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

Daniel Gaertner

March 2014

Executive summary

The model used for the Pacific Blue Marlin stock assessment (SA) is the Stock Synthesis (SS3). Stock Synthesis is a sex-specific, size-based, age-structured, integrated statistical stock assessment model. There is no doubt that SS3 shows most flexibility in comparison to other integrated models, allowing for instance to reduce the number of parameters to estimate by fixing some of them. This option was chosen by the Pacific Blue Marlin Working Group (WG) for several biological parameters (natural mortality and growth by sex, fecundity, length-weight relationship), previously estimated from robust meta-analysis methods and reintroduced with their associated uncertainty in the SS3 runs.

In addition to these biological parameters assumed to scale the population subcomponent of the system, basic inputs of SS3 are catch per year (including, if possible, discards), CPUEs used as apparent indices of abundance and size frequency data.

Due to the eastward expansion of the major fishing gear (specifically in the case of the Japanese longliners) used to calibrate the abundance index, and potential misidentification between two species of marlins reported in logbook catches, the SA WG decided to omit data before 1970. Consequently, the SA model did not assume that the first year of observations depict a virgin biomass but rather an equilibrium state at the beginning of the serie considered. The longline fisheries involved in the CPUEs component were the Japanese, Taiwanese and Hawaiian fleets. To account for changes in selectivity over time, these CPUEs series were segmented in homogeneous periods of years before the standardization procedure. There are however many questions pending with regard to the standardization of these three series of CPUEs.

With the aim of ensuring coherence between the trends over time depicted by the different series of CPUEs, correlation analyses between the series were done and a down-weighting process was applied by the model as recommended by Francis (2011). In case of doubts in some inputs, sensitivity runs omitting these data were also considered and the consequences on the final estimates were evaluated by the SA WG.

In light of SS3 outputs, the SA WG concluded that there is no evidence of overexploitation of Pacific Blue Marlin since the beginning of its exploitation. This perception of the state of the stock is not surprising if we take into consideration that standardized CPUEs do not show trend in the recent period and due to the fact that the new estimates of the natural mortality are larger than estimates used in previous SA. The accuracy in the quality of the inputs (mainly catch data and abundance indices) with respect to SS3 will be discussed in the different sections of this review.

It should be noted that traditional and useful tables and figures commonly used to support the description of the fishery are lacking in the Pacific Blue Marlin SA Report. With the exception of catch per gear, standardized CPUEs and size frequency, there are no fishery indicators (e.g., spatial distribution of catch per gear by decade, changes over time in nominal fishing effort for the main fleets/gears, changes in mean weight of Blue Marlin per fleet*gear, changes in nominal CPUEs, apparent movements from release-recapture events, etc.). Such data-driven indicators are very useful to describe objectively the current state of the fishery system, as well as to summarize the occurred historic changes. In the lack of this explanatory phase, the reader of the Pacific Blue Marlin SA Report has just in hand the model-driven perception of the stock.

Findings by Terms of Reference

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

It is argued that SS3 is highly scalable from data-weak situations. Furthermore, it accounts for complex model configurations such as multiple areas and multiple growth morphs and it can flexibly incorporate change over time in key parameters in response to environmental and fishery factors (Methot and Wetzel, 2013). There is no doubt that SS3 shows most flexibility in comparison to other integrated models, allowing for instance to reduce the number of parameters to estimate by fixing some of them. This option was chosen by the Pacific Blue Marlin WG for several biological parameters (natural mortality and growth by sex, fecundity, length-weight relationship). However, in practice because model-based management approaches dominate in tuna regional fishery management organizations (RFMOs), participants to SA Working groups are reluctant to introduce fixed parameters and let the model itself extracts from the data the underlying information assumed to depict accurately the fishery system. From my point of view, one of the risks in using sophisticated integrated model in weak data situation is to try to feed the model by substituting any lacking parameter by some proxies. This objective is suitable *a priori* but when much information is not available the modeling tool is constrained to navigate between the modeler's assumptions used to build the latent processes assumed to represent the fishery system. The Pacific Blue Marlin SA WG was aware of these difficulties and clearly highlighted the quantitative and qualitative limits of the available information in different parts of the Annex 10 "Stock assessment of Blue Marlin in the Pacific Ocean In 2013".

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.

Stock structure

Based on genetic studies only, the Pacific Blue Marlin SA WG notes that there is no evidence of population structuring and consequently assumes a single stock working hypothesis within the Pacific Ocean. This assumption is realistic and could be reinforced by tagging studies. There are low recapture rates in the Pacific Ocean but some long-range movements have been reported (Sippel et al., 2013), as commonly observed in other oceans. Notice than in chapter 2.1.4 (Movement), the description of the fact that most recaptures are taken in the vicinity of their original tag-release location should invalidate the assumption of long-range movements and the single stock working hypothesis. However, from figure 3 (right part) in Sippel et al. (2013) it might be hypothesized that Pacific Blue Marlin depict seasonal migration and come back regularly to the place they were released one or two years before (as suggested by the fact that the apparent distance travelled decreases systematically one year, two years, etc., after release in comparison with observed distances when recaptures are done along the year). For all of these reasons, it should be useful to include figures on tagging results in chapters 2.1.1 (e.g. Figure 1 from Sippel et al., 2013) and 2.1.4. (Figure 3 from Sippel et al., 2013) in the SA report.

Biological parameters

To account for the dimorphism in size between males and females, as well as to combine the variability in the growth parameters estimated by different studies, a meta-analysis of Blue Marlin growth in the Pacific Ocean was conducted within a Bayesian framework (Chang et al., 2013). The authors recognize the simplification of some assumptions (e.g., the choice of the Von Bertalanffy model as the base-case growth model and assuming independent prior distributions for the Von Bertalanffy parameters) but, at this level of knowledge, the results represent an improvement in growth rate predictions and their associated uncertainty when data from different ageing studies are used. It must be noted however that the majority of the growth parameters used as an input in the hierarchical growth meta-analysis were obtained from hard-part readings. A validation of the daily increment rate by tagging experiments, for which also oxytetracycline is injected into the fish, is required to be certain that growth rates estimates are not biased.

For a useful representation of differences in growth and in natural mortality by age between sexes, a figure of the sex-ratio per size class in the SA Report would have been helpful. It is unclear to me why the sex-dimorphism for Pacific Blue Marlin was showed for growth and natural mortality and not for growth only, but I am sure that the authors have good arguments for this. Natural mortality by age and sex was analyzed within the frame of a meta-analysis approach (Lee and Chang, 2013). Estimates of adult M were based on a random effects meta-analysis to synthesize M estimates from a range of estimators. An ad hoc mortality model based on the Lorenzen's relationship was used to rescale juvenile M (age 0) to account for size-dependent processes.

To estimate natural mortality, the authors claim that direct methods using the observed data (e.g. catch data, tagging data) rely on too many assumptions in the case of Pacific Blue Marlin (the paradox is that if this argument is true, the same criticism could be made for the inputs used in SS3). Consequently they prefer to use indirect methods based on estimates of maximum age, life history correlates, and evolutionary-ecology theory. From my point of view, with the exception of methods using only observed values to estimate M but which remain very imprecise (e.g., the relationship between M and the gonad index), the majority of these empirical relationships are derived directly from model-driven values (from estimates of the von Bertalanffy growth rate parameters, or from estimated ages at maturity, maximum age, etc.) which may have not been fully completed and validated. The authors are aware of all these aspects and there is no concern about the statistical meta-analysis which was thoroughly conducted with the objective to integrate different sources of uncertainty. Even if there is a potential bias in the final M estimates due to the potential biases in the inputs derived from these empirical relationships (the meta-analysis accounts for the uncertainty in parameter estimates but not the potential bias), the benefit of this study is to provide a reasonable L-shaped age- and sex-specific estimates of natural mortality rates.

The age-specific estimates of M used as fixed parameters in the assessment were 0.42 year^{-1} for age 0, 0.37 year^{-1} for age 1, 0.32 year^{-1} for age 2, 0.27 year^{-1} for age 3, and 0.22 year^{-1} for age above 4 for female and 0.42 year^{-1} for age 0, 0.37 year^{-1} for age above 1 for male. In previous Blue Marlin SA conducted with Multifan-CL, natural mortality (sex and age independent) was estimated at about 0.18 yr-1 and 0.38 yr-1 depending the year of the SA (Kleiber et al., 2002) which is lower than the values used in the current assessment. These changes may have an impact in the perception of the status of the stock which now may appear more productive than in the past.

With regard to fecundity, as far I know there is large discussion among the scientific community concerning the estimation of batch fecundity for large pelagic fish. To calculate the total number of batches per spawning season, Sun et al. (2009) indicated that the Blue Marlin spawned once every 2–3 days on average from May to September but there is apparent variability in the spawning season depending on the area. In addition, is there any information on the possibility that reproduction might not occur every year (as seen in Atlantic Bluefin tuna)? In the same way of idea, the large variability in the natural mortality for early life history stages in large pelagic fishes (see again Atlantic Bluefin tuna, in Simon et al., 2012) is another argument against the use of the spawning stock biomass as a proxy to predict recruitment levels.

To reinforce this aspect it can be seen from Figure 3 that recruitment depicts a relative stable trend but with an unusual yearly variability for a large pelagic species. Is there any biological reason to explain this type of inter-annual variability? Because of the lack of a marked trend and due to the large confidence intervals associated with yearly recruitment estimates, one can ask if it is wise to attempt to estimate the steepness parameter. An alternative could consist of assuming a constant recruitment (i.e., with no trend) and to fix the value of steepness (may be done in chapter 4.2.6 but unclear to me).

Fishery indicators

One of the major inputs in SS3 is the data catch series. For Pacific Blue Marlin the fishing gear which contributes the most to total catch is longline. There are many reasons to suspect that some longline catches and dead discards of Blue Marlin are non reported: (1) on average, blue marlin is a by-catch for longliners, with a low price value compared to the targeted bigeye tuna, (2) because Asian longliners were judged responsible for the decrease in billfishes in the world ocean by sportfishery associations and because discarding dead animal at sea is not considered to be a good fishing practice by many countries and environmental NGOs. It should be stressed that this assumption is in agreement with the very low report rate of recaptured tagged Blue Marlins observed for these fleets.

Another potential bias, which is not discussed in the SA Report, is the very low contribution to the total catch of the small scale fisheries. In both Atlantic and Indian Oceans, artisanal fisheries are not well evaluated in terms of annual catch of Blue Marlin but are known to contribute significantly to the total catch. To illustrate this point, an anchored FAD artisanal fishery began to catch Blue Marlin in the French Antillean islands in 1985. Catch increased continuously and were around 600-700 t/year in the 2000s for a total catch around 3000-4000 t. These data were not reported to the ICCAT secretariat until the Atlantic Blue Marlin SA was conducted in 2011, and consequently ignored in previous SA. It is suspected that other Antillean small scale fisheries, which may use also anchored FADs, as well as small-scale gillnet fisheries in both sides of the Atlantic Ocean do not report Blue Marlin catches. Similar situation may occur in the Pacific Ocean and the SA WG should have discussed assumptions about the proportion of non-reported Blue Marlin catch in the total catch and their consequences in terms of assessment of the stock.

As mentioned in the Executive Summary section of this review report, simple fishery indicators commonly presented in the executive summary of other tuna RFMOs are lacking. A part the indices cited previously, from size frequency data of the Beverton and Holt's Z estimate, generalized by Gedamke and Hoenig (2006) to allow mortality rate to change in nonequilibrium situations, may be useful to characterize contiguous blocks of years for which catchability is assumed to be constant (Gaertner, 2010). The underlying historic selectivity pattern could be compared with estimates given by other methods (chap. 4.3.1.).

Abundance indices

Assuming an ideal world, most stock assessments would focus narrowly on statistical uncertainties but do not consider whether key fishery indicators might be biased. An example is provided by large pelagic longline CPUEs series whose initial historic period is affected by the hyperdepletion phenomenon (Maunder et al., 2006). The Pacific Blue Marlin WG was aware of this feature and consequently omitted the historic period before 1970. It is unclear however if LL CPUEs up to 1970 reflect accurately the abundance of Blue Marlin in the Pacific Ocean.

Because in general large catch and effort time series area available only for longline fleets, CPUEs from this fishing gear represent the main information used to infer changes in apparent abundance in SA of tuna and tuna-like species. The Blue Marlin SA WG used abundance indices resulting from the standardization of Japanese, Taiwanese and Hawaiian longline CPUEs.

The explanatory variables used for standardizing the Taiwanese data are: year, month, latitude, longitude, and hooks per basket (Sun et al., 2013). Hooks per basket (HPB) is assumed to give insight with the exploration of deeper waters when the target species is bigeye. Even if this variable gives an idea of a such strategy, key information to calculate the maximum fishing depth reached by the mainline depends on (1) the gear configuration: sag ratio, mainline length per basket, (2) the fishing tactics: bearing of the setting, and (3) environmental variables characterizing water mass dynamics:

wind stress, current velocity and shear (Bach et al., 2009). Even HPB was found significant it is unclear whether the short-term strategies (e.g., by trip) exploring or not deep waters and the resulting effect on the catchability of Blue Marlin, is captured by the model. It should be noted than after breaking down the time series in three segments, the adjusted CPUEs of each segment remain very close to the nominal CPUEs (figure 6 in Sun et al., 2013). In such a situation, how can we gauge the benefits of the standardization process?

A general comment can be made about the use of GAM in CPUEs standardization. By allowing nonparametric smoothers to capture the shape of relations between response and the explanatory variables without restricting these relationships to a linear form, GAMs can be considered as an exploratory and visualization tool in complement with GLM analyses. However, it should be suitable to use GAM outputs to suggest parametric transformations of the variables that are substantively interpretable (i.e. when the relationship is as linear as possible) rather than to use directly the transformations estimated in their raw form.

For the Japanese CPUE standardization: delta-GLM, Habitat-based standardization (HBS) approaches were used (Kanaiwa et al., 2013). Since 1994 detailed gear configuration data (i.e., the length of branch lines, the length of the mainline between branch lines, and length of float lines) for each operation of the longline are requested on logbooks by the Japan Fishery Agency (JFA). Such information was used in the HBS to estimate the fishing depths of hooks using the catenary curve model (or a new model type but based on fixed parameters). Unfortunately difference between the maximum fishing depth according to catenary algorithms and the observed mean depths at the set level (i.e., the shoaling) is influenced by set by set conditions, as mentioned in the previous paragraph for the Taiwanese CPUE standardization. In conclusion, this method may give information only on the theoretical dominant fishing strategy by large areas in the Pacific Ocean. The based-GLM standardization (Delta-GLM) was divided in two time periods (1975-1993; 1994-2010) and considered factors as year, HPB, season, geographical region, latitude and longitude in both components of the model. Differences between nominal and standardized CPUEs are larger for the delta-GLM than for HBS (figure 4, 7 and 9, from Kanaiwa et al., 2013). I am not a statistician but the diagnostic plots for the Gaussian components of the Δ-GLM model for 1975 to 1993 (Figure A-6) show a trend in the variability of the residuals over the predicted values which suggest that something is wrong in the model (reinforced by the fact that the normal QQ plot does not show an exact straight line relationship). This could be due to a wrong response distribution. It appears that a group of observations may have a very strong influence on the model fitting as revealed by this plot as well as by the leverage plot. The interpretation of the residuals for the binomial component is much more difficult to analyze.

The Hawaii-based pelagic longline fishery CPUEs series was standardized for the period 1995-2011 (Walsh et al., 2013). The strong decrease of the nominal CPUE (69.9%) was due to an increase from 69.5% to 85.2% in zero catches and a decrease in positive catches from 30.5% to 14.8%. The significant explanatory variables used in different GLM were: years, calendar quarters, fishery sectors, fishing regions, bait types, sea surface temperature (SST) and vessel length. By contrast, hook types, leader materials and bathymetry were not found as significant. It was noted that this fleet which was primilarly targeting swordfish moved to bigeye areas in the beginning of 2000s due to the closure of the shallowest sector between 2001 and 2004. Based on the results of the zero-inflated negative binomial GLM (ZINB), it was concluded that in contrast to the decreasing trend showed by the nominal CPUEs, the standardized CPUEs, which accounted for changes in fishing strategy, remains relatively constant.

It was explained in this paper that the predicted standardized CPUEs were calculated with the R function "predict". I have no idea whether the results might be different but in some tuna RFMOs the LsMean function from SAS, or an R equivalent function, is preferred as it accounts for the non-equilibrium conditions in the sampling design. With regards to the data set used I have some comments. Analyzing the results of the ZINB, the authors showed that there are significant negative coefficients for several years between 2005 and 2011 in the zeros model; that is to say since the reopening of the shallow-set sector in 2004. In addition it is stated that 58.1% of the observed sets in the shallow-set sector since 2005 were deployed in the first and fourth quarters at a mean SST of 18.8°C. Why shallow-sets with SST < 20°C (i.e., assuming that 20°C is the lower limit of the range SST for Blue Marlin) were not omitted before the analysis? Does it make the analysis more objective by using the whole set of data rather than omitting sets with *a priori* no suitable habitat for Blue Marlin? In opposite, standardization methods accounting for changes in the spatial extent of the fleet and including predicted CPUE of unfished areas should be tentatively done in future Blue Marlin SA (Cao, et al., 2011; McKechnie et al., 2013).

In conclusion of the CPUEs standardization, important information to assess changes in catchability over time is lacking. The species of bait has been included only in the analysis of the Hawaiian fishery and was found as a factor affecting catches significantly (see Table A1 in Walsh et al, 2013), with all types of fish and "other" baits yielding significantly lower catches than squid baits. The fact that bait species and type (e.g., frozen, fresh, live) affects catchability and fishing power in longline fisheries was highlighted by Ward and Hindmarsh (2007). Live bait and the shifts in timing of the fishing set (more bait being available at dusk and dawn) are known to have a significant impact in the catchability of blue marlin (Ward, 2008). The same author noted that catch rates of blue marlin were also significantly lower on nylon leader commonly deployed by Japan's longliners in the 1990s. It is strange that these factors were not analyzed in the working documents presented during the Blue Marlin SA WG.

It can be stressed that the assumption that catchability be constant over time (chap. 4.3.2) is a strong assumption, generally not supported by the information concerning introduction of modern technology on board or new practices (Ward and Hindmarsh, 2007; Ward, 2008). However, as Blue Marlin is not targeted by tuna longliners, the different modifications of the gear, the introduction of new technology, different types of bait, etc., do not mean that catchability increased continuously over years. In such complex situation, a table of the date of introduction of the main technology, as proposed by Ward and Hindmarsh (2007; see Figure 2), should have been included in the SA report.

Another aspect that should be considered in the SA is the accuracy of the longline CPUEs to depict changes in abundance for large pelagic fish. This point was raised by Polacheck (2006) with regard to the fact that the initial longline catches were not responsible for a rapid depletion of the main tuna and billfish stocks but could be extended to other situation. In Figure A, I represented the concomitant evolution over time of total Blue Marlin catch (Table 3.2 of the Blue Marlin SA Report) and abundance indices from Japanese, Taiwanese and Hawaiian longliner fisheries (Table 3.3). As the series begins in 1971, the question of the initial decline of longline CPUEs, taken into account by the SA WG, is outside the scope of my comment but it can be seen that for some periods of years the CPUEs fluctuated in a similar way than total catch over time (e.g., index S1, JPN1 1975-1993 and index S6, TWN3 2000-2011) or in contrast are relatively independent (e.g., index S2, JPN2 1994-2011; index TWN1 1971-1978). Within a simple Pressure-State-Response (PSR) framework, it can be assumed that standardized CPUE is a proxy of the apparent abundance (i.e., the state of the natural population) and that total catch is a direct measure of the pressure. Consequently, one can expect that the apparent index of abundance behave differently than these observed patterns.

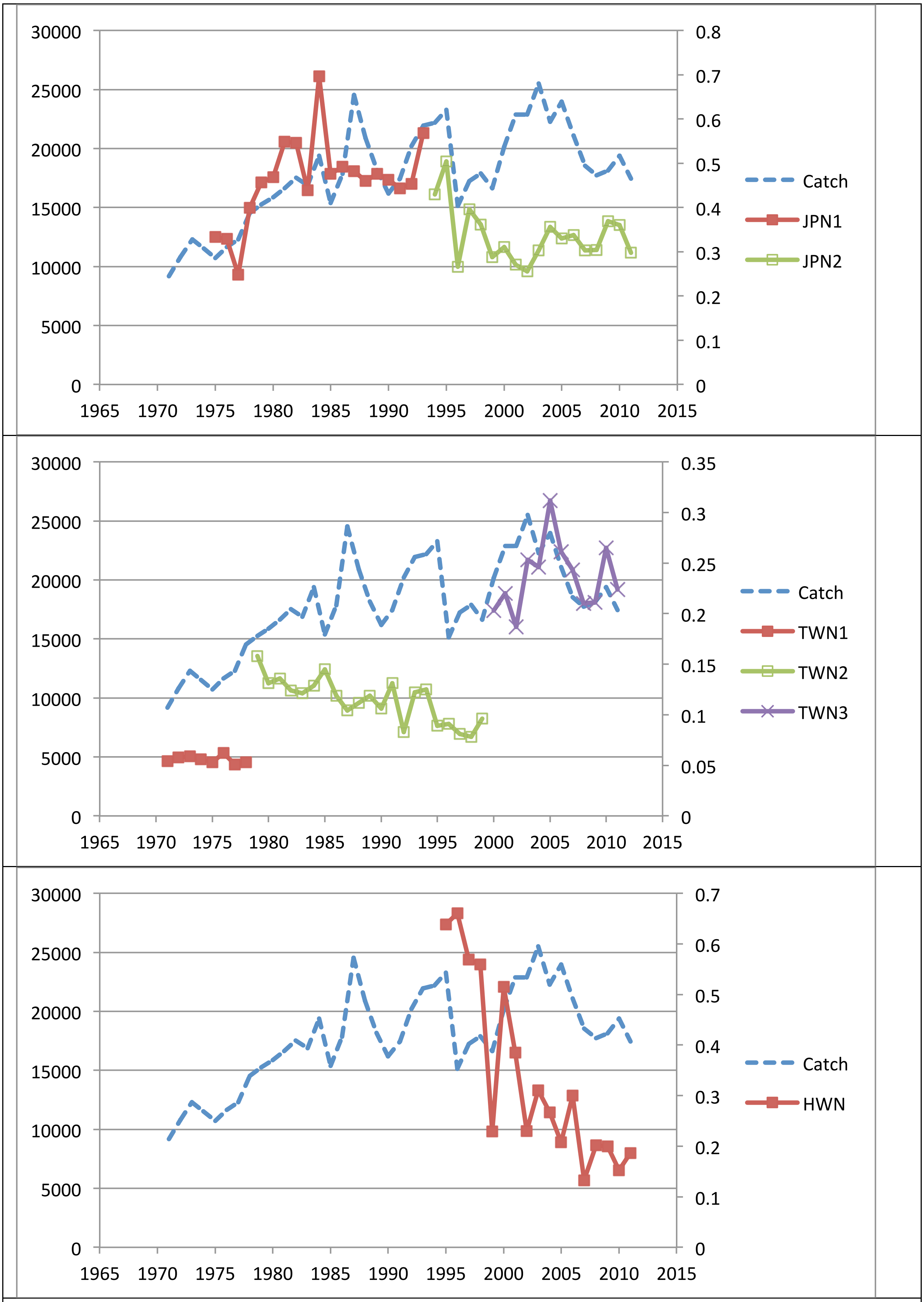

Figure A. Comparisons of the total catch of Blue Marlin in the Pacific Ocean and standardized CPUEs for longliners from Japan (upper panel), Taiwan (median panel) and Hawaii (lower panel). For the purpose of comparison a scaling factor of 0.1 was applied to the 2nd Japanese series.

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.

From Figure 4 (i.e., the Kobe plot) it can be seen that MSY was never reached during the historic period of the fishery. Consequently this Target Reference Point (TRP) was estimated by extrapolation outside the range of observed values which is more uncertain that in a case of stock which suffered a state of overexploitation during a period of its history and for which MSY is estimated by interpolation within the range of observed values. Obviously in the case of Pacific Blue Marlin there is not an alternative to tentatively estimate MSY, but caution should be kept in mind in terms of interpretation of the results, specifically when proposing projections. Due to the large uncertainties in the inputs (catch and effort data, growth and mortality by sex, migratory pattern by sex, changes in selectivity and catchability over time, etc.), it would be proper to consider management decisions based on simple Harvest Control Rules (HCR) and to use SS3 to evaluate the robustness of these decision rules. (Hilborn, 2003).

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

Even if SS3 allows one to project the trend of the population, I am not totally convinced that with good data information it is appropriate to project a population for a period of time longer that the number **of** exploited age classes of the species under study. In case of Blue Marlin for which fishery data are quantitatively and qualitatively weak, projecting female spawning stock biomass and total catch for the next 9 years (2012-2020), may be questionable (Fig. 6); even female Blue Marlin have a long lifespan.

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

The advantage in the use of integrated model is the gain from combining simultaneously multiple data sets through multi-objective likelihoods functions in order to avoid the loss of information, as seen in fishery studies carried out independently (Maunder and Punt, 2013). Nevertheless, one of the major concerns is that this efficiency is strongly related to the amount and the quality of the information available. The methods used to assess the status of the Pacific Blue Marlin focused mainly on quantification of the uncertainty in the integrated estimates of key management quantities, and in a certain way SS3 may account for this aspect. However, there are few, if any, references to potential bias.

I suggest the use of SS3 as a reference tool to explore how poor quality data and partial knowledge in each component of the fishery model (e.g., growth and mortality by length/sex, migration patterns, selectivity pattern, etc.) may affect the perception of the status of the stock. Then separately for each component for which biases in parameter estimates are suspected, attention should be paid to correct them and how to reintroduce the associated uncertainty, whenever possible, into the integrated model. As far I understand separate analyses combining external knowledge in parameter estimates and in their uncertainties were conducted within the framework of meta-analysis for natural mortality (Lee and Chang, 2013), length-weight relationship (Brodziak, 2013) and growth (Chang et al., 2013). This approach was very productive and should be encouraged for other stocks assessment studies but the question of bias remains (specifically for natural mortality by age and sex). Growth parameters obtained mainly by hard-part readings must be validated by tagging experiments. Simple fishery indicators as total mortality, mean weight by gear, etc., should be developed in association with SA models. The potential mis-reporting of catch by small scale fisheries and dead discards should be thoroughly evaluated. Due to potential sex-specific migrations, sex-ratio by spatio-temporal strata and resulting CPUEs analyses could be helpful.

References

Bach P., Gaertner D., Menkes C., Romanov E., and Travassos P. (2009) Effects of the gear deployment strategy and current shear on pelagic longline shoaling. Fisheries Research 95: 55-64

Brodziak J. (2013) Combining Information on Length-Weight Relationships for Pacific Blue Marlin ISC/13/BILLWG-1/01

Cao J., Chen X. J., Chen Y. (2011) Generalized linear Bayesian models for standardizing CPUE of *Ommastrephesbartramii* for Chinese squid-jigging fishery in northwest pacific ocean. ***Scientia Marina***, 75 (4): 679-689.

Francis R.I.C.C. (2011) . Data weighting in statistical fisheries stock assessment models. Can. J. Fish. Aquat. Sci., 68: 1124–1138.

Gaertner D. (2010) Estimates of historic changes in total mortality and selectivity for Eastern Atlantic skipjack (Katsuwonus pelamis) from length composition data Aquat. Living Resour. 23, 3–11

Gedamke T., and Hoenig J. M., (2006) Estimating mortality from mean length data in nonequilibrium situations, with application to the assessment of goosefish. Trans. Am. Fish. Soc. 135, 476–487.

Hilborn R. (2003) The state of the art in stock assessment: where we are and where we are going. Sci. Mar., 67 (Suppl. 1); 15-20

Kleiber P., Hampton J., Hinton M. G., and Uozimi Y. (2002) Update on Blue Marlin Stock Assessment. SCTB15 Working Paper BBRG-10

Maunder M. N., Sibert J. R., Fonteneau A., Hampton J., Kleiber P., and Harley S. J. (2006). Interpreting catch per unit effort data to assess the status of individual stocks and communities. ICES Journal of Marine Science, 63: 1373-1385.

Maunder M.N. and Punt A.E. (2013) A review of integrated analysis in fisheries stock assessment. Fisheries Research 142: 61– 74.

McKechnie S, Hoyle S., and Harley S. (2013) Longline CPUE series that account for changes in the spatial extent of fisheries. WCPFC-SC9-2013/SA-IP-05

Methot R. D. and Wetzel C. R. (2013) Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. Fisheries Research 142: 86-99

Polacheck T. (2006) Tuna longline catch rates in the Indian Ocean: Did industrial fishing result in a 90% rapid decline in the abundance of large predatory species? Marine Policy 30 : 470–482

Simon M, Fromentin J-M, Bonhommeau S, Gaertner D, Brodziak J, and Etienne M-P. (2012) Effects of Stochasticity in Early Life History on Steepness and Population Growth Rate Estimates: An Illustration on Atlantic Bluefin Tuna. PLoS ONE 7(10): e48583.

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. ISC/13/BILLWG-1/08

Ward P., (2008). Empirical estimates of historical variations in the catchability and fishing power of pelagic longline fishing gear. Rev. Fishe ih Biol. Fish. 18, 408–426.

Ward P. and Hindmarsh S. (2007) An overview of historical changes in the fishing gear and practices of pelagic longliners, with particular reference to Japan's Pacific fleet. Rev. Fish Biol. Fisheries, 17: 501-516.

Appendix 1: Bibliography of materials provided for the desk review

Anon. (2013) Stock assessment of Blue Marlin in the Pacific ocean in 2013. Billfish Working Group International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean

Chang, Y.J., Lee H.H., Brodziak J., DiNardo G., and Sun, C.L. (2013) A Bayesian Hierarchical Meta-Analysis of Blue Marlin (*Makaira nigricans*) Growth in the Pacific Ocean. ISC/13/BILLWG-1/02

Lee H.H. and Chang Y.J. (2013) Age-structured Natural Mortality for Pacific Blue Marlin Based on Meta-analysis and an Ad Hoc Model. ISC/13/BILLWG-1/07

Kanaiwa M., Kimoto A. and Hinton M. (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. ISC/13/BILLWG-1/05

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. Fishery Bulletin, 107(4):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. ISC/13/BILLWG-1/09

Sun, C.L., Su, N.J., and Yeh, S.Z. and Chiang W.C. (2013) Stock Assessment of Blue Marlin (*Makaira nigricans*) in the Pacific Ocean using an Age-Structured Model. ISC/13/BILLWG-2/07

Walsh, W. A. 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. ISC/13/BILLWG-1/13

Appendix 2: A copy of the CIE Statement of Work

Attachment A: Statement of Work for Dr. Daniel Gaertner

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.