PEER REVIEW OF 2013 NORTH AND SOUTH ATLANTIC ALBACORE STOCK ASSESSMENT

Adam Langley

SUMMARY

This document presents the report of the peer review of the North and South Atlantic albacore stocks assessment process conducted by the SCRS in 2013.

RÉSUMÉ

Ce document présente le rapport de l'examen par des pairs du processus d'évaluation des stocks de germon de l'Atlantique Nord et Sud conduit par le SCRS en 2013.

RESUMEN

Este documento presenta el informe de la revisión por pares del proceso de evaluación de los stocks de atún blanco del Atlántico norte y sur llevado a cabo por el SCRS en 2013.

KEYWORDS

Stock assessment, Peer review, Thunnus alalunga

Background

The peer review of the 2013 ICCAT north and south Atlantic Albacore stock assessment involved the attendance at two meetings of the Albacore working group, i.e. Data Preparatory (22nd to 26th April) and Stock Assessment (17th to 24th June) meetings, and a review of documentation from previous meetings of the Albacore working group available from the ICCAT website. The Data Preparatory meeting was held to collate and review the data inputs for the assessments of both albacore stocks.

The reviewer has 10 years of experience in the stock assessment of tuna fisheries. During 2003-2008, he was employed as the Principal Scientist of the Stock Assessment Section of the Oceanic Fisheries Programme of the Secretariat of the Pacific Community and was involved in the assessment of the four main tuna stocks in the western and central Pacific Ocean (skipjack, yellowfin, bigeye and south Pacific albacore). Since then the reviewer has maintained an involvement in the Pacific tuna assessments and has been routinely contracted by the Indian Ocean Tuna Commission to conduct assessments of yellowfin tuna and, more recently, bigeye tuna. These stock assessments are primarily implemented using the Multifan-CL (MFCL) and Stock Synthesis (SS) software.

The review addresses each of the items specified in the Terms of Reference for the peer review. The review highlights a range of issues associated with the current (2013) assessments. Many of these issues were identified by the wider assessment group and are documented in the assessment report. The main points are reiterated in this review to highlight the key outstanding issues associated with the current assessment. The inability of the assessment group to adequately address many of these issues should not be interpreted as a criticism of the individuals participating in the process. Instead, there a number of generic issues of relevance to the other tuna stock assessments that relate to the deficiencies of the available data (fisheries and biological data) and the uncertainty associated with the dynamics of the tuna stocks. There are also limited resources available to undertake the level of data analysis and modelling necessary to support and advance the assessments.

1. Evaluate the adequacy, appropriateness, and application of data used in the assessment.

i. There is a large amount of data available for the assessments of the north and south Atlantic albacore stocks, although the quality of the key data inputs is variable and, in some cases, key data sets provide conflicting trends. The assessments are highly dependent on the catch history from the fisheries and the CPUE indices derived primarily from the main longline fisheries.

ii. For all model options, the albacore catch history was assumed to be known without error. Thus, the historical catch history is highly informative regarding estimates of yield for the stock (and the associated confidence interval). No information is available to indicate that the established catch history is unreliable and certainly no data are available to enable the reconstruction of an alternative catch history. Nonetheless, there may be some uncertainty in the historical catches for some fleets, particularly associated with the degree of discarding of albacore taken as a bycatch of other target fishing operations.

iii. Analyses of the fleet CPUE data were provided by the respective CPCs and the presentation of the CPUE data sets represented one of the main agenda items for the Data Preparatory meeting. These analyses identified somewhat conflicting trends in the CPUE indices for some of the main longline fisheries (e.g. recent trends in the CPUE indices from the Chinese-Taipei and Japanese longline fisheries). The Albacore working group has developed criteria to review the quality of the respective CPUE indices and these criteria provide a basis for determining the utility of the individual indices as a primary abundance index. However, the nature of the catch and effort data sets, particularly historical changes in the spatio-temporal distribution in the fishing operations, mean that none of the CPUE indices are likely to represent an ideal index of stock abundance. Consequently, the criteria did not provide strong guidance for the selection of a specific primary CPUE index.

iv. The analysis of individual CPUE data sets by CPCs does not provide the opportunity for an integrated analysis of the entire fishery data set. For the main longline fleets, there have been considerable changes in the operation of the fishery, particularly the area of the fishing operation which relates to the degree of targeting of albacore. These spatial changes may not be adequately accounted for the GLM approaches used to standardize the CPUE data. An integrated analysis of the CPUE data from the main fleets may provide insights into the different trends in CPUE from the various fleets. The combined data set may also provide more reliable stock-wide indices by encompassing a larger proportion of the spatial domain of the albacore stock.

v. For the northern Atlantic albacore assessment, the working group gave greater emphasis to the longline CPUE indices derived from the fleets that primarily operated in the core area of the albacore fishery and, on that basis, were considered to provide a better index of the adult stock abundance i.e. the C-T longline CPUE indices in the northern albacore assessment. The Japanese longline CPUE indices were included in the model as a sensitivity analysis. This was the most appropriate treatment of the various longline CPUE indices that were available for the current assessment.

vi. Time series trends in catchability for the primary longline CPUE indices were not explicitly considered. However, the CPUE indices were apportioned into time blocks that were considered to represent periods of comparable fleet operation. Separate catchability coefficients were estimated for each time block, thereby, enabling the model to account for gross changes in catchability over the model period.

vii. One outstanding issue related to the application of the longline CPUE indices is the underlying fishery structure. There is a seasonal pattern in the historical operation of the longline fishery with the fishery operating in the higher latitudes during the summer. The albacore catch in that area tends to be of smaller, sub-adult fish while larger adult fish are caught in the subtropical areas. The current fishery definitions do not explicitly partition these two components of the fishery and catches and size data are aggregated for the entire stock area. The stock is being indexed by CPUE indices primarily from the core fishery area mediated by the composite fishery selectivity. Marked changes in the seasonal operation of the main longline fisheries could introduce significant biases into the assessment models. This issue appears to be most relevance to the northern albacore assessment.

viii. The northern albacore stock assessment also included CPUE indices from the surface fleets. The troll fishery indices revealed a long-term declining trend in CPUE from the fishery. The decline in the CPUE indices tended to coincide with a contraction in the spatial extent of the operation of the fishery. The current assessment assumes that the catchability of the fishery is constant and, therefore, the model interprets the decline in CPUE as a decline in albacore recruitment. In addition, inter-annual variation in these CPUE indices may relate to variation in the availability of albacore to the fishery driven by oceanographic conditions. Further analysis of the catch and effort data from the fishery should be undertaken to support the current assumptions relating to the catchability of the fishery.

ix. The northern assessment incorporates a large amount of length frequency data from the various fisheries. The assessment indicated that the integrated assessment models (MFCL and SS) are sensitive to the assumptions related to the length data from the longline fisheries (relative weighting and selectivity parameterization). However, very limited information are available to assess the reliability of these data and, particularly for the C-T longline fleet, there are large changes in the level of sampling over time and a high degree of inter-annual and decadal variation in the length composition of the sampled catch. A more thorough spatial analysis of these data is warranted; however, more crucially the protocols and methodology applied to the collection of the length samples is required to evaluate the utility of these data. In the current assessment, the appropriate decision was made to down-weight the length data from some fleets and, thereby, reduce the influence of these data in the final assessment.

x. These length data (CAS) are also fundamental in the derivation of the catch-at-age (CAA) for the northern stock which is a direct input to the VPA assessment model. If the size sampling data are not sufficiently representative of the annual catch from each fleet then the resulting CAA will be biased. Changes in the sampling approach have the potential to introduce substantial biases into the time-series of CAA estimates and, therefore, the VPA assessment model.

xi. During the Data Preparatory meeting, there were problems in replicating the time-series of CAA data derived by applying the Kimura-Chikuni length slicing algorithm to the CAS datasets. The sensitivity of the CAA estimates to minor changes in the fixed parameters assigned in the algorithm indicates that the CAS data are not sufficiently informative to determine precise estimates of CAA. However, the associated level of precision is unknown. An evaluation of the reliability of the generation of the CAA data sets is required to determine the suitability of applying these data in a VPA framework. Such a review would include an appraisal of the reliability of the historical and current collection of CAS data (by fishery) and the length slicing approach.

xii. Tag release/recovery data are available from the Spanish baitboat fishery from a period of tagging activity during the late 1980s and early 1990s. These data were incorporated in a single MFCL model sensitivity (the data could also be incorporated in the Stock Synthesis model). The resulting model estimated a low (approx. 10%) reporting rate from the fishery which contradicts the observation that the Spanish fishermen were active participants in the recovery phase of the tagging programme. The low recovery rates may be aliasing a number of other factors in the assessment model, e.g. high mortality of albacore following release, higher natural mortality of younger albacore, variability in fishery selectivity, and/or stock mixing assumptions. It is unknown whether a thorough analysis of these tag data has been undertaken; however, there is potential to apply these data to investigate the spatial distribution of releases/recoveries and, thereby, explore the current assumptions regarding stock mixing.

xiii. There remains some uncertainty related to the stock structure of albacore in both the north and south Atlantic. All the data sets were configured in a manner that was consistent with a single, discrete fish stock in each area. However, there is some indication, particularly in the south Atlantic, that the stock structure may be more complex. In that regard, it would have been useful to formulate alternative data sets to enable an evaluation of the sensitivity of the conclusions of the assessment to a number of different stock hypotheses.

2. Evaluate the adequacy, appropriateness, and application of methods used to assess the stock and if appropriate recommend alternative approaches to be accomplished in the future.

i. The current assessment was conducted using a range of modelling approaches implemented using established software packages that are routinely used in the tuna stock assessments conducted by the various RFMOs.

ii. For the northern stock, stock assessment models of differing levels of complexity (ASPIC, VPA, MFCL/SS) yielded very similar results. This does not necessary provide a validation of the range of approaches used. Instead, it highlights the relative importance of the two main sets of input data that are influential in the range of models, specifically the total catch history and the C-T longline CPUE indices. The MFCL/SS models have the flexibility to integrate more diverse data sets and while additional data were included in the model and information content from these data was either low (or down weighted) and/or consistent with the primary fishery observations. It is difficult to evaluate the benefits of adopting a model of increased complexity when there remains considerable uncertainty associated with some of the key input data sets. For example, the inclusion of a large time-series of length data from the longline fisheries have the potential to provide the model with information regarding recruitment and fishing mortality; however, these data have the potential to bias the model results if the sampling is not representative of the catch and/or if fishery selectivity has changed over time.

iii. Overall, the diagnostics from the main assessment models indicated a reasonable fit to the primary data sets. However, for the complex models it was difficult to fully evaluate the performance of the individual models due to the limited time available to develop the models during the course of the meeting. In addition, the complex models require considerable testing to ensure that the models have attained the optimum solution. Time constraint also limited the opportunity to undertake a range of supporting analyses using the model such as retrospective analyses and a greater range of sensitivity analyses. These analyses would have been informative about the reliability of the estimates of current stock status.

iv. For both southern and northern albacore, there have been changes in the proportion of the catch taken by the surface and longline fisheries over the history of the fishery. These changes are explicitly accounted for in the statistical age structured models; however, production models such as ASPIC do not account for changes exploitation pattern and the associated differences in yield may bias the estimates of stock production (and reference points). Further, production models are unable to take into account temporal variation in the productivity of the stock (recruitment). The results of the MFCL/SS models of the northern stock suggest that recruitment may deviate considerably from the basic stock-recruitment function.

v. The MFCL/SS platforms have a large number of features that enable the complexities in fishery structure to be accounted for explicitly within the modelling framework. This enables a more thorough evaluation of the level of model complexity required in the development of the main model used for the provision of management advice. While MFCL has been used as the main model platform for the assessment of the northern stock, I consider that the model is still in the development phase and further work is required to determine the optimum model structure (fishery definitions and structural assumptions such as the parameterization of fishery selectivity).

vi. The two platforms (MFCL and SS) also enable the integration of tagging data in the assessment model. These data have the potential to provide additional information regarding stock size and exploitation rates and the application of these data should be further evaluated.

vii. There is strong evidence of differential growth rates for the two sexes and differences in the sex ratio of the catch are suggestive of differential levels of natural mortality. Preliminary work was undertaken to develop a sex specific model for northern albacore using SS and further development and evaluation of this model formulation should be undertaken.

viii. Further, the MFCL/SS platforms provide a framework for evaluating the sensitivity of the assessment conclusions to a range of alternative assumptions, incorporating uncertainty associated with biological parameters (e.g. M, growth parameters, variation of length-at-age, SRR steepness), fishery structure (fishery definitions, selectivity and catchability assumptions), the reliability of individual data sets and, potentially, more complex spatial

structure of the fisheries and/or the stock units. A grid approach could be developed to quantify the structural uncertainty of these models and, thereby, highlight the most influential model assumptions.

ix. During the meeting, considerable effort was spent to reconcile differences between the MFCL and SS models. In general, the models produced similar results when configured in an equivalent manner. However, one outstanding issue was the estimation of initial, exploited conditions. The parameterization of initial F is different between the two platforms and there were marked differences in the estimated biomass level at the start of fishing. Further evaluation of the alternative modelling approaches is required to determine the reliability of the initial F estimates and the influence of these values on the estimation of stock status bench marks.

x. The assumptions regarding initial fishing mortality levels/depletion levels are also important in the application of the production models for both north and south Atlantic albacore. Most of the ASPIC models tend to fix the initial depletion level at an assumed value. This constraint results in a relatively low level of uncertainty associated with current stock status. The BSP approach applied to the southern albacore assessment enables the uncertainty associated with this parameter (expressed as a prior) to be more explicitly integrated into the estimates of current stock status.

xi. For the southern albacore assessment, the ASPIC and BSP production models incorporated four sets of longline CPUE indices. These CPUE indices exhibit quite different trends over the last 30 years and, consequently, the relative weighting of the indices was influential in the estimates of current stock status (esp. B/Bmsy). The group did not have sufficient information available to adequately assess the reliability of the individual CPUE indices and, thereby, select a preferred index of stock abundance. On that basis, a range of plausible alternative model options were used to formulate the current management advice and stock trajectories. The adoption of a range of model options resulted in a broad confidence interval for the current stock indicators. This confidence interval spans somewhat divergent sets of model estimates of current stock status, although the results of the individual model options are obscured when the range of indices were combined for the provision of management advice. The inclusion of all model options in stock projections also results in a broad range of outcomes for the constant catch scenarios. Consequently, to achieve a high probability of being above Bmsy the harvest (catch) must be set at a low level relative to current catches. However, there is still a reasonable likelihood of being above Bmsy with catches at the current level or higher. Future assessments should focus on resolving some of the inconsistencies in the various CPUE indicators and/or focus on specific indices that are more likely to be indicative of trends in stock abundance. This may reduce the suite of model options incorporated in the provision of management advice and, potentially, reduce the uncertainty of the current assessment results.

xii. The application of production models to the assessment of the south albacore stock does not utilize the large amount of additional data (primarily size frequency data) available from the fishery. There is potential to develop a statistical age structured model for the southern stock. However, given the limited resources available, priority should be given to the further development of the MFCL/SS models for the northern stock.

xiii. For comparative purposes, a VPA assessment approach was also applied to the northern albacore stock. While the results were generally comparable to the other assessment approaches, there is a great deal of uncertainty concerning the reliability of the CAA estimates that were included within the VPA model. It is recommended that a full evaluation of these input data be conducted to determine the adequacy of the VPA approach in future assessments.

3. Evaluate the methods used to estimate population benchmarks and stock status (e.g., MSY, FMSY, BMSY, or their proxies).

i. Stock status and MSY yields were estimated based on the equilibrium yield assumptions. For the MFCL/SS model options, the steepness parameter of the B-H stock-recruitment

relationship (SRR) was estimated during the fitting procedure. Model estimates of steepness were about 0.75-0.85. However, the observational data are unlikely to be informative regarding the nature of the SRR and, for that reason, tuna assessments conducted by other RFMOs tend to fix the steepness parameter at a range of plausible levels. Nonetheless, for the current assessment, the model estimates of steepness were generally consistent with that range of plausible values.

ii. The MFCL model estimates of steepness are influenced by the estimated decline in recruitment driven by the troll CPUE indices. An examination of the residuals to the fit to the SRR reveals a period of strong positive residuals during the 1960s and 1970s. The relationship between stock biomass and recruitment may not be adequately defined by the B-H SRR. An alternative hypothesis is that recruitment has declined due to environmental conditions (correlated with the NAO?). Further evaluation of the relationship between the recruitment estimates and key environmental variables is recommended. A large shift in the productivity of the stock due to environmental conditions will influence the reference biomass levels, estimates of yield and current stock status.

iii. For the northern albacore stock, estimates of MSY are considerably higher from the ASPIC models than from MFCL/SS. This reflects the deterministic recruitment assumption of the ASPIC model framework and the higher estimates of stock productivity to account for the higher catches in the 1960s and 1970s (contrasted with positive recruitment deviates for the MFCL/SS models).

iv. The MFCL estimates of stock status (esp. SB/SB_{MSY}) are likely to be sensitive to the assumed maturity OGIVE and the assumption of sex invariant population dynamics. As noted above, further development of a sex specific model is warranted.

v. Despite differences in the estimates of initial fishing mortality from the MFCL and SS models, the estimates of current stock status were relatively insensitive to the starting conditions.

vi. MSY based stock status was derived based on the recent exploitation pattern (F-at-age). The age specific exploitation pattern (from MFCL) has been relatively constant over the last five years so estimates of MSY are likely to be relatively insensitive to the period selected to determine the F-at-age matrix.

vii. For the ASPIC models, a more thorough evaluation of the model diagnostics would have been informative to assess the reliability of the yield estimates; for example, residual plots for the fit of observed catches to the estimated yield curve. As noted above, estimates of yield (and associated reference points) from the ASPIC models could potentially be biased by sustained periods of non-equilibrium recruitment during the model period.

4. Evaluate the adequacy, appropriateness, and application of the methods used to evaluate future population status, given the commissions objectives.

i. For the northern stock, no stock projections were undertaken using the MFCL and SS platforms. This decision was based on the inability of MFCL to incorporate the full model uncertainty into the stock projections. Thus, the projections would not adequately represent the probability associated with future stock status bench marks. The SS model was not sufficiently advanced to be adopted as the primary assessment model. SS could be applied to conduct stochastic stock projections and estimate uncertainty using a MCMC approach, however, this is likely to be a time consuming approach (probably requiring 1-2 days to attain a sufficiently large MCMC sample) and, hence, not very appropriate within the meeting time frame.

ii. A pragmatic decision was made to use ASPIC for the forward projections for the northern stock. This framework will provide reasonable advice regarding the relative performance of the various options of future catch considered in the stock projections. However, estimates of yield from the ASPIC models are generally higher than from the MFCL/SS models and, therefore, stock projections from ASPIC models are likely to be more optimistic than comparable projections undertaken using MFCL.

iii. Under the ASPIC framework, recruitment is essentially deterministic and, thus, there is no variability incorporated in recruitment in the projection period. Therefore, estimates of risk

associated with alternative fishing strategies will be under-estimated. In addition, the MFCL/SS models indicate there is considerable autocorrelation in recruitment (5-10 year cycle). To adequately determine levels of risk, similar temporal fluctuations in recruitment would (ideally) be propagated in the stock projections.

iv. The ASPIC models, for both south and north stocks, are also likely to under estimate the uncertainty in the projections due to the structural assumptions of the models, especially the constraint on the initial stock conditions (B1/K).

v. Considerable progress was made towards the formulation of a Harvest Control Rule for northern albacore during the meeting. A range of HCRs were evaluated during the projection phase for the ASPIC models. The performance of these HCRs, measured by the probability of rebuild (green quadrant), should be qualified by the generic issues relating to the ASPIC projections highlighted in the previous paragraphs. An evaluation of the HCRs should also include the probability of the stock falling below the limit reference point (interim level 0.4 Bmsy). The productivity of the stock at low biomass levels has not (cannot) been evaluated and candidate HCRs that minimise the risk of the stock breaching or approaching this level should be given preference over other candidates.

vi. For the southern albacore assessments, estimates of future stock status are highly uncertain due to the incorporation of the results from a broad range of model options (see paragraph 2xi).

5. Evaluate the adequacy, appropriateness, and application of methods used to characterize the uncertainty in estimated parameters. Comment on whether the implications of uncertainty in technical conclusions are clearly stated.

i. The meeting applied a range of approaches to estimate the statistical uncertainty of the assessment models (bootstrap, likelihood profiles, delta method), although the estimation of uncertainty of key derived parameters for the complex models (MFCL/SS) is time consuming and not well suited to the meeting environment. The approaches used to estimate uncertainty from the MFCL model(s) could not be applied to directly estimate the joint probability distributions for the indicators of current stock status (B/Bmsy and F/Fmsy).

ii. Statistical estimates of uncertainty quantify the uncertainty associated with a specific model formulation but do not reflect the structural uncertainty of the model(s). A range of plausible model options were developed, although in most cases, these options were simple one-change model sensitivities. The resulting suite of model options is considered unlikely to adequately represent the overall level of structural uncertainty of the assessment models. For the complex models, the estimation of structural uncertainty is not a trivial exercise and cannot be undertaken in the framework of the assessment meeting.

iii. For MFCL/SS and the ASPIC models, some additional analysis of the uncertainty associated with key model parameters is likely to be informative. For example, likelihood profiles of initial stock conditions (initial F, B1/K) and equilibrium recruitment are likely to be informative regarding the information content of various data sets. This information was presented for some of the model options but is particularly useful for understanding the interactions between the various data sets included within the more complex models. Again, this sort of detailed analysis cannot be undertaken during the framework of the assessment meeting.

iv. For the southern albacore model options, there was a thorough analysis of the uncertainty associated with the key model parameters (priors and fixed parameters) and the individual data sets. The main sources of uncertainty were encompassed in the selection of the model options used for the final assessment advice. For the ASPIC models, the constraints on key parameters (i.e. B1/K) resulted in considerably lower estimates of model uncertainty compared to the BSP model that incorporates uncertainty associated with the initial conditions.

6. Comment on whether the stock assessment results are clearly and accurately presented in the detailed report of the Stock Assessment.

i. Overall, the results of the individual assessments are well documented, particularly for the simpler production models. The more complex models are difficult to adequately document in the meeting report and more detailed papers are generally required to support these analyses. The final agreed MFCL model(s) differed from the model options presented in the preliminary SCRS document (58). It is recommended that this document be updated to fully document the final model results. This document should include the detailed model configuration, including the code used to set all the model options during the fitting procedure (doitall file). Considerable development was made in the implementation of the SS models for northern albacore. While not used in final management advice, the model results are useful, particularly in relation to the development of a sex specific assessment model. However, the documentation of the SS model is limited to the summary provided in the final report. A more comprehensive supporting document is warranted with particular focus on the elements unique to the SS platform.

ii. In summarising the assessment advice in the Executive Summary, it would be informative to identify and summarise the trends in the key model inputs and link those trends more directly to the current stock status. For the northern assessment, this would focus on the trend in the C-T longline and the troll CPUE indices; it is more difficult to succinctly summarise the inputs to the southern assessment but it would be informative to emphasise that much of the uncertainty in the assessment resulted from the influence of specific CPUE indices. Recent trends in stock status should be explicitly linked to recent catches to highlight that the improvement in stock status is attributable to a decline in the level of fishery catch (rather than an increase in recruitment).

iii. The Executive Summary outlook for the northern albacore is really quite optimistic, although I think this section should include a number of qualifiers. I am mainly concerned with the application of the ASPIC model for stock projections when the model dynamics differ considerably from the MFCL model, particularly relating to recruitment. The yield estimates from MFCL are lower than the ASPIC model and the MFCL model reveals a temporal trend in recruitment, with lower recruitment over the last 20 years. Consequently, model projections from the MFCL model may well be less optimistic and, with the inclusion of recruitment variability, considerably more uncertain.

iv. The summary of southern albacore projections would be more informative if it highlighted the contrasting trends in the catch limit stock projections for the different model options. The uncertainty in the southern assessment indicates that a more robust management approach is required to deal with the uncertainty (rather than a constant catch limit of F strategy). Some additional analysis using a HCR similar to the range of HCRs proposed for the northern stock would have been informative.

7. Comment on potential improvements on the stock assessment SCRS process (CPC participation, transparency, objectivity, documentation, uncertainty characterization, etc.) as applied to the reviewed assessments.

i. The data preparatory and assessment meetings were well supported by ICCAT secretariat staff and the range of participants from CPCs provided the technical skills required to conduct and review the various data sets and results of the stock assessment modelling. Both meetings were professionally conducted under the capable chairmanship of Dr Haritz Arrizabalaga (AZTI). The level of cooperation and collegial relationship among the meeting participants enabled the completion of the (ambitious) work plan and agenda at the assessment meeting.

ii. Participation at the data preparatory meeting included representatives from the main CPCs involved in the North Atlantic albacore fishery (Chinese Taipei, Spain, Ireland, Japan and USA) with the exception of France. The representation of the CPCs in the South Atlantic fishery included Chinese Taipei, Japan, Spain and Uruguay but did not include three of the

main fishing nations, in terms of recent annual catches (South Africa, Namibia and Brazil). The assessment meeting was well attended by the main CPCs, with the exceptions of South Africa and Namibia. There was a sufficient number of participants at the latter meeting to critically review the range of assessment models presented to the group and formulate the relevant management advice.

iii. A range of key data issues were identified during the preparatory meeting and the subsequent assessment meeting. Unfortunately, the meeting participants from the respective CPCs were not able to address the specific issues associated with these data during the two meetings or in the intervening period.

iv. The group has made considerable progress in the application of the relatively complex modelling approaches (MFCL/SS) utilised by many of the other tuna RFMOs to the assessment of the northern albacore stock. However, considerable resources are required to develop and support these more complex assessments. These models cannot be adequately developed during a short assessment meeting as they require considerable time to configure the data sets, conduct routine model testing, undertake a range of model sensitivities, evaluate model performance (goodness of fit criteria, etc) and estimate of model uncertainty. Experience from other tuna RFMOs indicates that the initial development of these assessment models typically takes an experienced analyst approximately 3-6 months. More routine updates of the assessment model probably require about one month of work in advance of an assessment meeting.

v. The 2013 assessment process was sufficient to update and maintain the previous MFCL assessment model, although there was not sufficient time or resources available to significantly progress the range of issues that require further analysis and investigation. It is unreasonable to expect that substantial improvements/developments in these complex models can be achieved in the framework of the assessment meeting without dedicated work being conducted inter-sessionally by a core group, including a lead analyst. This would enable a range of model options to be formulated and tested in advance of the meeting. The meeting could then focus on reviewing the range of model options and conducting a limited number of model projections based on the preferred candidate model(s).

vi. The development and testing of the MFCL/SS models during the assessment meeting meant that there was limited time to fully evaluate the current status quo models. This also limited the time available in the latter part of the meeting to undertake and review the range of associated model based analyses (estimates of uncertainty and stock projections) and reduced the time available for the formulation of the final stock status summaries.

vii. The current assessment approach, utilising both complex models and production models to assess a single stock, creates a large work load for the assessment group as the range of model. While the corroboration of the results from multiple model platforms may lead to increased confidence in the final assessment results, the development of multiple models reduces the time available for consideration of the detailed assessment results from the primary model option(s).

viii. During the current assessment there were some problems in the formulation of various model data sets (MFCL and VPA) by the ICCAT secretariat. This highlights the level of resourcing required to develop, maintain and support the more complex model platforms. In particularly, a high level of scrutiny is required in the configuration of the various data sets. These systems can be automated to some extent, although there is a considerable overhead in the development and maintenance of the necessary software.

ix. The application of the more complex model platforms means that the assessment is dependent on the expertise of a relatively small number of analysts with sufficient experience in the use of the software. The shared experience from other tuna RFMOs should be utilized either through direct collaboration at the assessment meetings or via an external review of the preliminary assessment results (preferably in advance of the assessment meeting). For the complex models, it would be useful to have draft SCRS documents available in advance of the assessment meeting. This would enable a review of the key data inputs and model assumptions and, thereby, streamline the assessment process during the meeting.

x. The group have developed a useful set of standards for the evaluation of the CPUE indices which represent the key inputs into the assessment models. This potentially provides a useful and objective framework for the weighting of the various CPUE indices and/or the selection of preferred model options. However, this approach needs to be supported by a more detailed analysis of the CPUE data from the range of fleets. A similar approach should also be developed for determining the reliability of the various CAS data sets.

8. Comment on the adequacy of the workplan for the assessment and whether it was adequately addressed by the Data or Assessment Working groups.

i. The data sets were all provided by the CPCs in advance of the Data Preparatory meeting. However, there were some technical issues that delayed the finalisation of the CAA estimates. Similarly, there were some delays in the finalisation of the MFCL input data sets.
ii. The group completed the range of assessment models specified in the work plan. In hindsight, the decision to apply three alternative modelling approaches to the northern albacore stock (and the addition of the ASPIC modelling approach), in addition to the assessment of southern albacore, placed a heavy work load on the group and insufficient time was available to thoroughly evaluate the complex models.
iii. Both the north and south assessments are highly dependent on the assumptions made regarding the various CPUE indices. Again, in hindsight, it would have been invaluable to have conducted a more detailed review of these data inputs at the Data Preparatory meeting. This was identified as an item of the work plan ("Evaluate the indices against the standards provided by the WGSAM") although such an evaluation requires a more integrated approach for comparing and contrasting the indices and there was insufficient time available during either meetings to undertake the required analyses.
iv. The candidate limit reference points and HCRs were not formally evaluated using an operating model consistent with MFCL (as specified in the work plan). Instead, the HCRs were evaluated using the projections of the range of ASPIC model scenarios. Most of the probability density of the bootstraps from the 7 scenarios was within the F<Fmsy; B<Bmsy quadrant. So the stock projections are essentially evaluating a rebuilding/recovery plan rather than evaluating the performance of the HCR for a (wide) range of stock conditions. A full evaluation of the candidate HCRs requires testing under a range of stock scenarios, particularly encompassing the range of recruitment assumptions (stochastic recruitment, auto correlation in recruitment) and harvesting assumptions (e.g. adherence to catch limits, auto correlation in fishing mortality levels).

9. Consider the research recommendations provided by the working group and suggest any additional recommendations or prioritizations warranted. Clearly denote research and monitoring needs that could improve the reliability of future assessments. Recommend an appropriate interval for the next assessment considering control rules or management strategy in effect.

i. The working group identified the primary research needs within the list of recommendations. Of the proposed research needs, the most immediate need to undertake a detailed review of the CPUE indices from the main longline fisheries. These indices are the primary input data in the assessment models and a relatively small study could result in considerable improvement in the quality of the current assessments and/or better quantification of the uncertainty associated with the assessments. Similarly, a detailed review of the various longline CAS data is also a high priority. These analyses may result in a restructuring of the input data sets for the current assessment models.
ii. The review of these data sets should be undertaken well in advance of the next assessment. A proposed timeframe would involve undertaking the work during 2013/14. The results could then be reviewed by the working group enabling detailed recommendations for the configuration of the next iterations of the MFCL/SS models. These models could be developed inter-sessionally for reporting to the working group in mid-2015. The assessment could be deferred for another year (until 2016) if there were no large changes in the perception

of the quality of various key data inputs from those applied in the current assessment, particularly if the current status quo level of catch, fishing effort and catch limit was maintained.

iii. Some progress has been made towards the development of a MSE for the northern albacore stock. The implementation of a MSE is a complex task and requires considerable technical resources from the group. However, it does provide a framework for evaluating alternative monitoring approaches for the stocks as well as testing alternative HCRs in a more thorough manner (including stochastic recruitment, autocorrelation in recruitment, consideration of the implementation error, etc). The MSE approach should be applied to determine the appropriate periodicity of future assessments under different management scenarios and the trade-offs/biases of various assessment approaches (MFCL/SS, VPA and production models).

iv. The MSE approach could be used to evaluate alternative monitoring approaches and the relative benefits of improved data collection. For example, there is potential to substantially improve the knowledge of the current status of the albacore stocks by obtaining a series of direct estimates of the annual age structure of the longline fishery. Alternatively, a tagging project may be the best way of determining the current stock size and exploitation rates. The relative merits of each programme are best tested using a MSE simulation approach, while ensuring the practical design elements of each monitoring approach are considered in sufficient detail.

v. In the south Atlantic, stock structure remains uncertain. It is highly unlikely that a single scientific approach will adequately resolve the stock relationships in this area. However, a useful starting point would be to obtain separate estimates of key biological parameters (growth, maturity, sex ratio) from the south-east Atlantic, south-west Atlantic and south-west Indian Ocean. In addition, similar age, growth, maturity information should also be routinely obtained from the main fishery area from the north-west Atlantic. These data may identify temporal and spatial differences in the growth rates of albacore similar to those recently documented for the south Pacific albacore stock.

Overall comment

The work plan for the group was ambitious, particularly the requirement to undertake assessments of the two stocks using a range of alternative modelling approaches. The implementation of the more complex models (MFCL and SS) is not well suited to the collaborative development of an assessment in the working group environment. These models require considerable development, evaluation and testing by dedicated assessment personnel. Ideally, the working group would be involved in providing direction during the development phase of the assessment and provide a thorough review of the final range of model scenarios. Further, for these complex models, the estimation of uncertainty associated with current and future stock status is complex and computationally intensive and cannot be undertaken during the meeting timeframe. The development of these complex models better represent the complexities of the tuna fisheries and the underlying stock dynamics; however, the reliability of these models is dependent on the availability of quality fisheries data. Simpler assessment techniques may be sufficiently reliable for the management of the albacore stocks; however, the simpler models should not be considered the default option until an evaluation of the potential biases of the various modelling approaches has been undertaken.

Overall, the Albacore working group conducted a reasonably comprehensive assessment of the northern and southern Atlantic albacore stocks, given the available resources. The conclusions regarding stock status are consistent with the assessment results and, particularly for the north Atlantic stock, these conclusions seem to be reasonably robust to the range of alternative assumptions and approaches investigated. For future projections for the north Atlantic stock should be considered indicative only and reflect the relative performance of the alternative HCRs considered. The current stock status of the southern Atlantic stock is considerably more uncertain and, consequently, the management advice (outlook) is much more equivocal than for the northern stock.