CIE Independent Peer Review Report

on

Stock assessment of albacore tuna, *Thunnus alalunga,*

in the North Pacific Ocean

Prepared by

Yong Chen

Professor of Fisheries Population Dynamics
School of Marine Sciences
University of Maine
Orono, ME 04469

Contents

SECTION **Page**

I. Executive Summary

Albacore tuna, *Thunnus alalunga*, support multinational and multi-gear fisheries in the Pacific Ocean. Two albacore stocks are defined in the Pacific Ocean: the north Pacific stock and the south Pacific stock. The north Pacific stock is the focus of this stock assessment which covers the time duration from 1966 through 2009 and the areas north of the equator from 10^0N to 55^0N latitude and from 120^0E to 120^0W longitude. The Albacore Working Group (ALBWG) of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean (ISC) is responsible for conducting regular stock assessment.

Stock Synthesis (SS) version 3.11b (Methot 2011) was used as a modeling platform for the current assessment to provide estimates of stock parameters, determine stock status, and develop scientific advice on fisheries management. Virtual Population Analysis (VPA) model, used in the 2006 stock assessment, was also used as a reference run to compare with the SS modeling results. Both models implicitly assume that “there is a single well mixed stock of albacore in the north Pacific Ocean”. The SS model uses season (quarter) as its time step and include sixteen fisheries. Eight abundance indices standardized using data from the logbooks were included in the SS. The growth model was estimated within the SS model. No fishery-independent data were available. The assessment results were then used to project the dynamics of future stock under different scenarios of stock productivity and fisheries regulations and quantify the management objective defined by the Northern Committee (NC) of the Western and Central Pacific Fisheries Commission (WCPFC). The ALBWG developed a base-case model for the assessment, and conducted some sensitivity analyses to evaluate impacts of different model configurations on SSB estimates. The ALBWG concludes that “the north Pacific albacore stock is healthy at the current level of fishing mortality and average historical recruitment” and recommends “maintaining present management measure.”

As a CIE reviewer, I independently evaluated this current stock assessment report with respect to a set of pre-defined Terms of Reference (ToRs). I conclude that the estimates of temporal trends of stock biomass, SSB, and recruitment are reliable, but absolute values of stock biomass estimates are subject to large uncertainty and not suitable for making management advice. *The current assessment of the albacore stock status is mainly based on the stock temporal trends, and I agree with the conclusion regarding stock status and management advice made in the stock assessment report. I conclude that overall this assessment is scientifically sound and adequately addresses management needs in determining the status of albacore stock in the North Pacific Ocean.* In particular, I would like to commend the efforts of the ALBWG to conduct multinational collaborative studies in compiling fisheries and biological data and addressing uncertainty regarding data quality and quantity in the assessment. I also in principle agree with the research recommendations outlined in the stock assessment report by the ALBWG. However, I believe some important questions have not received enough attention or have not been addressed in this assessment. These issues include a general lack of understanding of the SS model performance in quantifying the albacore stock dynamics; not considering uncertainty of all sources; need for more sensitivity analyses; need to explore more modeling options built-in in the SS such as dynamic binning to improve

model fitting; lack of a general guideline to determine data weighting scheme; and failure to explicitly define target, threshold and/or limit reference points for stock biomass and fishing mortality and relevant harvest control rules.

Accordingly, I recommend that future research be done in the following areas: (1) develop a management strategy evaluation (MSE) framework to evaluate the performance of the albacore stock assessment model and identifying key input data and model assumptions that may significantly influence the model performance; (2) conduct more studies to evaluate quality of the input data and consistency of the data from different fisheries and reduce the uncertainty in the data before they are used in modeling; (3) evaluate possible spatial and temporal variability in fish life history parameters (e.g., growth and maturation) and fisheries data (e.g., catch, catch size compositions and CPUEs) and coordinate research efforts to collect samples over a large spatial scale; (4) evaluate the coherence of CPUEs derived from different fisheries to identify factors that may influence the quality of the CPUE data and possible discrepancy among different sets of CPUE data in quantifying the overall stock biomass; (5) develop a general guideline to determine weighting scheme for each data set to consider both data precision and relative importance of the data in quantifying the albacore stock dynamics; (6) explore dynamic binning option in the SS to address issues of having too many zero observations for small and large size bins; (7) evaluate retrospective errors for the total stock biomass and recruitment because they are more likely to be subject to retrospective errors than SSB and need to use an early year (e.g., year 2000) as the reference year in the retrospective analysis; (8) explore the use of robust likelihood functions to identify outliers and then evaluate the identified outliers to determine if they should be removed (because of large measurement errors) or included (because of large process errors) in the assessment; (9) develop priors for key fishery and population parameters and apply the Bayesian estimator to better quantify uncertainty associated with modeling; (10) develop a habitat model to identify key environmental variables that regulate the spatial distribution of albacore; and (11) develop harvest control rules with target, threshold, and limit reference points being explicitly defined. More detailed recommendations and their justifications can be found in the sections of Summary of Findings and Conclusions and Recommendations.

II. Background

Albacore tuna, *Thunnus alalunga*, support multinational and multi-gear fisheries in the Pacific Ocean. Previous studies (Ramon and Bailey 1996; Takagi et al. 2001) suggest the existence of two albacore stocks in the Pacific Ocean: the north Pacific stock and the south Pacific stock separated by the equator with negligible mixing between the two stocks. The north Pacific stock is the focus of this stock assessment which covers the time period from 1966 through 2009 and the areas north of the equator from 10^0N to 55^0N latitude and from 120^0E to 120^0W longitude. The Albacore Working Group (ALBWG) of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean (ISC) is responsible for conducting regular stock assessment.

Juvenile albacore tend to grow fast, but the growth rates are reduced after reaching maturity. The estimated female length at maturity varies among studies, ranging from 83 cm (fork length) in the western Pacific (Chen et al. 2010a) to 93 cm north of Hawaii (Otsu and Uchida 1959). However, such differences may reflect spatial and temporal variability of length at maturity because the maturity studies covered a large spatial area and over a long time period (from the 1950s to recent). Von Bertalanffy growth function is usually used to quantify the growth of albacore. Because of lack of small and large fish in samples, large uncertainty exists in the estimates of size at young and old ages. A recent study suggests that size at age and longevity may differ between males and females (Chen et al. 2011). Understanding of spatial and temporal variability in the albacore growth is rather limited.

North Pacific albacore experience extensive migratory movement influenced by oceanic variables (Polovina et al. 2001; Zainuddin et al. 2008). The migrating population, inhabiting surface waters (< 50 m), mostly consists of juvenile fish younger than 5 years old and smaller than 85 cm FL. Westward migrations are more frequent than eastward movements in the North Pacific Ocean, which may contribute to the fisheries recruitment in the western and eastern Pacific Ocean. Maturing juveniles and mature fish tend to move to low latitude spawning grounds in the western and central Pacific Ocean. Spatial dynamics of albacore distribution can greatly influence the spatial distribution of fishing grounds for fisheries targeting different sizes of albacore.

Diets of albacore change with size, shifting from microplankton to macroplankton, fish and squids (Pusineri et al. 2005; Consoli et al. 2008). Diet compositions change with locations, time (day versus night), and seasons (Pusineri et al. 2005), which suggests possible variations in growth among locations and seasons.

Sixteen fisheries were defined in this assessment based on gear, spatial coverage, season, and measure of catch (i.e., number or weight) to minimize temporal variability in selectivity and catchability. These fisheries include: one troll (Canada/USA), two pole-line (Japan offshore and distant water), nine longline (USA, Japan coastal, offshore, distant water; Chinese Taipei offshore and distant-water; Korea and other), two gillnet (Japan, Chinese Taipei/Korea), and two miscellaneous fisheries. These fisheries vary greatly in their spatial and temporal coverage, selectivity and catchability; and differ in nature (albacore are targeted or bycatch species). Data were compiled by seasons (Jan-March, Apr-Jun, Jul-Sep, and Oct-Dec) for these fisheries from 1966 through 2009, including catches, length compositions estimated from subsampling the catches (port or sea sampling programs), abundance indices derived from logbooks, and age-at-length data. Not all these data are available for the sixteen fisheries. Catches and length compositions differ greatly among the fisheries (ALBWG 2011). No reliable estimates of discards over the time period covered in this stock assessment were available. Thus, discards were not included in the catch.

Stock Synthesis (SS) version 3.11b (Methot 2011) was used as a modeling platform for the current assessment to provide estimates of stock parameters, determine stock status, and develop scientific advice on fisheries management (ALBWG 2011). Previous assessment of albacore stock in the North Pacific Ocean was conducted in 2006 using a Virtual Population Analysis (VPA) model, which was also used as a reference run to compare with the SS modeling results. Both models implicitly assume that "there is a single well mixed stock of albacore in the north Pacific Ocean". The SS model uses season (quarter) as its time step and include sixteen fisheries. Eight abundance indices standardized using data from the logbooks were included in the SS. The growth model was estimated within the SS model. No fishery-independent data were available. The ALBWG developed a base-case model for the assessment based on extensive discussions, and conducted some sensitivity analyses to evaluate impacts of different model configurations on SSB estimates. Retrospective analysis was conducted for year 2005.

The ALBWG was asked to evaluate the status of the north Pacific albacore stock against $F_{SSB\text{-}ATHL\ 50\%}$, which is the fishing mortality that would lead to future lowest SSB having 50% chance to fall below SSB-ATHL threshold level for a 25 year projection period. The SSB-ATHL is an interim management objective to "maintain SSB of north Pacific albacore above the average of the ten historically lowest estimated points (ATHL) with a probability greater than 50%", defined in 2008 by the Northern Committee (NC 2008). In the stock assessment, $F_{SSB\text{-}ATHL\ 50\%}$ and SSB-ATHL were essentially used as limit F and biomass reference points, respectively, to define whether overfishing occurs and fish stock is overfished. Target F and biomass reference points were not explicitly defined in the assessment, and no harvest control rules were developed. No management strategy evaluation (MSE) was conducted for this tuna stock.

III. Description of the Individual Reviewer's Role in the Review Activities

As the SoW states that "…reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein", my role as a CIE independent reviewer is to conduct an impartial and independent peer review of stock assessment of Albacore tuna in the North Pacific Ocean which are fished by multiple nations with multiple gears in different regions, with respect to the pre-defined Terms of Reference.

This is a desk review. Thus, I had no opportunity for face-to-face discussion and questioning. I read the "Stock assessment of albacore tuna in the North Pacific Ocean in 2011" by the Albacore Working Group of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean and all other background documents that were sent to me (see the list in the Appendix II). I also read references relevant to the topics covered in the reports and the SoW. Based on the reading, I address each topic covered in the ToRs, evaluate the strengths and weaknesses of what was done in this assessment, and provide recommendations to improve future assessment. Based on these evaluations and analyses, I make research recommendations for future assessment of Albacore tuna in the North Pacific Ocean.

IV: Summary of Findings

1. Review the assessment methods: determine if they are reliable, properly applied, and adequate and appropriate for the species, fisheries, and available data.

   Overall, I conclude that the SS employed in this stock assessment provides an adequate and reliable modeling framework and is properly implemented for the assessment of albacore stock in the North Pacific Ocean, given the data available and complexity of the fisheries. However, I believe more work is needed to improve our understanding of the model performance in assessing the dynamics of albacore stock in the North Pacific Ocean.

   Stock Synthesis used in this assessment is an age- and size-structured model designed to accommodate both age- and size-structured data. One of the greatest strengths of SS is its flexibility to utilize a wide diversity of age-, size-structured, and/or aggregate data from fishery-dependent and fishery-independent monitoring programs. Because of its flexibility, SS can be used to develop simple to complex stock assessment models depending upon the data available. Using the SS framework, the ALBWG developed a seasonal, length-based, age-structured forward population projection model to predict fishery data which were then compared with observed data to formulate likelihood functions for the parameter estimation.

   Overall, I think the SS is a good choice for the assessment of the albacore stock in the North Pacific Ocean with data from such a diversity of fisheries. The stock assessment methods developed with the SS allow the modelers to incorporate data from fisheries with different gears and spatial and seasonal coverage and consider temporal changes in selectivity. The use of quarter as time step is consistent with seasonality of some fisheries. The estimation of growth model internally within the SS model allows for a better model fit and results in more realistic estimates of growth parameters. The use of standardized CPUE data removes the impacts of factors other than stock biomass. The use of projection feature built-in in the SS allows the modelers to develop simulation-based biological reference points, which depends on temporal trends of stock biomass rather than its absolute values, to quantify the management objective defined by the Northern Committee. The models developed in this study also allow the modelers to evaluate and quantify potential impacts of different fisheries on the dynamics of the albacore stock in the North Pacific Ocean. *The base-case assessment scenario developed by the ALBWG appears to represent the best knowledge we have had with respect to the fisheries and stock dynamics of albacore in the North Pacific Ocean.*

Having stated the strengths above, I would like to raise some concerns.

The weighting factors for some of the likelihood components were determined rather arbitrarily, and there is a need to develop some guidelines/principles for determining weighting factors for different data sources. For example, we may determine weighting factors of CPUEs of different fisheries based on their relative differences in catch, spatial/temporal coverage, and/or nature of the fisheries (i.e., target/bycatch). For size composition data, it is important to consider effective sample sizes, but I believe it is also important to consider the spatial/temporal coverage of the fisheries from which the size composition data were collected. A size composition data set can be very accurately estimated (large effective sample size), but may be collected from a small area over a short time period, which may not represent the overall size composition of the whole stock (limited spatial/temporal coverage of the fishery). I believe the guidelines/principles should be developed for weighting prior to a stock assessment so that weighting factors for different data sets can be determined consistently following these guidelines/principles to reflect both quality and relative importance of a data set.

There is a general lack of understanding of performance of the SS model in quantifying dynamics of albacore stock in the North Pacific Ocean which is subject to multinational fisheries with multiple gears and different spatial/temporal coverage. I would strongly suggest developing a management strategy evaluation (MSE) framework to evaluate the performance of the SS model. A well-defined simulation study can identify factors that may be critical in influencing the performance of the SS model. Although the estimates of temporal trends are rather robust in the assessment, sensitivity of estimating absolute stock biomass concerns me. Given how sensitive the modeling results are regarding different model configurations, I believe that absolute values of stock biomass estimates are not reliable for management advice. However, temporal trends of stock biomass estimates are robust, and are suitable for management advice. I think an extensive simulation study (Chen et al. 2005; Kanaiwa et al. 2008) is needed to improve our understanding of the model performance in assessing the albacore stock in the North Pacific Ocean.

2. <u>Evaluate the assessment model configuration, assumptions, and input 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.</u>

   I would like to commend the ALBWG to compile quarterly catch data of different fisheries, derive standardized CPUEs and assemble relevant biological and fisheries parameters for the assessment of multinational fisheries. This is a huge undertaking. Overall I conclude that the models are adequately configured, data are properly used, and assumptions are reasonably satisfied. However, I believe there is room for *improvement in accounting for sources of uncertainty and in selecting model parameters.*

For the stock assessment model of this complexity, there usually are a lot of assumptions with respect to model structure and statistical distribution of differences between the observed and model predicted data. Although some assumptions are listed in the report, no all the assumptions are spelled out. I suggest that ALBWG lists all model assumptions, explicit and implicit, in a table and describes if a particular model assumption is satisfied for a given assessment scenario. Potential consequence of violating these assumptions should also be described in the table.

Albacore fisheries cover a large spatial area with different gears and capture different sizes of fish for different reasons (target versus bycatch). It is highly likely that part of the catch is discarded at sea. Because of temporal changes in marketing and fishing practice, size composition and amount of discarded catch may vary over time. No estimates were derived for the discards. Discards were not included in this assessment. The model implicitly assumes that the catch is subject to no errors. Impacts of violating this assumption on the assessment are not evaluated and unknown.

Size composition data are available for the eight fisheries, and size ranges vary greatly among these fisheries because of their differences in fishing gear, grounds and seasons. The same size range from 40 to 140 cm FL was used to group size composition data. This results in a large number of zeros for small or/and large size bins, which may greatly affect the model fitting. I suggest that the dynamic binning option in the SS or robust multinomial likelihood functions (Fournier 1996; Chen et al. 2000) be used to reduce impacts of zero-observation in size composition data on the model fitting.

This assessment is based on maximum likelihood estimators. I suggest that the Bayesian estimators be used to incorporate uncertainty of all sources.

Outliers may exist in the stock assessment given multiple data sources and a large number of factors that may influence the data quality but cannot be controlled in sampling (Chen and Fournier 1999). Because both multinomial and normal or log-normal-based likelihood functions tend to be sensitive to outliers, the results may be biased by outliers. I suggest that the robust likelihood functions (e.g., Fournier 1996; Chen et al. 2000) be explored in the parameter estimation to identify outliers in data and evaluate possible causes to determine if the identified outliers should be removed in the assessment.

Estimating growth parameters within the assessment model provides the model with great flexibility in fitting size composition data. It appears that such an approach yields more biologically realistic estimates of growth parameters. However, it is suggested that size at maturity may vary spatially (Ueyanagi 1957; Otsu and Uchida 1959; Chen et al. 2010a) and possibly temporally, and maturation is likely to affect growth. This implies that growth may vary greatly spatially and temporally. This was not considered in the assessment. I suggest that impacts of spatial and temporal variability of growth on the assessment be evaluated in the MSE. I understand that it may take some time before this can be done. However, I think it is important to conduct such a study because of likely large variability in growth over time and space and sensitivity of the assessment results to the growth data. This also calls for more studies to evaluate and better quantify spatial and temporal variability in key life history parameters for albacore in the North Pacific Ocean.

Sexual dimorphism in size at age and longevity was suggested for the albacore by Chen et al. (2011). This implies differences in growth, maturation and natural mortality between females and males. More studies need to be done to test this hypothesis. If there is more evidence of sexual dimorphism, a sex-specific stock assessment model needs to be considered.

Choices of selectivity curves and time blocks assumed for different fisheries appear to be justifiable. These selectivity curves include both fish availability and gear selectivity. However, I did not see much discussion about catchability. For some fishing gears (e.g., longline and gillnet), gear saturation may be an issue. Whether CPUE is a good abundance indicator may be questionable even if it is standardized. A saturation parameter may be needed when CPUE is related to the stock abundance/biomass. Some fisheries have also experienced large changes in fishing practice as a result of changes in target species and management regulations, which may result in large changes in catchability over time. I understand changes in catchability may be reflected in changes in selectivity, but believe their impacts should be evaluated separately.

3. <u>Evaluate the adequacy of the sensitivity analyses in regard to completeness and incorporation of results.</u>

   The ALBWG evaluates impacts of some alternative model configuration and parameterization on the assessment in the sensitivity analysis. The analysis appears to be adequate in demonstrating sensitivity of assessment results with respect to alternative model settings, but I believe more systematic sensitivity analyses should be done to further improve our understanding of robustness of the stock assessment.

   Sensitivity analysis was conducted to evaluate robustness of the modeling results with respect to alternative values for (1) weighting factors of likelihood functions; (2) biological parameters; and (3) fishery parameters (i.e., selectivity in this assessment). Most sensitivity runs were conducted with just one parameter being given alternative values while other parameters were held same as those in the base case. Thus, the sensitivity analysis was done to evaluate impacts of a single factor on the assessment. Such a design is important in understanding of roles of each factor. However, limited efforts to change more than one factor at the same time may result in lack of understanding of interactions of these factors. For example, steepness h and natural mortality M are usually negatively correlated, and may not be estimated independently in the assessment. Likelihood profiles should be evaluated for a range of values for these two parameters to identify their relationship and to decide if they can be determined independently.

Current weighting factors only consider CV of CPUEs and effective sample sizes for catch size composition data, which only covers the precision of the data. However, CPUEs and catch size composition data were derived from different fisheries which have vastly different spatial/temporal coverage, and thus only represent the dynamics of partial stock. Such differences in the CPUEs and catch size composition data among different fisheries should be considered in weighting.

I did not see much description about the use of penalty functions, which are usually applied to constrain annual recruitment deviations and prevent the model from yielding biologically unrealistic values for model parameters. I think this issue should be clarified and importance of these penalty functions should be evaluated in the sensitivity analysis (Methot and Taylor 2011).

Sensitivity analysis was conducted for the estimation of reference points $F_{SSB\text{-}ATHL\text{-}50\%}$. In particular, I think running the analysis for two levels of recruitment dynamics is informative. However, I did not see such a sensitivity analysis was run with respect to alternative values for natural mortality, which may have greater impacts on the estimation of $F_{SSB\text{-}ATHL\text{-}50\%}$ compared with all other parameters considered in the analysis. I suggest that a sensitivity analysis be done with a combination of different values of h and M.

Retrospective analysis was conducted for year 2005 with the assessment done with the removal of the last four to one years of data from the assessment in sequence. No clear retrospective patterns were observed in SSB, suggesting that no directional biases as a result of retrospective errors (Mohn 1999). However, with such data set, retrospective analysis for the last four years may not yield stable results. I suggest that the retrospective analysis be done for year 2000 or 2001 with more years of data being removed to yield more stable and clearer results on retrospective errors associated with the SSB. I also suggest that the retrospective errors be carefully evaluated for recruitment and total biomass because these are two stock parameters that are prone to retrospective errors. I also suggest that Mohn's rho (Mohn 1993, 1999) be calculated to quantify the nature and magnitude of retrospective errors.

4. <u>Evaluate the adequacy, appropriateness, and application of the methods used to project future population status.</u>

Overall, I conclude that the methods used to project future population status are adequate and appropriate. However, I believe there is room for further improvement.

A projection model was developed to assess the impacts of current F on the dynamics of future stock. The probability that future SSB will fall below a threshold defined as "the average of the ten historically lowest SSB estimates (SSB-ATHL) in at least one year of a 25-yr (2010-2035)" was estimated. Given the sensitivity of absolute biomass estimates with respect to alternative model configurations, I believe this simulation-based reference point which is derived relative to the historical stock dynamics is a good choice and robust to the uncertainty associated with the stock dynamics. Avoidance of using most recent years of recruitment estimates in the projection because of possible retrospective errors is also a good choice to reduce errors in the projection.

There are no clearly and explicitly defined target, threshold, and limit reference points estimated in the projection model. I think the threshold reference points derived in this assessment is actually limited reference points. No harvest control rules are defined. Thus, the risk analysis of alternative management strategies included in the assessment is rather incomplete. *It seems that* $F_{SSB\text{-}ATHL\text{-}50\%}$ *and SSB-ATHL-50% were used as target reference points by the management (Northern Committee 2008), but as limit reference points in the assessment. Such a conflict should be addressed and different reference points and associated harvest control rules should be clearly and explicitly defined.*

Although considering two levels of recruitment dynamics in estimating $F_{SSB\text{-}ATHL\text{-}50\%}$ is good, both do not consider potential impacts of SSB on the recruitment. Although I understand SSB-recruitment relationship is hard to define here, their independency implicitly assumed in the projection implies that we really don't need to care about SSB, which contradicts to the logic of conservation and management proposed in the assessment. I would suggest that estimated SSB be grouped and frequency distribution of the corresponding recruitment be estimated. Recruitment of a given year in the projection can be randomly sampled based on the SSB of the year and recruitment frequency distribution.

The current assessment uses maximum likelihood estimator and does not incorporate uncertainty from all the sources. I suggest that Bayesian estimators be used in the assessment to incorporate uncertainty of all the sources. The posterior distributions derived from the Bayesian estimators can then be used in the projection. Such an approach can incorporate uncertainty of all the sources in the projection.

Random sampling of recruitment from historical estimates for the projection implicitly assumes that the recruitment dynamics behave in the same way as in the past, which may or may not be realistic given the length of time duration for the projection. I suggest evaluating how environmental variables may affect the distribution of albacore and developing habitat models to link environmental variables with albacore distribution and recruitment. Such habitat models can be used to project possible changes in albacore distributions and recruitment as a result of changes in the ecosystem (e.g., climate changes), which can be linked with the projection model.

5. Suggest research priorities to improve our understanding of essential population and fishery dynamics and improvements to the assessment model.

I believe the ALBWG has presented a rather comprehensive list of research recommendations on life history process, spatial structure, data quality and quantity, and model improvement. I support these research recommendations. I suggest that following studies (some of them may overlap with what are proposed in the report) be done to further improve albacore stock assessment and our understanding of population and fishery dynamics. Some of the research priorities are for long term, and some are for the short term (e.g., next assessment):

- Develop MSE framework based on our understanding and available information of albacore biology and fishery and use the MSE framework to evaluate the performance of the SS model used in the albacore stock assessment;

- Evaluate quality of catch data and conduct a sensitivity analysis to evaluate potential impacts of under-estimating landing data on the assessment as a result of at-sea discards;

- Need to produce a correlation matrix to evaluate correlations between every pair of CPUE data sets to evaluate their consistency in indexing the stock biomass and to justify the inclusion of each set of CPUE in the assessment;

- Need to evaluate correlations between each set of CPUE versus stock biomass estimates to identify possible gear saturation effects;

- Use dynamic binning approach to addressing issues of having large number of zeros in small and large size bins to improve model fitting;

- Need to evaluate retrospective errors for the total stock biomass and recruitment because they are more likely to be subject to retrospective errors than SSB and need to use an early year (e.g., year 2000) as the reference year in the retrospective analysis;

- Need to include the total stock biomass in the sensitivity analyses because juvenile albacore abundance tends to have large estimation errors and are important to some surface fisheries;

- Evaluate spatial and temporal variability in growth and maturation and if the information currently available is not sufficient for such a study new research program should be developed to improve our understanding of spatial and temporal variability in these key life history parameters;

- Need to evaluate spatial and temporal variability in fisheries data to identify the dynamics of spatial structure of fish size composition and CPUE to improve understanding of fisheries data quality and quantity and factors influencing them, which should be considered in determining data weighting factors;

- Evaluate the coherence of CPUEs derived from different fisheries to identify factors that may influence the quality of the CPUE data and possible discrepancy among different sets of CPUE data in quantifying the overall stock biomass, which can help determine relative weighting factors in the objective function;

- Evaluate possible differences in key life history parameters between females and males to determine if it is necessary for a sex-specific stock assessment;

- Explore the use of robust likelihood functions to identify outliers and then evaluate the identified outliers to determine if they should be removed (because of large measurement errors) or included (because of large process errors) in the assessment;

- Develop guideline/principle for determining weighting factors for likelihood functions of different data sets, which should consider both data precision and relative importance in representing the dynamics of the whole stock;

- Need to consider impacts of uncertainty in natural mortality on the estimation of $F_{SSB\text{-}ATHL\text{-}50\%}$ and SSB-ATHL-50%;

- Need to consider possible impacts of SSB on the recruitment in the projection;

- Need to develop a habitat model to identify key environmental variables influencing the distribution of albacore;

- Develop harvest control rules with target, threshold, and limit reference points being explicitly defined; and

- Develop priors for key fishery and population parameters and apply the Bayesian estimator to better quantify uncertainty associated with modeling;

V. Conclusions and Recommendations

Overall I believe this stock assessment provides rather robust assessment results on temporal trends for the albacore with respect to various uncertainties in data and models. *The current assessment of the albacore stock status is mainly based on the stock temporal trends, and I agree with the conclusion regarding stock status and management advice made in the stock assessment report. The assessment appears to be scientifically sound and adequately addresses management requirements.* In particular, I would like to commend the efforts of the ALBWG in addressing data quality issues and uncertainty in the assessment, in exploring alternative model configurations, and in evaluating impacts of different fisheries on the stock dynamics. However, I believe some important questions still need to be addressed and there is still room for improving the current stock assessment. I have made the following general comments and specific recommendations. Some may overlap with what I have described in the section of Summary of Findings.

General comments

In-depth analysis should be conducted to identify possible sources of uncertainty for a given set of data and relevant analysis should be done to reduce the uncertainty and improve data quality BEFORE the data are used in the stock assessment model. Trying to resolve all uncertainties within the SS model may complicate parameter estimation, resulting in difficulty in the model converging.

Given the flexibility and many choices that SS provides for functions quantifying life history and fishery processes, one needs to use background information and knowledge about fishery data collection, fish life history theory and local ecosystem to develop hypotheses to explain choices and resultant estimates. If a result cannot be justified in a reasonable way, the assessment should be evaluated (e.g., identify why scale of sock biomass changed so greatly when the model was re-configured in the sensitivity analysis).

It appears that assessment efforts were focused on accommodating requests for different model configurations. I believe more effort should be spent on model diagnoses to identify if the model assumptions, implicit and explicit, have been violated. This involves evaluating residual patterns for distributional assumptions, CVs of each estimated parameters to identify if an estimated parameter is significant, and the variance-covariance matrix to identify possible correlations between different parameters (and then to see if such a correlation can be justified biologically). For some parameters that have been shown to have high correlations in other stock assessments (e.g., steepness h and natural mortality M), likelihood profiles should be evaluated for their different combinations to identify if they can be determined independently in modeling.

Given the relative small number of samples, I suggest back-calculating length-at-age data using otoliths to derive length at each age for each fish with its respective otolith sample. A nonlinear random effects model explicitly assuming that an individual's growth parameters are samples from a multivariate distribution can then be applied to the back-calculated length at age data (Hart 2001; Pilling et al. 2002) to estimate between-individual and between-region variability.

Retrospective errors tend to result in large errors in estimates of current stock biomass. Current retrospective analysis is focused on the SSB. However, estimates of juvenile fish and recruitment tend to be subject to large retrospective errors, which may not be reflected in the retrospective errors of SSB. Some surface fisheries target juvenile albacore. Thus, I believe that retrospective errors should be quantified (using Mohn's rho; Mohn 1993, 1999) for the total biomass, SSB, and recruitment. In order to better quantify retrospective errors, I suggest using year 2000 or 2001 as reference year for the retrospective analysis.

The recruitment is currently measured as the number of age 0 fish in the assessment. I understand the number of age 0 fish is simply a reflection (discounted for natural mortality) of the number of fish in older ages (say 3) because there is no fishing mortality. However, given that age 0 implies larval stage and that there are no observations in survey and fishery, the biological meaning of the so-called recruitment is inappropriate and not well-defined. As it is defined, the current recruitment is neither representative of fishery recruitment nor the number of fish larvae. Rather, it is an index of the recruitment. Although this may not be an issue to fisheries stock assessment scientists, such a measure of recruitment may be mis-used by others who are not familiar with the stock assessment. I believe it is more appropriate to measure the fishery recruitment as the number of fish at an age group at which fish are subject to fishing mortality (e.g., number of fish at age 3).

The current assessment develops a base case scenario and then conducts sensitivity analyses to evaluate changes in the assessment results resulting from using alternative sets of parameters or data. This approach basically assumes that the base case is the most plausible model before the assessment modeling. Although this approach is commonly used in many stock assessments, I would like to propose a different approach which starts with the compilations of several competing models. Qualitative and quantitative criteria need to be defined for comparing the model performance to determine which model is most plausible after conducting model runs.

Outliers are likely to exist in input data used in the assessment, given that the data are derived from different sources and are subject to different levels of errors. They may bias parameter estimation in stock assessment. Robust likelihood functions can reduce impacts of outliers in size composition and survey abundance index (Chen et al. 2003). These functions also allow the modelers to identify outliers and evaluate them with the knowledge of fish biology and data quality.

Although SS is very flexible and has been tested and used in the assessment of many fisheries stocks, the results derived still need to be cross-validated to enhance the confidence in the assessment. It is good to see that ALBWG includes VPA model in the assessment, which yields similar temporal trends of stock biomass. However, the VPA results are not comparable with those from the SS in this study because they used different growth information. I recommend developing a MSE framework to evaluate the performance of the SS in quantifying the albacore stock dynamics. Although the SS has projection capacity, it has no built-in component for MSE. I believe research efforts to develop an MSE framework for the albacore can provide an important analytical tool to evaluate alternative management strategies and their associated risks.

A Bayesian approach has not been incorporated in the stock assessment. Thus, uncertainty in the assessment has not been fully incorporated in the assessment and stock projection under different harvest strategies. I would encourage future assessment to develop priors for key SS model parameters and utilize the Bayesian estimators in the SS.

The definition of biological reference points for quantifying the management objective defined by the Northern Committee is confusing. There is no clear and explicit definition about target, threshold and limit reference points, which makes the determination of stock and fishery status ambiguous. Harvest control rules are not defined, which makes the management advice unclear.

Specific recommendations

Although I have provided comments and recommendations under each TOR, I would like to re-iterate the following recommendations:

- There is a general lack of understanding of the performance of the SS model in assessing the albacore fishery in the North Pacific Ocean. I propose to develop an MSE framework based on our understanding and available information of albacore biology and fishery and then use the MSE framework to evaluate the performance of SS model used in the albacore stock assessment. Such a study can also allow us to evaluate and identify key variables and assumptions that are critical in influencing the quality of stock assessment results.

- Given the importance of the catch data in the assessment, I suggest conducting an extensive computer simulation study based on the data collected in the past to evaluate the effectiveness of the current sampling/reporting system in yielding catch estimates, to evaluate potential error sources and levels of catch estimates, and to identify alternative sampling/reporting program designs. I also suggest estimating uncertainty associated with catch estimates (e.g., discards) to develop a plausible range of catch estimates, which can then be used to evaluate impact of uncertainty associated with catch estimates on stock assessment.

- Ageing errors and variations should be estimated outside the SS3 model.

- More study needs to be done to explore the dynamic binning approach to addressing issues of having many zero observations for small and large size classes.

- The choice of time block for selectivity is not always justifiable. I believe that a random walk over a defined set of years may be a better choice. Once a model is run with random-walk selectivity over a defined set of years, the temporal trend of selectivity plots needs to be examined closely to identify any temporal pattern. The identified temporal pattern can be used in the future to decide the time block for selectivity. For multiple fleets, I suggest evaluating one fleet at a time for its temporal trend while holding others constant.

- I suggest that more effort be put towards model diagnosis and residual analysis.

- I suggest evaluating spatial and temporal variability in growth and maturation. If the information currently available is not sufficient for such a study, new research program should be developed to improve our understanding of spatial and temporal variability in these key life history parameters.

- I suggest evaluating spatial and temporal variability in fisheries data to identify the dynamics of spatial structure of fish size composition and CPUE to improve understanding of fisheries data quality and quantity and factors influencing them.

- I suggest evaluating the coherence of CPUEs derived from different fisheries to identify factors that may influence the quality of the CPUE data and possible discrepancy among different sets of CPUE data in quantifying the overall stock biomass, which can help determine relative weighting factors in the objective function.

- I propose to evaluate possible differences in key life history parameters between females and males to determine if it is necessary for a sex-specific stock assessment.

- I suggest back-calculating length-at-age data for each fish with its respective otolith sample and then apply a nonlinear random effects model explicitly assuming that an individual's growth parameters are samples from a multivariate distribution to the back-calculated length at age data to estimate between-individual and between-region variability.

- I propose to conduct a cross validation analysis that leaves some of the growth data out of the SS modeling for testing the growth model estimated within the SS.

- Given the quality of the data from different sources and potential errors in the data, it is likely to have outliers as a result of abnormal observational errors. It is also highly likely that outliers may arise as a result of abnormal process errors because of changes in the ecosystem over such a long time and over such a large area. I believe it is necessary to explore robust likelihood functions to identify outliers and then evaluate the identified outliers to determine if they should be removed (because of large measurement errors) or included (because of large process errors) in the assessment.

- Weighting factors for different data sets were determined rather arbitrarily and only consider the precision of the data in the current assessment. Little consideration is given to the representation and relative importance of a data set in its representation of the whole stock. I propose to develop a guideline/principle for determining weighting factors for different data sets, which should consider both precision and relative importance of a data set in representing the dynamics of the whole stock.

- I suggest developing harvest control rules with target, threshold, and limit reference points being explicitly defined and quantified.

- I propose to use Bayesian estimators for the parameter estimation to better capture uncertainty of all sources. I suggest analyzing the information collected in previous studies to define priors for key fishery and population parameters.

- Current assessment incorporates the model projection. I recommend that the performance of the projection done in the assessment be evaluated in the next assessment to evaluate their performance in achieving the management objectives.

- I suggest that the assessment model structure be kept relatively stable over time. If a new model needs to be used, it should be run in parallel to the old model to identify changes in stock assessment results resulting from changes in model configurations.

VI. References cited

Albacore Working Group (ALBWG). 2011. Stock assessment of albacore tuna in the North Pacific Ocean in 2011. Report of the albacore working group stock assessment workshop. International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean, 4-11 June 2011, Shizuoka, Japan, 424-8633.

Chen, K.-S., Crone, P.R., and Hsu, C.-C. 2010a. Reproductive biology of albacore Thunnus alalunga. J. Fish Biol. 77: 119–136. doi:10.1111/j.1095-8649.2010.02662.x

Chen, K.-S. Shimose, T., Tanabe, T., Chen, C.-Y., and Hsu, C.-C. 2011. Age and growth of Albacore, *Thunnus alalunga*, in the North Pacific Ocean. Working paper presented at the ISC Albacore Working Group Stock Assessment Workshop, 30 May-11 June 2011, Nat. Res. Inst. Far Seas Fish., Shimizu, Shizuoka, Japan. ISC/11/ALBWG/IP/01: 30 p.

Chen, Y. and Fournier, D. 1999. Impacts of atypical data on Bayesian inference and robust Bayesian approach in fisheries. Canadian Journal of Fisheries and Aquatic Sciences 56:1525-1533.

Chen, Y.,Jiao, Y., Chen, L. 2003. Developing robust frequentist and Bayesian fish stock assessment methods. FISH and FISHERIES, 2003, 4, 105-120.

Chen, Y., Kanaiwa, M., and Wilson, C., 2005. Developing and evaluating a size-structured stock assessment model for the American lobster, Homarus americanus, fishery. New Zealand Journal of Marine and Freshwater Research, 2005: 645–660.

Consoli, P. Romeo, T., Battaglia, P., Castriota, L., Esposito, V., and Andaloro, F. 2008. Feeding habits of the albacore tuna Thunnus alalunga (Perciformes, Scombridae) from central Mediterranean Sea. Mar. Biol. 155: 113–120. DOI 10.1007/s00227-008-1012-1.

Fournier, D.A. 1996. AUTODIFF: a C++ array language extension with automatic differentiation for use in nonlinear modeling and statistics. Otter Research Ltd., Nanaimo, B.C., Canada.

Hart, D. 2001. Individual-based yield-per-recruit analysis, with an application to the Atlantic sea scallop, Placopecten magellanicus. Can. J. Fish. Aquat. Sci. 58: 2351–2358.

Kanaiwa, M., Chen, Y. and Wilson, C. 2008. Evaluating a seasonal, sex-specific size-structured stock assessment model for the American lobster, *Homarus americanus*. Freshwater and Marine Research 59:41-56.

Methot, Jr., R.D. 2011. User manual for stock synthesis. Model version 3.20. January 2011. NOAA Fisheries, Seattle, WA, USA, 143 p.

Methot, R. D., and Taylor, I. G. 2011. Modeling the Variability of Recruitment in Fishery Assessment Models. Canadian Journal of Fisheries and Aquatic Sciences, in press.

Mohn, R. K. 1993. Bootstrap estimates of ADAPT parameters, their projection in risk analysis and their retrospective patterns. In Smith, S. J., Hunt, J. J., and Rivard, D., (eds.), Risk evaluation and biological reference points for fisheries management, Canadian Special Publication of Fisheries and Aquatic Sciences 120:173:184.

Mohn, R. K. 1999. The retrospective problem in sequential population analysis: An investigation using cod fishery and simulated data. ICES Journal of Marine Science, 56:473-488.

Northern Committee (NC). 2008. Fourth Regular Session, Tokyo, Japan, 9-11 September 2008, Summary Report. Western and Central Pacific Fisheries Commission, Kolonia, Pohnpei, Federated States of Micronesia, 41 p. http://www.wcpfc.int/system/files/documents/meetings/northern-committee/4th-regularsession/summary-report/NC4%20Summary%20Report%20%28Preliminary%20version%29-edited%20clean.pdf

Otsu, T., and Uchida, R.N. 1959. Sexual maturity and spawning of albacore in the Pacific Ocean. Fish. Bull. 50(148): 287-305.

Pilling, G. M., Kirkwood, G. P. and Walker, S. G. 2002. An improved method for estimating individual growth variability in fish, and the correlation between von Bertalanffy growth parameters. Can. J. Fish. Aquat. Sci. **59**: 424–432.

Polovina, J.J., Howell, E., Kobayashi, D.R., and Seki, M.P. 2001. The transition zone chlorophyll front, a dynamic global feature defining migration and forage habitat for marine resources. Prog. Oceanogr. 49: 469–483.

Pusineri, C., Vasseur, Y., Hassani, S., Meynier, L., Spitz, J., and Ridoux, V. 2005. Food and feeding ecology of juvenile albacore, Thunnus alalunga, off the Bay of Biscay: a case study. ICES J. Mar. Sci. 62: 116-122. doi:10.1016/j.icesjms.2004.09.004

Ramon, D., and Bailey, K. 1996. Spawning seasonality of albacore, Thunnus alalunga, in the South Pacific Ocean. Fish. Bull. 94: 725-733.

Takagi, M., Okamura, T., Chow, S., and Taniguchi, N. 2001. Preliminary study of albacore (Thunnus alalunga) stock differentiation inferred from microsatellite DNA analysis. Fish. Bull. 99: 697-701.

Ueyanagi, S. 1957. Spawning of the albacore in the Western Pacific. Rep. Nankai Reg. Fish. Res. Lab. 6: 113-124.

Wells, R. J.D., Kohin, S., Teo, S.L.H., Snodgrass, O.E., and Uosaki, K. 2011. Age and growth of North Pacific albacore (Thunnus alalunga). Working paper submitted to the ISC Albacore Working Group Stock Assessment Workshop, 30 May-11 June 2011, Nat. Res. Inst. Far Seas Fish., Shimizu, Shizuoka, Japan. ISC/11/ALBWG/02: 13 p.

Zainuddin, M., Saitoh, K., and Saitoh, S.-I. 2008. Albacore (Thunnus alalunga) fishing ground in relation to oceanographic conditions in the western North Pacific Ocean using remotely sensed satellite data. Fish. Oceanogr. 17: 61–73.

Appendix 1: Bibliography of materials provided for review

Albacore Working Group. 2010. Annex 4. REPORT OF THE ALBACORE WORKING GROUP WORKSHOP. International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean 12-19 October 2010, La Jolla, California, USA.

Albacore Working Group. 2006. Annex 5. REPORT OF THE ALBACORE WORKING GROUP WORKSHOP. International Scientific Committee for Tuna and Tuna-Like Species in the North Pacific Ocean (November 28 – December 5, 2006, Shimizu, Japan).

Albacore Working Group. 2011. Stock assessment of albacore tuna in the North Pacific Ocean in 2011. Report of the albacore working group stock assessment workshop. International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean, 4-11 June 2011, Shizuoka, Japan, 424-8633.

Chen, C. Y., F. C. Cheng and S. Y. Yeh. 2007. Standardized CPUE trend and age composition of North Pacific albacore exploited by Taiwanese longline fisheries, 1995-2008. ISC/10/ALBWG-2/07

Chen, C. Y., F. Cheng, and S. Yeh. 2008. Standardized CPUE and catch-at-age time series of North Pacific albacore exploited by Taiwanese longline fisheries 1995-2008. ISC/10/ALBWG-3/08.

Kiyofuji, H and Uosaki, K. 2007. Revision of standardized CPUE for albacore caught by the Japanese pole and line fisheries in the northwestern North Pacific. ISC/10/ALBWG-3/07

Kiyofuji, H., Iwata, S., Kai, M., Ichinokawa, M., Matsumoto, T., Uosaki, K. and Takeuchi, Y. 2011. Preliminary North Pacific albacore population analysis using VPA-2BOX and future projection using PRO-2BOX for 1966 – 2009. ISC/11/ALBWG/12

Matsumoto, T. 2004. Abundance indices of north Pacific albacore by Japanese longline for SS3 and VPA analyses. ISC/10/ALBWG-3/04

Teo, S. L. H, H. H. Lee, and S. Kohin. 2002. Time series associated with albacore fisheries based in the Northeast Pacific Ocean. ISC/10-3/ALBWG/02

Well, R. J. D., S. Kohin, S. L. H. Teo, O. E. Snodgrass, K. Uosaki. 2002. Age and growth of North Pacific albacore (Thunnus alalunga). ISC/11/ALBWG/02

Appendix 2: A copy of the CIE Statement of Work

Attachment A: Statement of Work for Dr. Yong Chen

External Independent Peer Review by the Center for Independent Experts

Stock assessment of albacore tuna in the North Pacific Ocean

Scope of Work and CIE Process: The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of NMFS scientific projects. The Statement of Work (SoW) described herein was established by the NMFS Project Contact and Contracting Officer's Technical Representative (COTR), and reviewed by CIE for compliance with their policy for providing independent expertise that can provide impartial and independent peer review without conflicts of interest. CIE reviewers are selected by the CIE Steering Committee and CIE Coordination Team to conduct the independent peer review of NMFS science in compliance the predetermined Terms of Reference (ToRs) of the peer review. Each CIE reviewer is contracted to deliver an independent peer review report to be approved by the CIE Steering Committee and the report is to be formatted with content requirements as specified in **Annex 1**. This SoW describes the work tasks and deliverables of the CIE reviewer for conducting an independent peer review of the following NMFS project. Further information on the CIE process can be obtained from www.ciereviews.org.

Project Description: Albacore tuna in the North Pacific Ocean (NPO) are harvested multi-nationally primarily using troll, pole-and-line, and longline gear. The U.S. has major troll and pole-and-line fleets based in Washington, Oregon, and California. An assessment of albacore tuna in the North Pacific Ocean was conducted by staff of the Pacific Islands Fisheries Science Center, the Southwest Fisheries Science Center, and collaborating scientists from members of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean (ISC), within the ISC's Albacore Working Group in FY 2011. Results of the 2011 assessment are more optimistic than the 2006 assessment because fishing mortality in 2006-2008 has decreased relative to 2002-2004 and because of new assumptions regarding growth in the 2011 assessment. Previous assessments used a VPA model with growth curve parameters derived from a study in the 1960s using relatively outdated aging techniques and a limited size range. The current assessment is based on an SS3 model with growth curve parameters derived from a recent study using modern aging techniques and a substantially larger range of fish sizes. The new growth curve, with a lower L_∞, resulted in a higher scaling of biomass and a more positive assessment of stock status.

The assessment provides the basis for scientific advice on the status of the North Pacific albacore stock and will be a key foundation for international management decisions of the Inter-American Tropical Tuna Commission and Western and Central Pacific Fisheries Commission and its Northern Committee, and domestic management decisions by the Pacific Fishery Management Council (PFMC). An independent peer-review of the assessment is essential. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**.

Requirements for CIE Reviewers: Three CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein. CIE reviewers shall have expertise, working knowledge, and recent experience in 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 Stock Synthesis and AD Model Builder. Scientists who are employed by or have significant interactions with the Inter-American Tropical Tuna Commission (IATTC) and the Western and Central Pacific Fisheries Commission (WCPFC), and the Secretariat of the Pacific Community (SPC), should not be considered as reviewers. Scientists associated with the ISC also should be excluded as reviewers. Each CIE reviewer's duties shall not exceed a maximum of 10 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review as a "desk" review of the necessary documentation of the current assessment of albacore tuna in the NPO. Therefore, no travel is required.

Statement of Tasks: Each CIE reviewer shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

Prior to the Peer Review: Upon completion of the CIE reviewer selection by the CIE Steering Committee, the CIE shall provide the CIE reviewer information (full name, title, affiliation, country, address, email) to the COTR, who forwards this information to the NMFS Project Contacts no later than the date specified in the Schedule of Milestones and Deliverables. The CIE is responsible for providing the SoW and ToRs to the CIE reviewers. The NMFS Project Contacts will provide the CIE reviewers with the background documents and reports of the current assessment and sensitivity analyses to be peer reviewed. Any changes to the SoW or ToRs must be made through the COTR prior to the commencement of the peer review.

Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE Lead Coordinator on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

Please note that supporting documentation for the review is confidential, and reviewers are not to circulate these documents.

Desk Review: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs can not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COTR and CIE Lead Coordinator.** The CIE Lead Coordinator can contact the Project Contact to confirm any peer review arrangements.

Contract Deliverables - Independent CIE Peer Review Reports: Each CIE reviewer shall complete an independent peer review report addressing each ToRs in accordance with the SoW. Each CIE reviewer shall complete the independent peer review according to required format and content as described in Annex 1. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in Annex 2.

Specific Tasks for CIE Reviewers: The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables**.

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contacts in advance of the peer review.
2) Conduct an independent peer review in accordance with the ToRs (**Annex 2**).
3) No later than 6 December 2011, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Mr. Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and to Dr. David Die, CIE Regional Coordinator, via email to ddie@rsmas.miami.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

Schedule of Milestones and Deliverables: CIE shall complete the tasks and deliverables described in this SoW in accordance with the following schedule.

10 November 2011	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact
15 November 2011	NMFS Project Contact sends the CIE Reviewers the report and background documents
16–30 November 2011	Each reviewer conducts an independent peer review as a desk review
6 December 2011	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
20 December 2011	CIE submits the CIE independent peer review reports to the COTR
23 December 2011	The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: Requests to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent substitutions. The Contracting Officer will notify the COTR within 10 working days after receipt of all required information of the decision on substitutions. The COTR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the CIE reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

Acceptance of Deliverables: Upon review and acceptance of the CIE independent peer review reports by the CIE Lead Coordinator, Regional Coordinator, and Steering Committee, these reports shall be sent to the COTR for final approval as contract deliverables based on compliance with the SoW and ToRs. As specified in the Schedule of Milestones and Deliverables, the CIE shall send via e-mail the contract deliverables (CIE independent peer review reports) to the COTR (William Michaels, via William.Michaels@noaa.gov).

Applicable Performance Standards: The contract is successfully completed when the COTR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards:
(1) Each CIE report shall completed with the format and content in accordance with **Annex 1**,
(2) Each CIE report shall address each ToR as specified in **Annex 2**,
(3) The CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

Distribution of Approved Deliverables: Upon acceptance by the COTR, the CIE Lead Coordinator shall send via e-mail the final CIE reports in *.PDF format to the COTR. The COTR will distribute the CIE reports to the NMFS Project Contact and Center Director.

Support Personnel:

William Michaels, Program Manager, COTR
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-713-2363 ext 136

Manoj Shivlani, CIE Lead Coordinator
Northern Taiga Ventures, Inc.
10600 SW 131st Court, Miami, FL 33186
shivlanim@bellsouth.net Phone: 305-383-4229

Roger W. Peretti, Executive Vice President
Northern Taiga Ventures, Inc. (NTVI)
22375 Broderick Drive, Suite 215, Sterling, VA 20166
RPerretti@ntvifederal.com Phone: 571-223-7717

Key Personnel - NMFS Project Contact:

Dr. Steve Teo
Southwest Fisheries Science Center, 8604 La Jolla Shores Drive, La Jolla, CA 92037
Steve.Teo@noaa.gov Tel: 1-858-546-7179, Fax: 1-858-546-7003

Dr. Hui-hua Lee

Pacific Islands Fisheries Science Center, 2570 Dole Street, Honolulu, HI 96822
Huihua.Lee@noaa.gov Tel: 1-808-983-5352; Fax: 1-808-983-2902

Annex 1: Format and Contents of CIE Independent Peer Review Report

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

2. The main body of the reviewer report shall consist of a Background and Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

3. The reviewer report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of the CIE Statement of Work

Annex 2: Terms of Reference for the Peer Review

Stock assessment of albacore tuna in the North Pacific Ocean

1. Review the assessment methods: determine if they are reliable, properly applied, and adequate and appropriate for the species, fisheries, and available data.

2. Evaluate the assessment model configuration, assumptions, and input 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.

3. Evaluate the adequacy of the sensitivity analyses in regard to completeness and incorporation of results.

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

5. Suggest research priorities to improve our understanding of essential population and fishery dynamics and improvements to the assessment model.

Please note that supporting documentation for the review is confidential, and reviewers are not to circulate these documents.