Peer review of stock assessment of Bigeye tuna in the Western and Central Pacific Ocean, including an analysis of management options

for

Centre for Independent Experts (University of Miami)

September 2007

Cefas Contract C3156

COMMERCIAL IN CONFIDENCE

Table of contents

Executive Summary

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, conducts assessments of bigeye tuna (*Thunnus obesus*). Results of the 2006 assessment indicate concerns over future stock status. While the assessment used applied a similar modelling 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 Pacific Islands Fisheries Science Center (PIFSC) therefore requested an independent review of the stock assessment of bigeye tuna in the Western and Central Pacific Ocean (WCPO).

This desk-based review was undertaken by Dr Graham Pilling at Cefas (Lowestoft, UK) over the period 3-30th September 2007. Comments on the assessment report are provided against the specific terms of reference provided (Appendix 1). This reviewer finds that the report authors should be commended for a well written and detailed assessment report. The clarity of writing and breadth of information presented greatly aided the process of reviewing. A summary of those findings and recommendations from this reviewer are presented here, by Terms of Reference, and the reader should refer to the main text for full context.

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

There is a considerable variety and time series of data available on which to base a stock assessment for bigeye tuna. The types of data used within the assessment are appropriate. There is general evidence of good sampling across most fleets and areas. The methodology used to standardise fishery-related information is appropriate. The report authors clearly indicate where sampling is lacking. These data issues do have implications for the types of assessment that can be performed and the development of the assessment model. Indeed, the lack of decent information for the Indonesian and Philippine fisheries is a concern, given the estimated magnitude of the catches.

While the analysis of exploitation data has been extensive (and these have strong weighting within the base case model), less emphasis appears to have been placed on uncertainties in the biological data.

Recommendations relevant to ToR 1 are made under the other ToR.

ToR 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.

Given the life history characteristics of bigeye tuna, and diversity of information available, the use of MULTIFAN-CL is appropriate. The model has been extensively tested, and can be considered to be reliable. The assessment is conducted in such way that allows maximum utilisation of information for simulation of the system of interest. Thus, the method is properly applied.

Although MULTIFAN-CL is an appropriate tool, given the limited knowledge of some processes that govern bigeye tuna dynamics, the use of another simpler model is

also recommended, thereby allowing results of alternative approaches to be compared. Retrospective analyses should be performed to identify persistent biases and uncertainties within the data or assessment.

Recommendation 1. A simpler assessment model should be developed in order to perform sensitivity analyses with greater efficiency. Given the limited level of mixing between certain regions, a simpler model could effectively simulate the population with fewer regions and assumptions.

Recommendation 2. Use catch curves to verify the perceived increase in total mortality over the period.

Recommendation 3. Routine performance of retrospective analyses of the stock assessments is strongly recommended.

ToR 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 base case model configuration makes use of the most up to date data on bigeye tuna dynamics and exploitation. Alternative model configurations are appropriate. The seven-region model needs further work before it can be used to supply management advice, and this development should be data-dependent. There is also a noted need to separate out the domestic fisheries of the Philippines and Indonesia within the model. To do this effectively, better information on catches and effort from this region is urgently required.

The natural mortality values used have significant impacts on the assessment and management reference points. There is some uncertainty over how the values of M-at-age used corresponded with those presented in Hampton (2000). Estimates of uncertainty in natural mortality rates of bigeye tuna are available, but not used.

The considerable variability around the estimated stock-recruitment relationship has implications for stock projections and management benchmarks. Recruitment can be predicted almost independently of spawning stock size. Indeed, the model may be able to increase or decrease recruitment to mimic increases or decreases in CPUE/catches. It is difficult to know how the model would react if recruitment was more closely constrained to the mean stock-recruitment relationship, but sensitivity analyses are recommended to examine this. Furthermore, the distribution of recruitment among regions appears independent of the local spawning stock size, which is a strong assumption that needs to be investigated.

Growth estimated by the model does not conform to current understanding. This could be due to the model configuration (e.g. where variability is assigned to annual recruitments rather than other model parameters). The use of informative priors would provide a more biologically realistic model of growth.

Assessment trends are driven primarily by trends in longline fleet catch and effort. These data appear to be the best available, and effort data are appropriately standardized. However, given their importance in defining model outputs, further exploration of the catch, effort and CPUE data for longline and purse seine fleets is recommended.

Uninformative priors for tag reporting rates for longline fisheries were used, but estimated values often hit the boundaries of the priors indicating that convergence had not been achieved. This warrants further investigation.

Recommendation 4. Pursue the development of the seven-region model, but only if a valid abundance index time series can be developed for region 7 of that model.

Recommendation 5. Confirm that the M-at-age values used are appropriate and valid. Assess the use of alternative approaches such as Lorenzen's M.

Recommendation 6**.** Consider the use of estimates of uncertainty in M-at-age available from Hampton (2000).

Recommendation 7. Examine alternative stock-recruitment parameterizations to understand how the model behaves.

Recommendation 8. Provide the posterior of the steepness parameter for the stock-recruitment relationship.

Recommendation 9. Investigate model sensitivity to assumptions made when distributing recruitment amongst geographic regions. Consider linking recruitment distribution to stock spatial distribution or to movement rates predicted for older fish. Provide biological justification of the approach used to distribute recruitment between model regions. Provide model predictions about recruitment distribution.

Recommendation 10. Develop informative priors for length-at-age, based upon existing knowledge (e.g. tagging and otolith studies).

Recommendation 11. Continue investigations into the standardization of the time series of longline effort, and the links between longline CPUE and bigeye abundance.

Recommendation 12. Investigate the potential for standardizing purse seine effort data, and refining effort, catch and biological data on the basis of the setting method (floating object, free school).

Recommendation 13. Improve information on tag reporting rates for longline fisheries, or develop alternative model configurations robust to this uncertainty.

ToR 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 performance of selected reference points will depend upon the fishery system in which they are placed. Therefore, extensive testing is required to identify candidate reference points that provide robust measures of stock status in the face of uncertainties present within the system.

MSY estimates are affected by the uncertainties in natural mortality and stock recruitment, and the fleet composition and their respective selectivities. Given the apparent robustness of recruitments to declining spawning stock sizes, an alternative limit reference point that acts in parallel to the target reference points could be developed.

A simpler reference level, where M is 'known' with some certainty, is to keep estimated fishing mortality at or below this level. This reference point would still require the recent high recruitments to be confirmed, as in the current situation F may be low purely as a result of this.

Recommendation 14. Candidate reference points should be tested through Management Strategy Evaluation (MSE) to evaluate their robustness to uncertainties within the system, against pre-agreed objectives.

Recommendation 15. Examine the utility of 'true' SSB_0 (i.e. the spawning stock biomass in the initial unexploited population) as an index of population depletion, compared to the current approach.

Recommendation 16. Consider the performance of management based upon a limit reference point in combination with the target reference points used, through MSE.

Recommendation 17. Consider the use of alternative, simple but robust reference points, such as levels of fishing mortality relative to natural mortality.

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

Projections appear appropriately configured to meet the requirements of the ECPFD-2 Conservation and Management Arrangements. The method and main assumptions are consistent with those within the stock assessment and therefore are appropriate. Earlier comments on the assessment model configuration also apply here, in particular on the distribution of recruitment between regions.

While stipulated requirements for projections may not specifically require it, future recruitments can be bootstrapped from stock-recruitment relationship residuals - particularly given that the fit to estimated values is poor. Alternatively, given that a projection with high recruitment (average of recent years) and one with moderate recruitment (based upon the estimated SRR) were performed, a projection with low recruitment would provide a more complete picture of plausible future outcomes.

Recommendation 18. Examine alternative scenarios for future recruitment levels used within the stock projections to bracket potential future scenarios.

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

Most recommendations made within this report are predicated upon the assumption that the current assessment model framework, being adequate and appropriate, will continue to be used. However, there are considerable benefits in also using simpler, more easily interpreted models in parallel, the results of which can be more readily communicated to stakeholders, and can be used to support MULTIFAN-CL findings. In turn, the impact of assumptions made within the model on management advice should be examined through MSE simulations.

Sex-specific aspects of bigeye biology may affect assessments. Investigation of whether sexual-dimorphism in growth is occurring is required, particularly given the impact of some gears on particular size classes. Furthermore, suggested higher mortalities on mature females than males cannot explicitly be taken into account, as the current model is sex-aggregated. Further investigation of the sensitivity of model predictions to changes in natural mortality and growth will require a decision on whether a sex-specific or female-specific model should be adopted. This is particularly important given the low estimates of relative SSB at MSY.

Recommendation 19. Use MSE simulations to identify key factors and uncertainties affecting model performance, thereby prioritising future research to fill these critical uncertainties.

Recommendation 20. Investigate whether sexual dimorphism affects growth and other biological characteristics, and hence the potential benefits of developing a sex-separated model.

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 assessment of bigeye tuna (*Thunnus obesus*). 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 ($B_{current}$ $B_{MSY} > 1$), 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 ($B_{current}$ $B_{current}$, F=0 = 0.28). While this assessment applied a similar modelling 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. The Pacific Islands Fisheries Science Center (PIFSC) therefore requested an independent review of the stock assessment of bigeye tuna in the Western and Central Pacific Ocean (WCPO).

This document represents the individual CIE Reviewer Report on the stock assessment of bigeye tuna in the Western and Central Pacific Ocean, at the request of the Center for Independent Experts (see Appendix 1). The author was provided with the stock assessment report (see bibliography) and developed this review report in light of this and other related research.

Description of review activities

The review was undertaken by Dr Graham Pilling at Cefas (Lowestoft, UK) over the period 3-30th September 2007. The documentation (see bibliography) was reviewed at Cefas, and this report to CIE subsequently completed.

Summary of findings

The findings of this reviewer are reported within relevant sections, addressing the six main areas of the Review Terms of Reference (Appendix 1). Numbered recommendations (in bold) refer to the correspondingly numbered items within the conclusions and recommendations section of this report.

The report authors should be commended for a well written and detailed assessment report. The clarity of writing and breadth of information presented greatly aided the process of reviewing.

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

The report shows that there is a considerable variety and time series of data available on which to base a stock assessment. The types of data used (e.g. catches, catch at size, etc.) are appropriate, and there is general evidence of good sampling across most fleets and areas. Indeed, the sources of data are more numerous than those often available for stock assessment. The methodology used to standardise fishery-related information is appropriate.

There are some areas where sampling is lacking, for example catch or effort time series for region 7 post early 1980s. Information from the Indonesian and Philippine fisheries is also relatively uncertain. The authors of the report make this clear. These data issues do have implications for the types of assessment that can be performed and the development of the assessment model. Indeed, the lack of decent information for the Indonesian and Philippine fisheries is a concern, given the estimated magnitude of the catches. This is discussed further, below.

Although an extended analysis of the exploitation data has taken place in order to secure their adequacy and appropriateness, less emphasis has been placed on uncertainties in the biological data. While the latter data appear generally appropriate, they may not be adequate enough to provide a complete picture of the system of interest. This is discussed further, below.

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

MULTIFAN–CL is a sophisticated model that allows scientists to increase realism in the simulation of population dynamics by explicitly accounting for age-specific characteristics of fish and fishery behaviour, movement, and area specific behaviour, for example. It can make use of a plethora of information, which again makes it a tool that allows the utilisation of all available sources of information. Given the life history characteristics of bigeye tuna, and the diversity of information available for this species in the assessment region, the use of MULTIFAN-CL for the stock assessment of this species is appropriate. The model has been extensively tested, and can be considered to be reliable. The assessment is conducted in such way that allows maximum utilisation of information for simulation of the system of interest. Thus, the method is properly applied.

The problem with the complex assessment models that can be developed with MULTIFAN-CL is that the great number of processes it simulates, and the interactions between parameters estimated, makes the interpretation of results difficult, and requires lengthy sensitivity analyses to identify processes that contribute to a given outcome. This can impede the thorough evaluation of all possible scenarios and model configurations. This is important particularly where uncertainty in data or process knowledge is considerable. Indeed, with a large number of parameters being estimated, care needs to be taken to ensure no confounding effects are occurring between those parameters. The Hessian matrix is available and examined, and it is assumed looked at for confounding between parameters. Even so, identifying the causes of some assessment trends remains difficult, as will be discussed.

Although MULTIFAN is an appropriate tool to assess bigeye tuna stocks, given the limited knowledge of some processes that govern bigeye tuna dynamics the application of another simpler model is recommended. The results of the alternative approaches can then be compared as a means of quality control. In turn, alternative scenarios could also be tested using simpler models, which would speed up the sensitivity analyses suggested (e.g. different recruitment scenarios, see below). **See recommendations 1 and 2**.

Although comparisons are made between the results of the current years' assessment and that of the previous assessment (although changes in the selectivity parameterisation makes this difficult), no formal retrospective analyses are performed. Retrospective analyses are invaluable to identify persistent assessment biases and uncertainties within the data or assessment performed. **See recommendation 3**.

TOR 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 are accounted for.

The configuration of the base case model makes use of the most up to date data on bigeye tuna dynamics and exploitation. The running of alternative model configurations (7 areas and high weighting to length data) is also appropriate since these models allow both the simulation of fisheries at a finer scale (7-areas) and the evaluation of alternative model structures that provide further information for the interpretation of the base-case model results.

I support the report author's views that the seven-region model needs further work before it can be used to supply management advice, and that in particular, efforts to develop an abundance index for region 7, where data ends currently in the early 1980s, is needed to allow this model to be taken forward. I also agree that there is a need to separate out the domestic fisheries of the Philippines and Indonesia within the model due to their unique characteristics, the level of catches in this area, and their associated uncertainty. To do this effectively, however, better information on catches and effort from this region is urgently required. **See recommendation 4.**

There are specific aspects of the assessment model configuration and assumptions that deserve further investigation and these are described below:

Natural mortality

A source of uncertainty within the model is the values selected for natural mortality. The values used have significant impacts both within the stock assessment, the estimated stock-recruitment relationship, and in the estimation of management benchmarks such as MSY and B_0.

Age-specific natural mortality values were used in the 2006 assessment. This was based upon sensitivity analyses performed in the 2005 assessment (Hampton et al., 2005), which showed a marginally better model fit (log-likelihood) compared to an assessment where a constant M value was estimated within the model, and upon an expectation of higher natural mortality on females post-maturity (as sex-specific differences in numbers of males and females at larger sizes were found).

There is some uncertainty over how the values of M-at-age used within the assessment compare with those presented in Hampton (2000). That paper states that there is an order of magnitude difference between the natural mortality for fish of sizes smaller than 40 cm and larger fish. The differences seen in the values of M used in the model is smaller than this. However, it is difficult to identify whether there is actually a problem with the estimates used within the assessment based on the documentation available, partly due to the different measures of 'time' within the two documents: Hampton (2000) being in terms of length, that of the assessment being in 'quarters' of age. **See recommendation 5**.

The model does not account for uncertainty in natural mortality. However, Hampton (2000) provides estimates of uncertainty in the natural mortality rate of bigeye tuna, showing that there is considerable uncertainty in the predictions of M for some size groups. **See recommendation 6**.

Stock-recruitment relationship

There is considerable variability around the stock-recruitment relationship estimated through the assessment, which has implications for stock projections and management benchmarks.

The model is allowed to predict changes in recruitment almost independently of spawning stock size, and this may have resulted in the strong recruitment estimates in recent years. However, there is little information as to why recruitment might be so weakly linked to spawning stock size. Indeed, the analysis suggests that recruitment drives model predictions about stock size - the concern is that with the current model configuration, the model can increase or decrease recruitment to mimic increases or decreases in CPUE/catches. Given the complexity of the model, it is difficult to know how the model would react if recruitment was more closely constrained to the mean stock-recruitment relationship. **See recommendation 7**.

A prior is provided for the steepness parameter of the Beverton-Holt stock recruitment relationship, but no posterior is provided, merely the mean value. It would be useful to see the posterior to understand how much information the data is contributing to the estimate. **See recommendation 8**.

It is not clear whether the distribution of recruitment among the regions is independent of the distribution of spawning stock amongst those regions. Even given that actual overall recruitment may be independent of stock size (when steepness of the relationship is estimated to be 0.95), the fraction of recruitment allocated to a given region should be related somehow to the fraction of the stock in that region. Otherwise, justification should be given why, for the same time period, the productivity in some regions is higher than others. The way recruitment is distributed affects models predictions about depletion levels, but also model predictions about future population sizes. It is therefore important to ensure that the predicted recruitment distribution is realistic biologically, and to test sensitivity of the model to assumptions used about recruitment distribution. Also, model predictions about recruitment distribution were not presented. It would be useful to see these. **See recommendation 9**.

Growth

While studies have suggested that the growth rate of bigeye tuna slows between 40 and 70 cm, the model fails to simulate such a trend (see figure 20). Indeed, it actually predicts a steep increase in length at young ages. This could be due to the configuration of the model (e.g. where variability is assigned to annual recruitments rather than other model parameters). The use of informative priors would provide a more biologically realistic model of growth. **See recommendation 10**. See also recommendation 7.

Catch and effort

As noted by the report authors, the assessment trends are driven primarily by trends in catch and effort of the longline fleet. The data for the longline fleets used within the assessment appear to be the best available, and appropriate approaches are used to standardize the effort time series. Further research on the standardization of longline effort is recommended, given its influence on model estimates. As the authors note, an exploration of the relationship between longline CPUE and bigeye abundance, and the methodology applied to standardise the longline CPUE data, particularly to account for temporal trends in fishing efficiency, is warranted. **See recommendation 11**.

The calculation of effort (and CPUE) for purse seine fisheries is uncertain, due to the nature of the fishing operation, the schooling of fish, and the desire to separate catches from floating objects and free schools, due to their different age and length structure. Furthermore, the effort data for these fleets has not been standardized and, as a result, the effort deviations allowed for that fishery are higher than those for longline fisheries. The impact of the effort and catch data from purse seines on the assessment is not clear, but is worth investigating further. **See recommendation 12**.

Tagging information

It was note clear whether the tag reporting rates provided in the report expressed the probability that a tag was reported once found, or was the product of the probability that a tag was reported once found and the probability of catching a tagged fish. The report (beginning of page 18) implies the latter?

The authors have chosen uninformative priors for tag reporting rates for longline fisheries given lack of information about those rates. However, the model appears not to have sufficient information to estimate values for some of these. As a result, estimated values often hit the boundaries of the priors (see figure 26) indicating that convergence has not been achieved. Further research is needed to develop more information on those rates, and/or develop a different model configuration that overcomes this problem. **See recommendation 13**.

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

A range of reference points is presented within the report, against which stock status is assessed. These, and alternatives, are discussed below. Ultimately, however, the performance of selected reference points will depend upon the fishery system in which they are placed, in combination with the levels of uncertainty within the system and the data collection regimes, fleet-based considerations, stock assessment procedures and any harvest control rules that define it. Therefore, extensive testing is required to identify reference points that provide robust measures in the face of the uncertainties present in the system. **See recommendation 14**.

The ratio of current biomass to that estimated as the initial ('true') unexploited biomass (B_0) is not used, to avoid biases that can be introduced by varying recruitment. Furthermore, the approach avoids the problems arising from the fact that the early catch data, which will drive perceptions on historical unexploited stock levels, will be more uncertain than recent data. However, the B_0 calculated from a particular year's population estimate would also suffer from uncertainties in estimated recruitment, the stock-recruitment relationship, and the natural mortality levels. If the use of unexploited levels is to be pursued for reference points, bias might actually be lower if the 'true' initial unexploited SSB level (rather than total population level) is considered. **See recommendation 15**.

The calculations of MSY as a reference point will be affected by a number of factors:

- There remains uncertainty in the estimates of natural mortality.
- MSY will depend upon the selectivity of fleets and the relative mix of fleets. It is not clear whether the overall selectivity used within the calculation of MSY is weighted by either the level of effort or fishing mortality from each fleet?
- The stock recruitment relationship is highly uncertain. The approach taken essentially assumes that recruitment is unaffected by the reduction in spawning biomass (steepness being 0.95). However, this general assumption may be highly risky with the decreases in spawning stock size noted.

Given the apparent robustness of recruitments to declining spawning stock sizes, an alternative limit reference point (for example a biomass limit; B_{lim}, similar to that used in ICES) may be appropriate. Set with an appropriate probability to ensure that stock collapse is unlikely (i.e. a B_{pa} in ICES parlance), this could act as an appropriate limit to work in parallel to the target reference points developed. This reference point still relies on the stock-recruitment relationship, and if recommendations made previously lead to a re-evaluation of that relationship and makes the stock dynamics clearer, this suggestion may be redundant. **See recommendation 16**.

A simpler reference level, where M is 'known' with some certainty, is to keep estimated fishing mortality at or below this level. This reference point would still require the recent high recruitments to be confirmed, as in the current situation F may be low purely as a result of this. **See recommendation 17**.

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

The projections appear appropriately configured to meet the requirements stipulated by the ECPFD-2 Conservation and Management Arrangements. The method and main assumptions used for projections are in line with the assumptions used in the stock assessment and, therefore, are appropriate. Comments provided above on the configuration of the assessment model also apply here, in particular comments on the assumptions made on recruitment and its distribution between geographic regions (see recommendation 9), on which the report authors highlight the sensitivity of model predictions.

The projections appropriately use different assumptions for future recruitment levels. While stipulated requirements for the projections may not specifically require it, the authors should consider bootstrapping future recruitments from the stock-recruitment relationship residuals, rather than taking the mean stock-recruitment relationship value - particularly given that the fit of that relationship to the estimated values appears marginal at best! Alternatively, given that a projection with high recruitment (average of recent years) and one with moderate recruitment (based upon the estimated SRR) were performed, a projection with low recruitment would provide a more complete picture of plausible future outcomes. **See recommendation 18**.

It is not clear in the report whether the same assumptions for recruitment were used in the calculation of $B_{MSYfinal}$ under the two scenarios for recruitment applied in the projections (see report page 24).

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

A number of recommendations have already been made under each of the preceding terms of reference. Many of these are expanded upon, and added to, in this section.

Most recommendations are predicated upon the assumption that the current assessment model framework, being adequate and appropriate, will continue to be used. However, as already recommended (see recommendation 1), there are considerable benefits in using simpler, more easily interpreted models in parallel with that approach, that can be more readily communicated to stakeholders, and can be used to support the findings of the more complex MULTIFAN-CL model. An example of use of models of varying complexity (including MULTIFAN-CL, VPA and production models) is the stock assessment of bigeye tuna in the Atlantic (ICCAT 2007). Although consideration of all these models is not suggested, these calculations provided an insight into the effects of model structure on model predictions that help the interpretation of complex models results.

The impact of the assumptions made within the model, and the impacts on management advice, should be examined through Management Strategy Evaluation simulations (e.g. Kell et al., 2005; Rademayer et al., 2007). This may involve the development of alternative (simpler) models (see above and recommendation 1). Management Strategy Evaluation would allow the impact and robustness of model assumptions to be examined, and hence allow the prioritisation of further research. **See recommendations 19 and 14**.

As noted, natural mortality at age is based upon the expectation of higher mortality on females after reaching maturity, sex-specific differences in numbers at larger lengths having been identified. This implies sex-specific aspects of bigeye biology may affect assessments. Investigation of whether sexual-dimorphism in growth is occurring is required, particularly given the impact of some gears on particular size classes. Furthermore, a higher mortality on mature females than males cannot explicitly be taken into account currently, as the current model is sex-aggregated. Further investigation of the sensitivity of model predictions to changes in natural mortality and growth will require a decision on whether a sex-specific or female-specific model should be adopted. This is particularly important given the low estimates of relative SSB at MSY. **See recommendation 20**.

Conclusions/Recommendations

The stock assessment scientists should be commended in developing a detailed and complete stock assessment report, which provides further information on the activities pursued to improve the assessment of bigeye tuna in the Pacific.

The assessment approach applied to the bigeye tuna stock data appears to provide the best available science, given the diversity of information available. Specific recommendations from this reviewer resulting from this review are presented here, and the reader should refer to the main text for full context.

ToR 1

Recommendations relevant to ToR 1 are made under the other ToR.

ToR 2

Recommendation 1. A simpler model should be developed in order to perform sensitivity analyses with greater efficiency. These models may be more robust to uncertainties in the data available. The level of mixing between certain regions is limited (figure 23 of the report) so a simpler model could effectively simulate the population with fewer regions and assumptions.

Recommendation 2. Given the extensive time series of length information available, the use of catch curves to verify the perceived increase in total mortality over the period is recommended. This will provide further support to the MULTIFAN-CL model outputs.

Recommendation 3. Routine performance of retrospective analyses of the stock assessments is strongly recommended.

ToR 3

Recommendation 4. Pursue the development of the seven-region model, but only if a valid abundance index time series can be developed for region 7 of that model.

Recommendation 5. Confirm that the values of natural mortality used within the assessment are appropriate and valid. Assess the use of alternative approaches such as Lorenzen's M (1996) to estimate natural mortality at age.

Recommendation 6. Where age-specific estimates of natural mortality developed from Hampton (2000) are used, use the estimates of uncertainty in natural mortality available from that paper within the model.

Recommendation 7. Given the importance of the recruitment process, further investigation is needed to explore the effects of using different scenarios about the stock recruitment relationship. Examine alternative stock-recruitment parameterizations to understand how the model behaves. For example, which parameters does the model try to change to follow trends in CPUE if recruitment is constrained more closely to the stock-recruitment curve? Would it, for example, predict a slower growth rate for young fish than the one estimated currently, in order to explain high catches of young fish in given years?

Recommendation 8. Provide the posterior of the steepness parameter for the stock-recruitment relationship.

Recommendation 9. Investigate the impact on assessment results of the assumptions made when distributing recruitment amongst geographic regions within the model. This includes considering linking recruitment distribution to the spatial distribution of the stock or to movement rates predicted for older fish. Provide biological justification of the approach used to distribute recruitment between model regions. Provide model predictions about recruitment distribution.

Recommendation 10. Develop informative priors for length-at-age, based upon existing knowledge (e.g. tagging and otolith studies).

Recommendation 11. Continue investigations into the standardization of the time series of longline effort, and the links between longline CPUE and bigeye abundance.

Recommendation 12. Investigate the potential for standardizing purse seine effort data, and refining effort, catch and biological data on the basis of the setting method (floating object, free school).

Recommendation 13. Improve information on tag reporting rates for longline fisheries, or develop alternative model configurations robust to this uncertainty.

ToR 4

Recommendation 14. Candidate reference points should be tested through Management Strategy Evaluation (MSE) to evaluate their robustness to uncertainties within the system, against pre-agreed objectives.

Recommendation 15. Examine the utility of 'true' SSB_0 (i.e. the spawning stock biomass in the initial unexploited population) as an index of population depletion, compared to the current approach.

Recommendation 16. Consider the performance of management based upon a limit reference point in combination with the target reference points used, through Management Strategy Evaluation.

Recommendation 17. Consider the potential to use alternative, simple but robust reference points, such as levels of fishing mortality relative to natural mortality.

ToR 5

Recommendation 18. Examine alternative scenarios for future recruitment levels used within the stock projections, including bootstrapping residuals and assuming a low-recruitment scenario, to bracket potential future scenarios.

ToR 6

Recommendation 19. Use Management Strategy Evaluation simulations to identify key factors and uncertainties affecting model performance, thereby prioritising future research to fill these critical uncertainties.

Recommendation 20. Investigate whether sexual dimorphism affects growth and other biological characteristics, and the potential benefits of developing a sex-separated model.

Bibliography

Primary documentation

Hampton, J., Langley, A. and Kleiber, P. (2006). Stock assessment of bigeye tuna in the Western and Central Pacific Ocean, including an analysis of management options. Western and Central Pacific Fisheries Commission, Scientific Committee 2nd Regular Session, 7-18 August 2006, Manila, Philippines. WCPFC-SC2-2006/SA WP-2.

Additional references for information used by the reviewer

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

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.

Hampton, J. (2000). Natural mortality rates in tropical tunas: size really does matter. Can. J. Fish. Aquat. Sci. 57: 1002−1010.

ICCAT (2007). Report of the 2007 ICCAT bigeye tuna stock assessment session. Madrid, Spain - June 5 to 12.

Kell, L.T., Pilling, G.M., Kirkwood, G.P., Pastoors, M., Abaunza, P., Aps, R., Biseau, A., Korsbrekke, K., Kunzlik, P., Laurec, A., Mesnil, B., Needle, C., Roel, B. and Ulrich, C. (2005). An evaluation of the implicit management procedure for some ICES roundfish stocks. ICES J. Mar. Sci. 62, 750-759.

Kleiber, P., Hampton, J., and Fournier, D.A. (2006). MULTIFAN-CL User's Guide

Lorenzen, K. (1996). The relationship between body weight and natural mortality in juvenile and adult fish: a comparison of natural ecosystems and aquaculture. Journal of Fish Biology 49, 627-647.

Maunder, M.N., and Hoyle, S.D. (2006). Status of bigeye tuna in the Eastern Pacific Ocean in 2005 and outlook for 2006. IATTC SAR7-BET

Rademayer, R.A., Plagányi, É.E. and Butterworth, D.S. (2007). Tips and tricks in designing management procedures. ICES J. Mar. Sci. 2007 64: 618-625

Appendix 1. Statement of work

CIE REQUEST

Subcontract between the University of Miami and CEFAS (Dr. Graham Pilling)

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.

Cefas