# External Independent Peer Review of the 2009 Stock Assessment of North Pacific Swordfish

**for**

**Center for Independent Experts**

**Feb 2010**

**Cefas Contract C3668**

**COMMERCIAL IN CONFIDENCE**

# Table of contents

# Executive Summary

Swordfish in the North Pacific are harvested multi-nationally, primarily using longline gear. An assessment of swordfish in the North Pacific Ocean was conducted by staff of the Pacific Islands Fisheries Science Center and collaborating scientists from members of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean (ISC). Results are key to international management decisions of the Western and Central Pacific Fisheries Commission and its Northern Committee, and domestic management decisions by the Western Pacific Regional Fishery Management Council (WPFMC) and Pacific Fishery Management Council.

This document represents the independent CIE Reviewer Report on the assessment. The review was performed at the consultant's offices at Cefas (Lowestoft, UK) as a desk-based study based upon the provided documentation. The Terms of Reference (ToRs) of the peer review are attached in Appendix 2 (Annex 2). Below is a summary of issues raised within each of those TORs.,

As a general comment, it is harder to review an assessment based upon a document, without the opportunity for the assessment authors to respond to questions and comments interactively. Therefore many of the comments and recommendations in this report may fall under the category of "been there, tried that, it didn't work"; but if so, it highlights the need for additional information to be provided to reviewers under the desk-based review format.

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

Bayesian implementations of the surplus production model appear appropriate for Pacific swordfish, where biological information are relatively limited but CPUE time series data are available. This implementation is reasonably well tested and understood within the scientific literature, while WinBUGS software is also well known and considered stable. The Bayesian implementation provides a number of advantages, such as the ability to provide probabilistic risk-based advice to managers. The approach required a number of assumptions, like all methods, but the sensitivity to those assumptions needs to be understood (see later). Ultimately, to help the reader fully understand the model implementation, the WinBUGS code should be provided as an appendix to the report (**Recommendation 1)**.

With the information presently available, the model appears to have fitted the main data sets reasonably well (Japanese CPUE time series in particular). The overall conclusions on stock status appear supported by recent trends in the CPUE data series. However, the report presents what appear to be the results from the 'final' model runs, and there is a general need to understand the performance of the model and robustness of outputs under alternative starting conditions and to uncertainty in input parameters (see recommendations below).

**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.**

Pragmatism requires that the best available data be used to derive stock status estimates and hence advice to inform management, and this appears to be the case within this assessment. To aid the review process, however, information on the methods used to develop the input data (e.g. CPUE time series standardisation) should be provided within the assessment report, particularly where uncertainty exists. This could be as a summary table for the assessment, including judgements on the reliability of data sets (**Recommendation 2**). Indeed, given the fact that results will be predicated on the CPUE data used, further efforts on their standardisation (**Recommendation 3**) and the consideration of catchability trends (**Recommendation 4**) is warranted. Given uncertainties in the standardised CPUE time series, the influence of alternative CPUE data set weightings within the model (e.g. based on their variance) should be examined. This should also allow the influence of CPUE time series outliers on stock assessment results to be identified (**Recommendation 5**). In turn, the potential influence of currently missing data sets from sub-area 2 of the two-stock model should be examined, and if influential and effort data allow, CPUE time series developed (**Recommendation 6**).

Two stock structure hypotheses are tested, appropriate given the available information, which allows the evaluation of the robustness of management advice to this uncertainty. However, the logic behind the parameterisation of these alternative models should be detailed (linked to Recommendation 2).

Appropriate and pragmatic methods have been used to generate prior distributions for key model parameters such as R and K, while that for catchability has been left uninformative so that the data can inform the outputs. However, the influence of the prior mean values selected for carrying capacity (K) under the alternative stock structure scenarios on the model output estimates is unclear, and some sensitivity analysis with alternative prior distributions of appropriate R and K mean values is recommended (**Recommendation 7**). In turn, the influence of the mean starting population status as a proportion of K (90%) on overall estimates of stock status given the trends in CPUE time series is also unclear and should be tested, in particular for stock 2 in the two-area model (**Recommendation 8**).

With regard model diagnostics, the use of convergence and goodness-of-fit criteria are appropriate, but the prior and posterior probability density functions (pdfs) for all parameters should also be presented to show how much influence the priors or the data had on the outputs from the model (**Recommendations 9** and **10**). Overall, model outputs presented appear appropriate given the input parameters, and the good fit to the Japanese CPUE data is reassuring, although poorer fits to shorter time series are noted. However, the authors do not discuss whether the output mean parameter and population estimates are biologically realistic (**Recommendation 11**). In turn, estimates from the 2-area model appear uncertain (stock 2), which has issues for the projection performed. The causes of this uncertainty should be examined (**Recommendation 12**).

While the results of two alternative structural scenarios are modelled (single- and two-stock formulations), there is a need to perform and/or present alternative runs with different prior distributions to allow the reader to judge the sensitivity of model outputs to the assumptions made when developing the priors (**Recommendation 13** and others).

**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.**

The methods used to estimate management parameters (MSY) are standard and appropriate based upon the surplus production method. The Bayesian approach allows probabilistic statements of stock status to be developed and these are clearly stated. Stock status appears robust to the alternative stock-structure assumptions presented. However, surplus production estimates of MSY are predicated upon the main parameter estimates, and hence the robustness of stock status statements to more uncertainties needs to be examined (see Recommendations above).

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

Projections were based upon the model fit of the 2 sub-area model, felt by the Working Group to represent the most plausible stock-structure scenario. The projection methodology is appropriate, and uses the variability estimated through the Bayesian model. Uncertainties in the most recent years (due to differing trends in the CPUE) may represent additional uncertainty. Given the uncertainty in model fits within the 2 stock scenario, projections should also be performed using the single-stock scenario to test robustness and provide further information to managers (**Recommendation 14**).

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

Many of the recommendations made above represent research priorities, in particular the need to examine the influence of assumptions and the impacts of specific uncertainties within the model, and investigate approaches for further CPUE standardisation. The reader is referred to the main report text for more details.

Given uncertainties in the CPUE data, and as recommended in the report, other assessment approaches are available, which could be used in parallel to the current approach as information allows. Before effort is expended in developing those methods and collecting new data for this purpose, the robustness and benefits of these alternative models should be tested through Management Strategy Evaluation (**Recommendation 15**).

# Background

Swordfish in the North Pacific are harvested multi-nationally, primarily using longline gear. The U.S. has a major fleet of swordfish longline vessels based in Hawaii and swordfish harpoon and longline vessels in California. An assessment of swordfish in the North Pacific Ocean will be conducted by staff of the Pacific Islands Fisheries Science Center and collaborating scientists from members of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean (ISC). The assessment will be conducted within the ISC's Billfish Working Group during FY 2009.

Results of the swordfish assessment will be key to international management decisions of the Western and Central Pacific Fisheries Commission and its Northern Committee, and domestic management decisions by the Western Pacific Regional Fishery Management Council (WPFMC) and Pacific Fishery Management Council. An independent peer-review of the assessment is essential. The Terms of Reference (ToRs) of the peer review are attached in Appendix 2 (Annex 2).

The review was performed at the consultant's offices at Cefas (Lowestoft, UK) as a desk-based study.

This document represents the individual CIE Reviewer Report on the results of the desk-based review of the document presenting the 2009 stock assessment of North Pacific swordfish, at the request of the Center for Independent Experts (see Appendix 1).

# Description of review activities

This review was undertaken by Dr Graham Pilling at Cefas (Lowestoft, UK) during the period 3rd January to 15th February 2010. The documentation (see bibliography) was reviewed at Cefas, and this report to CIE completed. The review focused upon the five areas of consideration detailed within the review Terms of Reference (see Appendix 1, Annex 2).

# Summary of findings against TOR

Below, my summary of findings is presented against each of the Terms of Reference (Appendix 1). Within these, generic and assessment-specific observations and recommendations are developed.

Numbered recommendations (in bold) refer to the correspondingly numbered items within the conclusions and recommendations section of this report.

As a general comment, it is harder to review an assessment based upon a document, without the opportunity for the assessment authors to respond to questions and comments interactively. Therefore many of the comments and recommendations in this report may fall under the category of "been there, tried that, it didn't work"; but if so, it highlights the need for additional information to be provided to reviewers under the desk-based review format.

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

### 1.1. General comments

It is appropriate to use Bayesian implementations of the surplus production model for stocks such as Pacific swordfish where biological information appears relatively limited but CPUE time series data are available. The Bayesian implementation of the surplus production approach has been used for a number of different species around the world and hence is reasonably well tested and understood. The WinBUGS software in which the model is implemented is also well known and stable. The Bayesian implementation provides a number of advantages, such as the ability to provide probabilistic risk-based advice to managers (as noted in Section 3 of this report).

A number of assumptions need to be made when using surplus production models, in particular that the CPUE data relate to stock abundance and assumptions on catchability, while estimation issues also include the degree of information (contrast) in the data. The performance of the assessment approach and model setup is discussed under Section 2.

It is noted that other assessment modelling approaches have been used for this species, and this is discussed further in Section 5 as a way of examining model uncertainty.

To help the reader fully understand the implementation of the model, it would be helpful to see the WinBUGS code as an appendix to the report – many of the questions raised below could be clarified with this addition. See **Recommendation 1**.

### 1.2. Estimated stock status

With the information presently available, the model appears to have fitted the main data sets reasonably well (Japanese CPUE time series in particular), and those time series (annual CPUE data for Japan and Chinese Taipei) show consistent trends (see Courtney and Wagatsuma, 2009) and appear to be driving the assessment results. The overall conclusions on stock status appear supported by the recent trends in the CPUE data series (under assumptions of constant catchability and that CPUE data realistically reflect stock abundance). However, the authors appear to present the results from the 'final' model runs, and there is a general need to understand the performance of the model and robustness of outputs under alternative starting conditions and to uncertainty in input parameters (see Section 2).

## 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.

It is always easy to pick holes in the data used within assessments. Ultimately, pragmatism requires that the best available data be used to derive stock status information and hence advice to inform management. This appears to be the case within this assessment. As a result, most of the issues I raise within this section should be considered within model development activities.

One general comment is the fact that the 2009 stock assessment is only using data up to 2006. Are more recent data not available?

In turn, to aid review process, information on the methods used to develop the input data (i.e. CPUE time series standardisation) should be provided within the assessment report, particularly where uncertainty exists. See **Recommendation 2**.

### 2.1. CPUE data

CPUE data series need to show sufficient contrast to provide information on the major parameters within the surplus production model.

CPUE data have been standardised across the time series (noted in Courtney and Wagatsuma, 2009), but only summary information on how this was performed was found in Annex 8 of ISC Billfish WG report. Issues with standardisation are noted in that report, and hence these issues need to be further examined. It is again noted that the time series of standardised CPUE data are generally consistent in their trends. See **Recommendation 3**.

It does not appear that catchability trends have been included within the standardisation approaches, which may reflect the difficulties in standardising the data series even without this factor being included (and the inherent difficulties in estimating this parameter), nor are attempts made to estimate this parameter within the surplus production model. Although a primary consideration is the improvement of the standardisation (see Recommendation 3), information needed to compensate the CPUE time series for changes in catchability should be examined and incorporated within the standardisation process or the model, where considered necessary and feasible. I note however that there is a trade-off between flexibility in inter-annual catchability changes and resulting reduction in information on abundance! See **Recommendation 4**.

The weighting of the different CPUE data sets within the model is unclear, although I assume they have not been weighted by their variance, for example? Given that the Japanese CPUE data time series appears highly influential on the model fit (and that there is some difficulty standardising this CPUE time series) sensitivity analyses of alternative CPUE time series weighting should be performed. This should also allow the influence of CPUE time series outliers on stock assessment results to be identified. See **Recommendation 5**.

It's notable from Courtney and Wagatsuma (2009) that for the eastern Pacific Ocean region, CPUE data for the Spanish and Chilean fleets are lacking. These countries could represent ~50% of the landings in region 2 of the two-stock model. Looking at the catches in recent years, that of Spain has increased over time, while that of Chile has fluctuated, and shown a general decline since 1991. The reasons behind this are not clear (Changes in effort? Different catch allocations?), but if the assumption is made that effort has remained constant, the catches of Spain appear to support the opinion of a generally sustainably exploited stock (although recent declines in overall catch and the declines of Chile are of some concern). See **Recommendation 6**.

### 2.2. Assumed stock structure

Two hypotheses are tested within the assessment to examine model uncertainty – a single stock across the northern Pacific, and a two-stock model based upon the work of Ichinokawa and Brodziak (2008). The examination of alternative stock structures within the model is appropriate given the available information, and ensures that potential over-exploitation of one sub-stock is not masked by the under-exploitation of another. It also allows the robustness of management advice to this uncertainty to be evaluated.

I realise that these two stock structures are based upon recommendations from the ISC Billfish Working Group and that the data have been appropriately re-arranged to take account of these hypotheses, but it is confusing why the 1 stock hypothesis does not equate to stock 1 + stock 2 in the 2-stock hypothesis?

I note that the Japanese longline CPUE for the two sub-stocks do not correlate – further supporting the division into two stocks - however, a significant correlation is found with the (much shorter) time series of Chinese-Taipei standardised CPUE data in the two populations. Given the fact that the Japanese CPUE data appear to be driving the assessment outputs (see below), this may not be a large issue, but the robustness of model outputs to these assumptions should be further examined through alternative CPUE weightings (see Recommendation 5).

## 2.3. Input parameter values

The development of the prior probability density functions (pdfs) can strongly influence the results obtained from the models in the face of uninformative data. Appropriate methods have been used to generate prior distributions for key model parameters such as R and K, while that for catchability has been left uninformative so that the data can inform the outputs. However, some comments are raised here. Further issues with the prior/posterior probability density functions are discussed in Sections 2.4.

Prior mean values for carrying capacity (K) under the alternative stock structure scenarios are based upon the levels required to sustain observed catches. The prior mean for the single-stock hypothesis is equal to that of the stock in sub-area 1 in the two-stock hypothesis, which implies that the carrying capacity in the overlap area with sub-area 2 is negligible. These differences appear to have some influence on the relative mean carrying capacity levels estimated in the posterior distributions. **See Recommendation 7**.

The starting conditions assume that the population is at 90% of the carrying capacity under each scenario (mean value). It is not clear whether this mean value combined with the lognormal variance selected allows biomass at the start of the time series to be greater than carrying capacity? It is also noted that in many cases the early years of a tuna longline fishery often exhibit strong declines in CPUE despite fishing below MSY levels; this phenomenon is not seen in the Japanese longline CPUE data. While this may be due to the fact this stock had been previously exploited (hence the 90% starting condition) it does raise uncertainty over this parameter. How this assumed value affects overall estimates of stock status given the trends in CPUE time series is not currently examined. For example, CPUE data for the stock in sub-area 2 begins in 1955 (the data for 1951-1955 being approximately zero; Courtney and Wagatsuma, 2009), but population size and harvest rate have been estimated by the model back to 1951. Perhaps as a result of the starting population being at ~90% of carrying capacity, this leads to a decline in exploitable biomass down to approximately $B_{MSY}$ at the start of the CPUE time series, and the impact of this on the assessment results is unclear. See **Recommendation 8**.

## 2.4. Model diagnostics

The main model diagnostics are the convergence criteria and goodness-of-fit criteria, which are appropriately used to evaluate the performance of the model fitting approach. The model has achieved a good fit to the ‘main’ (Japanese) CPUE time series.

My main concern with current report (not approach) is the limited information allowing the reader to identify how much influence the priors or the data had on the outputs from the model. See **Recommendation 9**. In turn, posterior pdfs for all the parameters with priors are not described. For example, what was the posterior pdf for catchability? See **Recommendation 10**.

I produced the table below from the reported information, in order to assess the relative impact of the prior information and CPUE data on the model fit. Overall, most posterior distributions are more precise than the priors provided, suggesting that there is some information in the CPUE data for parameter estimates. In some cases these improvements are somewhat marginal, while in the two-area model, the shape parameter for stock 2 is more uncertain in the prior than the posterior, suggesting that little information is available on this parameter in the data.

| Scenario | Parameter | Prior mean | Posterior mean | Prior CV | Posterior CV (%) |
|---|---|---|---|---|---|
| Single stock | R | 0.50 | 0.68 | 50 | 36 |
| | K (kt) | 150.0 | 113.6 | 50 | 23 |
| | M | 1 | 1.25 | 71 | 44 |
| 2-stock (1) | R | 0.50 | 0.58 | 50 | 38 |
| | K (kt) | 150.0 | 115.9 | 50 | 21 |
| | M | 1 | 1.02 | 71 | 47 |
| 2-stock (2) | R | 0.50 | 0.40 | 50 | 45 |
| | K (kt) | 75.0 | 54.6 | 50 | 25 |
| | M | 1 | 0.66 | 71 | 81 |

2.5. Population estimates

The general exploitable population biomass trajectories in the different models show similar patterns (single-stock and sub-area 1 of the two-stock scenario), driven by the CPUE trends. Mean posterior carrying capacity values are lower than the mean values input for the priors in all the models (see section 2.4) while mean posterior estimates of R are generally (but not always) higher. While the mean output values for R are within the range of values shown in the literature for this or similar species, the authors do not really discuss whether the mean parameter and population estimates are biologically realistic. See **Recommendation 11**.

Estimates for the sub-area stock 2 appear relatively uncertain. While the carrying capacity exhibits a more precise posterior pdf compared to the prior, those for the R parameter are shifted but not notably more precise, while that for M (S) is less precise than the prior. Inter-annual variation in estimated harvest rate is also relatively high compared to the estimates of exploitable biomass. This uncertainty has knock-on effects on the models used for the projection. See **Recommendation 12**.

The authors present two alternative structural scenarios with the single- and two-stock model formulations. These lead to different parameter and total catch estimates, although overall stock status appears robust to this uncertainty. However, it is not clear whether alternative runs with different prior distributions were performed to allow the reader to judge the sensitivity of the model outputs to the assumptions made when developing the priors. See **Recommendation 13**.

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.**

The methods used to estimate management parameters (MSY) are standard and appropriate based upon the surplus production method. The Bayesian approach allows clear probabilistic statements of stock status to be developed and these are clearly stated within the document.

Stock status appears robust to the alternative stock-structure assumptions made within the assessments. Surplus production estimates of MSY are predicated upon the main parameter estimates, and hence the robustness of the stock assessment results to more uncertainties needs to be examined through sensitivity analysis (see Recommendations).

I note that projected mean 2007 catch estimates under 'status quo' conditions in sub-area 1 are ~25% higher than that seen in 2006 (see Courtney and Wagatsuma, 2009), while that for sub-area 2 are ~30% higher, perhaps due to the use of the 3 year average F in a period of declining mortalities. I also note both estimates are within the range of the time series and catch CVs.

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

The projections are based upon the model fit of the 2 sub-area model, as this model structure was felt by the Working Group to represent the most plausible stock-structure scenario. The projection uses the variability estimated through the Bayesian model, and projects under status quo conditions for four years based on an average of the last three years of harvest rate. These are reasonably 'standard' settings for a projection, with the additional benefit of utilising the uncertainty captured by the Bayesian approach. The number of separate projections performed to develop confidence intervals around the projections is not specified, however. In turn, while the CPUE time series data are reasonably consistent, they do vary in their trends in the most recent years and hence there may be additional uncertainty in projections. I note that harvest rates are estimated to be generally declining in last 3 years, so the use of the 3 year average across this period might be considered a 'conservative' approach.

Given the uncertainty in some of the model fits within the 2 stock scenario, projections should also be performed using the single-stock scenario to provide further information to managers. See **Recommendation 14**.

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

In this section I concentrate on research priorities in both the fisheries dynamics of the species in question, and the assessment approach being used. Initial comments are aimed at the requirements for the current assessment approach. Later comments note that alternative assessment approaches could be used, *IF* the data to do so are available. Recommendations made under the other sections of this report may also form future research areas and only priority ones are repeated here. The reader is referred to the original recommendations for more details.

The model fit presented examines the robustness of advice to uncertainty in stock structure, but as noted in recommendations within the previous report sections, there is a need to examine the influence of assumptions and the impacts of specific uncertainties within the model (see Recommendation 13 in particular).

Further work to standardise CPUE time series, including consideration of how the gear types have changed over time and potential changes in catchability, should be concentrated upon as these are key to the current assessment approach (see Recommendations 3 and 4).

Given uncertainties with the CPUE data, and as noted (and recommended) in the report, other assessment approaches (e.g. CASAL, MULTIFAN-CL, Wang et al., 2007) are available, which have the potential to take into account spatial, fisheries and biological information. It is noted that there are always trade-offs between increased model complexity and hence increased data requirements and potentially marginal pay-offs in terms of improved management advice. See **Recommendation 15**.

## Conclusions/Recommendations

Based upon the information provided within the reports reviewed, this reviewer agrees with the presented stock status estimates. However, it should be noted that the assessment results may change as a result of the recommendations given within the current section of the report.

**Recommendation 1**. Provide details of the WinBUGS code used to implement the model within the report.

**Recommendation 2**. Provide a summary table for each assessment, noting each data set used along with its constraints and any treatments or modifications made. Judgements on the reliability of each data set should be included.

**Recommendation 3**. Given the critical role of the CPUE data underpinning the surplus production approach, ensure further effort is spent on the standardisation of CPUE time series, in particular that of the Japanese longline fleet which appears to drive the assessment.

**Recommendation 4**. Examine information relevant to fishing power in order to identify whether standardisation for changes in catchability is required, and if necessary include this within the standardisation process or assessment model.

**Recommendation 5**. Given the uncertainties in the CPUE time series, perform sensitivity analyses of model outputs to alternative CPUE time series weightings. See for example McAllister et al. (2001).

**Recommendation 6**. Where effort data allow, develop CPUE data for these major fishing nations in sub-area 2.

**Recommendation 7**. Given that the prior mean levels for the carrying capacity in the different stocks appear to have some influence on the posterior estimates, some sensitivity analysis with alternative prior distributions of appropriate R and K mean values is recommended.

**Recommendation 8**. Examine the sensitivity of model outputs to the value selected for the starting conditions of the population time series.

**Recommendation 9**. Present the prior and posterior pdfs to allow the reader to examine the degree of influence the priors and data have on the posterior outputs.

**Recommendation 10**. Present the prior and posterior pdfs for all major parameters assumed within the model.

**Recommendation 11**. Discuss the biological realism of the parameter estimates developed by the model.

**Recommendation 12**. Examine the causes of what appears to be relatively uncertain parameter estimates for the assumed stock in sub-area 2.

**Recommendation 13**. Perform appropriate sensitivity analyses on mean prior distribution values.

**Recommendation 14**. Given the uncertainty in the two-stock scenario results (particularly for the stock in sub-area 2), projections should also be performed for the single-stock scenario to test robustness.

**Recommendation 15**. Consider undertaking parallel assessments as data and parameter estimates allow.  Before effort is expended in developing and collecting new data to furnish alternative assessment models, test the robustness of alternative models through Management Strategy Evaluation.

# Appendix 1. Bibliography

## Primary documentation

Brodziak, J. and Ishimura, G. (2009). Stock assessment of North Pacific swordfish (*Xiphias gladius*) in 2009. 50p.

Courtney, D. and Wagatsuma, L. (2009). Input data for a North Pacific swordfish stock assessment using Bayesian production models. ISC/09/BILL WG-2/01.

Ichinokawa, M. and Brodziak, J. (2008). Stock boundary between possible swordfish stocks in the northwest and southeast Pacific judged from fisheries data of Japanese longliners. ISC/08/BILL WG-SS/04.

## Additional documentation used by this reviewer

Bremer, J.A.R., Hinton, M.G. and Greig, T.W. (2006). Evidence of spatial genetic heterogeneity in Pacific swordfish (*Xiphias gladius*) revealed by the analysis of ldh-A sequences. Bull. Mar. Sci. 79, 493-503.

Brodziak, J. (2007). An investigation of alternative production models to assess the Hawaiian bottomfish complex. Pacific Islands Fisheries Science Centre Administrative Report H-07-01. 63p.

McAllister, M; Babcock, EA; Pikitch, EK (2001). Evaluating the relative merits of alternative methods to weight different time series of abundance indices in stock assessment. Collect. Vol. Sci. Pap. ICCAT 52, 1094-1115

Meyer, R. and Millar, R.B. (1999). BUGS in Bayesian stock assessments. Can. J. Fish. Aquat. Sci. 56, 1078-1086.

Wang, S-P., Sun, C-L., Punt, A.E. and Yeh, S-Z. (2007). Application of the sex-specific age-structured assessment method for swordfish, *Xiphias gladius*, in the North Pacific Ocean. Fish. Res. 84, 282-300.

# Appendix 2. CIE Statement of work

## Attachment A: Statement of Work for Dr. Graham Pilling (CEFAS)

### External Independent Peer Review by the Center for Independent Experts

### Stock Assessment of North Pacific Swordfish

**Scope of Work and CIE Process:** The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract to provide external expertise through the Center for Independent Experts (CIE) to conduct impartial and independent peer reviews of NMFS scientific projects. This Statement of Work (SoW) described herein was established by the NMFS Contracting Officer's Technical Representative (COTR) and CIE based on the peer review requirements submitted by NMFS Project Contact. CIE reviewers are selected by the CIE Coordination Team and Steering Committee to conduct the peer review of NMFS science with project specific Terms of Reference (ToRs). Each CIE reviewer shall produce a CIE independent peer review report with specific format and content requirements (**Annex 1**). This SoW describes the work tasks and deliverables of the CIE reviewers for conducting an independent peer review of the following NMFS project.

**Project Description:** Swordfish in the North Pacific are harvested multi-nationally, primarily using longline gear. The U.S. has a major fleet of swordfish longline vessels based in Hawaii and swordfish harpoon and longline vessels in California. An assessment of swordfish in the North Pacific Ocean will be conducted by staff of the Pacific Islands Fisheries Science Center and collaborating scientists from members of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean (ISC). The assessment will be conducted within the ISC's Billfish Working Group during FY 2009.

Results of the swordfish assessment will be key to international management decisions of the Western and Central Pacific Fisheries Commission and its Northern Committee, and domestic management decisions by the Western Pacific Regional Fishery Management Council (WPFMC) and Pacific Fishery Management Council. An independent peer-review of the assessment is essential. The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**.

**Requirements for CIE Reviewers:** Three CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein. Each CIE reviewer's duties shall not exceed a maximum of 10 days to complete all work tasks of the peer review described herein. The CIE reviewers shall have the expertise, background, and experience to complete an independent peer review in accordance with the SoW and ToRs herein. CIE reviewer expertise shall include fish stock assessment, mathematical modeling, and statistical computing.

**Location of Peer Review:** Each CIE reviewer shall conduct an independent peer review during a desk review of a report on the stock assessment of North Pacific swordfish, whereby no travel shall be required.

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

Prior to the Peer Review: Upon completion of the CIE reviewer selection by the CIE Steering committee, the CIE shall provide the CIE reviewer information (name, affiliation, and contact details) to the COTR, 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 reviewer with the background documents, reports, foreign national security clearance, and information concerning other pertinent meeting arrangements. The NMFS Project Contact is also responsible for providing the Chair a copy of the SoW in advance of the panel review meeting. Any changes to the SoW or ToRs must be made through the COTR prior to the commencement of the peer 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 all 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 on where to send documents. The CIE reviewer shall read all documents in preparation for the peer review.

This list of background documents may be updated up to two weeks before the peer review. Any delays in submission of pre-review documents for the CIE peer review will result in delays with the CIE peer review process, including a SoW modification to the schedule of milestones and deliverables. Furthermore, the CIE reviewer is responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein.

Panel Review Meeting: The CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs. **Modifications to the SoW and ToRs can not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COTR and CIE Lead Coordinator.** The CIE reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified in the contract SoW. The NMFS Project Contact is responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The CIE Lead Coordinator can contact the Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

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

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

1) The CIE reviewer shall review all background material and reports provided by the NMFS Project Contact as part of the peer review;
2) The CIE reviewer shall conduct an independent peer review in accordance with the ToRs (Annex 2);
3) No later than 19 February 2010, the CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Mr. Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and to Dr. David Die, CIE Regional Coordinator, via email to ddie@rsmas.miami.edu. The CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in Annex 2;
4) The CIE reviewer shall address changes as required by the CIE review in accordance with the schedule of milestones and deliverables.

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

| | |
|---|---|
| 16 December 2009 | CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact |
| 16 December 2009 | NMFS Project Contact sends the CIE the background documents |
| 18 December 2009 – 19 February 2010 | The reviewer conducts an independent peer review |
| 19 February 2010 | The CIE reviewer submits a draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator |
| 5 March 2010 | CIE submits CIE independent peer review reports to the COTR |
| 19 March 2010 | The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director |

**Modifications to the Statement of Work:** Requests to modify this SoW must be made through the Contracting Officer's Technical Representative (COTR) who submits the modification for approval to the Contracting Officer at least 15 working days prior to making any permanent substitutions. The Contracting Officer will notify the CIE within 10 working days after receipt of all required information of the decision on substitutions. The COTR can approve changes to the milestone dates, list of pre-review documents, and Terms of Reference (ToR) of the SoW as long as the role and ability of the CIE reviewer to complete the SoW deliverable in accordance with the ToRs and deliverable schedule are not adversely impacted. The SoW and ToRs cannot be changed once the peer review has begun.

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

**Applicable Performance Standards:** The contract is successfully completed when the COTR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards: (1) the CIE report shall have 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 notification of acceptance by the COTR, the CIE Lead Coordinator shall send via e-mail the final CIE report in *.PDF format to the COTR. The COTR will distribute the approved CIE reports to the NMFS Project Contact and regional Center Director.

**Key Personnel:**

William Michaels, Contracting Officer's Technical Representative (COTR)
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-713-2363 ext 136

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

NMFS Project Contact:

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

Robert Moffitt, Project Contact
Pacific Islands Fisheries Science Center
2570 Dole Street, Honolulu, HI 96822-2396
Robert.Moffitt@noaa.gov Phone: 808-983-3742

**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.

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, and Conclusions and Recommendations in accordance with the ToRs.

   a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including providing a detailed summary of findings, conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each ToR even if these were consistent with those of other panelists, and especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the Summary Report that they feel might require further clarification.

   d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

   e. The CIE independent report shall be a stand-alone document for others to understand the proceedings and findings of the meeting, 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 as separate appendices as follows:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of the CIE Statement of Work
   Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

## Annex 2: Terms of Reference for the Peer Review

### Stock Assessment of North Pacific Swordfish

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.

Cefas