Center for Independent Experts (CIE) Desk Review of the 2013 Pacific Blue Marlin Stock Assessment

CIE Review prepared by:

John D. Neilson, Ph.D.
Independent Fisheries Consultant
10 Moss Drive
Chamcook, New Brunswick, Canada
E5B 3G7

Tel. 1 506 529-4922
Email: neilson@nbnet.nb.ca

Executive Summary

The 2013 Pacific blue marlin assessment contained much that was praiseworthy, with little to fault. In particular, the assessment Working Group employed several methods to determine stock status, and generally consistent results were found. Given the assessment methods used different sorts of input data (age and sex specific, and highly aggregated information), the consistency of the results offers assurance that the results and subsequent management advice are robust. It should also be noted that the overall assessment is generally consistent with the most recent (2003) assessment, although a long period has passed between the two evaluations. The assessment report was also well-written and easy to evaluate. Overall, the results of this desk review indicate that the report represents the best available science, given the current state of knowledge for this resource.

Notwithstanding the generally very positive review of the current assessment, there are some areas where the reviewer found potential areas for improvement in future assessments, as well as research recommendations. As noted by the report authors, the assumption of a single unit stock throughout the Pacific seems to be a simplifying assumption that is unlikely to reflect reality. The exclusion of the PIROP observer data needs a more convincing rationale, given the extensiveness and high quality (relative to logbook information) of the data series. It is also noteworthy that if these data were included in the assessment (as illustrated by sensitivity analyses), the outlook for the resource would be considerably less positive. For future assessments, methods to make better use of the observer data series should be evaluated. The sources of uncertainty in the assessment should be summarized. Given such uncertainties, clear research recommendations should also be enunciated for those parameters that are shown to be influential in the assessment, notably including the growth model, and natural mortality. Finally, the report should comment on the impact of the observed retrospective pattern on the conclusion made by the assessment Working Group that the stock is slightly increasing.

1. Background

Description of the Reviewer's Role

The author of this Desk Review was not involved previously with the 2013 blue marlin stock assessment at any stage. The Desk Review was completed at the request of the Center for Independent Experts (CIE), and submitted on February 3, 2014.

2. General Comments on the 2013 Pacific blue marlin Assessment

In general, this assessment seemed very competently completed, and the team is to be congratulated on what could stand as a model for graduate students studying fisheries science. Multiple stock assessment methods were used, and gave generally consistent results. The methods were state of the art. The assessment Working Group (WG) prepared a report that was obviously carefully constructed and edited, making the job of providing an external assessment much easier. A small example of this careful work can be found in the Figure and Table captions. They are complete, and informative. In my experience with other Regional Fisheries Management Organizations (RFMOs), reports are sometimes prepared in a rather hasty manner, and figure and caption labels are often very incomplete. In such circumstances, reconstructing what was done becomes problematic, particularly for someone not involved with the meetings. This was certainly not a problem with the 2013 Pacific blue marlin assessment. Again, congratulations and thanks to the assessment WG for their comprehensive work. I note that the WG met several times in 2013 to prepare this assessment, and the results reflect this extra effort.

However, there are a few general and specific areas where the report could be even stronger. The continuity between this assessment and the previous one could have been improved by a review of the previous methods used, and a discussion of why different assessment approaches were adopted in the current assessment. I am sure there are compelling reasons, but they are not clearly enunciated in the report.

As I reviewed the assessment, it became clear that there were important uncertainties in the model inputs. These include the absence of sex structure in the catch, stock structure, growth, and natural mortality. These uncertainties are not adequately discussed in the report.

A discussion of key uncertainties should naturally lead to research recommendations that could guide future research. However, a clear and complete list of recommendations appears to be lacking, or buried within the text. Given the long period that seems to pass between assessments, the lack of clear research directions that could be used by national scientists to help lever support for key applied research seems to be an unfortunate omission.

There was one puzzling aspect of the assessment for me that I could not reconcile. The Kobe plot indicates that in spite of substantial total annual removals that ranged from <10,000 t to >25,000 t, the stock has only been slightly overexploited for a few years relative to MSY targets (F_{msy} = 0.32). I found this somewhat surprising, given the scale and contrast in annual catch and I wondered if the productivity of the resource was being correctly specified in the model.

The rest of the report provides more detailed commentary on specific strengths and weaknesses I found in the 2013 assessment of Pacific blue marlin, following the terms of reference for this desk review (see Appendix 2).

3. *Review of the assessment methods*

The assessment makes appropriate and laudable use of multiple assessment methods, including state of the art approaches such as SS3. The base case model used catch information from 1971 to 2011. The WG also evaluated an alternative SS3 model variant that included a longer time series of catch information than the base case, and a pooled sexes, age structured model. Finally, the assessment WG also presented results from a hybrid Bayesian production model that combined results from an earlier analysis (BILLWG 2013).

While the assessment provided comprehensive results using a suite of relatively new methods, I expected to see more explanation of why such methods were adopted, and why they were preferable to older approaches. Such documentation make assessments more credible for decision-makers and other clients, at least those with long enough memories to recall the previous assessment. For example, Multifan CL was used as recently as 2003 (Hinton et al. 2003). Why was this approach not considered for the current assessment? What advantages are derived from the use of SS3?

As an aside, other RFMOs are considering approaches such as SS3, but are somewhat reticent to fully adopt them. The reluctance is due to the consideration that SS3 has a steep learning curve, and there are relatively few meeting participants that can contribute in a meaningful way to the modeling work. This has been the case with older methods as well, but it appears to me that the situation is not improving and getting worse. This is a somewhat peripheral concern, but I would recommend that the ISC consider the implications of adopting complex modeling approaches that limit the number of national scientists that are proficient with the method. Perhaps the organization has anticipated this, and taking the obvious mitigation steps of training initiatives in the new methods.

The assessment WG compared a variety of assessment methods, and the results are shown in Fig. 5.6. One logical inconsistency with the analyses was the length of time series included. To isolate the effect of methods on the interpretation of stock status, it seems appropriate to include the same input data in the various methods. But as shown in Fig, 5.6, this was not the case as the age structured and hybrid production models included longer time series than did the base model. Since I understand that there were problems identified with the earlier data, I had difficulty understanding why the Group decided to use the earlier information for this purpose.

The assessment WG also used the base SS3 model to complete a very comprehensive set of sensitivity analyses, examining the impacts of both data (including the length of catch time series), and key biological parameters. I found this part of the assessment to be very well done. The work was clearly documented, and the choice of the sensitivity runs was logical.

A within-model retrospective analysis was also completed, but not a comparison of estimates obtained from earlier assessments. Given the time interval between the assessments, this was an understandable decision by the WG.

4. *Assessment model configuration, assumptions, and input data and parameters*

Stock Structure

The assessment follows the convention established by the previous assessment (Hinton 2001), and treats the stock as a Pacific-wide management unit. While this is also consistent with the available genetic information (Graves and McDowell, 2003), there is increasing evidence from satellite archival tagging studies of other billfish species that fish sometimes follow complex and region-specific migration patterns that are at odds with the assumption of stock structure assumed by the responsible RFMO (Neilson et al. 2009). In addition, conventional tagging results available for Pacific blue marlin (Sippel et al. 2013) do not appear to be completely consistent with a well-mixed pan-Pacific stock structure.

Later in the desk review presented here, size at maturity is found to vary regionally. Some of the differences in growth rate could also conceivably reflect regional differences (but other explanations are possible, of course). If these regional differences are supported by future study, it implies that there are stock components with varying productivity. Such components may be over- or under-harvested by making the assumption of a single Pacific-wide stock.

As recognized by the assessment Working Group, further study of stock structure should be a high priority for Pacific blue marlin.

Catch and Effort Data

A major consideration for the assessment Working Group must have been the treatment of the Japanese C/E data prior to 1971. It seems like much thought has gone into the decision to abandon the older information, and given the concerns of changing fishery distribution and species identification, the action taken by the WG seems appropriate.

Typically, I expect to see some statements concerning the completeness of the catch data from the various fisheries, and whether major assumptions had to be made. This is often the case for the most recent year included in the assessment, as some countries may have failed to report their annual catch. Within the RFMO I am most familiar with, we usually spend considerable time explaining the gaps in the available catch and effort data, and developing substitutions to cover the deficiencies. Perhaps the data situation for Pacific marlin is better than in my experience, but it would be nice to have that confirmed.

Still on data gaps, I was not clear how the WG dealt with the apparent absence of catch information from Japan in 2011 (see Fig. 1). What assumptions were made to cover the lack of reporting of the 2011 information?

Fig. 2 seems to suggest that there was no longline catch reported in 2011, which is not consistent with Fig. 3.2(b). Probably Fig. 2 needs updating.

Since Pacific marlin are not subject to management measures, I assume that discards, either live or dead, are not a major part of the catch. This should be noted explicitly, however.

Maximum Age

The maximum age included in the catch matrix was 26 years, and that age also served as a “plus group”. As noted in the assessment WG report, longevity estimates remain tentative and unvalidated. A possible research recommendation would be to conduct bomb radiocarbon studies of archived hard parts (if available, particularly from the largest specimens) to ascertain both longevity and provide validation of age estimates. Such work is best performed with otoliths, if they are available, although based on some preliminary work done on bluefin tuna within ICCAT, there is some prospect of doing this work with fin spine sections as well.

Weight at length

This aspect of the assessment is straightforward and well-explained in Brodziak (2013), who provided sex-specific relationships.

Sex specificity

The assessment used sex-specific growth equations and weight-length relationships, but the sex of the catch was not known. This source of uncertainty was identified in the Executive Summary of the Assessment WG report. Because available size data were not identified to sex, the underlying assumption of selection by sex is that fish are equally vulnerable and taken by fisheries in a well-mixed ocean.

Apparently, the analyses assumed a 1:1 sex ratio in the catch regardless of size or region. In some billfish species such as swordfish, sex ratios in the catch can vary significantly as a function of area, fish size and season.

Recruitment and reproduction

The assessment provided estimates of recruitment at age 0. It is interesting that the largest annual recruitment occurred at very low biomass in 2009, an observation which is not made in the report. In fact, the stock and recruitment relationship is not shown in a plot, which would have been helpful.

Relative Abundance Indices

a. Japan longline (Kanaiwa et al. 2013)

This significant document included a catch rate standardization that included habitat availability, a relatively rare feature of such analyses (HBS model). The HBS model also included information on gear configuration. Given that blue marlin are a bycatch species, having such information is likely to be very important.

For stock assessment purposes, the authors recommended the use of a delta GLM standardization for the years 1975 through 1993, and the HBS approach from 1994 to 2011. The authors' recommendation seems to have been adopted without change by the assessment Working Group. Given the careful analyses evident in the Kanaiwa et al. (2013) document, this seems to be fully appropriate.

b. Taiwan LL (Sun et al. 2013)

The Taiwanese CPUE data are noteworthy, in that the time series covers the longest duration and covers the stock range well. Unlike other standardizations that I have seen which more typically include mixed models to better account for zero observations, this paper uses a GAM approach. The authors note that due to changes in targeting, their analyses was split into three discrete time periods. However, considering the diagnostic plots for the earliest time period, there appear to be some problematic aspects to the model fit (Fig. 5 of Sun et al. (2013) -- serially trended residuals, poor overall fit).

Notwithstanding these matters (which I don't consider to be major concerns), the assessment WG appeared to adopt the time series and recommendations by the authors with no change.

c. Hawaii LL (Walsh et al. 2013)

Significantly, this catch rate series was the only one derived from observer information, whereas the other series were obtained from logbook information. I would consider that observer information is generally of higher reliability compared with logbook data. Moreover, Walsh et al. (2013) describe the PIROP observer data series used in their analyses as being "now the largest pelagic observer program for longline fisheries in the Pacific Ocean".

However, the assessment WG apparently discounted this series in the base case model. Given the noticeable decline in the CPUE, this was a very important decision for the Group. Understanding the rationale for the exclusion is essential. Reading the report of the assessment WG, there is mention of the low rate of coverage (P.24) during the early years of the series. From Walsh et al. 2013, they note the coverage rate initially was about 5%, increasing to about 22%. While those coverage rates are not particularly high, I have seen worse. More importantly, such information should not be discounted unless there is reason to be believe that the observer deployments did not mirror overall fleet activity.

In my view, this is one of the most important potential concerns with this assessment. In my opinion, a better rationale should have been provided to justify this important and potentially influential exclusion. However, to the credit of the WG, they did examine the potential effect of including the HW data series. It is noteworthy that inclusion of Hawaii series resulted in a considerably more negative interpretation of stock status in recent years (Fig. 5.11).

Considering the available abundance indices for Pacific blue marlin, it is perhaps useful to recall a conclusion from the 2011 Atlantic blue marlin assessment completed by ICCAT (Anon. 2012): "These results highlight the fact that there is no information in the relative abundance indices that can be unequivocally used to determine how productive the stock is." This observation highlights the difficulty of using CPUE data for non-target fisheries, such as blue marlin. Unfortunately, there is no alternative, such as fishery-independent surveys. The observer data mentioned earlier does however comprise the next-best alternative, and those data should be further evaluated as part of the next assessment.

Natural Mortality (M)

The age-structured analyses completed by the WG used age- and sex-specific estimates of M provided in Lee and Chang (2013). Those authors found that "adult M estimates derived from a random effects inverse variance weighting of each method were 0.22 yr^{-1} (95% CI: 0.13-0.30) for female and 0.37 yr^{-1} (95% CI: 0.28-0.46) for male. M for both males and females was 0.42 yr^{-1} at age 0 and 0.37 yr^{-1} at age 1. M for adult males at age 1+ maintained at 0.37 yr^{-1} and female M decline linearly to 0.22 yr^{-1} at fully mature age 4 and maintained at this level afterward."

The Lee and Chang (2013) paper appears to be a comprehensive meta-analyses of existing studies of natural mortality. Moreover, there is a logical basis for the evolution of the study and its inclusion in the stock assessment, given recent improvements in the understanding of the growth and life history of Pacific blue marlin. However, given the range of the confidence interval estimates reported in the meta-analyses (see Table 3), it cannot be claimed that we know the true values of natural mortality with much precision. I also note that the age-specific M estimates were obtained by making certain assumptions of how mortality varies with age and size. Finally, I was not clear if the variance-weighted approach employed by the authors was well advised for comparing among the various studies, since it was not clear to me that the variance reported in the papers was computed in a comparable fashion. For these reasons, sensitivity analyses are clearly required (and this must have been recognized by the WG as well).

Moreover, for a very large teleost, I found those estimates to be surprisingly high. By comparison, the most recent assessment of blue marlin in the Atlantic was based on a relationship between natural mortality (M) and maximum age (Hoenig (1983)). For the purpose of the 2011 Atlantic blue marlin assessment, an M value of 0.139 (for both sexes) was used assuming a maximum age of 30 years (Anon. 2012). Lee and Chang provide a useful overview of various assumptions used in assessments (Table 4), but they incorrectly note that the 2011 ICCAT assessment assumes an M value, whereas it uses an indirect (Hoenig) method based on maximum age.

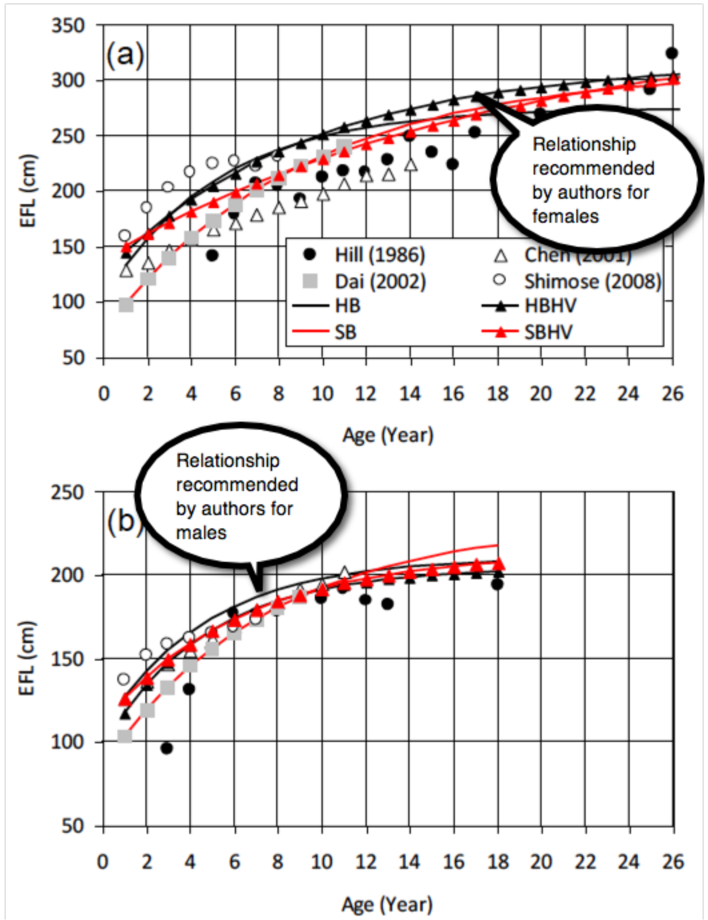
I would also note that previous assessments have found it difficult to estimate natural mortality with much precision. Hinton et al. (2002) estimated natural mortality to be approximately half of the previous year's estimate: 0.18 yr-1 in this analysis versus 0.38 yr-1 in the previous analysis.

It is noteworthy that the Pacific Blue Marlin WG completed a sensitivity analyses to evaluate the impact of different assumptions concerning M. It was found that the assessment is sensitive to the choice of M, and if a lower M schedule was included, stock status is less optimistic (the lower *M* scenario produced a lower level of spawning biomass and a higher level of fishing intensity (Figure 5.12.b)). This particular source of uncertainty does not appear to be well-communicated in the management advice.

Growth

The WG was faced with a number of incomplete studies of growth of Pacific blue marlin that presented quite disparate results (see Fig. 1 of Chang et al. (2013)). None of the available studies contained any attempts to validate the age interpretations using methods such as bomb radiocarbon (Neilson and Campana 2009).

In an attempt to synthesize the various sources of information into a single growth model for male and female blue marlins, Chang et al. (2013) used Bayesian hierarchical methods. Those authors concluded that the parameters of female HBHV (base hierarchical with heterogeneous variance) model and male HB (base hierarchical) model are appropriate for use in stock assessment modeling. However, when one examines the fit of the models to the sizes at ages reported in the various studies (see figure from Chang et al. (2013)) reproduced below, it is difficult to see how the authors reached those conclusions. Most of the data appear to lie below the model fit, leading to the impression that the majority of studies point to smaller average sizes at age than derived by Chang et al. (2013). It could also be argued that highly variable and incomplete results such as the Hill (1986) study should have been excluded from the exercise in the first place.



Review Fig. 1. Fig. 1 of Chang et al. (2013), modified to emphasize the authors' recommendations for the growth relationship to be used in the stock assessment.

The assessment WG adopted the recommendation by Chang et al (2013), noting that "Their hierarchical model with homogeneous variance (HBHV) for females was used in the assessment because the estimate of size-at-age one (144 cm) was very close to the estimated mean size (146 cm, CV = 7%) from Shimose (2008)". To me, this seems to be an incomplete basis for the adoption of the growth model. What about the size at asymptotic length predicted from the model, compared with the largest known individuals caught in the fisheries? For males, the assessment WG modified the recommendation of Chang et al. (2013) by fixing the length at age 1 to be the same as the females, based on otolith microstructure investigations by Shimose that indicated no sexual dimorphism in growth during the first year of life.

I also note that the estimates of $L_{infinity}$ were 226 and 316 cm for males and females respectively. In contrast, the 2011 Atlantic blue marlin assessment used values of ~180 and ~250 cm (read by eye from a plot) for males and females respectively. These seem rather large differences for conspecific stocks, even if they are in different oceans.

The WG performed sensitivity analyses where the $L_{infinity}$ was varied, and found that the results of the assessment were sensitive to the growth assumptions, thereby confirming the importance of a better understanding of this component of the assessment.

It seems clear to me that more work is needed to clarify the growth of Pacific blue marlin. In particular, validation studies should be pursued. This should be a highlighted recommendation.

Size at Maturity

Based on samplings from the western Pacific, Sun et al. (2009) recommended that the estimated sizes-at-maturity (EFL50) were 179.76 ±1.01 cm (mean±standard error) for females and 130±1 cm EFL for males. Those authors, however, noted that there are regional differences in this parameter throughout the Pacific. They attribute this to environmental, genetic or even methodological differences. I would note that genetic differences do not seem likely, given the genetic evidence available to date summarized by Graves and McDowell (2003).

The assessment WG adopted the values noted above for the assessment, albeit with re-parameterization to accommodate the logistic function for maturity used within the SS3 model. I could find no mention of potential regional differences in this value, and impacts on the assessment, nor was this identified as a research recommendation for future assessments.

5. *Proposed population benchmarks and management parameters*

As noted in the assessment WG report, a suite of candidate F-based biological reference points (F_{MSY}, $F_{20\%}$, SPR_{MSY}) were estimated in this assessment, and point estimates were provided along confidence intervals. The expected biomass levels corresponding to those benchmarks were also provided. No target or limit reference points have been established for the Pacific blue marlin stock by the responsible RFMOs.

6. *Adequacy, appropriateness, and application of the methods used to project future population status.*

Projections were completed for the base SS3 model only. The assessment WG made standard deterministic projections using various future harvest levels, and the results are displayed in "chicken-feet" plots.

The stock recruitment relationship was used to generate future recruitments, but I could find no figure showing the SR relationship. The average weight at age used in the projections was not documented, nor were any assumptions regarding possibly incomplete 2011 catches. In summary, some of the usual specifications of projections appeared to be missing.

Another concern is that presumably our confidence in the 2009 estimate of recruitment might not be that high, given that we have only three years of CPUE observations, and presumably the estimates at age 0 are highly variable. This is perhaps noteworthy, as this record-high recent recruitment will be responsible for the near-term optimistic projection results.

7. *Conclusions and Research priorities to improve the understanding of essential population and fishery dynamics necessary to formulate best management practices.*

This is one of the few areas where I could find significant fault in the assessment. In my view, the WG could have spent more time elucidating research priorities. Given that there is a very long time between assessments, I think it is critically important to identify research priorities.

a. *Assessment methods*

The methodologies employed by WG are complete and state of the art. I have no suggestions for improvement in this regard, apart from more complete documentation of what advantages new methods provide compared with approaches used in the earlier assessments.

b. *Assessment model configuration, assumptions, and input data and parameters*

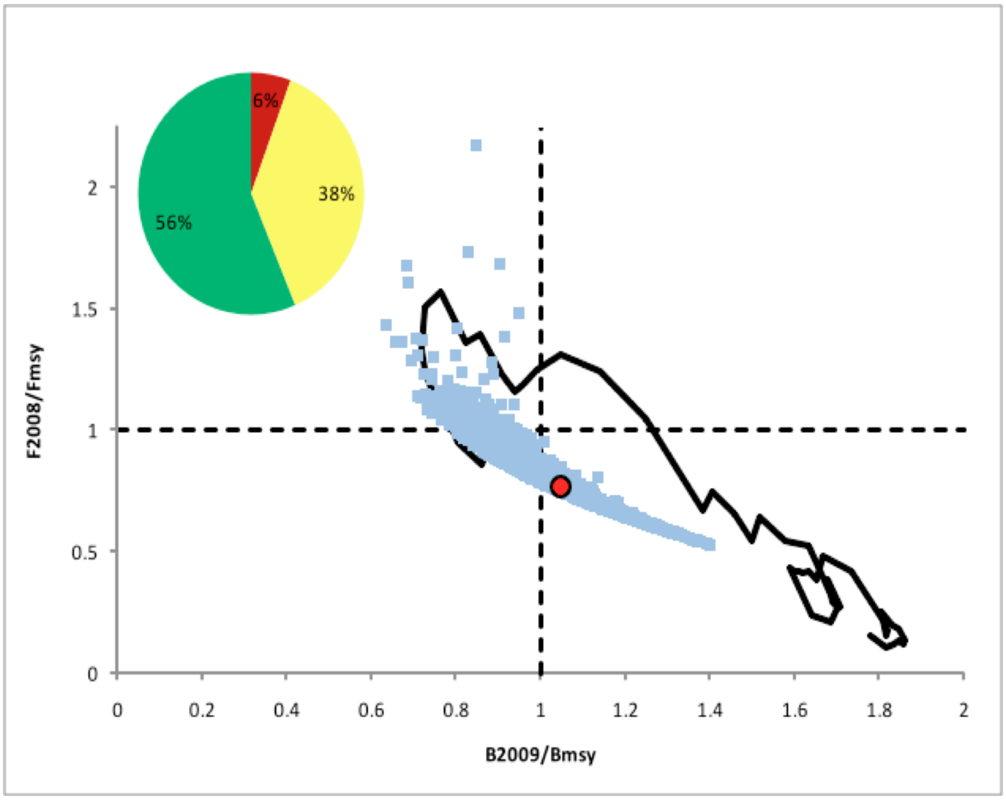
Confidence in the assessment results could be substantially improved by inclusion of tagging data and more complete as well as sex-specific sample data. I notice this is a long-standing recommendation (see Kleiber et al. 2002). It would be important for the WG to document why progress is not being made and present proposals to address this important deficiency. I would consider this to be a high priority item.

The assessment WG could consider developing a stronger rationale for the inclusion of the PIROP data from Hawaii. It seems very unfortunate that this rather large database from trained observers is not made better use of. I am also concerned that inclusion of these data would result in a considerably more pessimistic view of the resource. For these reasons, the deficiency should be addressed in the report if possible. The assessment WG should also consider how these data could be used in the next stock assessment. Again, this is a high priority area to address.

Given the considerable uncertainty in both growth and natural mortality, and the sensitivity of the assessment results to variations in those parameters, it is also clear that more work needs to be done on both parameters. I would suggest growth studies are a high priority, and investigations of natural mortality are an intermediate priority, given that studies of natural mortality are notoriously difficult, especially compared with growth.

c. Proposed population benchmarks and management parameters

The population trajectory was shown in relation to MSY targets in Kobe style plots (Fig. 4). A small improvement in presentation could be realized, I would suggest, by displaying the uncertainty in the terminal year of the assessment. I attach an example from the 2009 Atlantic swordfish assessments. Decision makers can see that while the point estimate for the terminal year of the assessment indicates the stock was rebuilt, there was considerable uncertainty concerning the conclusion. Fortunately, the 2013 swordfish assessment confirmed the interpretation!



Review Fig. 2. Example of Kobe plot including uncertainty in point estimates of terminal year in the assessment, taken from Neilson et al. 2013.

d. Adequacy, appropriateness, and application of the methods used to project future population status.

The Assessment WG report noted that "Fishing at the current level ($F_{23\%}$) or at the *MSY* level ($F_{18\%}$) should provide an expected safe level of harvest, where the average projected catches between 2012 and 2020 is close to *MSY*." Given the observed retrospective pattern, is this conclusion really accurate? I would have been more comfortable with this conclusion if the WG had specifically included the impact of the observed retrospective pattern in the short-term projections, as well as its conclusion that the stock is increasing slightly.

Literature Cited

Anon. 2012. Report of the 2011 Blue Marlin Stock Assessment and White Marlin Data Preparatory Meeting (*Madrid, Spain – April 25 to 29, 2011*). Collect. Vol. Sci. Pap. ICCAT, 68(4): 1273-1386. Available at:

http://www.iccat.int/Documents/CVSP/CV068_2012/no_4/CV068041273.pdf

Billfish Working Group (BILLWG). 2013b. Report of the Billfish Working Group Meeting, 20-28 May, 2013, Shizuoka, Japan. Annex 9. Report of the Thirteenth Meeting of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean, Plenary Session. 17-22 July, 2013, Busan, Korea. Available at:

http://isc.ac.affrc.go.jp/pdf/ISC13pdf/Annex%209%20-May_2013_BILLWG%20report_FINAL[1]_22JULY13.doC_.pdf

Graves, J.E. and J.R. McDowell. 2003. Stock structure of the world's istiophorid billfishes: a genetic perspective. Marine and Freshwater Research 54: 287-298.

Hinton, M.G. 2001. Status of blue marlin in the Pacific Ocean. IATTC Stock Assessment Report 1: 284-319.

Kleiber, P., M.G. Hinton, and Y. Uozumi 2003. Stock assessment of blue marlin (*Makaira nigricans*) in the Pacific using MULTIFAN-CL. Marine & Freshwater Research, 54(4): 349-360.

Neilson, J.D. and S.E. Campana. 2008. A validated description of age and growth of western Atlantic bluefin tuna (*Thunnus thynnus*). Can. J. Fish. Aquat. Sci. 65: 1523-1527.

Neilson, J.D. Smith, S., Royer, F., Paul, S.D., Porter, J., and M. Lutcavage. 2009. Investigations of horizontal movements of Atlantic swordfish using pop-up satellite archival tags. J.L. Nielsen et al. (eds.), Tagging and Tracking of Marine Animals with Electronic Devices, Reviews: Methods and Technologies in Fish Biology and Fisheries 9: 145-159.

Neilson, J.D., Arocha, F., Cass-Calay, S., Mejuto, J., Ortiz, M., Scott, G., Smith, C., Travassos, P., Tserpes, G., and Andrushchenko, I. 2013. The Recovery of Atlantic Swordfish: The Comparative Roles of the Regional Fisheries Management Organization and Species Biology, Reviews in Fisheries Science, 21:2, 59-97 Available at:

http://dx.doi.org/10.1080/10641262.2012.754842

Sippel, T., Wraith, J. and Kohin, S. 2013. A summary of blue marlin conventional tag recapture data from NMFS-SWFSC Cooperative Billfish Tagging Program in the

Pacific Ocean. Working paper ISC/13/BILLWG-1/08 submitted to the ISC Billfish Working Group Workshop, 16-23 January 2013, Honolulu, Hawaii, USA, 7pp. Available at:

http://isc.ac.affrc.go.jp/pdf/BILL/ISC13_BILL_1/ISC13BILLWG-1-08.pdf

Appendix 1: Bibliography of materials provided for review

Chang, Y.-J., Brodziak, J., Lee, H.-H., DiNardo, G., and Sun, C.-L. 2013. A Bayesian hierarchical meta-analysis of blue marlin (*Makaira nigricans*) growth in the Pacific Ocean. Working paper ISC/13/BILLWG-1/02 submitted to the ISC Billfish Working Group Workshop, 16-23 January 2013, Honolulu, Hawaii, USA, 23pp. Available at:

http://isc.ac.affrc.go.jp/pdf/BILL/ISC13_BILL_1/ISC13BILLWG-1-02.pdf

Kanaiwa, M., Kimoto, A., Yokawa, K., and Hinton, M.G. 2013. Standardized abundance indices for blue marlin (*Makaira nigricans*) in Pacific Ocean from analyses of catch and fishing effort from offshore and distance water longline vessels of Japan. Working paper ISC/13/BILLWG-1/05 submitted to the ISC Billfish Working Group Workshop, 16-23 January 2013, Honolulu, Hawaii, USA, 87pp. Available at:

http://isc.ac.affrc.go.jp/pdf/BILL/ISC13_BILL_1/ISC13BILLWG-1-05.pdf

Lee, H.-H. and Chang, Y.-J. 2013. Age-structured natural mortality for Pacific blue marlin based on meta-analysis and an ad hoc mortality model. Working paper ISC/13/BILLWG-1/07 submitted to the ISC Billfish Working Group Workshop, 16-23 January 2013, Honolulu, Hawaii, USA, 19pp. Available at:

http://isc.ac.affrc.go.jp/pdf/BILL/ISC13_BILL_1/ISC13BILLWG-1-07.pdf

Sun, C.-L., Chang, Y.-J., Tszeng, C.-C., Yeh, S.-Z., and Su, N.-J. 2009. Reproductive biology of blue marlin (Makaira nigricans) in the western Pacific Ocean. Fish. Bull. 107: 420-432.

Sun, C.-L., Su, N.-J., and Yeh, S.-Z. 2013. Standardizing catch and effort data of the Taiwanese distant-water tuna longline fishery for blue marlin (*Makaira nigricans*) in the Pacific Ocean, 1967-2011. Working paper ISC/13/BILLWG-1/09 submitted to the ISC Billfish Working Group Workshop, 16-23 January 2013, Honolulu, Hawaii, USA, 13pp. Available at:

http://isc.ac.affrc.go.jp/pdf/BILL/ISC13_BILL_1/ISC13BILLWG-1-09.pdf

Walsh, W., Chang, Y.-J., and Lee, H.-H. 2013. Catch statistics, size compositions, and CPUE standardizations for blue marlin (*Makaira nigricans*) in the Hawaii-based pelagic longline fishery in 1995-2011. Working paper ISC/13/BILLWG-1/13 submitted to the ISC Billfish Working Group Workshop, 16-23 January 2013, Honolulu, Hawaii, USA, 79pp. Available at:

http://isc.ac.affrc.go.jp/pdf/BILL/ISC13_BILL_1/ISC13BILLWG-1-13.pdf

Appendix 2: A copy of the CIE Statement of Work

Attachment A: Statement of Work for Dr. John Neilson

External Independent Peer Review by the Center for Independent Experts

Pacific Blue Marlin Assessment Desk Review

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 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: The International Scientific Commission (ISC) will be completing a Pacific blue marlin stock assessment in July 2013. The assessment provides the basis for scientific advice on the status of the Pacific blue marlin stock and will be the 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 Western Pacific Regional Fisheries Management Council (WPRFMC). An independent peer-review of the assessment will provide valuable feedback to the ISC in conducting future assessments. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**.

Requirements for CIE Reviewers: Three CIE reviewers shall have the necessary qualifications to complete an impartial and independent peer review in accordance with the statement of work (SoW) tasks and terms of reference (ToRs) specified herein. The CIE reviewers shall have expertise in population modeling, stock assessment, and billfish stock assessments to complete the tasks of the peer-review described herein. 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 participate and conduct an independent peer review as a desk review; therefore travel will not be 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 contact information 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 assessment and other pertinent background documents for the peer review. 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 shall 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 impartial and independent peer review in accordance with the tasks and ToRs specified herein, and each ToRs must be addressed (**Annex 2**).
3) No later than January 29, 2014, 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, via email to Dr. David

Die 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.

23 December 2013	CIE sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact
7 January 2014	NMFS Project Contact sends the CIE Reviewers the assessment report and background documents
11–25 January 2014	Each reviewer conducts an independent peer review as a desk review
29 January 2014	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
12 February 2014	CIE submits CIE independent peer review reports to the COR
19 February 2014	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) The CIE report shall completed with the format and content in accordance with **Annex 1**,
(2) The 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: 301-427-8155

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:

Gerard DiNardo
2570 Dole Street
Honolulu, HI 96822-2396
gerard.dinardo@noaa.gov Phone: 808-983-5397

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.

   The CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The CIE independent report shall be an independent peer review of each ToRs, and shall not simply repeat the contents of the summary report.

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

Pacific Blue Marlin Assessment Desk Review

1. Review of 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 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.
3. 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.
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 necessary to formulate best management practices.