the size composition data are recorded as fish weights. Weights have greater errors than length when converting from size to age. Fish condition factors, difficulties with accurate weighing (less of a problem with well-managed commercial sources) and the cubic relation reduce accuracy and make interpretation more difficult. Conversion factors from processed to whole weight was identified as a potential problem in the 2005 assessment and was addressed for the 2006 assessment. Size compositions are also generally not random, making them difficult to interpret. Given this, I would support the decision to downweight size compositions compared to the abundance index in the 2006 assessment.

Tagging data are less influential, and their overall influence will probably decrease as the tagging experiment was carried out over 10 years ago. However, the tagging experiment remains the primary source of information on movement.

Conclusions and Recommendations

The data appear to be of reasonable quality and adequate for stock assessment and scientific advice. The authors are clearly aware of the potential problems and are taking action to reduce errors and test sensitivity to results, as illustrated by the various reported research and discussion in WCPFC Scientific Committee Second Regular session from 7-18 August 2006 (WCPFC-SC2).

Given the uncertainty surrounding the Indonesian and Philippine catches, it would be useful to construct sensitivity analyses based on likely catch reporting scenarios. This would need to consider the degree to which trends in catches might be due to reporting. The countries concerned may have sufficient information to develop such scenarios.

The purse seine and pole and line catches from region 1 should be included in future even if only best-guess size compositions are used, such as borrowing the selectivity from other fisheries. Although it is likely that these catches are small and will have little overall impact, excluding them probably gives poorer results than using poor selectivity estimates. Also, using and pointing out poorer parts of the model can encourage action by industry and governments to come up with the relevant information.

Japanese Longline Standardisation

Summary of Findings

The primary index of abundance is the Japanese longline fleet. These data were standardised to correct for changes in catchability through the time series. As the abundance index is important in determining the state of the stock, the standardisation procedure was included in this review.

The standardisation does not change the general trends in the nominal CPUE index, but is an attempt to try to make the index more accurate. It accounts for the number of hooks set, hooks between floats (average hook depth), bigeye CPUE and area fished. The standardisation generates an effort time series which can be assumed to have a fixed q in the stock assessment. Therefore these data are very influential in driving the estimated abundance.

The aim of standardisation is to remove all changes to indices of abundance due to changes in catchability. To accompany any standardisation model, at the very least, a narrative is required as to why the model is the form it is, and why it is not removing abundance related information from the index. Clearly, there had been considerable discussion over the years, but I was unable to find full documentation on the background to current model in the time available.

Including area fished should generally raise the accuracy of the index as it removes the effect of permanent features within a region treated as a homogeneous stock. Fishermen use permanent oceanographic features (e.g. equatorial current interfaces) to raise their catch rates, although they may also make sets in different areas for operational reasons and reasons of cost. Accounting for catch rates in different areas generally helps remove these sorts of effects.

It is less clear why hooks between floats or bigeye CPUE should be included, but presumably this is an attempt to account for targeting yellowfin as opposed to bigeye. A more explicit attempt at this was used in the stat HBS model. The GLM should work because the covariates identify the different behaviour of the fishermen, rather than trying to identify different tuna habitat. The hooks between floats are reported as not explaining much variation in yellowfin CPUE. Presumably, HBF and location is confounded with bigeye CPUE.

Some of the covariates could be accounting for changes in abundance. In particular, inclusion of bigeye CPUE would seem to be dangerous, as the bigeye population has undergone a depletion over the same period.

The model does not include any explicit technology improvements, which are usually the main concern because they reduce the slope of the index as the stock is depleted, thereby underestimating the degree of depletion. There have been changes in the material and form of lines, hooks, haulers and engines, which ideally should be accounted for.

The general form of the model does not seem well-founded. The model is multiplicative, which makes sense, but this is achieved by transforming the data. This is not the usual generalized linear modelling approach, which instead would use a log-link function, enabling a separation between the link function and choice of likelihood.

The model is fitted presumably using least squares to the natural logarithm of the catch. I was unable to find a definition of the likelihood used. Using the log-catch means zero catches need to be excluded. This can be avoided by applying a log-link function, but choosing an alternative likelihood to the log-normal. I would suggest quasi-likelihood Poisson or binomial where the variance is assumed to be proportional to the mean. (The beta-binomial would probably be the best likelihood, but the GLM approach has no obvious method to estimate the dispersion parameter.) There are good theoretical reasons why these likelihoods are appropriate. The GLM can be fitted using iteratively re-weighted least-squares.

Using the logarithm of the catches weights the analysis heavily towards smaller catches, which is probably the reverse of what might be wished for longline. In general, the more hooks set, the better the estimate of CPUE. This seems to have been dealt with in part by removing cells with few sets, but the residuals are still skewed. The outliers when effort is low suggest that the choice of likelihood is a poor one. An alternative likelihood should give a better weighting scheme, so standardised residuals would not be skewed, but be normally distributed.

Polynomial terms generally should be avoided if possible, as polynomial parameters are systematically correlated and therefore poorly estimated. In this case, the square and cubic terms may fit better because of the non-linearity introduced by the transform. The model actually used relates catch and effort through a complex function:

$$Ln(Catch) = f\ Hooks + g\ Hooks^2 + \cdots$$

which has no obvious theoretical interpretation. It is possible that the polynomial in this case could be correcting for the non-linear form of the link function introduced by the model. Ln(*Hooks*) has been used as an independent variable in previous working papers, and would seem more appropriate.

Data are provided in aggregated form. This clearly limits the options for the analysis. It is unfortunate data were not available as individual longline sets, as this could allow removal of more direct factors, like specific changes in moon phase, line type, hook type and bait, all of which genuinely could be attributed to catchability.

It is not clear why the regional scaling abundance indices require a separate model of CPUE, unless it is difficult to fit all regions simultaneously. Ideally the estimates weighting region by proportion biomass should be integrated into the same model.

I would agree with the general conclusion, that while a model accounting for tuna habitat could improve understanding of catch rates, current habitat covariates are probably inadequate. The “stat HBS” model uses covariates to define habitat that are not entirely convincing. While temperature and oxygen are likely to be contributing factors, ocean fronts, seamounts, and other factors affecting prey abundance will also determine tuna densities. The “habitat” variables of stat HBS were shown not to explain much more than a straightforward area effect.

Even if the habitat is adequately described, it is not clear why this model would necessarily improve CPUE much as an index of abundance. It needs to be assumed that habitat effects contribute to catchability alone, not overall abundance. A trend of increasing or decreasing habitat over time, due to climate change for example, could produce a trend in abundance. More importantly, it is likely that fishermen already are fully aware of “tuna habitat” and distribute their effort accordingly. Removing such effects will have a limited improvement on the indices, compared with removing technological change, which can be achieved through a GLM.

Conclusions and Recommendations

Overall, there is insufficient support for the linear model used to make it clear that the resulting indices are much better than the nominal CPUE. Technological change increasing fishing power, as considered through the sensitivity analysis, would be the main concern and needs to be addressed.

Individual longline set data does exist and should be made available to the scientists, if at all possible. It seems that the assessment is very limited in what it can do to improve the longline catch and effort time series. Individual longline set data would allow the scientists to consider standardisation in relation to date (moon phase), hook depth, bait type and set location in relation to oceanographic effects, such as the equatorial currents, ENSO and other factors identified by the oceanographic/ecological research. Other covariates are needed to address fishing power. Covariates on vessel characteristics should be available from the WCPO vessel register.

In any case, the model structure needs to be reconsidered. A good general form for the model linear predictor of longline CPUE would be:

$$lp = a \ln(Hooks) + \cdots$$

In this case, a could be either estimated close to 1.0 or forced to be 1.0, making catch proportional to the number of hooks in a multiplicative model. Estimating a value slightly less than 1.0 allows for local depletion and hook interference, and greater than 1.0 allows for targeting higher fish density. Values different from 1.0 are dangerous, however, as they may be attempting to explain changes in CPUE due to abundance where there has been a continuous decline in CPUE, as in this case. After ln(hooks), the remaining linear predictor would be the catch-per-hook and directly interpretable as q. Any fit should also be presented with residuals plotted against expected values and other diagnostics.

The fish capture on hooks could be alternatively be modelled using a multinomial. This model would estimate the catch of yellowfin conditional on the total catch (all species) and the total catch conditional on the number of hooks. However, given only a small proportion of hooks usually have any fish on, the Poisson likelihood should be a good approximation.

Stock Assessment Model

Summary of Findings

The Multifan CL assessment model and software were specifically designed for these fisheries. The modelling approach makes use of size compositions and converts to age based on size frequencies used to identify cohorts. The approach should work well where cohort modes can be identified in the data. The model also makes use of the available tagging data and incorporates a simple spatial structure to allow for movement and different densities across the WCPO area.

The model is required to make a series of assumptions which rely on expert opinion. The Scientific Committee of the Western and Central Pacific Fisheries Commission meets annually to review the details of the stock assessment. For example, biomass was redistributed away from the non-equatorial regions to the more heavily exploited tropical regions, increasing the level of exploitation. The decision to apply this, largely based on the analysis of longline CPUE, was made by the Scientific Committee, not individual scientists. This increases confidence in the assessment.

Where catchability is allowed to change through time, the CPUE is rendered uninformative as indicators of stock size. This is appropriate for the fleets considered. For all non-longline fleets, effort is either unavailable (Indonesian and Philippine surface fisheries), or it is unlikely that CPUE (all surface fisheries) is a good indicator of abundance. This leaves the longline fisheries (LL All 1-6) as the main driving force indicating changes in abundance. I believe that this is correct. However, some minor longline fleets have not been standardised (Hawaiian, Australian), which could form good indices in the regions they exploit.

The first step of testing the sensitivity of the results of the assessment to various factors has been undertaken. Where the results are found to be sensitive, improved methods of estimation should be attempted. In particular, the trend in longline catchability influences the result. I accept that the increase of fishing power of longline is likely to be around 1% or less per year, but some attempt needs to be made to estimate this value.

It is slightly worrying that the fishing power sensitivity analyses fit the data better. A better fit is not the only criterion for choosing the base case, and I can understand that the assessment scientists were not able to justify the values used or the assessment results. However, this rejection applies only to the assumed rate of change used and lower rates of change may indicate both greater depletion and realistic assessment results.

Conclusions and Recommendations

The model appears to be reliable and properly applied. The method is complex consisting of several tweaks and penalty functions to allow adaptation to the available data. However, the model clearly follows the general trends and behaviour indicated by the data. This indicates the assessment at the very least is consistent with the data and provides a reasonable interpretation.

The model incorporates all the available data and adequately describes the changes in the fisheries. The assessment method is properly applied and appropriate for this species and the available data. The model configuration, assumptions and priors are reviewed by the Scientific Committee of the Western and Central Pacific Fisheries Commission, and appear to be satisfactory, subject to on-going research and sensitivity analyses.

The 2005 sensitivity analyses are well chosen, and suggest a focus of research on the longline fishing power change. The sensitivity analyses are identifying important issues which the assessment scientists need to consider and have led to a shift in the base model – the model on which current scientific advice is based.

Retrospective analyses should be developed as part of the diagnostics of the assessment. An important test of models is their ability to predict changes in the stock size in response to management actions. Retrospective analyses provide a useful tool for checking how well a model does this, and should be relatively easy to develop based on the 2006 assessment projections.

Another useful diagnostic would be to match fluctuations in abundance indices with indices of recruitment. This would help in seeing how well the indices and size compositions match. Sharp increases in abundances can only be explained by changes in catchability (which should be removed by standardisation), or recruitment (or immigration), which should be marked by an increase proportion of smaller fish in the catches.

The documentation could be improved if it is planned that external reviews will be carried out regularly. The rationale for the base model and detailed methodology is spread among a large number of documents. It may be worth considering keeping a single detailed document for the current base model for easier and more rapid review. This document could compiled from current sources and be updated regularly.

Management Advice

Summary of Findings

The assessment suggests that overfishing is occurring and current catches are unsustainable. The assessment reports the stock status (total and spawning stock biomass) and fishing mortality indicators, which are appropriate indictors for this fishery. These indicators are presented in a clear and appropriate way (i.e. $B_{current}$ /B_{MSY} $SB_{current}$ / SB_{MSY}, $F_{current}$ / F_{MSY}) and as a range of values emphasizing the uncertainty in the advice.

The 2005 assessment does not provide specific management advice. The conclusions as to the state of the stock and the fishery seem well founded and consistent with the information presented. Given the international nature of the fishery, it would not have been clear what actions management would undertake collectively to manage the stock, making specific advice difficult. The 2006 conservation and management arrangements (WCPFC-2 Report) broadly aim to maintain current effort levels.

It is not clear exactly how the sensitivity analyses are being used for management advice. They are useful for scientists to assess the sensitivity of the results to different assumptions, and to identify an appropriate base model and where future research needs to focus. However, this information is not being presented in a way which managers can easily use.

The 2005 assessment document does not present any projections of future population status. The advice relies on comparing current indicators with long term equilibrium reference points. The 2006 assessment presents a projection based on the proposed management arrangements which should maintain current fishing effort. This has been interpreted as maintaining the final year fishing mortality-at-age, which is probably optimistic given likely future increases in catchability.

Conclusions and Recommendations

The management advice provided in the 2005 assessment was adequate for a broad assessment of the status of the stock. The 2006 assessment gives a much improved assessment for the purposes of management advice, showing a time series of the development of the fishery in terms of indicators and reference points.

A risk analysis would be a better way to develop and present structural sensitivity. This would integrate the sensitivity analyses with the management advice. Decision tables are probably the best way to present qualitative “states-of-nature”, and should be considered. I would suggest three base cases covering the range of assumptions from worst to best case, with the current base case set between the two. These would need to be developed by the Scientific Committee, based on past individual sensitivity analyses and likely scenarios. Management actions would require setting overall fishing mortality to various levels, from precautionary to risky, based on the scientific advice. This would help management to discuss by how much catches might need to be reduced over the coming years.

The current method, as demonstrated in the 2006 assessment, is adequate and appropriate for projecting the population status over 2 to 5 years. One of the greatest uncertainties is the projection of fishing effort or catches. This needs to be dealt with by the management authority defining what projections they wish to be applied as alternative management actions.

Without a reliable stock recruitment relationship, projections of more than 5 years, particularly under heavy exploitation, will always be inaccurate. There is no evidence of a decline in recruitment from the current data, so essentially a long term average is used in the form of a Beverton and Holt stock recruitment function. If the steepness parameter can be estimated, the minimum spawning stock biomass can be defined, and the current projection method would then prove adequate for medium term projections. However, this requires that the functional relationship adequately describing recruitment. In point of fact, this stock recruitment relationship is very poor in explaining past recruitments.

Future Research

The following priority research would probably provide the greatest improvement in the assessment in the short term.

1. Improve the method for standardising CPUE: The current model can be improved and there exist more covariates which can be used to explain changes in catch rates not due to changes in population size. It is recommended that longline set data are used, so that date, time and location can all be used directly or indirectly (for example, through a factor classifying moon phase). This would also allow habitat to be defined as a covariate and included in a genuine GLM rather than some hybrid model.
2. Improve catch estimates for Indonesia and the Philippines: Presumably all that can be done is being done currently to build the best catch estimates possible. The only further research activities might be to look at data collection initiatives that have been introduced and identify changes in catches that might be explained by the recording procedures rather than change in fishing practices. It may be possible, where new procedures have not been rolled out uniformly, to estimate the proportional improvement an intervention may have had using GLMs. Research could form the basis for developing standard scenarios that might be used to bracket the uncertainty as well as improve estimates.
3. Apply adaptive management: Much more information on the response of the resource to fishing will be obtained if the exploitation rate is reduced so that there is an observable increase in biomass; in this case that is standardised longline catch-per-hook increases. The biomass recovery will provide a lot more contrast in the data for the model to fit to. Clearly, this depends on action taken by management rather than the scientists, but managers' attention needs to be brought to the value of allowing biomass to increase.
4. The processed weight conversion factors should probably be updated. This can be done relatively easily if an observer is present during fish processing. This

should ideally be done regularly, as the fish condition factors may change from year to year.

In the medium term, another tagging experiment could be conducted. Since it is over a decade since the previous experiment, the data being used may not accurately represent current dynamics. If this is the case, updating the data may improve the model, but also may indicate that unless frequent experiments are undertaken, the data are of limited value. It is unclear the degree to which the assessment is sensitive to the tagging information. Before undertaking a new experiment, the current model should be used to simulate tagging data to test how influential they might be on the assessment.

In the longer term, understanding tuna interaction with their habitat would be most useful in developing the assessment models further. Research on tuna behaviour will improve general understanding of distribution, but can are very demanding with respect to data. It seems unlikely that this approach with its explicit fine-scale spatial modelling could be used directly in assessments. While continued progress is likely in general understanding, converting the outputs from the research to a form which the assessment can use should take priority. There are two areas where this might be done:

1. Habitat identification: Permanent oceanographic features, such as the equatorial currents, islands and seamounts, might help explain catch rates through aggregating fish. Research might be used to classify locations as particular habitat types to be used as factor covariates in a CPUE standardising GLM (or GAM) above. Dynamic oceanographic features also might define habitat, but it may prove too difficult to generate a time series of the data necessary for standardisation.
2. Recruitment: Dynamic oceanographic data might be used to explain changes in productivity which could affect recruitment. Fecundity might vary with overall productivity, which also might be linked to female condition if such information is available. Various oceanographic variables might then be used to explain density dependent and independent larval mortality, which in future might be combined into an empirically-based recruitment model including spawning biomass as one of the variables. Such a model might explain past recruitment better than the current model. Improving models of recruitment is difficult, but would have the greatest impact on the assessment and management advice.

Annex I Bibliography

The main relevant document describing the stock assessment was sent by email. Other background documents were supplied during the review. The main document to be reviewed was:

Hampton, J., Kleiber, P, Langley, A., Takeuchi, Y. and Ichinokawa, M. (2005) Stock assessment of yellowfin tuna in the western and central Pacific Ocean. WCPFC–SC1 SA WP–1. 1st Meeting of the Scientific Committee of the Western and Central Pacific Fisheries Commission WCPFC–SC1 Noumea, New Caledonia. 8–19 August 2005.

This document made reference to several others which had consulted to complete the review. These were provided on request or obtained direct from the Secretariat of the South Pacific Community Ocean Fisheries Programme website (http://www.spc.int/oceanfish/Html/SCTB/index.htm):

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.

Langley, A. 2003. Standardised analysis of yellowfin and bigeye CPUE data from the Japanese longline fleet, 1952−2001. WP RG-2, SCTB 16, Mooloolaba, Australia, 9−16 July 2003.

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.

Annex II Statement of Work

STATEMENT OF WORK
Consulting Agreement between the University of Miami and Dr. Paul Medley
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 written report[1] 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.

Each written report will undergo an internal CIE review before it is considered final.