External Peer Review of

GARM-III "Biological Reference Point" Meeting

April 28 – May 2, 2008, NEFSC, Woods Hole, Massachusetts

Report to

Center for Independent Experts

Dr Michael C. Bell
Velzian, Rendall
Orkney,KW17 2EZ, UK
bandm.bell@virgin.net

16 May 2008

Contents

Executive Summary

- This report evaluates the science underpinning the definition of biological reference points (BRPs) for 19 fish stocks managed by the New England Fishery Management Council. This was undertaken during the third meeting of GARM-III.
- The main purpose of the meeting was to provide updates, where appropriate, of the current BRPs for each stock in the light of new data and analyses. The meeting also examined related issues, including initial conditions for forecasting and rebuilding strategies and putting BRPs for individual species in the context of ecosystem productivity.
- The views of the six member external review panel are included in a consensus summary report, which should be read in conjunction with this individual CIE reviewer report.
- The meeting met its Terms of Reference in all important respects. Updated BRP values were accepted for all 19 stocks, although the values for American plaice are provisional upon updating the assessment with new data.
- A method was proposed for specifying initial conditions for stock projections in the presence of patterns of retrospective bias in terminal estimates from age-based assessments. The best approach appears to be age-specific adjustments to population numbers at age. The use of this approach is supported, but further simulations and management strategy evaluations will be needed to determine the best techniques for adjustment before the method can be applied in a formal context.
- Trends in length- and weight-at-age and in age-at-maturity are identified in a number of stocks. These have been taken forward into assessments and BRP calculations as appropriate. Further research is needed into techniques for forward projection of these trends, particularly in the context of environmental effects.
- The meeting considered three different approaches to addressing the question of whether the Gulf of Maine / Georges Bank can support the BRPs summed across the different stocks. Cross-system comparisons suggested that the B_{MSY} values are reasonable in terms of overall biomass density; aggregate surplus production modeling suggested that aggregate B_{MSY} is lower than the sum of individual species B_{MSY} values, and energy budget contextualization was inconclusive. Great progress has been made with this approach, but given the uncertainty inherent in the analyses it is unclear whether or not the differences in aggregate and summed B_{MSY} values are significant.
- Before the ecosystem approach can deliver workable benchmarks for management, two further elements are needed: (i) guidance on management objectives at an ecosystem level, determined by policy makers and managers; and (ii) knowledge of what is operationally achievable given the composition and nature of the fleets pursuing the fisheries.
- BRPs were derived from age-based assessments for 14 stocks out of 19 stocks. Stock-recruitment relationships were unconvincing so that non-parametric BRPs were preferred. $F_{40\%MSP}$ ($F_{50\%MSP}$ for redfish) was used as the proxy for F_{MSY}, and B_{MSY} calculations were based on samples of recruitment considered representative of stock productivity at undepleted SSB levels.

- B_{MSY} proxy values were generated as the median biomass resulting from stochastic projections at the F_{MSY} proxy. This was to address a perceived inconsistency with the requirement for B_{MSY} to be achieved with at least 50% probability in rebuilding scenarios. B_{MSY} values generated in this way were accepted by the meeting, but with some reservations expressed. Mean stochastic values may be preferred to medians, although the two are likely to be close.
- Dome-shaped selectivity curves for Gulf of Maine cod were discussed. Declining selectivity at older ages suggests hidden biomass of older fish, and has a strong effect on the perception of stock status. Discussions were inconclusive about the likelihood of such patterns, but some lines for further research are suggested. Following the recommendations of the previous GARM-III meeting, the default is to accept assessment results and BRPs based on a flat-topped selectivity curve unless there is compelling biological evidence of a decline in selection over older ages.
- Appropriate values for natural mortality deserve greater attention for the GARM stocks. Analyses under the ecosystem approach may shed light on likely changes in M. This information could also be used to inform approaches to deal with retrospective patterns in assessments.
- With respect to the index-based methodology, some suggestions are made for alternative formulations for the relationship between Relative F and Replacement Ratio.
- The Term of Reference concerned with models for forecasting and evaluating rebuilding scenarios was not addressed during the meeting, but the AGEPRO software is well tried and tested and contains all the appropriate options. It is suggested that there be some investigation into the performance of two-stage stock-recruitment models in the projections.

Background

The Groundfish Assessment Review Meeting (GARM) process is a system of peer review for stock assessments and the development of fisheries management benchmarks for 19 important fish stocks managed by the New England Fishery Management Council (NEFMC). The current round of GARM-III meetings, taking place over October 2007 to August 2008, comprises four workshops: (i) a *Data Methods* meeting, reviewing the fishery-dependent, survey and biological data to be used in the assessments; (ii) a *Modeling* meeting, determining the most appropriate methods of assessment for each stock; (iii) a *Biological Reference Point* meeting, updating and refining the management benchmarks for each stock; and (iv) the *Final* meeting, determining the historical and current status of each stock.

This report considers the third of these meetings, the *Biological Reference Point* workshop, the main purpose of which was to provide updates, where appropriate, of the current values of biological reference points (BRPs) for each stock in the light of new data and analyses. The meeting also examined various issues relating to the BRPs, including the definition of initial conditions for forecasting and rebuilding strategies, and putting BRPs for individual species in the context of biological productivity at the ecosystem level. Work presented at the meeting took account of results and recommendations of the previous *Data Methods* and *Modeling* meetings. Full Terms of Reference (ToR) for the *Biological Reference Point* meeting are set out at Annex 1 of the Statement of Work (Appendix III).

The *Biological Reference Point* meeting was attended by the Northeast Fisheries Science Center (NEFSC) scientists responsible for undertaking the assessments of the 19 fish stocks, representatives of the NEFMC, fishing industry representatives and other interested parties. There was also an external peer review panel consisting of a non-CIE chairman invited by NEFSC (Robert O'Boyle), three CIE reviewers (Vivian Haist, Stuart Reeves and myself) and two non-CIE reviewers invited by NEFSC (Stratis Gavaris and Grant Thompson). The review panel were collectively responsible for producing a consensus report, summarizing the views of the panel on each ToR and setting out the final values of BRPs accepted by the meeting. The present report is an individual CIE reviewer report representing my own views on the analyses presented during the meeting. The emphasis is on issues relating to generation of BRPs rather than assessments and data preparation, since these have already been comprehensively reviewed after the previous two GARM-III meetings. The report does not attempt to duplicate comments and recommendations contained in the Consensus Reports, but provides additional emphasis and elaboration where necessary.

Description of Review Activities

Documents from the GARM-III *Data Methods* and *Modeling* meetings and other background materials were made available on the NEFSC website around a month before the meeting. Working papers to be reviewed at the *Biological Reference Point* meeting (see Appendix I) were made incrementally available during the two weeks prior to the meeting, and during the meeting itself. This allowed sufficient time for reviewers to become familiar with the overall context of the GARM process and with the material to be covered at the meeting. Terms of Reference and a draft agenda were also available before the meeting.

The GARM-III *Biological Reference Point* meeting was held at the Northeast Fisheries Science Center (NEFSC) at Woods Hole, Massachusetts, starting at 9.00 am on Monday 28 April and finishing at 5.00 pm on Friday 2 May 2008. The final agenda for the meeting is attached at Appendix II. The first day of the meeting was taken up with presentations on ToRs 2 (trends in stock productivity) and 3 (ecosystem approaches) and with the more general aspects of ToR 4 (biological reference points), including an overview of the current BRPs. Tuesday started with talks on Working Papers relating to ToR 1 (retrospective patterns and initial conditions for forecasting) before the meeting embarked on a set of presentations on the development of BRPs for each of the 19 GARM stocks. This part of the meeting lasted until the middle of Thursday, and included alternative perspectives on assessments for Gulf of Maine cod and white hake offered by Doug Butterworth on behalf of the fishing industry. Each presentation was followed by extensive discussion amongst the review panel and the wider audience. Following the stock-by-stock presentations the panel either selected a set of BRPs for acceptance by the meeting, or recommended further analyses from which more acceptable values could be derived. These further analyses were reviewed on Thursday afternoon, at which point the final values of BRPs were adopted by the meeting. Friday was devoted to a closed review panel session during which a consensus was reached on the main points relating to the BRPs for each stock. Panel discussions were also held at intervals during the main meeting, during which consensus was reached on all ToRs.

A summary of the panel consensus was available at the end of the meeting, and responsibilities for drafting sections of the consensus report were assigned to individual panel members. During the two weeks following the meeting these contributions were collated and reviewed by all panel members. Consensus was reached on all the main issues discussed by the meeting. Additional caveats and individual viewpoints are included in the individual reports submitted by the CIE panelists.

The meeting agenda was described as ‘ambitious’, but the skilled chairmanship of Robert O’Boyle and the coordination of Jim Weinberg as SAW chairman ensured that due consideration was given to all the Working Papers and to all the topics covered. A huge amount of very high quality work was presented by scientists of the Populations Dynamics Branch at Woods Hole, under the leadership of Paul Rago. The meeting was conducted in a spirit of openness and cooperation by all participants.

Summary of Findings

General

The GARM-III *Biological Reference Point* meeting met its Terms of Reference in all important respects:

1. For relevant stocks, determine the influence of retrospective patterns in parameter estimates (e.g., fishing mortality, biomass, and/or recruitment) from assessment models on the computation of BRPs and on specification of initial conditions for forecasting. The second part of this ToR, relating to the influence of retrospective patterns on the specification of initial conditions for forecasting, was addressed in WP 1.1, but the first part of the ToR was not addressed.

2. Trends in Stock Productivity: (a) for relevant stocks, identify trends in biological parameters (i.e., life history and/or recruitment) and assess their importance for the computation of BRPs and for specification of rebuilding scenarios; (b) if possible, summarize trends in pertinent environmental variables that might be related to the trends in those biological parameters relevant to BRPs. Both parts of this ToR were fully addressed in WP 2.1-2.

3. Ecosystem approaches to Gulf of Maine/Georges Bank fisheries: (a) determine the production potential of the fishery based on food chain processes and estimate the aggregate yield from the ecosystem; (b) comment on aggregate single stock yield projections in relation to overall ecosystem production, identifying potential inconsistencies between the two approaches. This ToR was fully addressed, to the extent possible given available science, in WPs 3.1-6

4. Biological Reference Points (B_{target}, $B_{threshold}$, F_{target}, $F_{threshold}$): (a) for each stock, list what the current BRPs and/or BRP Proxies are (e.g., B_{MSY}, B_{MAX}, F_{MSY}, $F_{40\%MSP}$, historical survey catch per tow, etc.), and give their values (i.e., typically from GARM II); (b) for each stock, update or redefine BRPs or BRP proxies that will be used for stock status determination, and compute their expected values and precision. Note: These BRPs and their proxies must be comparable and consistent with outputs from the recommended assessment models from the GARM III "Modeling" Meeting. Existing BRPs and proposals for new BRPs were given for each stock in WPs 4.A-S. The meeting adopted new BRPs for each stock based on presentations of these WPs, with supplementary analyses where appropriate. The only stock for which a full update was not possible was Gulf of Maine / Georges Bank American plaice, for which the assessment data were not available in time to perform an assessment before the meeting. Provisional new BRPs were put forward for this stock based on the previous assessment, and the analysis will be fully updated in August.

5. For each stock, identify appropriate models for forecasting and for evaluating rebuilding scenarios. No presentations were made that explicitly addressed this ToR, but the well-tried and tested AGEPRO software was used for all projections shown at the meeting.

One very positive aspect of the GARM process is that it is based on a very well developed system of BRPs and approaches to estimating them. The methods are set out clearly in NEFSC (2002), and there is an extensively tested and validated suite of software (NOAA Fisheries Toolbox) for implementing the assessments and further analyses. Methods, software and their application have been comprehensively reviewed in the past, both internally and externally.

A second positive aspect is the existence of extensive and high quality data sets, fit for the purpose of the GARM-III analyses. Spatially comprehensive standardized surveys dating back to 1963, a long history of commercial data gathering, collection of information on discards and recreational removals, measurements of age- and length-compositions, all lead to a high level of confidence that removals and trends in population structure and abundance have been well characterized and that assessments and estimation of BRPs have a strong scientific basis.

Thirdly, the GARM process itself is very comprehensive. It involves a huge amount of data collation and analysis, but the series of four meetings sequentially taking on board the recommendations and views of external reviewers ensures that the work proceeds in a logical and productive way and that the end results will be scientifically excellent and, importantly, defensible.

ToR 1 – Influence of retrospective patterns

In the context of retrospective bias in terminal estimates from age-based assessments, five types of assessment outcome can be identified:

(1) no systematic bias in terminal estimates, giving confidence in the interpretation and projection of stock status indicators;

(2) systematic biases from identifiable sources such as changes in survey catchability, natural mortality or partial recruitment patterns, and which can be rectified by appropriate modification of the assessment model;

(3) systematic biases from sources which remain unidentified, but which can be reduced by judicious model adjustments;

(4) systematic biases from unidentified sources, which remain resistant to model adjustments and which must be taken into account before using the parameter estimates in further quantitative analyses such as stock forecasts;

(5) systematic biases of such pattern and magnitude that alternative assessment approaches must be sought, e.g. index-based methods.

The ideal assessment is clearly of type 1, but in practice most assessments fall somewhere in between the extremes of types 1 and 5. As shown in analyses presented at the GARM-III *Modeling* meeting (NEFSC, 2008a), it is not possible to identify sources of retrospective patterns from model diagnostics alone, so that if causes are not known or suspected *a priori* the final assessment will be of type 3 or 4. In these circumstances the model is considered the best basis for inference, but it is acknowledged that the results cannot be interpreted uncritically. Several of the GARM-III assessments fall into type 3. The VPA for Georges Bank yellowtail flounder is a good example, where splitting the survey series in the mid-1990s results in much reduced retrospective patterns compared with an unsplit VPA (WP 4.C). A difficulty with type 3 adjustments is that the basis for removing retrospective patterns is not understood. The assessment is improved in the sense that it appears better behaved. The problem is that the converged VPA estimates are treated as an absolute criterion of reality, when in fact there is no guarantee that this is so.

The same problem with 'reality' arises with type 4 assessments, where the retrospective pattern remains but the estimates are adjusted for the presumed bias before further quantitative use. WP 1.1 describes some of the ways in which either population numbers or fishing mortality can be adjusted to account for retrospective bias before being used in rebuilding projections. The question is, can we use information on the likely direction and magnitude of biases to improve our picture of the current stock position for the purposes of conducting projections? The results presented in WP 1.1 suggest that the best way to do this is to make age-specific adjustments to population numbers at age.

The review panel strongly supported the use of this approach, but agreed that the method was not yet ready to be included in the GARM-III assessment process. The panel's preferred rationale and defaults for adopting this approach are clearly set out in the consensus report (NEFSC, 2008b). Without seeking to duplicate what is already set out in the consensus report, I would offer the following comments and recommendations:

- The first priority should always be to identify sources of retrospective bias and to modify the assessment model to take account of the processes that have generated the data observations. NEFSC (2008a) indicates that this will not be possible on the basis of model diagnostics alone. However, independent analyses of spatio-temporal patterns in survey and catch data, of data on the abundance and distribution of potential predators, supplementary information on fishing gears and practices, and anecdotal information from fishermen may all be at least indicative of changes that may have occurred in availability to survey and commercial gear, natural mortality regimes and partial recruitment patterns.
- Obviously, even if it is not possible to assign causes to retrospective patterns, pragmatic model adjustments that reduce their severity are preferable to posterior adjustment of values for the purposes of stock forecasts. In some cases this might simply involve a shorter time-series of data for which assessment assumptions are more likely to be met.
- Projections and other analyses involving posterior adjustments for presumed bias are inherently more uncertain than analyses where no adjustment was thought necessary. When such adjustments are deemed necessary, it will be advisable to update the assessments, benchmark statistics and projections more frequently than for other stocks.
- Management strategy evaluation and extensive simulations based on diverse examples will be needed to determine the best techniques for adjustments. These might include adjustments applied to historical data for stocks for which the later stock trends are now known.
- The adjustments shown in WP 1.1 involved a single step back in time. It would be worth investigating the extent to which it is possible to develop models of the way bias changes as more years are added to the assessment. Is it possible to adjust towards the asymptotically 'correct' values rather than to a set of values which are only somewhat less biased.
- Some investigation is also needed into the implications of retrospective bias for the estimation of BRPs, i.e. the targets towards which the adjusted or unadjusted projections are aiming. For some of the GARM stocks the BRPs appear to be relatively robust to the presence of retrospective patterns, with similar values emerging from assessments with and without modifications to reduce the bias (splitting the survey series). However, this is not necessarily always true. It is easy to envisage, for example, that a mis-specified partial recruitment

vector could have a large effect on values emerging from yield and spawning biomass per recruit analyses.

ToR 2 – Trends in stock productivity

Eighteen out of 20 stocks examined showed trends in length- and weight-at-age over recent years, mostly declines (12 stocks), while 13 out of 21 stocks showed trends in female age at 50% maturity, mostly increases (11 stocks), i.e. delayed maturation (WP 2.1). Other points to emerge from the analyses were:

- length and weight tended to show similar trends, indicating that trends are due to growth patterns rather than changes in condition;
- some stocks show opposing trends of weight and abundance, which may be evidence of density-dependence of growth;
- in most cases (other than haddock) the trends in biological parameters showed year effects rather than cohort effects, suggesting that environmental factors may be responsible;
- a number of environmental variables were compiled, but no attempt was made to correlate these variables with the biological parameters for the GARM stocks.

Clearly, these trends have serious consequences for the estimation of BRPs and for stock forecasting (WP 2.2). It is clear that a great deal of care has been taken to ensure that these trends were taken into account in the assessments and in the estimation of BRPs for stocks considered by GARM-III. The panelists were supportive of the use of five-year averages of partial recruitment, maturity and weight at age in calculating the BRPs when trends are present. This approach is adequate to represent current stock productivity. However, it is important to recognize that, when strong trends are present, significant changes in stock productivity may occur within the medium- to long-term time horizons of stock projections involved in rebuilding and other management scenarios. Some research is needed into techniques for forward projection of trends in biological parameters over longer than short-term time-scales.

Ideally, such forward projection would take account of prevailing environmental conditions (and density-dependence). Obviously, future environmental conditions are unknown, but investigation of the extent to which biological parameters of fish stocks co-vary with environmental factors may shed light on the likely transience or otherwise of current trends. Uncertainty about environmental trends could be incorporated into an overall framework of risk assessment for evaluation of management strategies.

It may be particularly important to take account of the effects of environmental variation on stock-recruitment relationships. As noted in the consensus report, the default assumption should be that low recruitment is a result of low stock size. However, especially for stocks at the distributional limits of a species' range (e.g. Southern New England – Mid-Atlantic yellowtail flounder) which may be vulnerable to the effects of environmental change, current conditions may preclude rebuilding to levels implied by historic levels of productivity. Inclusion of environmental variables as covariates in stock-recruitment relationships, or as the main

explanatory factors for productivity regime, may aid the future development of BRPs which are meaningful under current circumstances.

ToR 3 – Ecosystem approaches to Gulf of Maine / Georges Bank fisheries

Five Working Papers were presented, addressing the key question of whether the NE shelf can support the BRPs summed across the individual stocks. Several different approaches were taken: cross-system comparisons in terms of energy budget densities at B_{MSY} for all species (WP 3.1); contextualization of B_{MSY} values in terms of energy budgets for the broader ecosystem (WP 3.2); aggregate surplus production modeling (WP 3.3 and WP 3.4); and estimates of system-wide fishery production potential (WP 3.5). Whilst the cross-system comparisons indicate that it is "not unreasonable" for the shelf area to support the biomass density of GARM species (together with elasmobranchs and pelagic species) at their B_{MSY} levels, the energy budget contextualization is inconclusive, and the surplus production modeling and estimates of fishery production potential suggest that the single species B_{MSY} levels are beyond the capacity of the ecosystem to support.

The first comment that should be made is that the authors of the five Working Papers have made impressive progress in addressing ToR 3. Clearly, the modeling results are not yet at the point when they can be used to inform management decisions, but the fact that some of these calculations can be performed at all is remarkable. The review panelists were very supportive of this research program, and recommended that further research be encouraged, particularly to validate the finding (from WPs 3.4-5) that the ecosystem may not be able to support all species simultaneously at their B_{MSY} levels.

The second point is that, given the diversity of modeling approaches and the assumptions, simplifications and uncertainties inevitably associated with their application, it is heartening that the results were even in the same ball park, and that it is the same ball park as the sum of single species B_{MSY} values. Without a rigorous evaluation of the overall statistical uncertainty involved in the model outcomes it is difficult to judge whether, for example, an aggregate B_{MSY} that is only one half the level of the sum of single species B_{MSY} values really represents a significant difference.

What is not clear is, if we had a meaningful value of aggregate B_{MSY} for an ecosystem, what would we do with it? The idea appears to be that it would serve as a check on the sum of the individual species B_{MSY} values. If the aggregate B_{MSY} is significantly below the species sum, this seems to suggest that the single species values are too high. But how should they be reduced? Proportional reductions, i.e. $B'_{MSY}(i) = B_{MSY}(i) \times B_{MSY}(\text{agg.}) \Big/ \sum_j B_{MSY}(j)$, would be an obvious approach, but there is no guarantee that a stable aggregate B_{MSY} would have the same species proportions as the sum of individual species values. Overfishing as a concept is reasonably straightforward to define for a stock of an individual species – it damages the future productive capacity of the stock, i.e. it leads to an overfished stated where the stock is not as productive as it has the potential to be given its biological properties and external circumstances. The same definition could be defined for an aggregate of multiple species, but this leads to problems in defining the units of productivity. Treating all units of biomass as equivalent leads to a kilogram

of redfish or halibut being considered as identical to a kilogram of ocean pout or pollock. Unfortunately, there is no obvious way to construct a measure of equivalent units, and the whole notion of ecological 'value' raises the issue of defining management objectives for the ecosystem approach and making value judgments that lie in the realm of policy makers and managers rather than scientists. The policy-level question is: what do we want out of an ecosystem? The scientific approach to dealing with disparities between benchmarks for individual species and benchmarks for the system as a whole depends on the answers to this question.

For now, it is probably enough to leave the individual species BRPs to work themselves out. The question of whether all species can be held simultaneously at or around their current B_{MSY} values is somewhat academic, since many species are well below this level. The word 'current' is important in this context. An aggregate B_{MSY} that is lower than the sum of current species B_{MSY} values implies that there will be limitations, in terms of species interactions and overall energy flow, on the productivity of individual species at higher biomass levels. In other words, current productivity, on which current BRPs are based, will not continue to apply as stocks rebuild. If BRPs are frequently updated to reflect changes in productivity, then in principle, as stocks rebuild, the sum of single species B_{MSY} values should eventually converge with the aggregate B_{MSY}. This, of course, places the onus on scientists to properly understand and characterize the nature of current productivity for individual species – not entirely straightforward, since perceptions of current productivity are usually based on historical observations.

In the context of historical observations it is relevant to ask whether the NE shelf ecosystem has in the past supported all the GARM species at (current perceptions of) B_{MSY} levels. If so, and we believe that this is not now possible, what has changed? Is the ecosystem less productive, or else has its structure changed in some fundamental way?

In addition to a sound quantitative understanding of the production of fish biomass in the context of the broader ecosystem, multi-species fishery management requires information on the targeting behavior of the fishing fleets. Catches in demersal fisheries are rarely composed of single species. Even when fishing is primarily directed at a particular species, others are taken as by-catch or are discarded. Individual species are often taken by more than one fleet, operating different fishing methods. Thus management measures taken to modify the fishing mortality experienced by one species will usually have repercussions for a suite of other species taken alongside it. In terms of management, the most precise tool for manipulating fishing mortality is the 'métier' – a fishery segment characterized by gear and mode of fishing and by the assemblage of species targeted. Identification of métiers involves analyses of gear-specific catch composition at the most disaggregated level – preferably haul-by-haul, since fishermen may target more than one species in a single fishing trip. Spatial and temporal considerations are also important, since when and where fishing occurs is at least as important as the method used.

On the evidence of the GARM-III Working Papers, the science of ecosystem productivity is well advanced, and may soon be in a position to make more definitive statements about whether current single species targets are achievable in the context of energetic constraints for the system as a whole. However, before this can progress to producing workable benchmarks for ecosystem approaches to management two further elements are needed: (i) management objectives at an ecosystem level, determined by managers on the basis of value judgments; and (ii) a knowledge

of what is operationally achievable given the composition and nature of the fleets pursuing the fisheries.

ToR 4 – Biological Reference Points

Biological reference points have been developed for 19 stocks. Most of the stocks are obvious and important components of the overall Gulf of Maine / Georges Bank fishery, but inclusion of at least two species in GARM-III is surprising: there is virtually no market for ocean pout, most if not all the fishery removals are discards and these are relatively small in scale (WP 4.O); Atlantic halibut appears to be heavily depleted, but GARM-III considers only a portion of a much bigger stock (WP 4.S). Given that these stocks have been included, the assessments on which the proposed BRP values are based have been commendably rigorous and based on the best available information, as indeed is the case for all the GARM-III stocks.

It proved possible to use age-based assessments (VPA or ASAP) for 14 out of the 19 stocks. For two stocks this represented a considerable advance over the previous index-based (Gulf of Maine Haddock, WP 4.R) or surplus production (Georges Bank winter flounder, WP 4.K) models used previously. Where appropriate, survey indices were used to hindcast recruitment estimates for years prior to the assessment period, providing a longer perspective on historical stock-productivity. Both parametric and non-parametric BRPs were estimated from the age-based assessments, and parametric BRPs were estimated both internally to the assessment model, and externally based on assessment model outputs. The review panel were unconvinced by fits of stock-recruitment curves (usually Beverton-Holt) to the data, and instead favored the non-parametric BRPs, based on $F_{40\%MSP}$ ($F_{50\%MSP}$ for redfish) (fishing mortality at 40% or 50% of the maximum spawning potential from a spawning biomass per recruit curve) and recruitment values considered representative of the productivity of undepleted stocks. In some cases (e.g. Georges Bank haddock, WP 4.B) recruitments from 'bonanza' year-classes were excluded, the philosophy being that BRPs should be based on typical productivity levels, with exceptional year-classes treated as a bonus for the fishery. During the meeting a variance minimizing criterion (the 'Rago Razor') was developed as an aid to determining an appropriate cut-off value for SSB in determining the stock sizes over which recruitment should be included in BRP calculations.

The final values of BRPs accepted by the meeting were agreed by all panel members and are set out in the consensus report (NEFSC, 2008b). In my opinion, the accepted values are the best possible given the available data and scientific knowledge of the GARM stocks. Comments, caveats and recommendations are listed for each stock in the consensus report. These are comprehensive and reflect the consensus of all panel members, including my own input, and I do not intend to duplicate that material in this, my individual CIE report. Instead, I will confine my comments to some of the generic issues raised in relation to estimation of BRPs.

GARM-III adopted an important change in the approach to estimating B_{MSY}. This arose from the observation that the deterministic value of B_{MSY} derived by multiplying the spawning biomass per recruit at F_{MSY}[1] by average recruitment is not the same as the median biomass after long-term fishing at F_{MSY} obtained in stochastic projections. This leads to a perceived inconsistency

[1] F_{MSY} is here taken to mean F_{MSY} or its proxy, e.g. $F_{40\%MSP}$.

between the biomass reference point and the statutory requirement that stock rebuilding should lead to a 50% chance or greater of SSB exceeding B_{MSY} within the prescribed time frame (WP 4.2). The review panel noted that guidelines require B_{MSY} to be equal to the mean rather than the median stock size after fishing at F_{MSY} and that there is a straightforward conversion between the deterministic B_{MSY} and the stochastic mean, without the need for projection. However, in most cases the median and mean values are expected to be very close to one another, and therefore the panel chose to adopt the estimates of B_{MSY} provided by the median of long-term projected biomass values.

Despite cautiously accepting the B_{MSY} values derived from the projections, the panel also suggested that NEFSC consider estimating B_{MSY} by the mean rather than the median. I concur with this view and also offer the following points for consideration:

- It does not seem necessary that there should be a relationship defined between $F_{rebuild}$, the fishing mortality required for rebuilding the stock with 50% probability by the prescribed deadline, and F_{MSY}, the threshold reference point for fishing mortality. Projections at F_{MSY} are modeling a scenario that will never occur, i.e. the overfishing definition does not represent a harvest strategy in relation to rebuilding.
- In a hypothetical case where the updated assessment provides an identical perspective on stock productivity to the previous assessment, the BRPs would be expected to remain the same. If, in this case, the biomass reference point is revised because of the new method of estimation (all other things remaining equal), then it must be possible to demonstrate that the revised value is closer to what is intended in the Guidelines than is the old value.

A topic that received a great deal of attention during the meeting was the possibility of dome-shaped fishery selection curves, particularly for the Gulf of Maine cod stock. Dome-shaped selection means that, rather than rising to a plateau at full recruitment to the fishing gear, there is a curve of declining partial recruitment for the older age-classes. This is of more than mere academic interest, since if there is sufficient survivorship to the older age-classes, this means that there is an 'invisible' stock of large fish, not available to the fishery or to the survey gear. An alternative perspective on the Gulf of Maine cod stock was provided by Doug Butterworth and Rebecca Rademeyer on behalf of the fishing industry. Using a statistical catch-at-age model in which selectivity was allowed to vary freely for the older age-classes, they showed that, on statistical grounds, the data are most consistent with a dome-shaped curve (WP 4.F.1). They presented a similar paper for white hake (WP 4.L.1). Compared with the NEFSC VPA assessment (WP 4.F), this provides a radically different perception of cod stock status in relation to B_{MSY} – the alternative analysis suggests that biomass is at least 80% of its B_{MSY} level, compared with around a third of B_{MSY} indicated by the NEFSC assessment with a flat-topped curve.

The position adopted at the GARM-III *Modeling* meeting was that the burden of proof lies with providing compelling evidence of the existence of large fish unavailable to the fishing gear; otherwise the default is to accept the more conservative NEFSC assessment. At present there is no such evidence, and the *Biological Reference Point* review panel took the same view as the previous meeting. Analyses of tagging data showed no evidence of reduced capture probability at older ages (WP 4.7). Possibilities to account for the disparity of results between analyses include:

- The dome could be an artifact of the assessment method. WP 4.F.1 reported only a slight bias towards dome-shaped selectivity in the statistical catch-at-age model. Analyses presented during the meeting indicated that the curve became less dome-shaped if a higher value of M was used. There are some indications of a high M value from the tagging analyses, although in this case M is likely to be confounded with other loss processes.
- Natural mortality could be higher in the older age-classes, so that the assumption of constant M causes lower abundance to be viewed as lower selection in the assessment model. Although senescence is a possibility, it seems unlikely to account for the dome, since the downward curve sets in around age 6, well below the expected life-span.
- Older fish may move to habitats or areas not accessible or simply not visited by the surveys or the majority of the fishing fleet. This possibility cannot be discounted, and it would be interesting to seek the views of fishermen on this.
- Capture probability could be heterogeneous in the population, causing differential survival. This would mean that, above the age of full selection, older age-classes might appear progressively less catchable. This pattern would not be apparent in the tagging data, since the tagged fish would be likely to be a biased sample with respect to catchability. This is perhaps an unlikely possibility, but it would be useful to explore the implications in a simulation context.
- In the tagged samples, reporting probability could be higher in the older fish, confounding the effects of reduced selection in the estimates of capture probability. Again, this is perhaps unlikely, but it would be worth exploring the possibility of differential reporting rates between age-classes in the tagging model.

These possibilities are not exhaustive, but suggest some directions for future research. The meeting was unable to reach a conclusion on the issue, and it is one that is likely to arise again at the GARM-III *Final* meeting in August.

Questions about natural mortality were raised for two other species than Gulf of Maine cod. These were redfish, for which an M of 0.05 is assumed (WP 4.N) and Atlantic halibut for which an M of 0.06 was assumed (WP 4.S). It was questioned whether these values are too low, particularly in the latter case when stock assessments for comparable halibut stocks use larger values. More generally, the topic of natural mortality deserves attention for all the GARM stocks. As is almost universal in fish stock assessments, sweeping assumptions are made about the values of M and its invariance with time and age. All the attention to refining and updating the analyses has focused on other (more tractable) biological parameters and assessment options, yet assumptions about M potentially have strong implications for BRPs and the outcome of assessments. It is not clear how many M values have actually been estimated, or derived from life-history invariants, assumed life-spans or relationships with environmental factors. What ever the answer to this question, it is highly recommended that this topic be revisited in the near future. WP 3.4 highlighted the importance of species interactions for the GARM species, a more important factor than fishery removals. Natural mortality, particularly by predation, would be an important component of these interactions. It is suggested that the implications of these multi-species models be explored for the individual species stock assessments. Whilst it may not be possible to provide defensible time-varying parameters for the assessments, it may be possible to

provide a range of values for sensitivity analyses and which could be carried forward into management strategy evaluation in a risk assessment context.

Changes in natural mortality are particularly relevant for retrospective patterns in the assessments. As noted above under ToR 1, assessments are modified to remove retrospective patterns, e.g. by splitting the survey series, under the assumption that the converged, historical part of the assessment represents an absolute criterion of reality. It would be worth considering the implications of species interactions for levels, or at least stanzas of natural mortality in the assessment models – based on the species composition of GARM and other fish stocks on the NE shelf, is it possible to draw any conclusions about natural mortality regimes for individual species?

Finally under this ToR, the AIM analyses deserve some comment. The derivation of this index-based method in NEFSC (2002) is rigorous, resulting in expressions for assessment indices that are reassuringly intuitive. AIM was used to generate BRPs for four of the GARM species. Selection of the B_{MSY} proxy is based on best scientific judgment – inevitably there is a subjective element in the visual interpretation of index time-series, but in practice the selected values are robustly defensible in terms of their consistency with the observed patterns in the data. Estimation of the F_{MSY} proxy is more objective in the sense that it is based on a statistical procedure – the equilibrium point from the relationship of Relative F and Replacement Ratio. However, this relationship is rarely well-defined, and it would be worthwhile exploring alternative methodologies for estimating Relative F at replacement. This might include different functional relationships (e.g. logistic or log-linear rather than log-log) and type II regressions, such as reduced major axis, that assume error in the *x* as well as the *y* variable.

ToR 5 – Models for forecasting and evaluating rebuilding scenarios

This ToR was not formally addressed during the meeting, although the user manual for the AGEPRO projection software was made available as WP 5.1. AGEPRO was universally used as the stochastic projection tool for deriving the B_{MSY} proxy from age-based assessments (see ToR 4 and WP 4.2) and in the analyses examining how to take account of retrospective patterns in specifying initial conditions for forecasting (see ToR 1 and WP 1.1). The software package appears to be user-friendly, well developed and tested and to have all the options necessary for forecasting and evaluating rebuilding scenarios.

There was no discussion of how stock-recruitment considerations would be taken forward into rebuilding scenarios, although it was clear that these would have to be consistent with the approaches used to estimate BRPs. This raises the issue of how to model recruitment at low stock sizes, i.e. below the SSB cut-off values applied in compiling the sample of recruitment values for use in estimating B_{MSY}. Presumably a two-stage model would be used, with different recruitment samples used for SSBs above and below the cut-off. If so, some investigation is warranted into how the recruitment discontinuity at the SSB cut-off affects the projections.

Recommendations

- The proposed method of dealing with retrospective patterns in specifying initial conditions for rebuilding projections is to make age-specific adjustments to population numbers. It is recommended that this method be applied after further investigation of the best techniques for adjustment. This would involve extensive simulations and management evaluations. It would be worth investigating models of the way bias changes as more years are added to an assessment, with a view to adjusting towards asymptotically 'correct' values.
- The first priority should be to identify sources of retrospective bias and to modify the assessment model to take account of the processes that have generated the data observations. Failing this, pragmatic model adjustments that reduce retrospective patterns are also preferable to posterior adjustment of model outcomes.
- Frequent updating of rebuilding scenarios is necessary if they are based on posterior adjustments to assessment model outcomes.
- Further investigation is needed into the implications of retrospective bias for estimation of BRPs.
- Research is needed into techniques for forward projection of trends in biological parameters over longer than short-term time-scales.
- It is recommended that there be investigation of the extent to which biological parameters of fish stocks co-vary with environmental factors. Where possible, forward projection of biological parameters should take account of prevailing environmental conditions.
- Environmental variation may be particularly important for stock-recruitment relationships. Inclusion of environmental variables as covariates in stock-recruitment relationships, or as the main explanatory factors for productivity regime, may aid the future development of BRPs which are meaningful under prevailing conditions.
- It will be important to refine the current estimates of aggregate B_{MSY} for the Gulf of Maine / Georges Bank ecosystem. This should include a rigorous evaluation of statistical uncertainty.
- Further consideration is needed into the appropriate response to disparity between the aggregate B_{MSY} and the sum of single species B_{MSY} values. This is strongly linked to overfishing definitions at the ecosystem level. Judgments of ecological 'value' may be needed from policy makers and managers. The default response should perhaps be to allow single species B_{MSY} values to converge with the aggregate, by frequent updating to reflect changes in stock productivity as the system capacity is approached.
- Some investigation is needed into whether the Gulf of Maine / Georges Bank ecosystem has in the past supported all the GARM species at B_{MSY} levels.
- It is recommended that there be research into technical interactions between fisheries, specifically the identification of fishery métiers, which form the most precise tool for manipulating fishing mortality. This would involve analyses of gear-specific catch composition at the most disaggregated level, preferably on a haul-by-haul basis.
- Further consideration is needed into the use of stochastic rather than deterministic B_{MSY} values. It is recommended that the mean be considered as an alternative to the median.

- More research is needed into the possibility that gear selectivity can decline with age over the older age-classes in cod. This might involve seeking the views of fishermen on the possibility of movements of large fish into areas or habitats where they are less accessible to survey and commercial gear. Further tagging analyses may also be warranted, examining the effects of age-specific reporting rates and the possibility of heterogeneous capture probabilities.
- The topic of natural mortality deserves attention for all GARM stocks. This includes values for M and age- and time-specific patterns. The implications of species interactions in multi-species models should be explored for individual species stock assessments.
- The implications of multi-species models should also be explored in relation to retrospective patterns in the single species assessments. On the basis of species interactions it may be possible to infer likely changes in M, even if absolute levels cannot be estimated.
- For the AIM assessments, it is recommended that there be some exploration of alternative formulations for the relationship between Relative F and Replacement Ratio. This might include consideration of type II regression models, such as reduced major axis regression.
- For rebuilding scenarios, some investigation is warranted into how recruitment discontinuity at an SSB cut-off level would affect projections using a two-stage stock-recruitment model.

References

NEFSC, 2002. Re-Evaluation of Biological Reference Points for New England Groundfish. NEFSC Reference Document 02-04.

NEFSC, 2007. Report of the Groundfish Assessment Review Meeting (GARM III). Part 1. Data Methods. NEFSC Report.

NEFSC, 2008a. Report of the Groundfish Assessment Review Meeting (GARM III). Part 2. Models. NEFSC Report.

NEFSC, 2008b. Report of the Groundfish Assessment Review Meeting (GARM III). Part 3. Biological Reference Points. NEFSC Report.

APPENDIX I – Bibliography of materials provided during the workshop

WP 1.1 Legault C, Terceiro M. Specifying Initial Conditions for Forecasting When Retrospective Pattern is Present.

WP 1.2 Legault C, Seaver A, Brooks L. A Simulation Study to Evaluate Estimation of Biological Reference Points from VPA and ASAP.

WP 2.1 O'Brien L. Trends in Average Length, Weight and Maturity at Age for Relevant Stocks.

WP 2.2 Rago et al. Implications of Biological Trends for Estimation of Biological Reference Points and Rebuilding Schedules.

WP 3.1 Overholtz W, Link J, Fogarty M, Col L, Legault C. US Northeast Shelf LME Biomass, Target Biological Reference Points for Fish and Worldwide Cross-System Comparisons.

WP 3.2 Link J, Overholtz W, Legault C, Col L, Fogarty M. Energy Budget Contextualization of Fish Biomasses at B_MSY

WP 3.3 Overholtz W, Fogarty M, Link J, Legault, Col L. Estimates of Aggregate Surplus Production for the GARM and Other Stock Groups for the US Northeast Shelf LME.

WP 3.4 Link J, Gamble R, Overholtz W, Legault C, Col L, Fogarty M. An Aggregate and MS Production Model: A Simulator Tool

WP 3.5 . Fogarty M, Overholtz WJ, Link J. Fishery Production Potential of the Northeast Continental Shelf of the United States.

WP 3.6 Link et al. Synthesis of Ecosystem Considerations.

WP 4.1 Rago et al. Overview of Current BRPs Methods and Estimates.

WP 4.2 Legault C. Setting SSBmsy via Stochastic Simulation Ensures Consistency with Rebuilding Projections.

WP 4.3 Palmer M, Legault C. Sensitivity of the Long-term Observation-error Survey Series (LOSS) Model to Variable Stock-Recruit Steepness and Stock Depletion Inputs: A Test Case using Gulf of Maine haddock

WP 4.4. Palmer M. (Supplementary Paper): A Method to Apportion Landings with Unknown Area, Month and Unspecified Market Categories Among Landings with Similar Region and Fleet Characteristics

WP 4.5. Palmer M, Wigley S, O'Brien L, Mayo R, Rago P. (Supplementary Paper): A Description of Discard Estimation Methods Where Observer Coverage is Unavailable

WP 4.6 Legault C, Palmer M, Wigley S (Supplementary Paper): Uncertainty in Landings Allocation Algorithm at Stock Level is Insignificant.

WP 4.7 Miller T, Hart D. (Supplementary Paper): Analysis of Tagging Data for Evidence of Decreased Fishing Mortality for Large Gulf of Maine Cod.

WP 4.8	Butterworth D, Rademeyer R. (Supplementary Paper): Implications of Tagging Analyses for the Shape of Selectivity at Age of GoM cod.
WP 4.8a	Butterworth D. (Supplementary Paper). Further Runs of ASPM/SCAA for GoM cod
WP 4.A.	Georges Bank Cod . O'Brien L.
WP 4.B.	Georges Bank Haddock. Brooks L.
WP 4.C	Georges Bank yellowtail flounder. Legault C
WP 4.D	Southern New England-Mid Atlantic yellowtail flounder. Legault C, Cadrin S.
WP 4.E	Cape Cod-Gulf of Maine yellowtail flounder. Legault C, Cadrin S, King J, Sherman S.
WP 4.F.	Gulf of Maine Cod. Mayo R
WP 4.F.1	Gulf of Maine Cod, Butterworth D
WP 4.F.1a	Gulf of Maine Cod Addendum, Butterworth D, Rademeyer R
WP 4.G.	Witch Flounder. Wigley S
WP 4.H.	Gulf of Maine/Georges Bank American Plaice. O'Brien L
WP 4.I.	Gulf of Maine Winter Flounder. Nitschke P
WP 4.J.	Southern New England Winter flounder. Terceiro M
WP 4.K.	Georges Bank Winter Flounder. Hendrickson L
WP 4.L.	White Hake. Sosebee K
WP 4.L.1	White Hake, Butterworth D
WP 4.M.	Georges Bank/Gulf of Maine Pollock. Mayo R
WP 4.N.	Gulf of Maine/ Georges Bank Acadian Redfish. Miller T
WP 4.O.	Ocean Pout . Wigley S
WP 4.P.	Gulf of Maine/Georges Bank Windowpane Flounder. Hendrickson L
WP 4.Q.	Southern New England – Mid-Atlantic Windowpane Flounder . Hendrickson L
WP 4.R.	Gulf of Maine Haddock. Palmer M
WP 4.S.	Atlantic Halibut. Col L
WP 5.1.	Rago P, Brodziak R. (Supplementary Paper): Overview of age-based projection model (AgePro) for reference point estimation and scenario analyses.

APPENDIX II

Meeting Agenda (last revised April. 27, 2008, noon)

GARM III Biological Reference Points Meeting: April 28-May 2, 2008

Date /Day	*Start*	*End*	*Duration (min)*	*Topic*	*Presenter*	*Rapporteur*
28-Apr	9:00	9:10	10	Introduction	Weinberg	
1	9:10	9:30	20	Overview of GARM/ meeting objectives	GARM Chair	
				TOR #4 Biological Reference Points: a.Current values and proxies		
1	9:30	9:45	15	WP 4.1 Overview of current BRPs methods and estimates	Rago	Brooks
1	9:45	10:00	15	Discussion		
1	10:00	10:30	30	WP 4.2 Setting SSBmsy via Stochastic Simulation Ensures Consistency with Rebuilding Projections. Chris Legault	Legault	Brooks
1	10:30	10:45	15	Break		
1	10:45	11:00	15	Discussion		
				TOR #2: Trends in Stock Productivity		
1	11:00	11:45	45	WP 2.1 Trends in Average length, weight and maturity at age for relevant stocks and trends in environmental variables.	*O'Brien*	Blaylock
1	11:45	12:00	15	Discussion		
1	12:00	12:15	15	WP 2.2 Implications of biological trends for estimation of biological reference points and rebuilding schedules.	*Rago et al*	Blaylock
1	12:15	12:30	15	Discussion		
1	12:30	13:30	60	Lunch		

Date /Day	*Start*	*End*	*Duration (min)*	*Topic*	*Presenter*	
				TOR #3 Ecosystem Approaches to Gulf of Maine/Georges Bank Fisheries		
1	13:30	13:50	20	WP 3.1 US Northeast Shelf LME Biomass, target biological reference points for fish and worldwide cross-system comparisons. Overholtz, Link, Fogarty, Col, Legault.	Overholtz	Chute
1	13:50	14:00	10	Discussion		
1	14:00	14:20	20	WP 3.2 Energy Budget contextualization of fish biomasses at B_MSY	Link	Chute
1	14:20	14:30	10	Discussion		
1	14:30	14:50	20	WP 3.3 Estimates of aggregate surplus production for the GARM and other stock groups for the US Northeast Shelf LME. Overholtz, Fogarty, Link, Legault, Col.	Overholtz	Chute
1	14:50	15:00	10	Discussion		
1	15:00	15:15	15	Break		
1	15:15	15:35	20	WP 3.4 An Aggregate and MS Production Model: A Simulator Tool	Link	Jacobson
1	15:35	15:45	10	Discussion		
1	15:45	16:10	25	WP 3.5 Fishery Production Potential	Fogarty	Jacobson
1	16:10	17:00	50	Discussion—WP 3.6 Synthesis: Implications for single species reference points	Link/Fogarty	Jacobson
				TOR #4 Biological Reference Points:		
1	17:00	17:15	15	WP 4.3. Sensitivity of the Long-term Observation-error Survey Series (LOSS) model to variable stock-recruit steepness and stock depletion inputs: A test case using Gulf of Maine haddock (Palmer and Legault).	Palmer/Legault	Shepherd
1	17:15	17:25	10	Discussion		
1	17:25	17:40	15	WP 4.7 (Supplementary WP) Size-specific tag recovery rates of cod and implications for estimation of fishing mortality in analytical models. Miller and Hart	Miller/Hart	Shepherd
1	17:40	17:50	10	Discussion		
1	17:50	18:00	10	Summary/Followup (Chair)		

Date /Day	Start	End	Duration (min)	Topic	Presenter	Rapporteur
29-Apr	9:00	9:15	15	Progress review and Order of the Day (Chair)	Chair	
				TOR #1 Influence of retrospective patterns on parameter estimates and specification of initial conditions for forecasting.		
2	9:15	9:35	20	**WP 1.1** Specifying Initial Conditions for Forecasting When Retrospective Pattern is Present.	Legault/ Terceiro	Miller
2	9:35	9:50	15	Discussion		
2	9:50	10:10	20	**WP 1.2** A simulation study to evaluate estimation of biological reference points from VPA and ASAP.	Brooks/ Legault/ Seaver	Miller
2	10:10	10:25	15	Discussion		
2	10:25	10:40	15	Break		
				TOR #4 Biological Reference Points: b. Update by stock		
2	10:40	11:25	45	**WP 4.A** Georges Bank Cod	O'Brien	Wigley
2	11:25	11:55	30	Discussion		
2	11:55	12:55	60	Lunch		
2	12:55	13:40	45	**WP 4.F** Gulf of Maine Cod	Mayo	Wigley
2	13:40	14:05	25	Discussion		
2	14:05	14:30	25	**WP 4.F.1** Gulf of Maine Cod	Butterworth	Wigley
	14:30	14:40	10	Discussion		
2	14:40	15:30	50	**WP4.B** Georges Bank Haddock	Brooks	Mayo
2	15:30	15:55	25	Discussion		
2	15:55	16:10	15	Break		
2	16:10	17:05	55	**WPs 4.C** Georges Bank + **4.D** Southern New England + **4.E** Cape Cod-Gulf of Maine Yellowtail Flounder	Legault	Hendrickson
2	17:05	17:50	45	Discussion		
2	17:50	18:00	10	Summary/Followup	Chair	
Date /Day	Start	End	Duration (min)	Topic	Presenter	Rapporteur
30-Apr	9:00	9:15	15	Progress review and Order of the Day (Chair)	Chair	
3	9:15	10:00	45	**WP 4.N** Gulf of Maine/ Georges Bank Acadian Redfish	Miller	Brooks
3	10:00	10:15	15	Discussion		
3	10:15	11:00	45	**WP 4.K** Georges Bank Winter Flounder	Hendrickson	Sosebee
3	11:00	11:15	15	Break		
3	11:15	11:30	15	Discussion		
3	11:30	12:30	60	**WP 4.I** Gulf of Maine Winter Flounder	Nitschke	Sosebee
3	12:30	12:45	15	Discussion		
3	12:45	13:45	60	Lunch		
3	13:45	14:30	45	**WP 4.J** Southern New England Winter flounder	Terceiro	Alade
3	14:30	14:45	15	Discussion		
3	14:45	15:30	45	**WP 4.G** Witch Flounder	Wigley	Col
3	15:30	15:45	15	Discussion		
3	15:45	16:00	15	Break		
3	16:00	16:45	45	**WP 4.H** Gulf of Maine/Georges Bank American Plaice	O'Brien	Richards
3	16:45	17:00	15	Discussion		
3	17:00	17:30	30	**WP 4.M** Georges Bank/Gulf of Maine Pollock	Mayo	Richards
3	17:30	17:45	15	Discussion		
3	17:45	18:00	15	Summary/Followup	Chair	
	19:30	22:30		Social/Dinner --British Beer Company, Falmouth Heights		

Date /Day	*Start*	*End*	*Duration (min)*	*Topic*	*Presenter*	*Rapporteur*
1-May	*9:00*	9:15	15	Progress review and Order of the Day	Chair	
4	9:15	10:05	50	**WP 4.L** White Hake	Sosebee	Palmer
4	10:05	10:20	15	Discussion		
4	10:20	10:35	15	Break		
	10:35	10:55	20	**WP.4.L.1** White Hake alt	Butterworth	Palmer
	10:55	11:05	10	Discussion		
4	11:05	12:00	55	**WP 4.R** Gulf of Maine Haddock	Palmer	Mayo
4	12:00	12:15	15	Discussion		
4	12:15	13:15	60	Lunch		
4	13:15	13:35	20	**WP 4.O** Ocean Pout	Wigley	Col
4	13:35	13:45	10	Discussion		
4	13:45	14:05	20	**WP 4.P** Gulf of Maine/Georges Bank Windowpane Flounder	Hendrickson	Chute
4	14:05	14:15	10	Discussion		
4	14:15	14:35	20	**WP 4.Q** Southern New England – Mid-Atlantic Windowpane	Hendrickson	Chute
4	14:35	14:45	10	Discussion		
4	14:45	15:05	20	**WP 4.S** Atlantic Halibut	Col	Alade
4	15:05	15:15	10	Discussion		
4	15:15	15:30	15	Break		
4	15:30	17:50	140	Review/Revisions/Follow-up	TBD	
4	17:50	18:00	10	Summary/Followup (Chair)	Chair	
2-May	*9:00*	9:30	30	Progress review and Order of the Day	Chair	
5	9:30	10:30	60	Review of Outstanding Issues as necessary	TBD	
5	10:30	10:45	15	Break		
5	10:45	12:00	75	Report Development [CLOSED]		
5	12:00	13:00	60	Lunch		
5	13:00	16:00	180	Report Development, Summary and Assignments [CLOSED]		
5	16:00	16:00	0	Adjourn		

APPENDIX III Statement of Work for Dr. Michael Bell

External Independent Peer Review by the Center for Independent Experts

GARM-III "Biological Reference Point" Meeting:

Meeting Date: April 28 – May 2, 2008

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 Methods" 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. The runs of the models will be used to evaluate diagnostics of model fit and appropriateness, including retrospective analyses.

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 historical and current 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 "Biological Reference Point (BRP)" Meeting, which will take place at the Woods Hole Laboratory of the Northeast Fisheries Science Center (NEFSC) in Woods Hole, Massachusetts, from April 28 – May 2, 2008. The meeting will have a chairman (non-CIE) as well as external panelists, three of whom will be provided by 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 shall 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 GARM panel review meeting, editorial assistance to the GARM 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. Reviewers should have experience in development of biological reference points that includes knowledge for the varying quality and quantity of data available to support estimation for individual fish species living within the ecosystem. Expertise should include statistical catch-at-age, traditional VPA approaches, and index-based methods. Desirable background includes life-history theory, risk analyses, stock-forecasting methodology, and ecosystem fisheries ecology. 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 6. 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 representation of the GARM panel viewpoint on the quality and soundness of the science, methods and results 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 results 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 only minor alterations in procedures previously established at the peer review of the "Data Methods" Meeting in October 2007 and the "Modeling" Meeting in February 2008. 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 or Biological Reference Points (BRP), the panel should explain why they 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 and/or BRPs are the best available at this time.

Roles and responsibilities

(1) Prior to the meeting

(GARM Chair and CIE panelists)
Background reports will be provided to the CIE reviewers in advance of the GARM review meeting.

(2) During the Open meeting

(GARM Chair)
Act as chairperson, where duties include control of the meeting, coordination, 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)
Participate in panel discussions on the quality and soundness of the science, methods and results 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 Chair, CIE and non-CIE panelists)
The GARM Chair will lead preparing, editing, and completing the GARM Panel Summary Report, based on contributions from the panelists (CIE and non-CIE). This report (see Annex 3 for information on contents) is to comment on the quality and soundness of the science, methods, and results with regard to each Term of Reference. If any modeling approaches and/or BRPs 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 and/or BRPs are the best available at this time.

The panelists and the chair will discuss whether their views on each Term of Reference can be summarized into a consensus conclusion. In cases where multiple, differing views exist on a given Term of Reference, the GARM Panel Summary Report will note that

there was no consensus and will summarize the various opinions and the reason(s) for these.

(GARM Chair)

The Chair's role during GARM Panel Summary Report development will be to facilitate rather than to force consensus from the panel.

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

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

(GARM CIE panelists)
Each CIE 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 results with regard to each Term of Reference.

If any modeling approaches and/or BRPs 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 and/or BRPs are the best available at this time.

During the meeting, questions which are not in the Terms of Reference but are directly related to the meeting may have been raised. Questions not explicitly referenced in the TOR but relevant to its intent can be documented and addressed.

Schedule of Milestones and Deliverables

The milestones and schedule are summarized in the table below. No later than May 16, 2008, the CIE panelists should submit their CIE independent peer review reports to the CIE for review[2]. 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**
CIE reviewers attend GARM workshop to conduct peer review at Northeast Fisheries Science Center (NEFSC) in Woods Hole, MA, USA	April 28 – May 2
GARM Chair and CIE panelists work at the NEFSC drafting reports. Report writing starts during the meeting. Panelists leave meeting with at least the summary bullets.	May 1 - 2
Draft of GARM Panel Summary Report, reviewed by all panelists, due to the GARM Chair **	May 16
CIE panelists submit CIE independent peer review reports to CIE for approval	May 16
GARM Chair sends Final GARM Panel Summary Report, approved by CIE panelists, to NEFSC contact (i.e., SAW Chairman)	May 23
CIE provides reviewed CIE independent peer review reports to NMFS COTR for approval	May 30
COTR notifies CIE of approval of CIE independent peer review reports	June 6 *
COTR provides final CIE independent peer review reports to NEFSC contact	June 6

* 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 chairman 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 dissemination 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

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

COTRs at the Office of Science and Technology have the responsibility for the distribution of the final CIE reports to the Project Contacts.

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

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 "Biological Reference Point (BRP)" Meeting

(Last Revised: 1/11/08; A final draft will be distributed to the Panel prior to the meeting.)

1. For relevant stocks, determine the influence of retrospective patterns in parameter estimates (e.g., fishing mortality, biomass, and/or recruitment) from assessment models on the computation of BRPs and on specification of initial conditions for forecasting.

2. Trends in Stock Productivity:

   a.) For relevant stocks, identify trends in biological parameters (i.e., life history and/or recruitment) and assess their importance for the computation of BRPs and for specification of rebuilding scenarios;

   b.) If possible, summarize trends in pertinent environmental variables that might be related to the trends in those biological parameters relevant to BRPs.

3. Ecosystem approaches to Gulf of Maine/Georges Bank fisheries:

   a.) Determine the production potential of the fishery based on food chain processes and estimate the aggregate yield from the ecosystem;

   b.) Comment on aggregate single stock yield projections in relation to overall ecosystem production, identifying potential inconsistencies between the two approaches.

4. Biological Reference Points (B_{target}, $B_{threshold}$, F_{target}, $F_{threshold}$):

   a.) For each stock, list what the <u>current</u> BRPs and/or BRP Proxies are (e.g., B_{MSY}, B_{MAX}, F_{MSY}, $F_{40\%MSP}$, historical survey catch per tow, etc.), and give their values (i.e., typically from GARM II);

   b.) For each stock, update or redefine BRPs or BRP proxies that will be used for stock status determination, and compute their expected values and precision. Note: These BRPs and their proxies must be comparable and consistent with outputs from the recommended assessment models from the GARM III "Modeling" Meeting.

5. For each stock, identify appropriate models for forecasting and for evaluating rebuilding scenarios.

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 results 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.

2. If any modeling approaches and/or BRPs 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 and/or BRPs 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 meeting can be addressed. This section need only be included if additional questions were raised during the GARM meeting.

5. The report shall include a copy of the Statement of Work with Terms of Reference and meeting agenda attached as appendices.

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 results 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 and/or BRPs 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 and/or BRPs are the best available at this time.

3. The report shall also include: a.) the bibliography of all materials provided during the meeting and any papers cited in the GARM Panel Summary Report; and separate appendices with b.) a copy of the CIE Statement of Work; c.) the assessment with the Terms of Reference used for the GARM BRP Meeting, including any changes to the Terms of Reference or specific topics/issues directly related to the assessments and requiring Panel advice; d.) a list of participants; e.) the meeting agenda, f.) a list of working papers; and g.) Presentation Highlights and Meeting Discussion Summary for each working paper. The Highlights and Discussion Summary are to be written by the assessment scientists and rapporteurs, respectively, with editing and oversight by the GARM Chairman.