CIE Reviewer's External Independent Report on the assessment of Hawaii deepslope bottomfish.

Prepared by

Neil Klaer

Prepared for

Center for Independent Experts (CIE)

Desktop Review
January/February 2011

Contents

Executive Summary

A Western Pacific Stock Assessment Review (WPSAR) examined the 2008 Hawaii deep slope bottomfish assessment update in June 2009. The review found that the assessment required further work on data filtering, investigation of the cause of change in catchability, alternative CPUE standardization and Bayesian model structure, and improvement to the presentation of model selection, diagnostics and sensitivity tests. As a consequence, the assessment was revised and has now been made available for further independent external review. This desktop review of the revised assessment was carried out during the period 18 January to 7 February 2011.

Findings by term of reference

1. Determine if recommendations from the June 2009 WPSAR/CIE review have been adequately addressed within the assessment update.

    - The assessment was significantly revised and attempted to address the primary review recommendations. This was partly successful, but the WPSAR highlighted, in particular, the need to examine questions of technological change and historical catch levels in collaboration with Hawaii Division of Aquatic Resources and fishers. As this collaboration is yet to take place, the choices made in the assessment on base case and sensitivity scenarios were not sufficiently justified to be relied upon for management purposes.
    - There were a number of additional review recommendations that remain to be addressed.

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

    - The WPSAR/CIE review found that the Bayesian stock production model is an appropriate assessment method that could provide a sound basis for the provision of management advice. I also agree with this view.

3. Evaluate the implementation of the assessment model: configuration, assumptions, and input data and parameters (fishery life history); more specifically determine if data are properly used, if choice of input parameters seem reasonable, if models are appropriately specified and configured, if assumptions are reasonably satisfied, and primary sources of uncertainty accounted for.

    - An objective procedure was developed to determine cutoff values in data filtering that requires additional documentation.
    - The Bayesian production model implementation was acceptable, but insufficient detail was given to justify the choices made in setting prior values and distributions for a number of parameters.

4. Comment on the scientific soundness of the estimated population benchmarks and management parameters (e.g. MSY, F_{msy}, B_{msy}, MSST, and MFMT) and their

potential efficacy in addressing the management goals stated in the relevant FMP or other documents provided to the review panel.

- The MSY-based methods used for reference point calculations are appropriate, and can provide useful management inputs.
- There is a need for harvest control rules to be specified that are likely to attain the management objectives of the fishery in the long-term. Particular difficulties with the main Hawaiian Islands (MHI) Deep7 bottomfish is that the complex combines multiple species with different distributions and biology, the different regional abundance and exploitation patterns within the MHI, and the existence of various fishery closures. There is also a suggestion that species within the Deep7 complex (onaga and ehu) may be overfished. Under these conditions, the development of specific harvest control rules would benefit from simulation testing to determine those that meet management objectives given these difficulties.

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

   - The projection method appears to be appropriate and adequately applied, but requires further documentation.

6. Suggest research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices. Include guidance on single species models, and whether this is possible given the current nature of this multispecies fishery, and difficulties in partitioning fishing effort between species.

   - A number of research priorities have been listed.

1 Introduction

1.1 Background

Hawaii deepslope bottomfish are managed under the Magnuson Stevens Act and stock assessments are examined within the Western Pacific Stock Assessment Review (WPSAR) process. The goal of the WPSAR is to provide independent review of input data and stock assessments to ensure the best available stock assessment advice is available for fishery management – particularly the Western Pacific Regional Fishery Management Council (WPRFMC).

The first review under WPSAR that examined the 2008 Hawaii deep slope bottomfish assessment was held from 15-19 June 2009. That 2008 assessment for the period 1948-2007 (Brodziak *et al*. 2009) updated a previous 2006 assessment for the period 1948-2004 (Moffitt *et al*. 2006). The review found that the 2008 assessment required more work on data filtering, investigation of the cause of change in catchability, alternative CPUE standardization and Bayesian model structure, and improvement to the presentation of model selection, diagnostics and sensitivity tests. Reports that summarized the review findings were by the WPSAR (Skillman 2009) and the Center for Independent Experts (CIE) (Stokes 2009).

As a consequence, the assessment has been revised (Brodziak *et al*. 2011) and has now been made available for further independent external review.

1.2 Review Activities

The statement of work for the current review is given in Appendix 2. My role was to provide an external independent desktop peer review in accordance with the statement of work and the terms of reference (TOR) of the review. The review commenced 10 days later than originally planned and took place during the period 18 January to 7 February 2011.

2 Review of the Hawaii deepslope bottomfish assessment

2.1 Terms of reference

The Review considered the assessments in light of the terms of reference provided as follows:

1. Determine if recommendations from the June 2009 WPSAR/CIE review have been adequately addressed within the assessment update.

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

3. Evaluate the implementation of the assessment model: configuration, assumptions, and input data and parameters (fishery life history); more specifically determine if data are properly used, if choice of input parameters seem reasonable, if models are appropriately specified and configured, assumptions are reasonably satisfied, and primary sources of uncertainty accounted for.

4. Comment on the scientific soundness of the estimated population benchmarks and management parameters (e.g. MSY, F_{msy}, B_{msy}, MSST, and MFMT) and their potential efficacy in addressing the management goals stated in the relevant FMP or other documents provided to the review panel.

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

6. Suggest research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices. Include guidance on single species models, and whether this is possible given the current nature of this multispecies fishery, and difficulties in partitioning fishing effort between species.

2.2 Findings by term of reference

2.2.1 TOR1 Determine if recommendations from the June 2009 WPSAR/CIE review have been adequately addressed within the assessment update.

The assessment was significantly revised and attempted to address the primary review recommendations. This was partly successful, but the WPSAR/CIE review highlighted in particular the need to examine questions of technological change and historical catch levels in collaboration with Hawaii Division of Aquatic Resources (HDAR) and fishers. As this collaboration is yet to take place, the choices made in the assessment on base case and sensitivity scenarios are not sufficiently justified to be relied upon for management purposes. There were a number of additional review recommendations that remain to be addressed.

There were a large number of recommendations made in the WPSAR and CIE review documents. I do not have information on what level of resources were available to address the recommendations, but I can see that addressing all of them would be a considerable undertaking, and not possible in the time between the last and this review.

The CIE reviewer (Stokes 2009) noted which research priorities could be addressed immediately (within a year), and which were short or medium-term. The assessment team has taken those recommendations deemed as immediate to address in particular in the current assessment, which I agree was a reasonable approach. Specifically, those recommendations were:

1. Comprehensively explore main Hawaiian Islands (MHI) CPUE data and qualitative information in close collaboration with HDAR and fishers throughout the process. Develop credible CPUE standardization, including if appropriate alternative indices.
2. Attempt to reconstruct non-commercial catch histories, possibly in the same collaborative process used for (1).
3. Consider using meta-data to develop informative prior on $R_{\max}$. Develop prior for B_{init} in collaborative process above (1).
4. Assess MHI as single stock to develop population benchmarks and management parameters. Ensure appropriate sensitivity testing to CPUE uncertainty.

Rather than a comprehensive exploration (including input from fishers), the input data for standardization was subjected to a re-audit. Rules used to filter the data were re-examined, and an objective method for selecting bottomfish trips from the handline data was developed. Development of a credible CPUE standardization was approached primarily through the development of a number of alternative series that accounted for uncertainty in changes in fishing power through time, inclusion of 1958-1960 data, and alternative model structural approaches (delta-GLM, quasi-likelihood Poisson GLM). Interaction between model terms was examined, and a base model chosen that included an area-season interaction. The model with negligible change in fishing power through time was selected as the base model because the WPSAR review panel viewed models with increases as unsupported by data, and ad hoc, and WPRFMC used this model in post hoc modeling of the 2008 assessment. There is still a need to examine alternative scenarios in close collaboration with DAR and industry (including recreational fishers), and I believe this should also include group sign-off on the most appropriate base case, and acceptable alternatives for sensitivity testing. Group sign-off was not specifically recommended by the WPSAR/CIE review, but is a logical outcome of such collaboration, and greatly assists with the provision of justification for a base case and plausible sensitivities, and also with industry acceptance of the assessment and any management consequences.

The second recommendation was examined through the development of four scenarios of unreported catch to characterize that uncertainty, ranging from none to 550,000lb/y in recent years. The catch scenario used for the base case assumed moderate recent unreported catch of 240,000lb/y. A consultative reconstruction was not included in the documents, so I assume has not yet been undertaken, although a new review of alternative scenarios (Courtney) was provided. While it appears that a useful range of scenarios have been constructed, reasons for selection of one of them as the best available estimate were not developed. I believe such a decision still needs to be made in a collaboration as above, and the reasons documented.

The third recommendation was addressed through concentration of the assessment on MHI Deep7 bottomfish species only, thereby somewhat reducing the uncertainty in the development of priors for *R* and initial biomass levels. The value for *R* for Deep7 bottomfish was revised based on recent information about the longevity of the species that forms the greatest portion of that complex, and mean values for the 1949 biomass as a proportion of carrying capacity were derived from posterior mean values.

The MHI Deep7 complex was assessed following the fourth recommendation. Sensitivity testing to CPUE uncertainty was carried out.

A more detailed examination of recommendations and responses to individual issues follows.

I have extracted recommendations common to the WPSAR and CIE reports, and also recommendations made in each individual report, or where the recommendation on an item differed in the details. In the detailed examination below, specific recommendations from the WPSAR and CIE reports are shown as italicized text, and the response in the most recent assessment as normal text.

Common recommendations

<u>CPUE data set creation</u>

Create a data set with catches > 0 and < 1500 lbs (as currently filtered).

This was used previously as an initial filter to handline trips to exclude those that caught no bottomfish, and multi-day trips.

Compute the ratio $Catch_{BMUS}/Catch_{all\ species}$ *as a possible index of targeting and add to data set.*

The current standardization examined CPUE of the combined Deep7 species only (and additional zero catches as a sensitivity). Targeting was not examined specifically.

For the assessment catch, remove ta'ape, kahala, armorhead and any non-BMUS.

Only Deep7 species were examined, so the above species were excluded.

<u>CPUE data exploration</u>

Explore all data comprehensively using, for example, regression trees to help identify factors that could be included in the data standardization model. Possible factors might include depth (inshore/offshore), targeting, technology changes, spatial variability due to aggregating statistical fishing areas, and environmental effects.

A specific exploration of additional factors was not reported in the documentation. Factors from the data included in the CPUE standardization were those that were reported consistently in all years – the HDAR area and month (aggregated in various ways).

<u>CPUE standardization model</u>

Use Fishing Year, Month, and some scientifically defensible definition of fishing area and interactions between those factors.

Standardization models that included the factors year, month or quarter, and HDAR fishing area or island group were examined under CPUE scenario I and AIC used to select the best model. The structure chosen using that approach included year, HDAR area, quarter and HDAR area x quarter interactions.

Investigate aggregating the HDAR statistical areas into 4 main MHI island groups.

The model selection process above concluded that aggregation of HDAR areas 100-128, 300-333, 400-429 and 500-528 into a simplified 4 island grouping was sub-optimal according to the AIC criteria. Uncombined HDAR areas were therefore used for the base model.

If using the C parameter for technological changes in catchability, use this as an offset in the standardization.

The change in fishing power was applied directly to the observations in the assessment, which is comparable to using an offset.

Preferably put factors/variables for fishing power directly in the CPUE standardization model even if applied to all records in a year.

Change in fishing power was dealt with under three CPUE scenarios. The first assumed that there was no change in fishing power through time (CPUE I), the second with no change 1949-1970, 0.25%/y 1971-1980, 0.5%/y 1981-1990, 0.25%/y 1991-2000 and no change 2001-2010 (CPUE II), and 1.2%/y since the 1950s (CPUE III). Rates for CPUE II were derived from values collected from fishers during interviews and the subsequent CPUE workshop (Moffitt et al. 2008), and CPUE III was similar to the scenario used for the 2006 and 2009 assessments. The annual power factor was multiplied by the CPUE observations and then used as input to the standardization.

Investigate various environmental factors, for example SST and SSH.

A separate investigation (Lee and Brodziak) was made of the correlation of MHI deep slope bottomfish CPUE (CPUE I from the assessment report) and the southern oscillation index, Pacific decadal oscillation (PDO), sea surface temperature and sea surface height. The report concluded that the CPUE was significantly negatively correlated with PDO, and that both CPUE and PDO showed autocorrelation. There was also a negative association between the CPUE and 1-year lag PDO. No effort has been made to incorporate environmental variables in the stock assessment at this stage, although it would be difficult to do so based on these initial results. The demonstration of correlation would normally need to be accompanied by some understanding of the linkage mechanism before it could be incorporated in a base case assessment, although sensitivities could be investigated.

Flag the years 1958-1960 inclusive as bad due to data errors and treat as outliers (i.e., fit dummy variables [1,0] to identify bad years).

The specific recommendation was not followed, but sensitivity to the inclusion of the 1958-1960 data was examined using CPUE standardization CPUE 1b, and data from these years were excluded for all other CPUE analyses.

<u>Stock assessment</u>

Stock assessment documents should follow a standard format (such as the recently developed for the Bering Sea and Aleutian Islands Crab SAFE Report).

The assessment generally followed the structure of the previous report. I agree that a standard such as the Crab Safe (2009) is a good one, and that there is still a need to adopt such a standard for the assessment documentation.

Combined plots of priors and posteriors for model parameters should be provided, as well as posterior plots of management indicators (e.g. $B_{current}/B_{msy}$, $F_{current}/F_{msy}$). MCMC evidence of model convergence should be given.

These plots were not provided. It would have been useful if they were, especially for the baseline model from Table 19 of the assessment report. The posterior plots of management indicators, or at least indicators of the error range should be provided (e.g. in Table 13). The errors for parameters directly estimated are given, but it is the derived quantities that are important for management.

Model convergence was investigated using the Geweke convergence diagnostic, the Gelman and Rubin diagnostic and the Heidelberger and Welch stationarity and half-width tests for key model parameters. Overall, these diagnostic tests supported the convergence of the base case and all of the sensitivities. These results were discussed in the text of the assessment report, but it would have been useful to provide the values for the diagnostic results (particularly the Geweke) in a table. No diagnostic MCMC plots were produced.

As a potentially independent measure of stock status, undertake length frequency sampling and use past data to calculate SPR or an SPR proxy by species.

This has not been investigated.

Investigate the use of a hierarchical Bayesian stock production model, with a multilevel R_{max} to account for different biological characteristics of species in the complex, and time-varying R_{max} to account for changing species composition through time (worthy of investigation but not the highest priority).

The Deep7 complex was dealt with as a single group, so species-specific R values were not investigated at this stage. A model that estimated R values using a multi-level prior for annual and overall mean values was examined. However, that model was rejected using Bayesian information criteria (over a range of sensitivity tests) in favor of the simple model with a single R value.

New information from bomb radiocarbon analysis for the most abundant Deep7 species opakapaka (A. Andrews, PIFSC, unpublished data) suggested a maximum age in the order of 40 years, and a corresponding natural mortality rate of about 0.1. This new information was used to revise the R_{max} estimate used in the 2008 assessment downwards from 0.45 to a range 0.05-0.15 for the Deep7 complex.

There is a need to improve documentation and especially explanations as to how and why decisions have been made in the past (e.g. values for the technology change parameter C).

Changes in fishing power was dealt with as a number of alternative scenarios (CPUE I – III). It remains difficult to judge the relative plausibility of those scenarios because of a lack of documentation of the procedures used to convert observations of methods used through time (as in Moffitt *et al*. 2008) to annual indices of fishing power.

Include sensitivity analyses of the technology coefficient, habitat ratios and initial biomass. Show key management indicators such as B_{2007}/B_{msy}, F_{2007}/F_{msy} *in the sensitivity analysis table.*

Sensitivity analysis results were given for alternative CPUE (accounting for different technology changes) and catch history (accounting for various levels of unreported catch) scenarios. Habitat ratios were not relevant to the analysis as it was MHI only. For the CPUE x catch scenario sensitivities (Table 13 – 14.3), management indicators relH_{2010} and relB_{2010} were given.

As the fishery will shortly cease in the NWHI region, the assessment should now focus on the MHI as a single stock.

The 2010 assessment update applies to the MHI region only.

In the medium term, a model that spatially resolves island-specific population structure may be developed. Further research is required to parameterize the spatial aspects of such a model.

An island-specific model has not yet been investigated.

WPSAR

Generalized Linear Mixed Models (GLMM) and Generalized Additive Models (GAM) that are formulated to include spatial and temporal effects, technological changes and other factors affecting catchability are recommended in preference to Generalized Linear Models (GLM). Investigate including a vessel size effect, and alternative spatial schemes.

Log-linear models were the primary standardization method used for the assessment. Alternative CPUE standardization models were examined as a sensitivity analysis using the observed Deep7 single-trip handline data under CPUE Scenario I (no change in fishing power through time). The first included data from 1958-1960 in the analysis (CPUE 1b), the second applied a two-stage delta model to data that included zero catches for Deep7 species (CPUE IV), and the third used a quasi-likelihood Poisson-GLM approach (CPUE V). GLMMs or GAMs were not examined.

Given the biological differences in the bottomfish species, separate Bayesian assessments should be carried out for (a) the two species most likely to be overfished (onaga and ehu) and the remaining deep slope bottomfish, (b) combined Deep7 bottomfish, and remaining deep slope bottomfish (c) fast growing snappers and slow growing snappers.

The assessment only considers combined Deep7 species.

Account for potentially significant levels of unreported catch.

Four scenarios of unreported catch were developed using published estimates and other sources, which were assumed to characterize the uncertainty. Catch scenarios I-IV were in order of decreasing levels of non-reporting, from 550,000lb/y to 0lb/y recently. Catch scenario II was used by the assessment team for the base case.

Use distributions rather than point priors for habitat ratios (MHI, Mau and Ho'omalu), undertake a meta-analysis for priors for R, K and the habitat ratios. Investigate other possible data sources such as oral histories or auction data for priors on initial biomass.

Only the MHI region was used in the analysis. Additional biological information for opakapaka was used to develop a prior distribution for *R* for that species, which was then used for the Deep7 complex. The assessment states that there is limited quantitative information on life history parameters for the Deep7 bottomfish. A meta-analysis is still justified however – if only to look at related species, and for information on initial biomass.

The Council and SSC should provide guidance to the assessment team on control rules and therefore assessment outputs.

Additional documentation on specific guidance from the Council or SSC was not available. There is considerable scope for the testing and development of specific harvest control rules for the conversion of MHI Deep7 assessment advice into catch recommendations.

CIE

A credible CPUE model is likely to require either Generalized Linear Mixed Models (GLMMs) or Generalized Additive Mixed Models (GAMMs) formulated to include spatial and temporal effects, technological changes, and other quantitative and qualitative factors affecting catchability. This is because a mixed model will allow investigation of correlation structure in the year effects as well as interactions, not just main effects of Year, Month, and Area. Some concern was noted in the review that attempts to include interactions, especially with area, would lead to unbalanced models. The simple way to deal with this issue, if it is important, is to bin data differently by area or, for example, to bin by quarter instead of month. Exploration is required.

GLMMs or GAMs have not yet been examined.

Future workshops should include more commercial and non-commercial fishers to link qualitative information and quantitative models to better understand the catch rates and catch history.

There does not appear to have been any workshops on catch rates or catch history with fisher participation since the last assessment.

Fully explain the rationale and limitations of the species complex selected for assessment.

The assessment now concentrates on the MHI Deep7 complex alone. Assessment documentation describes the lack of comprehensive biological information for all 7 species, and the reasoning for the use of a low *R* for the complex.

Consideration should be given to whether it is feasible to undertake separate assessments (including relevant CPUE analyses) of species most likely to be overfished (onaga and ehu). Alternatively, or possibly in addition, an assessment of the Deep7 complex could be attempted to set a Deep7 total allowable catch (TAC) directly, rather than by taking a multiplier from the BMUS assessment derived catch limit. It is not at all clear that working directly to derive a Deep7 TAC would result in a higher or lower TAC. Alternative assessment splits might also be explored based on life history and fishery characteristics as well as data availability.

The assessment now concentrates on Deep7 species. The problem of having species that may be overfished that form a minor part of the Deep7 complex is a difficult one. There needs to be clear management objectives for species or groups of species within the Deep7 complex, and either the ability to carry out more species-specific assessments, or a demonstration that the current management procedure for the complex should achieve management objectives for the component species or groups.

Both in the assessment documents and presentations, only limited diagnostics were provided on CPUE standardization or the stock assessment. For the future, it is essential to see good diagnostics of both the CPUE standardization and the Bayesian stock assessment. For the assessment, it is important to see how the likelihood components are affected by model assumptions and parameterisation; to compare directly posterior and prior distributions (graphically on the same plots); and convergence performance, graphically as multiple traces and as standard convergence statistics. It is always useful to see the convergence diagnostics not just for main parameters but also for the derived parameters (F_{status} and B_{status}).

Most of these recommendations still apply, and there is a need to provide additional diagnostic information – particularly for the selected base case, and evidence of convergence at least as tables for all cases including sensitivities.

While analyzing data to develop a credible CPUE index, it would be worthwhile exploring in detail whether it is possible to disaggregate data by area sufficient to *develop area specific CPUE and assessments. Although this would be a useful exploration, the priority should be on development of a credible single area MHI.*

Area specific CPUE indices within the MHI area have not been examined at this stage, although HDAR area and island groupings of them were included as factors in the standardization.

2.2.2 TOR2 Review the assessment methods used: determine if they are reliable, properly applied, and adequate and appropriate for the species, fisheries, and available data.

The WPSAR and CIE review reports both agree that the Bayesian stock production model is an appropriate assessment method that could provide a sound basis for the provision of management advice. I also agree with this view.

2.2.3 TOR3 Evaluate the implementation of the assessment model: configuration, assumptions, and input data and parameters (fishery life history); more specifically determine if data are properly used, if choice of input parameters seem reasonable, if models are appropriately specified and configured, assumptions are reasonably satisfied, and primary sources of uncertainty accounted for.

Implementation of the assessment methods can be divided into data preparation, CPUE standardization and Bayesian stock production model for discussion on each of these components separately.

Data preparation

An aspect of data filtering that did not attract specific review recommendations previously was the cutoff fraction used to determine which handline trips were bottomfish ones. Earlier values for this fraction were 90% in the 2005 assessment and 50% in the assessment update. An objective procedure was developed for the current assessment that maximized total catch weight and value captured by filtered records, while minimizing catch and value variation. The new cutoff values obtained were 29% for bottomfish and 17% for Deep7 species. The development of such a procedure is acceptable, but the documentation of it requires further work. To me it is unclear how the catch value was calculated (e.g. were historical sales records used, or recent values assumed to apply for all years?), and whether an objective function that treats proportions of total catch/value and CV ratios as equally important can be justified (and what alternatives may also be appropriate). The importance of change to the cutoff value is one that needs to be included as a sensitivity analysis of the assessment results.

CPUE standardization

An issue in the standardization was the necessary removal of zero catches because it was recognized that reporting practices had changed through time. Both the WPSAR and CIE reviews agreed that this was not optimal, but the approach was appropriate. Additional work was carried out to examine the effect of zero catches (Piner and Lee) using a two-stage delta standardization procedure. It was noted that 98% of the MHI observations from 1948-2011 used in the analysis were non-zero for the Deep7 species, so the effect of non-zero catches was small.

Until further investigation of factors other than year, month or quarter, and HDAR fishing area or island group (e.g. depth, factors affecting technological change and catchability) take place, the simple log-linear model and the model selection process used in the assessment is appropriate. I agree with the previous review that a more rigorous investigation of additional factors that affect the catch rate is required and that it should also take place in consultation with HDAR and industry. I am unable to predict whether additional useful factors that apply to all/most of the data would result from such a process, and how the analysis would be improved by the implementation using GLMM or GAM. I believe that it is uncertain whether this issue can be resolved in the short-term, but should be attempted if resources are available.

Bayesian stock production model

The 2008 update assessment was found to be unsatisfactory primarily because of poor quality of input data – particularly the standardized CPUE and total catch series and limited stock assessment diagnostics and sensitivity testing.

The current assessment update was limited to the MHI Deep7 complex alone, which addresses a number of difficult issues in examining the wider management regions, and additional species. The remaining major difficulties of estimation of unreported catch, and accounting for technological change affecting catchability have been dealt with as sensitivity scenarios.

A number of additional changes have been made to the production model structure, some of which were possible or necessary because the assessment concentrated on the Deep7 species complex:

- The assessment method included a production shape parameter *M* (and an associated prior), where values other than 1 move the surplus production peak away from 0.5*K*.

- Previously the prior for *K* was essentially uninformative, but the most recent assessment imposed a mean value of 18.0 million pounds and CV of 50% for high catch scenarios and a mean of 9.0 million pounds and CV of 25% for low catch scenarios.

- A different distribution was used for the prior for *R* (now lognormal rather than beta). The mean value for *R* was set to 0.1 based on previously published recommendations about the low productivity of Deep7 species, and new information about the expected life span of opakapaka, and the CV set to 25% - changes that essentially restricted *R* to the range of 0.05 to 0.15. The previous assessment used a mean value of 0.46 and a CV of 26%.

- The prior distribution for catchability was changed, but remained essentially uninformative as previously.

- The prior for process error variance was increased substantially compared to that used previously, and observation error was set to 10 times process error rather than +40% previously.

While all of these changes seem acceptable, the assessment documentation needs to include convincing reasons for making those choices rather than alternative ones – particularly where the choice has a significant effect on the results. For example, the discussion about the rejection of a multi-level *R* value had a sufficient reason and detail to be convincing.

Tables 13 – 14.3 generally indicate that *M* was very poorly estimated which suggests that a fixed value for this parameter is probably required (and sensitivity investigated if values other than 1 are used).

The sensitivity of the new cutoff fraction was not investigated.

Overall, the Bayesian production model implementation was acceptable, but insufficient detail was given to justify the choices made in setting prior values and distributions for a number of parameters.

2.2.4 TOR4 Comment on the scientific soundness of the estimated population benchmarks and management parameters (e.g. MSY, Fmsy, Bmsy, MSST, and MFMT) and their potential efficacy in addressing the management goals stated in the relevant FMP or other documents provided to the review panel.

The management objectives of the WPRFMC bottomfish fishery management plan are to (from Ralston *et al*. 2004):

- maintain opportunities for fishing experiences by small-scale commercial, recreational, and subsistence fishermen, including native Pacific islanders
- protect stocks from environmentally destructive fishing practices
- improve quality and quantity of data available for fisheries management
- maintain year-round supply of fresh fish in Hawaii
- maintain a balance between harvest capacity and harvestable fishery stocks to prevent over-capitalization

The WPRFMC has relied on a two main assessment approaches previously to provide information on stock status and to assist in setting catch limits for Hawaiian bottomfish – CPUE based SPR, and dynamic production models. The current assessment does not examine SPR, and provides management advice in the form of MSY-based reference points. Specifically these are $B_{\text{current}}/B_{\text{MSY}}$, $H_{\text{current}}/H_{\text{MSY}}$, and short-term probability of the exploitable biomass exceeding the limit of $0.7B_{\text{MSY}}$, and probability of the harvest rate exceeding H_{MSY} under different future catch scenarios. The methods used for the stock assessment and reference point calculations are appropriate, and can provide useful management inputs. However, as the current assessment requires further work, it can not currently be used for management advice.

I could not find specific control rules used by management to convert the reference point information from the current assessment method (e.g. short-term probability of being overfished) into a TAC (recognizing that these are guided by legislation such as the Magnuson-Stevens Act Overfishing Provisions, and the WPRFMC Bottomfish Fishery Management Plan). There is a need for harvest control rules to be specified that are likely to attain the management objectives of the fishery in the long-term. Particular difficulties with the MHI Deep7 bottomfish is that the complex combines multiple species with different distributions and biology, the different regional abundance and exploitation patterns within the MHI, and the existence of various fishery closures. Of note also is the suggestion that species within the Deep7 complex (onaga and ehu) may be overfished. Under these conditions, the development of specific harvest control rules would benefit from simulation testing to determine those that meet management objectives given these difficulties. An often used framework for such testing is management strategy evaluation.

2.2.5 TOR5 Evaluate the adequacy, appropriateness, and application of the methods used to project future population status.

The analytical approach was closely tied to the assessment procedure, and involves the forward application of the production model with known catches. This procedure is adequate for projection purposes. However, the assessment documentation does not describe the projection method in detail, and does not describe how the various underreporting scenarios are implemented within the projections.

2.2.6 TOR6 Suggest research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices. Include guidance on single species models, and whether this is possible given the current nature of this multispecies fishery, and difficulties in partitioning fishing effort between species.

I have compiled a set of research priorities, some of which reiterate those made by the previous review. I have also categorized them into those that could be undertaken in the short-term to produce an assessment that could be used to make management recommendations, and more medium-term recommendations.

Short-term

1. Form a data advisory group that consists of HDAR personnel, fishers with knowledge of the fishery history, other relevant government MHI Deep7 data and CPUE experts and the assessment team or representatives. Whether formally constructed or not (from formal meetings/workshops to an email group), this group should have the capacity to sign off on the best available scenario to be used for a base case stock assessment, and also a plausible range of alternatives to be used for sensitivity testing. Such decisions should include the participation of fishers especially, both to give credibility to assessment inputs, and to assist in acceptance of assessment results and any subsequent management actions.

2. Use the data advisory group to review in particular the methods used to filter data going into the CPUE analysis, the scenarios of technological improvement, and the scenarios for underreporting. Such a group is best placed to decide which of these are the best estimates currently available and should be used for a base case assessment. The group should also consider whether the alternative

scenarios developed by the assessment team reasonably characterize the uncertainty.

3. The assessment documentation should be improved to better justify the choices made in setting model priors, to better demonstrate model convergence particularly for the base case, to document the projection procedure, to show errors values for quantities of interest to management in the results, and to show the effect on management quantities of all sensitivities investigated. A summary table of sensitivity results should include for the base case and all sensitivities, as columns: principal direct estimates (*K*, *R*, *M*, *Q*), principal derived quantities (e.g. B_{MSY}, TAC 25%OF2012, TAC 50%OF2012, B_{2013}/B_{MSY}) and individual objective function components (perhaps as a separate table if getting too wide). A comprehensive sensitivity table should include CPUE and catch scenarios as well as alternative model assumptions about model formulations, parameter means and priors. Developing national standards in stock assessment documentation should be implemented.

Medium-term

4. Investigate additional factors that affect the catch rate in consultation with HDAR and industry, and consider standardization using GLMM or GAM.

5. Develop specific harvest control rules for the conversion of MHI Deep7 assessment advice into TACs. Simulation-test the control rules for robustness to uncertainty, and achievement of management goals, particularly given the multispecies nature and spatial heterogeneity of the assessed stock.

6. The problem of having species that may be overfished that form a minor part of the Deep7 complex requires addressing. There needs to be clear management objectives for species or groups of species within the Deep7 complex, and either the ability to carry out more species-specific assessments, or a demonstration that the current management procedure for the complex should achieve management objectives for the component species or groups.

7. Develop systems to improve the collection of non-commercial catch information.

8. Continue research to develop informative priors for *R*, *K* and 1949 proportions of *K*.

9. Consider length frequency sampling for the development of species-specific stock indicators, or movement of the assessment to length/age based approaches.

10. Investigate the utility of a metapopulation assessment model, with a spatially resolved island-specific structure (Hawaii, Maui complex, Oahu, Kauai), to better address island-specific fisheries risk as well as local and regional management options.

11. If the management measures are shown to be sensitive to dispersal rate, then get better species dispersal information to support the potential use of meta-population assessment model.

References

Crab Plan Team 2009. Stock Assessment and Fishery Evaluation Report for the King and Tanner Crab Fisheries of the Bering Sea and Ateutian Regions. North Pacific Fishery Management Council, Anchorage, AK.

Appendix 1: Bibliography of materials provided for review

Hawaii Deepslope Bottomfish
Review Document List

Title	Authors
Updated Assessment report	
Stock Assessment of the Main Hawaiian Islands Deep7 Bottomfish Complex Through 2010	Jon Brodziak, Dean Courtney, Lyn Wagatsuma, Joseph O'Malley, Hui-Hua Lee, William Walsh, Allen Andrews, Robert Humphreys and Gerard DiNardo
Reference documents	
Analysis And Expansion Of The 2005 Hawai'i State/Western Pacific Regional Fishery Council Bottomfish Fishermen Survey	Megan R. Lamson, Blake Mcnaughton and Craig J. Severance
Review of Unreported to Reported Catch Ratios for Bottomfish Resources in the Main Hawaiian Islands for Application to Stock Assessment	Dean Courtney
Estimation of Bottomfish CPUE using the Delta Method and HDAR Logbooks 1948-2010.	Kevin Piner and Hui-Hua Lee
CPUE Standardization Workshop Proceedings August 4-6, 2008	Robert B. Moffitt , Gerard DiNardo, Jon Brodziak, Kurt Kawamoto, Michael Quach, Minling Pan, Karl Brookins , Reginald Kokubun, Clayward Tam, and Mark Mitsuyasu
Status and Trends of the Hawaiian Bottomfish Stocks: 1948-2004	Steven J.D. Martell, Josh Korman, Meaghan Darcy, Line B. Christensen and Dirk Zeller
Investigation of the association between Hawaii deep slope bottomfish CPUE and environmental variables	Hui-hua Lee and Jon Brodziak
Bottomfish Stock Assessment Workshop January 13-16, 2004	Western Pacific Fishery Management Council
Status of the Hawaiian Bottomfish Stocks, 2004	Robert B. Moffitt, Donald R. Kobayashi, and Gerard T. DiNardo
Hawaiian Bottomfish Assessment Update for 2008	Jon Brodziak, Robert Moffitt and Gerard DiNardo
Agenda: Bottomfish CPUE Standardization Workshop 2007	
Report of the Western Pacific Stock Assessment Review 1 Hawaii Deep Slope Bottomfish	Robert A. Skillman, Chair
Report on the Western Pacific Stock Assessment Review 1 Hawaii Deep Slope Bottomfish (WSPAR 1)	Kevin Stokes

Appendix 2: A copy of the CIE Statement of Work

Statement of Work for Dr. Neil Klaer

External Independent Peer Review by the Center for Independent Experts

Hawaii Deepslope Bottomfish

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: A peer review of the Hawaiian multispecies deepslope bottomfish resource is required using the CIE process. The scientific information and assessment for Hawaiian deepslope bottomfish was peer reviewed in June 2009 providing recommendations to increase the accuracy of the assessment. The objective of this review is to conduct a follow-up peer review to determine if the recommendations have been adequacy addressed and adequacy of the revised assessment for management purposes. The assessment has a large potential impact on a valuable fishery important to commercial and recreational fishers in Hawaii and fish consumers in the state. It forms the basis of bottomfish management decisions by the Western Pacific Regional Fishery Management Council (WPFMC), NMFS, and the State of Hawaii. 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 as a desk review, therefore no travel is required.

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

<u>Prior to the Peer Review</u>: Upon completion of the CIE reviewer selection by the CIE Steering Committee, the CIE shall provide the CIE reviewer information (full name, title, affiliation, country, address, email) to the COTR, who forwards this information to the NMFS Project Contact no later the date specified in the Schedule of Milestones and Deliverables. The CIE is

responsible for providing the SoW and ToRs to the CIE reviewers. The NMFS Project Contact is responsible for providing the CIE reviewers with the background documents, reports, and other pertinent information. Any changes to the SoW or ToRs must be made through the COTR prior to the commencement of the peer review.

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

This list of pre-review 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, each CIE reviewer is responsible only for the pre-review documents that are delivered to the reviewers in accordance to the SoW scheduled deadlines specified herein.

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

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

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

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Conduct an independent peer review in accordance with the ToRs (**Annex 2**).
3) No later than 28 January 2011, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Mr. Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and Dr. David Die, CIE Regional Coordinator, via email to ddie@rsmas.miami.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

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

4 January 2011	CIE sends each reviewer contact information to the COTR, who then sends this to the NMFS Project Contact
7 January 2011	NMFS Project Contact sends the CIE Reviewers the pre-review background documents
13 January 2011	Project contact provides the CIE reviewers with the report to be peer reviewed
14-28 January 2011	Each reviewer conducts an independent peer review as a desk review
28 January 2011	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
11 February 2011	CIE submits CIE independent peer review reports to the COTR
Feb. 15 2011	The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: Requests to modify this SoW must be 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 reviewers 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 report by the CIE Lead Coordinator, Regional Coordinator, and Steering Committee, this report 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 report 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, Program Manager, COTR
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-713-2363 ext 136

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

NMFS Project Contact:

Gerard DiNardo
Pacific Islands Fisheries Science Center
2570 Dole Street
Honolulu, HI 96822-2396
Robert.Moffitt@noaa.gov
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.

3. The reviewer report shall include the following appendices:

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

Annex 2: Terms of Reference for the Peer Review

Hawaii Deepslope Bottomfish

1. Determine if recommendations from the June 2009 WPSAR/CIE review have been adequately addressed within the assessment update. .

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

3. Evaluate the implementation of the assessment model: configuration, assumptions, and input data and parameters (fishery life history); more specifically determine if data are properly used, if choice of input parameters seem reasonable, if models are appropriately specified and configured, assumptions are reasonably satisfied, and primary sources of uncertainty accounted for.

4. Comment on the scientific soundness of the estimated population benchmarks and management parameters (e.g. MSY, Fmsy, Bmsy, MSST, and MFMT) and their potential efficacy in addressing the management goals stated in the relevant FMP or other documents provided to the review panel.

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

6. Suggest research priorities to improve our understanding of essential population and fishery dynamics necessary to formulate best management practices. Include guidance on single species models, and whether this is possible given the current nature of this multispecies fishery, and difficulties in partitioning fishing effort between species.