CIE Independent Peer Review Report

on

Stock Assessment of Striped Marlin, *Tetrapturus audax,*

in the Western and Central 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

Striped marlin, *Tetrapturus audax*, supports multinational and multi-gear fisheries in the Pacific Ocean. Two stocks are defined in the North Pacific Ocean: the Western and Central North Pacific (WCNPO) stock and Eastern North Pacific stock. The WCNPO stock is the focus of this stock assessment covering the time duration from 1975 through 2010 and the waters to the west of 140^0W in the North Pacific Ocean. The Billfish Working Group (BILLWG) of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean (ISC) is responsible for conducting regular assessment for the WCNPO stock.

Stock Synthesis (SS, version 3.11b; Methot 2011) was used as the stock assessment modeling platform to provide estimates of stock parameters and an age-structured projection model was used to project the stock dynamics from the time period from 2010 through 2017 under six different levels of constant fishing mortality and two levels of constant catch. Based on the stock assessment and projection, the WCNPO striped marlin stock status was determined and scientific advices on fisheries management were provided. Both the assessment and projection models implicitly assume that there is a single well mixed stock of striped marlin in the WCNPO. The assessment model uses seasons (quarters) as its time step and includes eighteen fisheries and the projection model uses year as its time step.

Based on temporal and spatial variability in fishing operation, hook-per-basket distribution, targeted species and length distribution of fish, fifteen sets of standardized annual relative abundance indices were developed for eight fisheries, including ten Japanese longliner indices, two driftnet indices, two Taiwanese longliner indices, and one Hawaii-based longliner index. The timing (i.e., season) of each index was determined by the timing when the most landings were recorded for the fishery. No fishery-independent data were available. The BILLWG developed a base-case model for the assessment, and conducted a well-planned and structured sensitivity analysis to evaluate impacts of different input data and model configurations and parameterization on stock assessment. The BILLWG concluded that the WCNPO striped marlin stock was overfished and overfishing occurred and recommended that the current fishing mortality be reduced to allow for increased stock biomass. The BILLWG further projected how the stock biomass might change with different levels of fishing mortality/catch.

I independently evaluated the WCNPO striped marlin stock assessment report with respect to a set of pre-defined Terms of Reference. *I conclude that overall this stock assessment is based on the best science available. I conclude that this assessment is scientifically sound and adequately addresses needs for management advice. I agree with the conclusion regarding the WCNPO striped marlin stock status and management advice made in the stock assessment report.* In particular, I would like to commend the efforts of the BILLWG for compiling updated fisheries and biological data for the multinational and multi-gear fisheries and conducting a well-planned and structured stock assessment and sensitivity analysis to evaluate and address uncertainty regarding data quality and quantity and model configuration and parameterization. However, I believe some important questions did not receive enough attentions or were not addressed in the assessment. These issues include lack of retrospective analysis; lack of considering all uncertainty in the projection; and failure to explicitly define target and limit reference points for stock biomass and fishing mortality and relevant harvest control rules. I also believe more studies are needed to further improve the quality of fisheries data and biological information on the stock spatial structure and spatial variability in key life history processes such as growth and maturation.

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 striped marlin stock assessment model and identify key assumptions that may significantly influence the model performance; (2) conduct more studies to evaluate the quality of the input data and the 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) conduct retrospective analysis to evaluate possible retrospective errors associated with stock biomass, recruitment, and fishing mortality estimated for the recent years and calculate Mohn's rho to explicitly describe and quantify the nature and magnitude of retrospective errors; (5) explore the use of the dynamic binning option in the SS or robust multinomial likelihood functions to reduce impacts of non-informative zero-observation or outliers in size composition data on the model fitting; (6) plot each set of CPUE against the estimated stock biomass to evaluate if gear saturation exists; (7) evaluate likelihood profiles for a range of values for steepness h and natural mortality M to determine if they can be assumed independently in input data; (8) evaluate the roles of recruitment deviation penalty functions in estimating annual recruitment deviation; (9) evaluate the performance of MSY-based biological reference points in the management of the WCNPO striped marlin stock and identify alternative biological reference points (e.g., some historical fishing mortality and stock biomass); (10) examine the uncertainty associated with biological reference points and its impacts on the determination of stock status; (11) evaluate potential impacts of discrepancies between the stock assessment and projection model outlined in Table 5 (BILLWG 2012) on the evaluation of stock status; (12) consider more measures for comparing performance of different management options; (13) consider an alternative management time period, other than eight years, in the projection in order to identify possible differences in long-term and short-term projections; and (14) develop priors for key fishery and population parameters and apply the Bayesian estimator to better quantify uncertainty in stock assessment and projection. More detailed recommendations and their justifications can be found in the sections of Summary of Findings and Conclusions and Recommendations.

II. Background

Striped marlin, *Tetrapturus audax*, supports multinational and multi-gear fisheries in the Pacific Ocean. One unit stock was assumed in the North Pacific Ocean in the 2007 stock assessment. More recent studies suggest that there are two distinct genetic populations in the North Pacific Ocean (McDowell and Graves 2008, Purcell and Edmands 2011). As a result, two stocks are defined in the North Pacific Ocean for the assessment: the Western and Central North Pacific (WCNPO) stock and Eastern North Pacific stock. The WCNPO stock is the focus of this stock assessment covering the time duration from 1975 through 2010 and the waters to the west of 140^0W in the North Pacific Ocean. The Billfish Working Group (BILLWG) of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean (ISC) is responsible for conducting regular stock assessment.

Estimating growth of billfish species is difficult because of minute sizes of their otoliths and the challenge to obtain samples of good size coverage. The first growth study based on calcified parts in the North Pacific was done in 2003 (Melo-Barrera et al. 2003). The result was used in the last striped marlin stock assessment in 2007. However, the growth estimates for young striped marlin may not be reliable because of lack of young fish in the samples. A more recent study with samples of juvenile striped marlin suggests much faster growth rates for young striped marlin (Sun et al. 2011a,b), which is considered to more realistically quantify the growth of striped marlin. However, the studies are limited in their spatial and temporal coverage, and there is little information available on spatial and temporal variability in the growth.

Sexual dimorphism of striped marlin is related to spawning season and body size (Wang et al. 2006). Male striped marlin tend to mature at a smaller size than females (Kopf et al. 2009, and Sun et al. 2011c,d). Large spatial-temporal variability is evident in the maturation and reproduction of striped marlin in the North Pacific Ocean (Kopf et al. 2009).

Analyses of temporal and spatial variability in the fishery CPUE and catch size composition indicate that the striped marlin population in the North Pacific moves to higher latitudes during summer, but there is no evidence to support trans-ocean movement. Various tagging studies suggest the lack of trans-Pacific and trans-equator movement by striped marlin (Sippel et al. 2011), with ambient oceanographic current being one of main factors influencing individual movement. Vertical movement was mainly limited from the surface to the mixed layer above 90 m depth and regulated by relative changes in water temperature with depth. More studies are still needed to have a better understanding of the extent of movement in the northwest Pacific and into the Hawaiian region.

Most of striped marlin catch is harvested using longline, driftnet and harpoon by Japan, USA, and Taiwan in the WCNPO. Japanese fishing fleets dominated the fishery in the 1950s and 1960s, and striped marlin were caught in longline fisheries targeting albacore and were targeted in harpoon fisheries in coastal waters of Japan. Longline catches of striped marlin reached the highest level in the late 1960s. During the 1970s and 1980s, longline fisheries moved into deeper waters in more tropical waters targeting adult bigeye tunas, where striped marlin were less abundant. This shift of spatial distribution of fishing effort might result in the reduced striped marlin catch in the 1970s. Catches have continued to decline from approximately 6,000 mt per year in the 1990s to 4,200 mt per year in the early 2000s and 3,500 mt per year during 2005-2008. Reported catch of 2,560 mt in 2009 was the lowest catch reported since 1952. However, this did not necessarily result from the decline in fish abundance; but rather it reflects changes in spatial distribution of the fishery which expanded eastward in WCNPO in the 1950s and 1960s but then reduced in the 1990s and 2000s (MAROWG 2006).

Seventeen fisheries were initially identified, however, a preliminary analysis suggested that a residual pattern and quarterly size composition data from the "Japan other fishery" showed a strong seasonal pattern. Thus the "Japan other fishery" was divided into two separate fisheries: early (seasons 1-2) and late (seasons 3-4) fisheries. Subsequently eighteen fisheries were defined in this assessment based on country/regions, gear, spatial coverage, and season to minimize spatial/temporal variability in selectivity and catchability. These fisheries include: nine longline (USA, Japan coastal, Japan offshore and distant water by area, Japan other seasons 1-2, Japan other seasons 3-4, Chinese Taipei offshore and distant-water, and Korea), two driftnet (Japan high sea and coastal large-mesh and Japan squid), one bait (Japan), one trap (Japan), one set net (Japan), two harpoon (Japan), one coastal fishery (Taiwan offshore and coastal gillnet, coastal harpoon, coastal set net and other) and one miscellaneous longline (WCPO data). These fisheries vary greatly in their spatial and temporal coverage, selectivity and catchability, and differ in nature (striped marlin are targeted or bycatch species).

Catch and size composition data were estimated and compiled by seasons (Jan-March, Apr-Jun, Jul-Sep, and Oct-Dec) for these fisheries from 1975 through 2010. The 2010 catch was assumed to be the same as the 2009 catch because the 2010 catch data were incomplete. Strong seasonal patterns were observed in catch and such seasonality in catch differs among the 18 fisheries defined in this stock assessment. Quarterly length composition data, measured as lower jaw fork length (LJFL) and compiled in 5-cm size bins from 55 to 230 cm, were available for eleven fisheries. The length frequency data represent actual number of striped marlin measured. Because of large spatial and temporal variability in timing of recruitment and rapid growth in early ages, the first size bin was set at 120 cm LJFL, which essentially acts as an accumulation for fish smaller than age 1 size. Based on temporal and spatial variability in fishing operation, hook-per-basket distribution, targeted species, and length distribution of fish, fifteen sets of standardized annual relative abundance indices were developed for eight fisheries, including ten Japanese longliner indices, two driftnet indices, two Taiwan longliner indices, and one Hawaii-based longliner index. The timing (i.e., season) of each index was determined by the timing when most landings were recorded for the fishery. Although different in timing and magnitude of decline, these abundance indices tend to suggest a decreased abundance in the 2000s. Except for the two Taiwan longliner indices for which constant CVs of 0.2 and 04 were used for all the years, the coefficients of variation (CVs) of the standardized indices for other indices were derived from GLM models. The CVs were used to essentially quantify the quality of these data in the assessment.

Stock Synthesis (SS, version 3.20b; Methot 2011) was used as a modeling platform for the development of a seasonal, length-based, age-structured, forward-simulation population model for the current assessment of the WCNPO striped marlin stock (BILLWG 2012). A previous

ISC striped marlin assessment was done in 2007 using Stock Synthesis 2. Seven major differences between the 2007 and current stock assessments were identified including different assumptions on stock structure, selectivity and initial stock condition; different assumed values for steepness parameter (h), natural mortality (M), and growth and maturity parameters; and different time period covered.

The BILLWG conducted the stock assessment to estimate key fishery and population parameters. Various sensitivity analyses were done to evaluate potential impacts of assumptions made implicitly and explicitly in modeling. Various weighting schemes were evaluated and iterative modeling approaches were used in identifying relative weights for CPUE and size composition data of different sources and different quality to improve internal model consistency. Various analyses were done to compare model fitting of different selectivity functions and time blocks to identify optimal selectivity functions and time blocks for different fisheries. Population projection was done using an age-structured model (Punt 2010) for the time period from 2010 through 2017 (but with a year starting on July 1 and ending on June 30) to evaluate the performance of six levels of constant fishing mortality rates and two levels of constant catch in conserving SSB (measured as the ratio of SSB values between 2017 versus 2010). The population projection model is different from the SS in model structure (see Table 5 in BILLWG 2012 for details). Uncertainties in recruitment dynamics were considered in the projection. However, no probability distribution for the ratio of SSB values between 2017 versus 2010 was estimated. No target and limit biological reference points were explicitly specified and no harvest control rules were developed.

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

As the SoW states that "*Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs*", my role as a CIE independent reviewer is to conduct an impartial and independent peer review of stock assessment of striped marlin in the Western and Central North Pacific Ocean which are fished by multiple nations with multiple gears, with respect to the pre-defined Terms of Reference.

This is a desk review. Thus, I have no opportunity for face-to-face discussion and questioning. I read the "Stock assessment of striped marlin in the Western and Central North Pacific Ocean" by the Billfish 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. 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 striped marlin in the Western and Central 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.

   Given the complexity of the fishery with multiple fishing fleets and gears from multiple nations and large temporal-spatial variability in the fishing operation and striped marlin life history processes, I conclude that the SS employed is the most appropriate assessment platform which can provide an adequate modeling framework for the assessment of the WCNPO striped marlin. I further conclude that the stock assessment models are properly implemented with caution being paid to the evaluation of data quality and quantity, implicit and explicit assumptions, and alternative modeling options through preliminary analyses, sensitivity analyses, iterative weighting in modeling, and careful evaluation of model fitting and residual patterns. The base-case assessment scenario developed by the BILLWG appears to represent the best knowledge available with respect to the WCNPO striped marlin fisheries.

   However, I could not find that retrospective analyses were done for evaluating possible retrospective errors in the estimation of SSB, recruitment and fishing mortality. The projection model differs from the stock assessment model in time step, model structure, timing for calculating SSB, selectivity and timing for applying M, which may result in inconsistency and add extra uncertainty in the population projection. The assessment only considered the management options of constant fishing mortality rates or constant catch and no target and limit biological reference points were explicitly defined for fishing mortality and stock biomass. Not all uncertainty was considered in the population projection and in the determination of stock status.

   Stock Synthesis used in this assessment is a seasonal age-based, size-structured model. One of the greatest strengths of SS is its flexibility to utilize a wide diversity of age/size-based data and aggregate data of different sources and to account for temporal variability in catchability and selectivity. Using the SS framework, the BILLWG developed a seasonal, length-based, age-structured forward population projection model to predict fishery data (CPUEs and size composition data) which were then compared with corresponding observed data to formulate likelihood functions for the parameter estimation.

   I support the BILLWG's choice of using the SS for the assessment of the WCNPO striped marlin stock with data from such diverse sources. 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 catchability and selectivity. The use of quarters as time step is consistent with seasonality of some fisheries described in the stock assessment report (BILLWG 2012). The updated information on the stock area, steepness parameter, natural mortality, growth and maturity parameters, selectivity, assessment duration and setting of the initial stock condition since the last stock assessment in 2007 improved the assessment. Iterative estimation of effective sample sizes and data variances improved the internal consistency of model fitting. The use of standardized CPUE data removed the impacts of factors other than stock biomass.

Although I am impressed by this well-thought out, planned and structured stock assessment, I was surprised that no retrospective analysis was done for evaluating possible retrospective errors that are often associated with estimates of stock biomass, recruitment, and fishing mortality in the recent years in a stock assessment of this nature. Retrospective analyses are routinely done in stock assessment and I believe should be done in this assessment. I was also surprised at the inconsistency of the model structure and parameterization between the stock assessment model and projection model. Although the projection model is designed for the output from the SS model and widely used on the west coast, I think such an inconsistency may raise some issues in interpreting the results of the stock projection, complicating the interpretations of the ratio of stock biomasses of 2010 and 2017; which is used as the measure to compare the performance of different management options. This would be especially true when uncertainty is considered in the population projection.

2. <u>Evaluate the assessment model configuration, assumptions, and input data and parameters (fishery, life history, and spawner recruit relationships): determine if data are properly used, input parameters seem reasonable, models are appropriately configured, assumptions are reasonably satisfied, and primary sources of uncertainty accounted for.</u>

The BILLWG compiled quarterly catch and size composition data of different fisheries, developed standardized CPUEs and assembled updated biological and fisheries parameters for the assessment of this multinational and multiple-gear fishery. This is a major undertaking, and I commend the BILLWG for their efforts to compile such a comprehensive and updated data base for the WCNPO striped marlin stock assessment. I conclude that the stock assessment models are adequately configured, data are properly prepared, screened and used, and assumptions are reasonably satisfied.

However, I believe uncertainty has not been fully considered. My concern is the failure to conduct retrospective analyses to estimate the magnitude and nature of possible retrospective errors that are often associated with the estimates of stock biomass, fishing mortality and recruitment in the recent years in the stock assessment. A lack of understanding of retrospective errors limits our understanding of the quality (and uncertainty) of the key fisheries parameters in the most recent years, which are used in the projection model for evaluating the performance of various management options. I suggest that retrospective analysis be done and that Mohn's rho (Mohn 1993, 1999) be calculated to explicitly describe and quantify the nature and magnitude of retrospective errors.

The BILLWG has explicitly described some assumptions with respect to assessment models, projection models, and statistical analyses. However, most of the assumptions are embedded in the texts and are sometimes hard to evaluate if such assumptions were violated. The possible consequences if some of the assumptions are violated are also unclear. I suggest that the BILLWG summarize all the assumptions about the model, data, and statistical analyses, explicit and implicit, in a table and describe if a particular assumption is satisfied for a given assessment scenario. Potential consequences of violating these assumptions should also be described in the table. I believe this can greatly help understand potential sources of uncertainty and improve the design of the sensitivity study in the assessment.

Eighteen fisheries were defined in the assessment. These fisheries have different spatial-temporal coverage, tend to target different components of the WCNPO striped marlin, and have different impacts on the stock dynamics. Relative weights of different data sets from different fisheries in the model fitting were mainly determined by the data quality measured by CVs for CPUE data and effective sample sizes for size composition data. I am curious if this is sufficient to reflect the relative importance of different fisheries in driving the dynamics of striped marlin stock. Maybe weighting factors for different likelihood functions should have been considered for the spatial coverage of relevant fisheries, instead all were set at 1.

Size compositions data are available for the eleven fisheries, and size ranges of the data tend to vary among the fisheries because of differences in gear, spatial, and temporal coverage of these fisheries. A size range of 120 to 230 cm LJFL was used to group size composition data from all the fisheries for which the size composition data were available. This might result in a large number of zeros for small and/or large size bins for some fisheries, which might result from limited sampling efforts or spatial coverage, rather than no fish in these size classes. This may 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 non-informative zero-observation in size composition data on the model fitting.

Choices of selectivity curves and time blocks assumed for different fisheries appear to be reasonable. The selectivity essentially includes both fish availability and gear selectivity. Impacts of changing fishing operations on catchability were considered in the model parameterization. The BILLWG also considered alternative choices of selectivity functions for different fisheries, and carefully evaluated and compared patterns of residuals in fitting size composition data with different selectivity functions to justify the choices of selectivity functions. However, I do not see much discussion about gear saturation. For longline and gillnet, gear saturation might be an issue, which can affect the reliability of CPUE as abundance indices even if it is standardized. A saturation parameter may be needed when CPUE is related to the stock abundance/biomass.

Sensitivity analysis was conducted to evaluate robustness of the modeling results with respect to alternative values for (1) quality of different data sets, which determines weights of different data sets; (2) biological parameters; and (3) fishery parameters (i.e., selectivity and catchability in this assessment). Most sensitivity runs were conducted with just one parameter being given alternative values while other parameters were held constant as in the base case. Thus, the sensitivity analysis was essentially done for evaluating 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 should not be assumed independently in the assessment. Likelihood profiles should be evaluated for a range of values for these two parameters to identify their relationship and if they can be determined independently.

I did not see the description about the use of recruitment deviation 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 the penalty functions should be evaluated in the sensitivity analysis (Methot and Taylor 2011).

The CPUE data were weighted in model fitting using CVs estimated from the GLM-based CPUE standardization. The GLM-estimated CVs were small (actually much smaller than what I have seen for other fisheries). Because of limited spatial and temporal coverage of a given fishery, the GLM-estimated CV for CPUE of the fishery is more suitable for describing local variability, but not for the whole stock range. When the CPUE is used to describe the population dynamics of the whole stock range, this level of CV under-estimates the variability. Thus, I believe the GLM-estimated CVs for CPUE data are not appropriate for being used in the likelihoods. However, the BILLWG used an iterative weighting approach to adjust weighting for the CPUEs, which reduces my concerns on this issue.

3. <u>Comment on the proposed population benchmarks and management parameters (e.g., MSY, Fmsy, Bmsy, MSST, MFMT); if necessary, recommended values for alternative management benchmarks (or appropriate proxies) and clear statements of stock status.</u>

   The MSY-based reference points are estimated in the SS and are implicitly assumed to be limit reference points in the determination of stock status. However, I believe more research is needed in the evaluation of these reference points in their effectiveness of managing the fishery. Uncertainty of the estimated biological reference points and its implications should also be carefully evaluated. Overall I agree with the conclusion regarding the WCNPO striped marlin stock status. However, I believe uncertainty should be considered in the determination of stock status (although the conclusion regarding the stock status would be the same).

   Although the MSY-based biological reference points are estimated in the stock assessment, no limit and target reference points were explicitly defined in the risk analysis of alternative management strategies. However, the B_{MSY} and F_{MSY} were implicitly defined as limit reference points because they were used in the assessment to conclude that the WCNPO striped marlin stock was overfished and the fishery was in the status of overfishing. Although the use of MSY-based biological reference points for limit reference points is rather common, the performance and implication of using them need to be carefully evaluated.

I recommend that management strategy evaluation (MSE) be developed for the WCNPO striped marlin stock. The MSE can then be used to evaluate and identify target and limit biological reference points. Only the constant fishing mortality (catch) harvest control rules were considered in this assessment. Such control rules lack the ability to adjust fishing mortality based on the updated status of fisheries. The MSE can be used for the evaluation and development of different harvest control rules.

Uncertainty associated with biological reference points was not considered and evaluated in this assessment. Given the uncertainty associated with fisheries and biological parameters, I suspect that the uncertainty may be considerable for biological reference points. The uncertainty associated with biological reference points is important in determining the status of stock.

The biological reference points were estimated in the SS, and used in the projection model which has different model structures and parameterization. I think it is necessary to evaluate the consistency in the performance of fishing mortality and stock biomass biological reference points in the projection model.

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. However, discrepancies of the assessment model and projection model need to be addressed. In particular, I recommend evaluating potential impacts of the discrepancies on the evaluation of different management options. I suggest that more measures, in addition to the ratio of stock biomass in 2011 and 2018, are used in evaluating management options. I also suggest that management strategies that can adjust exploitation rates based on stock status should be considered, rather than just considering constant fishing mortality or constant catch strategy.

   The projection model used in the assessment was developed to incorporate the output from the SS and was used in the assessment of some groundfish populations on the west coast. However, I believe this is not sufficient to justify the use of this model in this study. The discrepancies of assessment and projection models were outlined in Table 5 in the assessment report (BILLWG 2012), but the potential impacts of such discrepancies on the evaluation of stock status were not evaluated and discussed.

   Only constant fishing mortality (6 levels) and catch (2 levels) management strategies were considered in this study. This type of harvest control rule does not adjust fishing mortality based on the stock status. I suggest developing MSE for the WCNPO striped marlin and evaluate and identify target and limit references for both fishing mortality and stock biomass and to consider and evaluate more forms of harvest control rules which allow the adjustment of fishing mortality based on stock status.

Two types of recruitment dynamics were used in the projection: recruitment randomly drawn from recent recruitment pattern (i.e., 1994-2008) and recruitment determined from the SR curve estimated based on SSB and R from 1975 to 2008 in the assessment model with annual deviation randomly drawn from residuals in SR modeling. The approach is sound and commonly used. However, an implicit assumption associated with the approach is that recruitment variability for the time period covered by the projection would be similar to the recruitment variability in the past. With non-random environmental changes (e.g., climate changes) more frequently observed in marine systems, such an assumption may be violated, which may introduce extra errors in the projection for a given set of management option.

I suggest that the BILLWG evaluate the performance of the projection model for the time period since the last stock assessment (i.e., 2007 to 2010). The BILLWG can use the projection model to project population dynamics from 2007 when the last assessment was done and compare the results derived in the current stock assessment for the same time period to see if the projection model yields feasible results. This can be repeated in a few years when the next stock assessment is conducted to evaluate the discrepancy of the stock biomass projected in the current assessment and stock biomass estimated in the next stock assessment for the time period from 2010 to the years when the next assessment is done.

I only found one measure was used to evaluate the performance of different management options. The measure used is ratio of stock biomass in the beginning and end of the projection period, which may not be sufficient to measure the performance of a management option. I suggest considering more measures (e.g., total catch, lowest stock biomass in the project period, among-year variability in catch). I also suggest quantifying the uncertainty associated with the ratio of stock biomasses at the beginning and end of the projection period to develop a probability distribution of this ratio.

I am a little puzzled by the use of eight years as the projection time period. This seems too short for the long-term projection but too long for the short-term projection. In any case, I suggest evaluating possible differences in long-term and short-term projections.

Maximum likelihood estimator was used in this assessment. I suggest that Bayesian estimators be used in the assessment to better quantify the uncertainty. The posterior distributions for key fisheries parameters derived from the Bayesian estimators can be used in the projection.

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

I suggest the following research priorities to improve the WCNPO striped marlin stock assessment and our understanding of population and fishery dynamics. Some of the research priorities are for the long term, and some are for the short term (e.g., next assessment).

- Evaluate relationships between each pair of CPUE data sets to evaluate their consistency in indexing the temporal trend of stock biomass and to 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, which can help determine relative weighting factors in the objective function;

- Plot each set of CPUE versus stock biomass estimates to identify possible gear saturation effects;

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

- Continue 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;

- Continue evaluating differences in key life history parameters between females and males to determine if a sex-specific stock assessment is necessary;

- Explore a dynamic binning approach to address potential issues of including too many size bins with non-informative zero to improve model fitting;

- Conduct retrospective analysis to evaluate retrospective errors for stock biomass, recruitment, and fishing mortality estimates;

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

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

- Develop MSE for the WCNPO striped marlin stock to evaluate (1) alternative biological reference points, harvest control rules and management strategies; (2) impacts of data quality and quantity on the quality of stock assessment; and (3) impacts of violating some key assumptions on stock structure (e.g., possible meta-populations structure or large spatial variability in key life history process) and fisheries (e.g., selectivity and catchability).

V. Conclusions and Recommendations

The assessment appears to be well-planned and structured, scientifically sound, and adequately addresses needs for management advice for the WCNPO striped marlin stock. Uncertainties in the input data, fisheries and biological processes, and model parameterization were carefully evaluated in the stock assessment and projection. The conclusion on the stock status appears to be robust to uncertainty in the assessment and projection. In particular, I would like to commend the efforts of the BILLWG in addressing data quality issues, designing and conducting a well structured sensitivity analysis, exploring alternative model configurations and parameterization, and evaluating impacts of different fisheries on the assessment of stock dynamics.

However, I do have concerns that I hope the BILLWG could address to further improve the assessment of the WCNPO striped marlin stock. I made the following general comments and specific recommendations.

General comments

Although the new growth data were used in this study, they were derived based on samples collected in limited areas. I believe more study is still needed to improve the estimation of growth for the WCNPO striped marlin stock. Given the difficulty in sampling a large number of striped marlin over a large spatial area, I suggest that research efforts be devoted to develop approaches to back-calculate length-at-age data to derive length at each age for each fish. 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.

I am concerned with the failure to conduct retrospective analyses to estimate magnitude and nature of possible retrospective errors that are often associated with the estimates of stock biomass, fishing mortality and recruitment in the most recent years in stock assessment. A lack of understanding retrospective errors limits our understanding of the quality of the key fisheries parameter estimates in the recent years, which are used in the projection model for evaluating the performance of various management options. I suggest that Mohn's rho (Mohn 1993, 1999) be calculated to explicitly describe and quantify the nature and magnitude of retrospective errors for SSB, recruitment and fishing mortality.

Although the BILLWG considered uncertainty in data, model configuration and parameterization, and fisheries processes, some important sources of uncertainty were not considered in the assessment. No uncertainty associated with initial stock structure and biomass was considered in the projection. No uncertainty was considered in the estimation and use of biological reference points. No probability distribution was estimated for the ratio of stock biomass at the beginning and end of the projection, which is used as the measure in the evaluation of the performance of alternative management strategies.

Although the assessment model and projection model used in this study are well known and widely used, I believe the discrepancy of these two models for the WCNPO striped marlin stock may cause some concerns in the consistency and comparability of stock dynamics between the time period covered by the stock assessment model and that covered by the projection model.

No target and limit biological reference points are explicitly defined for fishing mortality and stock biomass. No harvest control rule was explicitly described, which makes the management advice unclear. Although the SS is 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. I believe that the development of MSE needs to be a future research priority to evaluate the performance of the SS in quantifying the striped marlin stock dynamic, improve the risk analysis of alternative management strategies, and help identify key factors that may influence the quality of stock assessment and projection.

A Bayesian approach was not used in the assessment, and uncertainty in the assessment was not fully quantified in the stock assessment and projection for different harvest strategies. I encourage future assessments to develop priors for key model parameters and utilize the Bayesian estimator in the SS to incorporate uncertainty in the assessment and projection.

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.

<u>Specific recommendations</u>

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

- I suggest evaluating spatial and temporal variability in growth and maturation. If the information currently available is not sufficient for such a study, a 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 suggest continuing evaluation of possible differences in key life history parameters between females and males to determine if it is necessary for a sex-specific stock assessment.

- I advise conducting 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 (Chen and Fournier 1999) 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.

- 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 occurring from changes in model configurations.

- I suggest that retrospective analysis be conducted to evaluate possible retrospective errors associated with stock biomass, recruitment and fishing mortality estimated for the recent years and that Mohn's rho (Mohn 1993, 1999) be calculated to explicitly describe and quantify the nature and magnitude of retrospective errors.

- I suggest that all the assumptions about the model, data, and statistical analyses, explicit and implicit, be summarized in a table and potential consequences of violating these assumptions be described in the table to help understand potential sources of uncertainty and improve the design of sensitivity study in the assessment.

- I suggest that the dynamic binning option in the SS or robust multinomial likelihood functions (Fournier 1996) be used to reduce impacts of non-informative zero-observation in size composition data on the model fitting.

- I suggest that each set of CPUE used be plotted against the stock biomass estimated in the assessment to evaluate possibility of gear saturation. If gear saturation is found to exist I suggest that a saturation parameter be added in the observational models linking the observed CPUE with stock abundance/biomass.

- I suggest that likelihood profiles be evaluated for a range of values for steepness h and natural mortality M to identify impacts of different combinations of these two values on the assessment results and to determine if they can be assumed independently in input data (which was essentially how their values were determined in the assessment).

- I suggest that a description of the use of recruitment deviation penalty functions, which are usually applied to constrain annual recruitment deviations and prevent the model from yielding biologically unrealistic values for model parameters, be added to clarify the roles of the penalty functions.

- I suggest that the use of MSY-based biological reference points for limit reference points be evaluated for their performance and implication in the management of the WCNPO striped marlin stock and that alternative biological reference points (e.g., some fishing mortality and stock biomass in the past) be considered, evaluated, and compared with the MSY-based reference points.

- I recommend that management strategy evaluation (MSE) be developed for the WCNPO striped marlin stock. The MSE can then be used to evaluate and identify target and limit biological reference points and different types of harvest control rules.

- I suggest that the uncertainty associated with biological reference points be evaluated and its impacts on the determination of stock status be examined.

- I suggest that discrepancies (Table 5 in BILLWG 2012) be examined and discussed between the assessment and projection models with respect to their impacts on the evaluation of alternative management strategies.

- I suggest that MSE be developed for the WCNPO striped marlin to include both target and limit reference points for both fishing mortality and stock biomass and to consider and evaluate more forms of harvest control rules for allowing the adjustment of fishing mortality based on stock status in fisheries management.

- I suggest that the performance of the projection model be evaluated. The BILLWG can use the projection model to project population dynamics from 2007 when the last assessment was done and compare the results derived in the current stock assessment for the same time period to identify if the stock assessment model and project model yield consistent results and if the projection approach yields feasible results.

- I suggest that more measures be considered in evaluating performance of different management options (e.g., total catch, lowest stock biomass in the project period, among-year variability in catch).

- I suggest that alternative management time period, other than eight years, be considered in the projection to identify possible differences in long-term and short-term projections.

- I suggest that Bayesian estimators be used in the assessment to better quantify the uncertainty. The posterior distributions for fisheries parameters derived from the Bayesian estimator can be used in the projection.

VI. References cited

Billfish Working Group. 2012. Stock Assessment of Striped Marlin in the Western and Central North Pacific Ocean. International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean. Document prepared by Hui-Hua Lee, Kevin R. Piner, Robert Humphreys, and Jon Brodziak. NOAA NMFS Pacific Islands Fisheries Science Center, 2570 Dole St., Honolulu, HI 96822, USA.

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

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.

Kopf, R.K., Pepperell, J., and Davie, P.S. 2009. A preliminary report on age, growth, and reproductive dynamics of striped marlin (kajikia audax) in the southwest Pacific Ocean. WCPFC-SC5/BI-WP-01, 5th Meeting of the Scientific Committee of the Western and Central Pacific Fisheries Commission (WCPFC-SC5), 10-21 August 2009, Port Vila, Vanuatu.

Marlin Working Group (MAROWG). 2006. Report of the Marlin Working Group Meeting, 15-21 November, 2005, Honolulu, HI, USA. Annex 5. Report of the Sixth Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean, Plenary Session. 23-27 March, 2006, La Jolla, CA, USA.

McDowell, J.R. and Graves, J.E. 2008. Population structure of striped marlin (*Kajikia audax*) in the Pacific Ocean based on analysis of microsatellite and mitochondrial DNA. Can. J. Fish. Aquat. Sci. 65(7): 1307–1320. doi:10.1139/F08-054.

Melo-Barrera, F.N., Félix-Uraga, R., and Quinonez-Velazquez, C. 2003. Growth and length-weight relationship of the striped marlin, *Tetrapturus audax* (Pisces: Istiophoridae), in Cabo San Lucas, Baja California Sur, Mexico. Ceinc. Mar. 29(3): 305-313.

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

Pilling, G. M., G. P. Kirkwood, and S. G. Walker. 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.

Punt, A.E. 2010. SSC Default Rebuilding Analysis: Technical specifications and User Manual. Jan. 2010.

Purcell, C.M. and Edmands, S. 2011. Resolving the genetic structure of striped marlin, *Kajikia audax*, in the Pacific Ocean through spatial and temporal sampling of adult and immature fish. Can. J. Fish. Aquat. Sci. 68: 1861–1875. doi:10.1139/F2011-104.

Sippel, T., Holdsworth, J., Dennis, T. and Montgomery, J. 2011. Investigating behavior and population dynamics of striped marlin (*Kajikia audax*) from the Southwest Pacific Ocean with satellite tags. PLoS ONE 6(6): 1-13.

Sun, C.-L., Hsu, W.-S., Chang, Y.-J., Yeh, S.-Z., Chiang, W.-C., and Su, N.-J. 2011a. Age and growth of striped marlin (*Kajikia audax*) in waters off Taiwan. Working paper submitted to the ISC Billfish Working Group Meeting, 19-27 January 2011, Honolulu, Hawaii, USA. ISC/11/BILLWG-1/09: 15p.

Sun, C.-L., Hsu, W.-S., Chang, Y.-J., Yeh, S.-Z., Chiang, W.-C., and Su, N.-J. 2011b. Age and growth of striped marlin (*Kajikia audax*) in waters off Taiwan: A revision. Working paper submitted to the ISC Billfish Working Group Meeting, 24 May-1 June 2011, Taipei, Taiwan. ISC/11/BILLWG-2/07: 12p.

Sun, C.-L., Hsu, W.-S., Chang, Y.-J., Yeh, S.-Z., Chiang, W.-C., and Su, N.-J. 2011c. Reproductive biology of female striped marlin (*Kajikia audax*) in the waters off Taiwan (preliminary). Working paper submitted to the ISC Billfish Working Group Meeting, 19-27 January 2011, Honolulu, Hawaii, USA. ISC/11/BILLWG-1/11: 14p.

Sun, C.-L., Hsu, W.-S., Chang, Y.-J., Yeh, S.-Z., Chiang, W.-C., and Su, N.-J. 2011d. Reproductive biology of male striped marlin, *Kajikia audax*, in the waters off Taiwan. Working paper submitted to the ISC Billfish Working Group Meeting, 24 May-1 June 2011, Taipei, Taiwan. ISC/11/BILLWG-2/09: 17p.

Wang, S.-H., Sun, C.-L., Yeh, S.-Z., Chiang, W.-C., Su, N.-J., Chang, Y.-J., and Liu, C.-H. 2006. Length distributions, weight-length relationships, and sex ratios at lengths for the billfishes in Taiwan waters. Bull. Mar. Sci., 79(3): 865-869.

Appendix 1: Bibliography of materials provided for review

A. STOCK ASSESSMENT REPORT (MAIN REVIEW DOCUMENT)

Billfish Working Group. 2012. Stock Assessment of Striped Marlin in the Western and Central North Pacific Ocean. International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean. Document prepared by Hui-Hua Lee, Kevin R. Piner, Robert Humphreys, and Jon Brodziak. NOAA NMFS Pacific Islands Fisheries Science Center, 2570 Dole St., Honolulu, HI 96822, USA.

B. ISC BILLFISH REPORTS

Billfish Working Group (BILLWG). 2011a. Report of the Billfish Working Group Meeting, 19-27 January, 2011, Honolulu, HI, USA. Annex 7. Report of the Eleventh Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean, Plenary Session. 20-25 July, 2011, San Francisco, CA, USA. Available at: http://isc.ac.affrc.go.jp/pdf/ISC11pdf/Annex_7_ISC11_BILLWG_FINAL.pdf

Billfish Working Group (BILLWG). 2011b. Report of the Billfish Working Group Meeting, 24 May-1 June, 2011, Taipei, Taiwan. Annex 8. Report of the Eleventh Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean, Plenary Session. 20-25 July, 2011, San Francisco, CA, USA. Available at: http://isc.ac.affrc.go.jp/pdf/ISC11pdf/Annex_8_ISC11_BILLWG_FINAL.pdf

Billfish Working Group (BILLWG). 2012a. Report of the Billfish Working Group Meeting, 6-16 December, 2011, Honolulu, HI, USA. Annex 5. Report of the Twelve Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean, Plenary Session. 18-23 July, 2012, Sapporo, Japan. Available at: http://isc.ac.affrc.go.jp/pdf/ISC12pdf/Annex%205%20-%20Report%20of%20the%20BILLWG%20Workshop%20(Dec%202011).pdf

Billfish Working Group (BILLWG). 2012b. Report of the Billfish Working Group Meeting, 2-9 April, 2012, Shanghai, China. Annex 7. Report of the Twelve Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean, Plenary Session. 18-23 July, 2012, Sapporo, Japan. Available at: http://isc.ac.affrc.go.jp/pdf/ISC12pdf/Annex%207%20-%20Report%20of%20the%20BILLWG%20Workshop%20(Apr%202012)%20FINAL.pdf

C. CPUE STANDARDIZATION

Kanaiwa, M., Kimoto, A., and Yokawa, K. 2011. Standardized CPUE of striped marlin in north western central Pacific Ocean by using GLM. Working paper submitted to the ISC Billfish Working Group Meeting, 24 May-1 June 2011, Taipei, Taiwan. ISC/11/BILLWG-2/06: 16p.

Yokawa, K. and Kimoto, A. 2011. Standardized CPUE of striped marlin caught by Japanese coastal large-mesh drift fishery in the northwest Pacific in the periods between 2001 and 2009. Working paper submitted to the ISC Billfish Working Group Meeting, 24 May-1 June 2011, Taipei, Taiwan. ISC/11/BILLWG-2/05: 5p.

Sun, C.-L., Su, N.-J., Yeh, S.-Z., and Chang, Y.-J. 2011d. Standardized catch-rates for striped marlin (Kajikia audax) for Taiwanese distant-water longline fishery in the North Pacific Ocean for 1967-2009. Working paper submitted to the ISC Billfish Working Group Meeting, 19-27 January 2011, Honolulu, Hawaii, USA. ISC/11/BILLWG-1/07: 45p. Available at: http://isc.ac.affrc.go.jp/pdf/BILL/ISC11_BILL_1/ISC11BILLWG1_WP07.pdf

Walsh, W. and Lee, H.-H. 2011. Standardization of striped marlin, Kajikia audax, CPUE with generalized linear models fitted to pelagic longline observer fata from the Hawaii-based fishery: 1995-2009. Working paper submitted to the ISC Billfish Working Group Meeting, 19-27 January 2011, Honolulu, Hawaii, USA. ISC/11/BILLWG-1/08: 45p. Available at: http://isc.ac.affrc.go.jp/pdf/BILL/ISC11_BILL_1/ISC11BILLWG1_WP08.pdf

D. BIOLOGICAL ASSUMPTION

Brodziak, J. and Mangel, M. 2011. Probable values of stock-recruitment steepness for north Pacific striped marlin. Working paper submitted to the ISC Billfish Working Group Meeting, 24 May-1 June 2011, Taipei, Taiwan. ISC/11/BILLWG-2/11: 13p.

Piner, K.R. and Lee, H.-H. 2011a. Meta-analysis of striped marlin natural mortality. Working paper submitted to the ISC Billfish Working Group Meeting, 19-27 January 2011, Honolulu, Hawaii, USA. ISC/11/BILLWG-1/10: 09p. Available at: http://isc.ac.affrc.go.jp/pdf/BILL/ISC11_BILL_1/ISC11BILLWG1_WP10.pdf

Piner, K.R. and Lee, H.-H. 2011b. Correction to Meta-analysis of striped marlin natural mortality. Working paper submitted to the ISC Billfish Working Group Meeting, 24 May-1 June 2011, Taipei, Taiwan. ISC/11/BILLWG-2/08: 01p. Available at: http://isc.ac.affrc.go.jp/pdf/BILL/ISC11_BILL_2/ISC11BILLWG2_WP08.pdf

Sun, C.-L., Hsu, W.-S., Chang, Y.-J., Yeh, S.-Z., Chiang, W.-C., and Su, N.-J. 2011a. Reproductive biology of female striped marlin (Kajikia audax) in the waters off Taiwan (preliminary). Working paper submitted to the ISC Billfish Working Group Meeting, 19-27 January 2011, Honolulu, Hawaii, USA. ISC/11/BILLWG-1/11: 14p. Available at: http://isc.ac.affrc.go.jp/pdf/BILL/ISC11_BILL_1/ISC11BILLWG1_WP11.pdf

Sun, C.-L., Hsu, W.-S., Chang, Y.-J., Yeh, S.-Z., Chiang, W.-C., and Su, N.-J. 2011c. Age and growth of striped marlin (Kajikia audax) in waters off Taiwan. Working paper submitted to the ISC Billfish Working Group Meeting, 19-27 January 2011, Honolulu, Hawaii, USA. ISC/11/BILLWG-1/09: 15p. Available at: http://isc.ac.affrc.go.jp/pdf/BILL/ISC11_BILL_1/ISC11BILLWG1_WP09.pdf

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 Striped Marlin

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 Representative (COR), and reviewed by CIE for compliance with their policy for providing independent expertise that can provide an 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: Striped marlin (*Tetrapturus audax*) is one of six species of billfishes commonly harvested multi-nationally from commercial and recreational fisheries in the western and central Pacific Ocean regions. Fishery management requires high quality science to effectively manage and conserve our living marine resources, and the scientific peer review of stock assessments by external CIE expertise is an important process in the determination of best scientific information available. 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 the application of fish stock assessment, mathematical modeling, and statistical computing. 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, therefore no travel is required.

Statement of Tasks: Each CIE reviewers 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 COR, who forwards this information to the NMFS Project Contact no later 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 Contact is responsible for providing the CIE reviewers with the background documents, reports, and other pertinent information. Any changes to the SoW or ToRs must be made through the COR 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.

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 COR 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 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 Contact in advance of the peer review.
2) Conduct an independent peer review in accordance with the ToRs (**Annex 2**).
3) No later than **03 December 2012**, 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 CIE Regional Coordinator, 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.

22 October 2012	CIE sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
25 October2012	NMFS Project Contact sends the CIE Reviewers the report and background documents
1-16 November 2012	Each reviewer conducts an independent peer review as a desk review
3 December 2012	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
17 December 2012	CIE submits the CIE independent peer review reports to the COR
21 December 2012	The COR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: This 'Time and Materials' task order may require an update or modification due to possible changes to the terms of reference or schedule of milestones resulting from the fishery management decision process of the NOAA Leadership, Fishery Management Council, and Council's SSC advisory committee. A request to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent changes. The Contracting Officer will notify the COR within 10 working days after receipt of all required information of the decision on changes. The COR 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 COR 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 COR (William Michaels, via William.Michaels@noaa.gov).

Applicable Performance Standards: The contract is successfully completed when the COR 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 COR, the CIE Lead Coordinator shall send via e-mail the final CIE reports in *.PDF format to the COR. The COR will distribute the CIE reports to the NMFS Project Contact and Center Director.

Support Personnel:

William Michaels, Program Manager, COR
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:

Gerald DiNardo, Stock Assessment Program Leader (NMFS Project Contact)
Pacific Islands Fisheries Science Center
2570 Dole Street, Honolulu, HI 96822-2396
Gerard.DiNardo@noaa.gov Phone: 808-983-5397

Kevin Piner
Pacific Islands Fisheries Science Center
2570 Dole Street, Honolulu, HI 96822-2396
Kevin.Piner@noaa.gov Phone: 858-546-7003

Jon Brodziak
Pacific Islands Fisheries Science Center
2570 Dole Street, Honolulu, HI 96822-2396
Jon.Brodziak@noaa.gov Phone: 808-983-2964

Hui-Hua Lee
University of Hawaii, Joint Institute for Marine and Atmospheric Research
2570 Dole St., Honolulu, HI 96822
Huihua.Lee@noaa.gov Phone: 808-983-5352

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, Description of the Individual Reviewer's Role in the Review Activities, 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 Striped Marlin

6. Review of the assessment methods: determine if they are reliable, properly applied, and adequate and appropriate for the species, fisheries, and available data.
7. Evaluate the assessment model configuration, assumptions, and input data and parameters (fishery, life history, and spawner recruit relationships): determine if data are properly used, input parameters seem reasonable, models are appropriately configured, assumptions are reasonably satisfied, and primary sources of uncertainty accounted for.
8. Comment on the proposed population benchmarks and management parameters (*e.g., MSY, Fmsy, Bmsy, MSST, MFMT*); if necessary, recommended values for alternative management benchmarks (or appropriate proxies) and clear statements of stock status.
9. Evaluate the adequacy, appropriateness, and application of the methods used to project future population status.
10. Suggest research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices.