GARM-III "Models" Meeting, 25-29 February 2008:
Center for Independent Experts peer review report

José De Oliveira

Executive Summary

I am in full agreement with the GARM Panel Summary Report with regard to the quality and soundness of the science, and the methods used relating to each of the Terms of Reference (TOR). Additional comments and elaborations, which may merit further consideration, include the use of 2-stage population models in "data poor" situations (TOR 1 and 5), alternative error assumptions for proportion at age data (TOR 2), two suggestions for dealing with zero or close-to-zero data (variable aggregation for catch- or index-at-age data, and density-dependent additional variance; TOR 3), the use of an existing VPA-ADAPT framework for estimating missing catch (TOR 4), the use of additional variance for weighting likelihood components (TOR 5), and modelling spatially-dependent survey *q*s (TOR 5). Further suggestions are made (not linked to specific TORs) with regard to potentially useful pre-assessment diagnostic tools.

Comments

I have taken each TOR in turn and provided comments. Additional comments not directly related to the TOR appear at the end of the report.

TOR 1
Applicability of one or more of the following modelling approaches to assess stock status: index methods; production models; age- or length-based models

My comments are adequately expressed by the GARM Panel Summary Report, but I have an additional comment. A class of models not considered in great detail at the meeting, but which may merit consideration (see TOR 5 below) are those that fall between production/replacement yield models on the one hand and age-based models on the other. These include delay difference models (Deriso 1980, but see also Dichmont *et al.* 2003 for recent application), and 2-stage population models (the approach of Collie and Sissenwine 1983 was mentioned, but see also Mesnil 2003 and Roel and Butterworth 2000).

References:

Collie, J.S., Sissenwine, M.P., 1983. Estimating population size from relative abundance data measured with error. Can. J. Fish. Aquat. Sci. 40, 1871–1879.

Deriso, R.B., 1980. Harvesting strategies and parameter estimation for an age-structured model. Can. J. Fish. Aquat. Sci. 37, 268-282.

Dichmont, C.M., Punt, A.E., Deng, A., Dell, Q., Venables, W., 2003. Application of a weekly delay-difference model to commercial catch and effort data for tiger prawns in Australia's Northern Prawn Fishery. Fish. Res. 65, 335-350.

Mesnil, B. 2003. The Catch-Survey Analysis (CSA) method of fish stock assessment: an evaluation using simulated data. Fisheries Research 63, 193-212.

Roel, B. A., Butterworth, D. S. 2000. Assessment of the South African chokka squid *Loligo vulgaris reynaudii*. Is disturbance of aggregation by the recent jig fishery having a negative impact on recruitment? Fisheries Research 48, 213-228.

TOR 2

For certain stocks that are aged, utility of statistical catch-at-age vs. VPA based models with respect to: retrospective patterns; flexibility to account for alternative parameterisations; ability to incorporate external sources of information, especially tagging and environmental data; and ability to estimate parameters incorporating prior, external information

My comments are adequately expressed by the GARM Panel Summary Report, but I have some further comments. An alternative error assumption for proportion at age data, which avoids the need to specify a value for effective sample size (often done on an *ad hoc* basis, and needed when a multinomial assumption is made) is to use an adjusted normal (after log-transform) or lognormal formulation. The "adjustment" refers to an adjustment to the variance, so that proportions that make up a small fraction of the catch are not assigned too much weight (Punt and Kennedy 1997).

An example of an adjusted normal (after log-transform) is given in WP2.2a (equation A2.19), and for the adjusted lognormal in Punt and Kennedy (1997, equation A3.6). Apart from the normal/lognormal difference between the formulations, a further difference is the adjustment made to the variance: in the former the observed proportions are used, but in the latter the model estimated proportions are used. Simulation work conducted by Ernst (1992), in which he compared an observation-based variance to a model estimate-based variance revealed better performance for the former. Therefore, more recently, preference has been given to the observation-based variance formulation (Punt pers. comm.), even though Ernst's conclusions were based on a different likelihood formulation (the robust normal likelihood of Fournier *et al*. 1990).

Ernst (1992) also showed that Fournier's robust likelihood formulation performed better than a multinomial likelihood in situations when the data were contaminated with outliers (but comparably when there was no contamination, see also Maunder and Langley 2004, section 2.2).

References:

Ernst B. 1992. An investigation on length-based methods used in quantitative population modelling. Ph.D. Dissertation. University of Washington, Seattle, WA.

Fournier, D.A., Sibert, J.R., Majkowski, J., Hampton, J. 1990. MULTIFAN a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using data for southern bluefin tuna (Thunnus maccoyii). Can. J. Fish. Aquat. Sci. 47, 301-317.

Maunder, M.N., Langley, A.D. 2004. Integrating the standardization of catch-per-unit-effort into stock assessment models: testing a population dynamics model and using multiple data types. Fish. Res. 70, 389-395

Punt, A.E., Kennedy, R.B. 1997. Population modelling of Tasmanian rock lobster, Jasus edwardsii, resources. Mar. Freshwater Res. 48, 967-980.

TOR 3

Implications of zeros in the evaluation of fishery independent indices

My comments are adequately expressed by the GARM Panel Summary Report, but I wish to expand on one of the ideas presented, namely different aggregation for the catch- or index-at-age data. One idea to get around the problem of zero data, while still keeping the lognormal error assumption, is to aggregate the data so as to have a minus- and plus-group in the data (different from the model plus-group), so that the data potentially has a varying minus- and plus-group (i.e. the data now appear as a ragged array). The corresponding model estimates should then be aggregated in the same way for the purpose of fitting these data in the objective function.

A possible approach to modelling survey data of depleted resources is to have a likelihood formulation with a density-dependent additional variance component, so that as the resource becomes scarcer, the total variance associated with survey data is inflated (De Oliveira *et al*. 2007).

References:

De Oliveira, J.A.A., Boyer, H.J., Kirchner, C.H. 2007. Developing age-structured stock assessment models as a basis for management procedure evaluations for Namibian sardine. Fisheries Research 85, 148-158.

TOR 4

Potential factors responsible for retrospective patterns

My comments are adequately expressed by the GARM Panel Summary Report, but I have an additional comment. Unrecorded changes in catch reporting are presented as one of the possible causes for retrospective patterns. This is the case for North Sea cod, for which an assessment method is used that estimates misreported catch within a VPA-ADAPT framework (called BADAPT, Darby 2005). It relies on a period of reliable catch data, and survey data being available that has sufficient overlap between the period of reliable catch data, and the period where misreporting occurs. BADAPT is used as a basis for advice for North Sea cod, and could be made available, on request, for use as an exploratory tool for the stocks considered by GARM-III.

References:

Darby, C.D. 2005. Appendix 4. Estimating systematic bias in the North Sea cod landings data. In: Report on the assessment of demersal stocks in the North Sea and Skaggerrak. 7-16 September 2004, Bergen, Norway. ICES CM 2005/ACFM:07.

TOR 5

For each stock, define assessment models that will be used to determine stock status and productivity characteristics until the next "benchmark" assessment is conducted. Where possible, apply the models to data (probably through 2006), to obtain current and historical estimates of F and B, and estimates of uncertainty

My comments are adequately expressed by the GARM Panel Summary Report, but I have a few additional comments.

With regard to "weighting of model components", the concept of additional variance marries the internal and external approaches to weighting likelihood components. The internal approach derives weighting iteratively or by maximum likelihood estimation, whereas the external approach uses input weights (which could be sampling CVs derived externally to the assessment model). A disadvantage of the internal approach, when more than one index is used, is that unrealistically high precision can be accorded certain indices, particularly when there are conflicting trends. On the other hand, the external approach requires variance estimates that are reliable enough to be used, and the variance estimates for the various indices may not be comparable (e.g. some may account for more explanatory variables than others in a GLM, simply because of data availability, Geromont and Butterworth 2001). The use of additional variance gets around some of these issues, and combines the best features of these two approaches. When applying the additional variance approach, externally derived input variances are used, but the additional variance that is estimated internally for each index series allows for the fact that the input variances (usually a sampling CV) may not capture all the variance associated with the index (Punt et al. 1997, Geromont and Butterworth 2001). Because the additional variance is not allowed to be less than 0, the approach avoids the internal weighting problem of according unrealistic weight to certain of the indices.

With regard to Georges Bank yellowtail flounder, the GARM Panel Summary Report states that "Yellowtail might be occupying more preferred habitat due to environmental or other influences, causing survey catchability to increase. This hypothesis should be investigated." Since no further detail is given on this suggestion in the report, I provide more detail. Say a survey index I covers the entire stock area, but can be disaggregated into three sub-areas: $I = I_1+I_2+I_3$. In the aggregated case, a single survey q is estimated and the objective function takes the form:

$$\sum_y \left[\ln(\sum_i I_i(y)/q) - \ln(B(y)) \right]^2$$

In the disaggregated case (i.e. when investigating the above hypothesis), three survey qs are estimated, and the objective function takes the form:

$$\sum_y \left[\ln(\sum_i (I_i(y)/q_i)) - \ln(B(y)) \right]^2$$

With regard to the windowpane flounder stocks, an alternative to SCAA to fit data partitioned into recruitment and plus-group components is to make use of delay-difference or 2-stage biomass models (see TOR 1)

References:

Geromont, H.F., Butterworth, D.S. 2001. Possible extensions to the ADAPT VPA model applied to western north Atlantic bluefin tuna, addressing in particular the need to account for "additional variance". Col. Vol. Sci. Pap. ICCAT 52(5), 1663-1705.

Punt, A.E., Cooke, J.G., Borchers, D.L., Strindberg, S. 1997. Estimating the extent of additional variance for Southern Hemisphere minke whales from the results of the IWC/IDCR cruises. Rep. int. Whal. Comm. 47, 431-434.

TOR 6

Sufficiency of the assessment models to estimate measures of stock status consistent with Biological Reference Points

My comments are adequately expressed by the GARM Panel Summary Report.

Additional comments outside the TORs

There are a number of pre-assessment diagnostic tools that we have found helpful at ICES Working Groups (see e.g. ICES 2008) that stock assessors may be interested in, and which would supplement the information provided in WP1.1 (data summaries for the 19 stocks). These include a look at catch curves with associated negative gradients over a reference age range (as a proxy to total mortality; Fig. 14.10), and catch-at-age data expressed as proportions at age, and then standardised across years (see Fig 14.9 for details) – for the latter, strong year classes are clearly visible, as is the severe contraction in the age range. Another useful diagnostic is within-cohort correlation plots, which gives some indication of consistency in the data (in terms of year class strength; Fig. 12.3.3). These diagnostics could be repeated for survey data, including not only within-survey cohort correlation plots, but also between-survey correlation plots by age (to check how consistent the data are between these surveys). The R code functions for these plots are available, and can be supplied if there is interest.

References:

ICES 2008. Report of the Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak – Spring and Autumn (WGNSSK), 1-8 May, ICES Headquarters, Copenhagen, and by correspondence. 960pp.

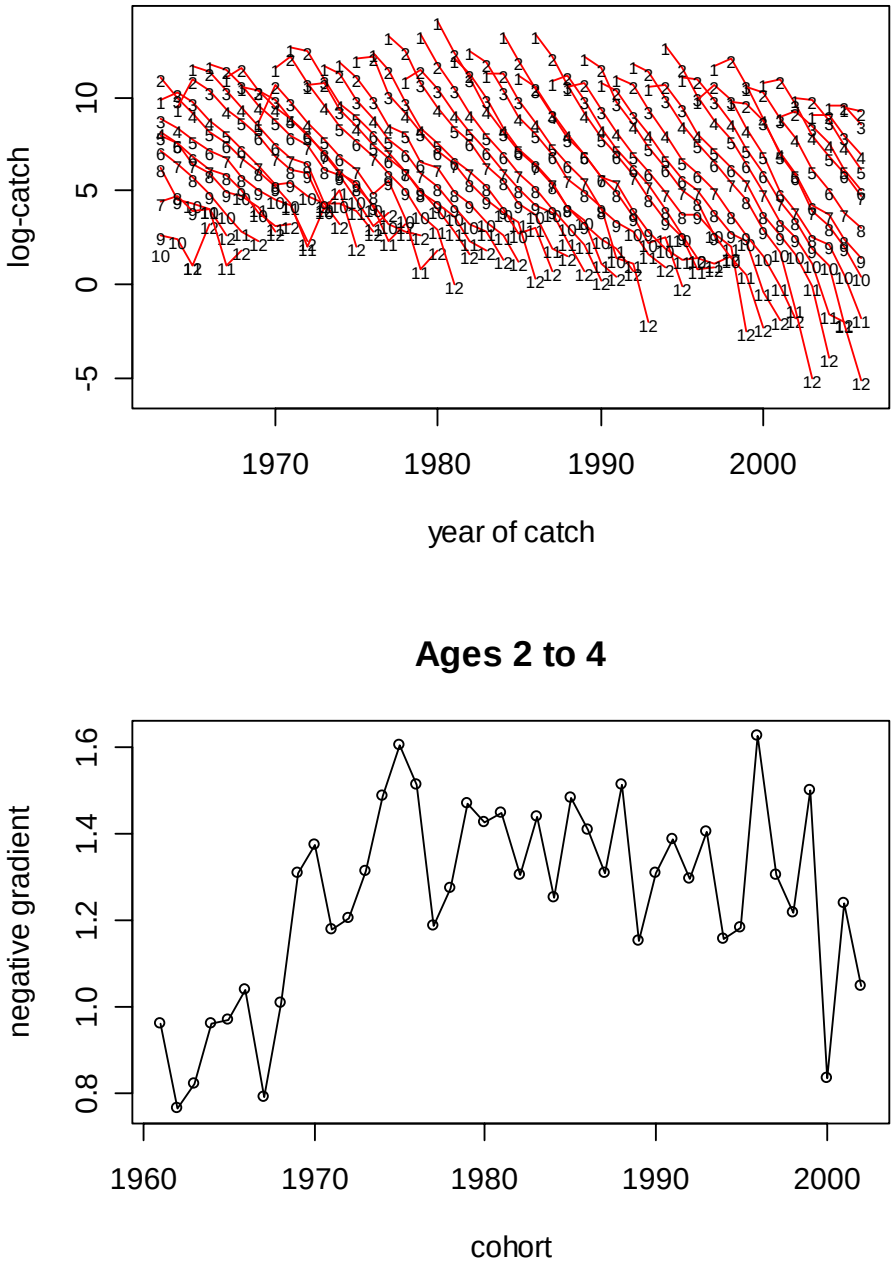

Figure 14.10 Cod in Subarea IV and Divisions IIIa (Skagerrak) and VIId. Log-catch cohort curves (top panel) and the associated negative gradients for each cohort across the reference fishing mortality of age 2-4. Taken from ICES 2008.

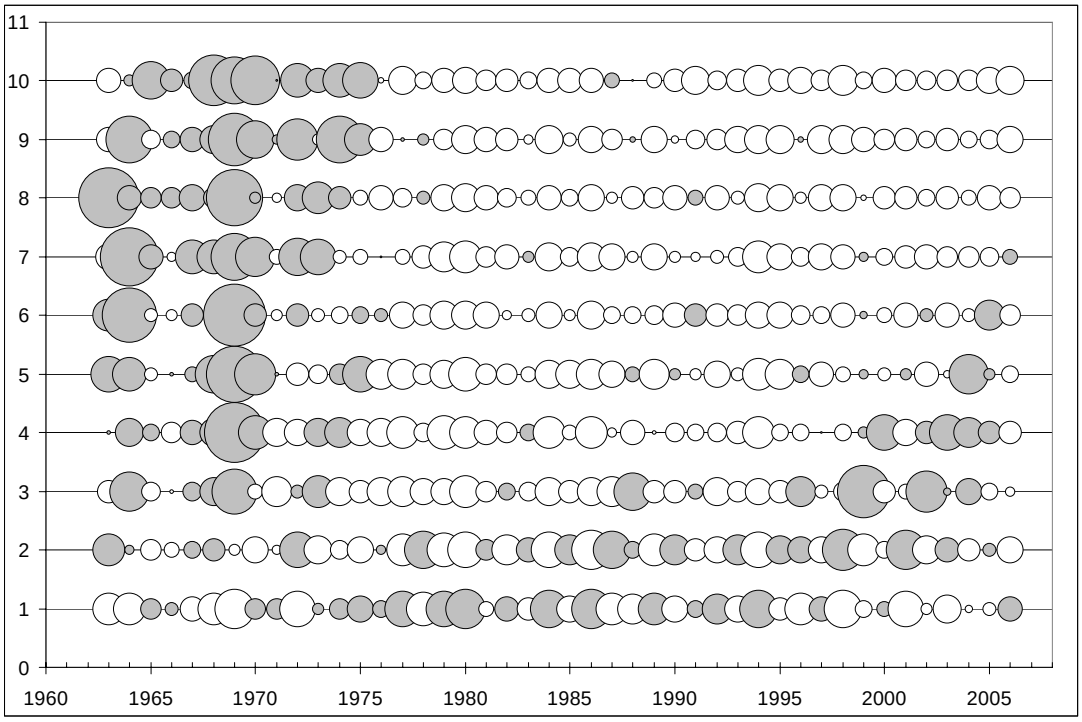

Figure 14.9 Cod in Subarea IV and Divisions IIIa (Skagerrak) and VIId. Total catch-at-age matrix expressed as proportions-at-age which have been standardised over time (for each age, this is achieved by subtracting the mean proportion-at-age over the time series, and dividing by the corresponding variance). Grey bubbles indicate proportions above the mean over the time series at each age. Taken from ICES 2008.

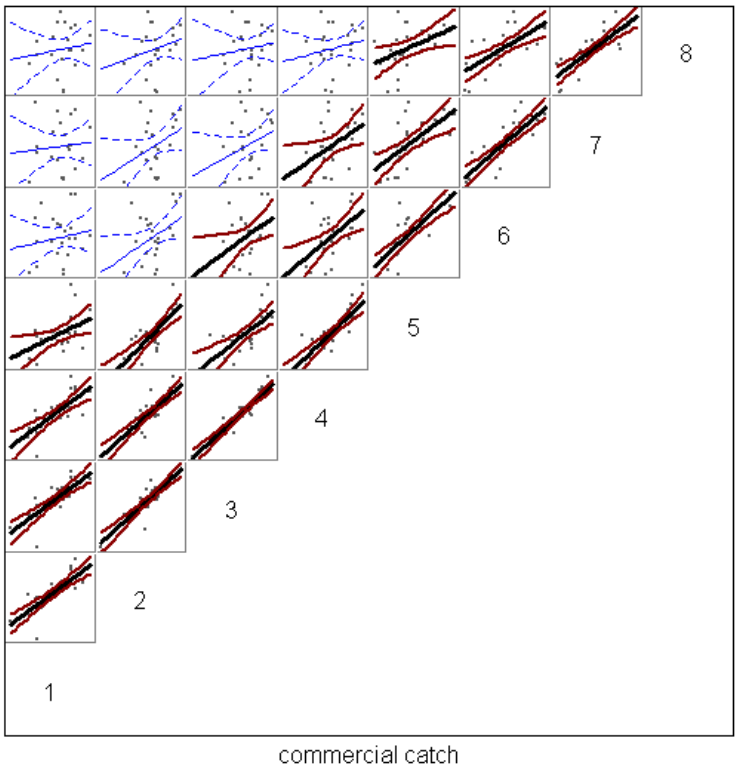

Figure 12.3.3 Whiting in IV and VIId. Correlations in the catch at age matrix (log numbers). Individual points are given by cohort (e.g. a point in the age 1-age 2 plot reflects numbers at age 1, and 1 year later in the same cohort at age 2), the line is a normal linear model fit. Thick lines represent a significant ($p<0.05$) regression. Taken from ICES 2008.

Appendix 1: Background material

1.1 Rago, et al. 2008. Data Summary for Nineteen Groundfish Stocks in the Northeast U.S.

1.2 Rago, et al. 2008. Overview of Assessment Methods and Model Selection Criteria for Nineteen Groundfish Stocks in the Northeast U.S.

M.1 Mayo R, Col L, Traver M. 2008. Data Summary of Catch and Abundance Measures. (for Working Papers 1.1 & 1.2)

1.3 Hendrickson L, Col L. 2008. Maps Showing NEFOP Sampling Coverage and Management Areas

2.1 Jacobson L, Legault C, O'Brien L, Sosebee K. 2008. Utility of Statistical Catch-at-age Models for Assessing Northeast Groundfish Stocks (a workshop report)

2.2a Butterworth DS, Rademeyer RA. 2008. Statistical Catch-at-age Analysis *vs* ADAPT-VPA: The Case of Gulf of Maine Cod

2.2b Rademeyer RA, Butterworth DS. 2008. Retrospective Analysis for the Gulf of Maine Cod ASPM Reference Case Assessment

2.3 Shepherd G. 2008. Comparison of ADAPT VPA and ASAP Models for Gulf of Maine Cod

2.4 Brooks L, Legault C, Nitschke P, O'Brien L, Sosebee K, Rago P, Seaver A. 2008. Evaluation of NMFS Toolbox Assessment Models on Simulated Groundfish Data Sets

2.5 Butterworth DS, Rademeyer RA. 2008. Application of and Age-Structured Production Model to the Georges Bank Yellowtail Flounder

3.1 Legault C, Seaver A. 2008. Simulation Studies of Issues Associated with Filling Zeros in VPA Tuning Indices

3.2 Terceiro M. 2008. The Treatment of "Zero" Observations in the Summer Flounder ADAPT VPA Calibration

3.3 ICES Resource Management Committee. 2007. Report of the Working Group on Methods of Fish Stock Assessment (WGMG)

4.1 Legault et al. 2008. Report of the Retrospective Working Group

5.1 Rago et al. 2008. Recommended Modeling Approaches by Stock Initial Proposals for Consideration by the GARM III Review Panel

6.1 Rago et al. 2008. Sufficiency of Models for Biological Reference Points

Supplementary Papers

TOR 1 Gulf of Maine Winter Flounder SCALE Run 2

TOR 2 Miller T, Hart D, Cadrin S, Jacobson L, Legault C, Rago P. 2008. Analyses of Tagging Data for Evidence of Decreased Fishing Mortality for Large Yellowtail Flounder

TOR 2 Butterworth DS, Rademeyer RA. 2008. On Drawing Inferences Concerning Trends in Selectivity with Age from Tag-Recapture Information

TOR 2 Jacobson L. 2008. Questions about the Adjusted Lognormal Error Distribution used in Calculating Goodness of Fit for Survey and Commercial Age Composition Data in Preliminary ASPM Models.

TOR 2 Jacobson L. 2008. Pope vs Baranov

Appendix 2: Statement of Work for Dr. José DeOliveira (CEFAS)

External Independent Peer Review by the Center for Independent Experts

GARM-III "Models" Meeting:

Statement of Work (SOW) for CIE Panelists
(including description of GARM-III Chairman's duties)

General

The Groundfish Assessment Review Meeting (GARM) brings together stock assessment experts to peer review work on the status of 19 important fish stocks that are managed by the New England Fishery Management Council. GARM-III takes place in 2007-2008, and it will consist of four meetings that are cumulative in nature (i.e., successive meetings incorporate methods and results that were accepted at previous GARM-III meetings). Each meeting will have a chair as well as external panelists. A brief description and dates of the four GARM-III meetings are given below:

1. "Data" Meeting (October 29 – November 2, 2007)
Review the commercial and survey data that will be used in the stock assessments. Identify appropriate statistical methods for analyzing those data (including bycatch and discard issues, changes in growth rates and other life history traits, issues related to merging databases, etc.). Other sources of data to be considered are tagging programs for cod and yellowtail flounder, and Industry-Based Surveys. Candidate sources of data relevant to ecological and ecosystem considerations will also be described.

2. "Modeling" Meeting (February 25 – 29, 2008)
Determine the most appropriate stock assessment methods and models for each of the 19 stocks. Perform runs of those models to obtain results (historical and current estimates of F and B) based on commercial and survey data, probably through calendar year (CY) 2006. Evaluate retrospective patterns and their importance for status determination.

3. "Biological Reference Point (BRP)" Meeting (April 28 – May 2, 2008)
Update or redefine BRPs for each of the 19 stocks. Use data available through CY2006. Consider whether the BRPs are reasonable in light of results from the "Modeling" Meeting. Define the appropriate initial conditions for forecasting and rebuilding strategies, particularly with respect to trends in biological attributes, recruitment and survival rates. Comment on relevant ecosystem considerations as they relate to rebuilding strategies.

4. GARM-III "Final" Meeting (August 4 - 8, 2008)
Use all of the methods proposed from the previous three meetings, along with survey and catch information through CY2007, to estimate fishing mortality rates and biomass for each stock. Based on procedures from the BRP Meeting, finalize the BRPs, appropriate initial conditions, and

biological assumptions related to forecasts. Determine the status of each stock.

This SOW applies specifically to the GARM-III "Modeling" Meeting, which will take place at the Woods Hole Laboratory of the Northeast Fisheries Science Center (NEFSC) in Woods Hole, Massachusetts, from February 25 - 29, 2008. The meeting will have a chairman (non-CIE) as well as external panelists, three of whom will be from the Center of Independent Experts (CIE).

Overview of CIE Peer Review Process:

The Office of Science and Technology implements measures to strengthen the National Marine Fisheries Service's (NMFS) Science Quality Assurance Program (SQAP) to ensure the best available high quality science for fisheries management. For this reason, the NMFS Office of Science and Technology coordinates and manages a contract for obtaining external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of stock assessments and various scientific research projects. The primary objective of the CIE peer review is to provide an impartial review, evaluation, and recommendations in accordance to the Statement of Work (SoW), including the Terms of Reference (ToR) herein, to ensure the best available science is utilized for the National Marine Fisheries Service management decisions.

The NMFS Office of Science and Technology serves as the liaison with the NMFS Project Contact to establish the SoW which includes the expertise requirements, ToR, statement of tasks for the CIE reviewers, and description of deliverable milestones with dates. The CIE, comprised of a Coordination Team and Steering Committee, reviews the SoW to ensure it meets the CIE standards and selects the most qualified CIE reviewers according to the expertise requirements in the SoW. The CIE selection process also requires that CIE reviewers can conduct an impartial and unbiased peer review without the influence from government managers, the fishing industry, or any other interest group resulting in conflict of interest concerns. Each CIE reviewer is required by the CIE selection process to complete a Lack of Conflict of Interest Statement ensuring no advocacy or funding concerns exist that may adversely affect the perception of impartiality of the CIE peer review. The CIE reviewers conduct the peer review, often participating as a member in a panel review or as a desk review, in accordance with the ToR producing a CIE independent peer review report as a deliverable. The Office of Science and Technology serves as the COTR for the CIE contract with the responsibilities to review and approve the deliverables for compliance with the SoW and ToR. When the deliverables are approved by the COTR, the Office of Science and Technology has the responsibility for the distribution of the CIE reports to the Project Contact.

Requirements for CIE Reviewers:

Three CIE reviewers are requested to conduct an impartial and independent peer review in accordance with the Terms of Reference (ToR) herein. Each CIE reviewer's duties shall not exceed a maximum of 14 days conducting pre-review preparations with document review, participation on the SARC panel review meeting, editorial assistance to the SARC Chair, and completion of the CIE independent peer

review report in accordance with the ToR and Schedule of Milestones and Deliverables. CIE reviewers shall have working knowledge and recent experience in the application of modern fishery stock assessment models. Expertise should include both the use of statistical catch-at-age and traditional VPA approaches. Experience with comparative studies of these approaches is especially valuable. Reviewers should also have experience in evaluating measures of model fit, identifiability, uncertainty, and forecasting. Some experience with groundfish (such as cod, haddock, flounder) population dynamics would be useful.

Specific Activities and Responsibilities

The CIE's deliverables shall be provided according to the schedule of milestones listed on page 5. The GARM Chair will use contributions from the CIE panelists as well as from other external panelists, to produce the GARM Panel Summary Report. In addition, each CIE panelist will write an individual independent report. These reports will provide peer-review information for a presentation to be made by NOAA Fisheries at meetings of the New England and Mid-Atlantic Fishery Management Councils in 2008. The GARM Panel Summary Report shall be an accurate and fair representation of the GARM panel viewpoint on the quality and soundness of the science, methods and data with regard to each Term of Reference (see Annex 1). The report shall also contain recommendations for improvement that might be implemented in a future GARM meeting.

Charge to GARM panel

The panel is to determine and write down its viewpoint on the quality and soundness of the science, methods and data with regard to each Term of Reference (see Annex 1). Criteria to consider include whether: (1) the data are adequate and were used properly; (2) the analyses and models were appropriate and correctly accomplished; and (3) the conclusions are correct/reasonable. Where possible, the chair shall identify or facilitate agreement among the panelists regarding each Term of Reference.

During the course of the review, the panel is allowed limited flexibility to deviate from the results and recommendations of earlier GARM-III meetings. This flexibility may include minor alterations in procedures previously established at the peer review of the Data Methods Meeting in October 2007. Large scale changes, such as changing a stock definition would not be possible in view of the difficulties of implementing these changes in time available before the final GARM meeting in August 2008.

Furthermore, if the panel rejects certain assessment models, the panel should explain why those particular models are not suitable, and the panel should recommend suitable alternatives. If such alternatives cannot be identified, then the panel should indicate that the existing (status quo) models are the best available at this time.

Roles and responsibilities

(1) Prior to the meeting

(GARM Chair and CIE panelists)
Review the reports produced by the Working Groups, and read background reports.

(2) During the Open meeting

(GARM Chair)
Act as chairperson, where duties include control of the meeting, coordination, control, and facilitation of the presentations and discussions, and ensuring that all Terms of Reference of the GARM are reviewed and completely addressed.

During the question and answer periods, provide appropriate feedback to the assessment scientists on the sufficiency of the analyses and when possible, suggest improved approaches. It is permissible to discuss the working papers, and to request additional information to clarify or revise existing analyses, if that information can be produced rather quickly.

(CIE panelists)
For each model approach, participate in panel discussions on the quality and soundness of the science, methods and data with regard to each Term of Reference (see Annex 1).

During the question and answer periods, provide appropriate feedback to the assessment scientists on the sufficiency of the analyses. It is permissible to request additional information if it is needed to clarify or revise existing analyses, if that information can be produced rather quickly.

(3) After the Open meeting

(GARM CIE panelists)
Each panelist shall prepare a CIE independent peer review report (see Annex 2). This report should comment on the quality and soundness of the science, methods, and data with regard to each Term of Reference.

If any modeling approaches are considered inappropriate, the CIE independent peer review report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing modeling approaches are the best available at this time.

During the meeting, additional questions that are not in the Terms of Reference but which are directly related to the assessments may be raised. Comments on these questions should be included in a separate section at the end of the CIE independent peer review report prepared by each panelist.

If a panelist feels that his/her comments are adequately expressed in the GARM Panel Summary Report, it will not be necessary to repeat the same comments in the Independent CIE Report. In this case, the CIE independent peer review report can be used to provide greater detail on specific Terms of Reference or additional questions raised during the meeting.

(GARM Chair)
The GARM Chair shall prepare a document summarizing the background of the work to be conducted as part of the review process, and summarizing whether the process was adequate to successfully address the Terms of Reference. If appropriate, the chair will include suggestions on how to improve the process. This document will constitute the introduction to the GARM Panel Summary Report.

(GARM Chair, CIE and non-CIE panelists)
The GARM Chair will take the lead in preparing, editing, and completing the GARM Panel Summary Report, based on contributions from the external panelists (CIE and non-CIE). The panelists and the chair will discuss their views on each Term of Reference and whether their opinions can be summarized into a single conclusion for all—or only for some—of the Terms of Reference. For TORs where a consensus view can be reached, the GARM Panel Summary Report will contain a summary of such views. In cases where multiple and/or differing views exist on a given Term of Reference, the GARM Panel Summary Report will note that there was no agreement and will specify—in a summary manner—what the various opinions are and the reason(s) for the different opinions.

The Chair's objective during this Summary Report development process will be to identify or facilitate the finding of an agreement, rather than forcing the panel to reach an agreement if this is not possible.

The GARM Panel Summary Report (please see Annex 3 for information on contents) should comment on the quality and soundness of the science, methods, and data with regard to each Term of Reference.

If any modeling approaches are considered inappropriate, the GARM Panel Summary Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing modeling approaches are the best available at this time.

The contents of the draft GARM Panel Summary Report will be approved by the CIE panelists by the end of the Summary Report development process. The GARM chair will finalize all editorial and formatting changes prior to approval of the contents of the draft GARM Panel Summary Report by the CIE panelists. The GARM chair will then submit the approved GARM Panel Summary Report to the NEFSC contact (i.e., SAW Chair).

Schedule of Milestones and Deliverables

The milestones and schedule are summarized in the table below. No later than March 14, 2008, the CIE panelists should submit their CIE independent peer review reports to the CIE for review[1]. The CIE reports shall be sent to "University of Miami Independent System for Peer Review," and sent to Dr. David Sampson, via e-mail to David.Sampson@oregonstate.edu and to Mr. Manoj Shivlani via e-mail to mshivlani@ntvifederal.com

Milestone	**Date**
Open workshop at Northeast Fisheries Science Center (NEFSC) (begin writing reports, as soon as open Workshop ends)	Feb. 25 – 29, 2008
GARM Chair and CIE panelists work at the NEFSC drafting reports	Feb. 28 – 29
Draft of GARM Panel Summary Report, reviewed by all CIE panelists, due to the GARM Chair **	March 14
CIE panelists submit CIE independent peer review reports to CIE for approval	March 14
GARM Chair sends Final GARM Panel Summary Report, approved by CIE panelists, to NEFSC contact (i.e., SAW Chairman)	March 21
CIE provides reviewed CIE independent peer review reports to NMFS COTR for approval	March 28
COTR notifies CIE of approval of CIE independent peer review reports	April 4 *
COTR provides final CIE independent peer review reports to NEFSC contact	April 4

* Assuming no revisions are required of the reports.
** The GARM Panel Summary Report will not be submitted, reviewed, or approved by the CIE.

The SAW Chairman will assist the GARM chair prior to, during, and after the meeting in ensuring that documents are distributed in a timely fashion. NEFSC staff and the SAW Chairman will make the final GARM Panel Summary Report and CIE independent peer review reports available to the public. Staff and the SAW Chairman will also be responsible for production and publication of the collective Working Group papers.

Acceptance of Deliverables:

Upon review and acceptance of the CIE reports by the CIE Coordination and Steering Committees, CIE shall send via e-mail the CIE reports to the COTRs (William Michaels William.Michaels@noaa.gov and Stephen K. Brown Stephen.K.Brown@noaa.gov) at the NMFS Office of Science and Technology by the date in the Schedule of Milestones and Deliverables. The COTRs will review the CIE reports to ensure compliance with the SoW and ToR herein, and have the responsibility of approval and acceptance of the deliverables. Upon notification of acceptance, CIE shall send via e-mail the final CIE report in *.PDF format to the COTRs. The COTRs at the Office of Science and Technology have the responsibility for the distribution of the final CIE reports to the Project Contacts.

[1] All reports will undergo an internal CIE review before they are considered final.

Key Personnel:

Contracting Officer's Technical Representative (COTR):

William Michaels
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

Stephen K. Brown
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
Stephen.K.Brown@noaa.gov Phone: 301-713-2363 ext 133

Contractor Contacts:

Manoj Shivlani, CIE Primary Coordinator
10600 SW 131st Court, Miami, FL 33186
mshivlani@ntvifederal.com Phone: 305-383-4229

Roger Peretti, NTVI Regional Director
Northern Taiga Ventures, Inc., 814 W. Diamond Ave., Ste. 250, Gaithersburg, MD 20878
rperetti@ntvifed.com Phone: 301-212-4187

Project Contact:

James Weinberg, NEFSC Contact person and SAW Chairman
NMFS Northeast Fisheries Science Center, 166 Water Street, Woods Hole, MA 02543
James.Weinberg@noaa.gov Phone: 508-495-2352

Request for Changes:

Requests for changes shall be submitted to the Contracting Officer at least 15 working days prior to making any permanent substitutions. The Contracting Officer will notify the Contractor within 10 working days after receipt of all required information of the decision on substitutions. The contract will be modified to reflect any approved changes. The Terms of Reference (ToR) and list of pre-review documents herein may be updated without contract modification as long as the role and ability of the CIE reviewers to complete the SoW deliverable in accordance with the ToR are not adversely impacted.

ANNEX 1:

Draft Terms of Reference for the GARM-III "Models" Meeting

(Last Revised: Oct. 31, 2007; A final draft will be distributed to the Panel prior to the meeting.)

1. For each stock, consider the applicability of one or more of the following modeling approaches to assess stock status:
    - Index methods
    - Production Models
    - Age- or Length-based Models

2. For certain stocks that are aged, compare and contrast the utility of statistical catch-at-age vs. VPA based models with respect to the following criteria:
    - Retrospective patterns
    - Flexibility to account for alternative parameterizations
    - Ability to incorporate external sources of information, especially tagging and environmental data
    -Ability to estimate parameters incorporating prior, external information.

3. Address the implications of zeros in the evaluation of fishery independent indices.

4. Examine potential factors responsible for retrospective patterns.

5. For each stock, define the assessment model that will be used to determine stock status and productivity characteristics until the next "benchmark" assessment is conducted. Where possible, apply the models to data (probably through 2006), to obtain current and historical estimates of F and B and estimates of uncertainty.

6. Evaluate the sufficiency of the assessment models to estimate measures of stock status consistent with Biological Reference Points.

ANNEX 2: Contents of GARM-III CIE independent peer review report

1. The Independent CIE Report should comment on the quality and soundness of the science, methods and data with regard to each Term of Reference. CIE panelists should consider whether the work provides a scientifically credible basis for developing fishery management advice. Scientific criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable.

If a panelist feels that his/her comments are adequately expressed in the GARM Panel Summary Report, it will not be necessary to repeat the same comments in the Independent CIE Report. In that case, the Independent CIE Report can be used to provide greater detail on specific Terms of Reference or additional questions raised during the meeting.

2. If any modeling approaches are considered inappropriate, the Independent CIE Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing modeling approaches are the best available at this time.

3. Any independent analyses conducted by the CIE panelists as part of their responsibilities under this agreement should be incorporated into their Independent CIE Reports. It would also be helpful if the details of those analyses (e.g., computer programs, spreadsheets etc.) were made available to the respective assessment scientists.

4. Additional questions that were not in the Terms of Reference but that are directly related to the assessments should be addressed. This section should only be included if additional questions were raised during the GARM meeting.

ANNEX 3: Contents of GARM-III Panel Summary Report

1. The first section the report shall consist of an introduction prepared by the GARM chair that will include the background, a review of activities and comments on the appropriateness of the process in reaching the goals of the GARM. The next section will contain comments on the quality and soundness of the science, methods and data with regard to each Term of Reference. The GARM Panel should consider whether the work provides a scientifically credible basis for developing fishery management advice. Scientific criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable.

If the CIE panelists, the non-CIE panelists and GARM chair do not reach an agreement on a Term of Reference, the report should explain why. It is permissible to express majority as well as minority opinions.

2. If any modeling approaches are considered inappropriate, the GARM Panel Summary Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing modeling approaches are the best available at this time.

3. The report shall also include the bibliography of all materials provided during the meeting and any papers cited in the GARM Panel Summary Report, along with a copy of the CIE Statement of Work.

The report shall also include as a separate appendix the Terms of Reference used for the GARM Models Meeting, including any changes to the Terms of Reference or specific topics/issues directly related to the assessments and requiring Panel advice.