Pacific Whiting
The Joint U.S.-Canada STAR Panel Report

Pacific Fishery Management Council
Seattle, Washington
February 3-6, 2009

Review Panel Members:

David Sampson (Chair), Oregon State University and SSC representative
Norman Hall, Center for Independent Experts
Jon Helge-Volstad, Center for Independent Experts
Tom Carruthers, University of British Columbia

STAT Team Members Present:

Ian Stewart, Northwest Fisheries Science Center
Owen Hamel, Northwest Fisheries Science Center

Advisors:

John Wallace, GMT representative
Dan Waldeck, GAP representative
Jeff Fargo, Department of Fisheries and Oceans, British Columbia
Chris Grandin, Department of Fisheries and Oceans, British Columbia
Greg Workman, Department of Fisheries and Oceans, British Columbia
John DeVore, PFMC representative

Overview

During 3-6 February 2009, a joint Canada-U.S. Pacific hake / whiting Stock Assessment Review (STAR) Panel met in Seattle, Washington, to review a draft stock assessment document that had been prepared by Hamel & Stewart (2009). The Panel operated under the U.S. Pacific Fishery Management Council's Terms of Reference for the Groundfish Stock Assessment and Review Process for 2009-2010 (PFMC 2008). As in previous years, the Panel attempted to adhere to the spirit of the Canada-U.S. Treaty on Pacific hake / whiting, with the Panel including a member from Canada and Canadian assessment scientists providing data to the stock assessment team. The revised stock assessment and the STAR Panel Report will be forwarded to the Pacific Fishery Management Council (PFMC) and its advisory groups, and to the Canadian Department of Fisheries and Oceans (DFO) managers and the Groundfish Sub-committee of PSARC (Pacific Scientific Advice Review Committee).

Both members of the stock assessment team (STAT, Drs. Owen Hamel and Ian Stewart) attended and actively participated in the meeting. Public comment was entertained throughout the three-and-a-half-day meeting, which was held at the Hotel Deca in Seattle. The STAR Panel members were able to receive all draft assessments and supporting materials via an ftp site two weeks prior to the meeting, which was sufficient time to adequately prepare for the review of the assessment. Although data from the 2008 Canadian fishery were not incorporated into the model developed for the draft assessment document, the STAT had developed a new preliminary base model employing these data shortly before the start of the STAR Panel meeting (discussed further below).

The meeting convened at 09:00 on Tuesday February 3rd. Dr. Elizabeth Clarke (U.S. National Marine Fisheries Service, NMFS) welcomed the group and provided an overview of the U.S. process for ratifying the new treaty, currently stalled pending changes in the implementing language. Dr. Sampson then opened the meeting with a brief review of the agenda (Appendix A), explanation of the Terms of Reference, and discussion of the review and reporting process, followed by self-introductions by Panel members and others in attendance. Mr. John DeVore and Mr. Barry Ackerman reviewed management needs for the U.S. and Canadian fisheries, respectively. The Canadian Advisors stated that their science staff were working on updating last year's assessments and would appreciate the opportunity to provide input to the PFMC process, by either addressing the SSC regarding science and modeling issues or possibly addressing the Council regarding management of the fishery. Mr. John Devore indicated this would be possible. Dr. Stewart and Mr. Grandin then presented overviews of the 2008 whiting fisheries for the U.S. and Canada, respectively, Dr Stewart presented details of the input data used by the STAT in the 2009 stock assessment, Dr Chu presented the plans and described progress made in improving acoustic estimates of Pacific Hake, and Drs Hamel and Stewart presented details of the approaches used and results obtained when applying Stock Synthesis III (SS3) to the data for Pacific hake.

The STAT advised the STAR Panel that they had received and processed the 2008 Canadian length and age composition data after they had already completed the draft assessment document provided to the STAR Panel in advance of the meeting. Large catches by the Canadian fishery late in 2008 had delayed acquisition and processing of data and biological material. The STAT had developed a slightly revised assessment for the STAR Panel to review. The STAR Panel acknowledges the extraordinary achievements of Mr. Chris Grandin and the STAT in bringing these data into the assessment at the eleventh hour so that they could be considered by the review.

Based on discussion of the stock assessment documents and related presentations, the Panel requested 21 clarifications (described below), many involving additional model runs, to help identify the most appropriate base model, explore opportunities for improved model simplification, and evaluate the uncertainty of the stock assessment results and the assessment's sensitivity to assumptions. This iterative process of making additional runs and discussing the results continued until mid-morning of Friday 6 February. The Panel spent the remainder of that morning reviewing the outline structure of its report. The meeting was adjourned at noon. A draft Panel report was distributed by email to all Panel members for continued development and finalization.

After careful review of the model diagnostics and results, but with concern that the reliability of the model predictions is compromised by structural inadequacy and over-parameterization, the Panel recommended a particular configuration of the SS3 model as the final base model. The basic data sets used by this model consisted of the following: total catches from the US and Canadian fisheries for 1966-2008; length compositions from the US fishery (1975-2008) and the Canadian fishery (1988-2008); standard age-composition data (derived from age-length keys) from the US fishery (1973-1974) and the Canadian fishery (1977-1987); conditional age-at-length compositions from the US fishery (1975-2008) and the Canadian fishery (1988-2008); biomass indices, length-composition data, and conditional age-at-length composition data from the joint US-Canadian acoustic/midwater trawl surveys (1977, 1980, 1983, 1986, 1989, 1992, 1995, 1998, 2001, 2003, 2005, and 2007); ageing-error matrices based on 2,820 cross-read otoliths with adjustment for the "lumping effect" of strong cohorts; seasonal length-composition data from Santa Barbara for 1963-1970; plus biological data relating to growth, maturity at length, and natural mortality.

The final base model differed from the preliminary base model (presented by the STAT during the first day of the STAR) in that it assumed a slightly simpler blocking structure for the parameters determining the selectivity curves of the US and Canadian fisheries (four-year rather than two-year blocks after 2001). It also differed from the preliminary model in that it estimated the value of the acoustic survey catchability coefficient rather than keeping this parameter set to a fixed value. When the STAR Panel meeting finished, the STAT had not had sufficient time to conduct a Monte Carlo Markov Chain (MCMC) run to confirm convergence of the final base model and to develop a decision table for the assessment. This work was completed during the week following the STAR Panel and the decision table, based on preliminary converged MCMC results for the final base model, was distributed to the STAR Panelists by email.

Suggestions for future reviews of Pacific hake / whiting assessments.

When it is fully implemented, the Pacific Hake / Whiting Agreement between the U.S. and Canada will establish for this important transboundary stock a process for developing and reviewing stock assessments and providing management advice. The current STAR process for assessment review is not adequate to meet the requirements of the Agreement.

- Ratification and implementation of the U.S. / Canada treaty needs to occur quickly so that there is a more coherent process for addressing unresolved issues.
- Given the definite possibility that the assessment review next year (2010) may again operate under the STAR Terms of Reference, the PFMC's Scientific and Statistical Committee (SSC) should consider altering the STAR Terms of Reference to better accommodate alternative stock assessments developed by Canadian scientists. In particular, simpler assessment models with fewer parameters than the base model employed in 2009 should also be evaluated.
- The process for future assessments of Pacific hake should ensure that the STAT has adequate time to undertake the assessment. Late arrival of data and a compressed schedule to resolve the assessment result in a rushed assessment that can lead to incorrect results. A different assessment and review process is needed given the expectation that this situation will re-occur with late-season fishing in both countries. For example, a partial release of catch quota could be made to accommodate the early season, with a later release based on a new assessment that is completed in March or April.

Requests by the STAR Panel and Responses by the STAT

Request #1: Provide the inverse Hessian matrix for the STAR base run (with the 2008 Canadian data) to show parameter correlation which might reveal possible model over-parameterization.

Response: The STAT presented a graphical (bubble plot) representation of the cross-correlations for most of the estimated parameters (excluding the forecast recruitment deviations). Many of the recruitment deviations had high positive correlation, as did some blocks of selectivity parameters. A strong correlation between two parameters generally indicates that the model is over-parameterized and that comparable fits to the data could be obtained from a different model structure with fewer parameters.

Request #2: Provide a likelihood profile across survey q[1] for all the likelihood components to expose tensions among the data.

Response: The STAT produced Fig. 1 showing the changes in negative log-likelihood (NLL) for a series of fixed values of the survey q parameter. In the preliminary base model the survey q was fixed at 0.7 on the grounds that the available data did not provide sufficient information to reliably estimate this parameter. The likelihood profile shows abrupt changes in the NLL at survey q ≈ 0.85, with a large trade-off between the length-composition component versus the conditional-age-composition component. This pattern suggests that there are complexities in the NLL surface that bear further investigation. Complex trade-offs among data sources could be artifacts of the blocking structure or other assumptions of the model.

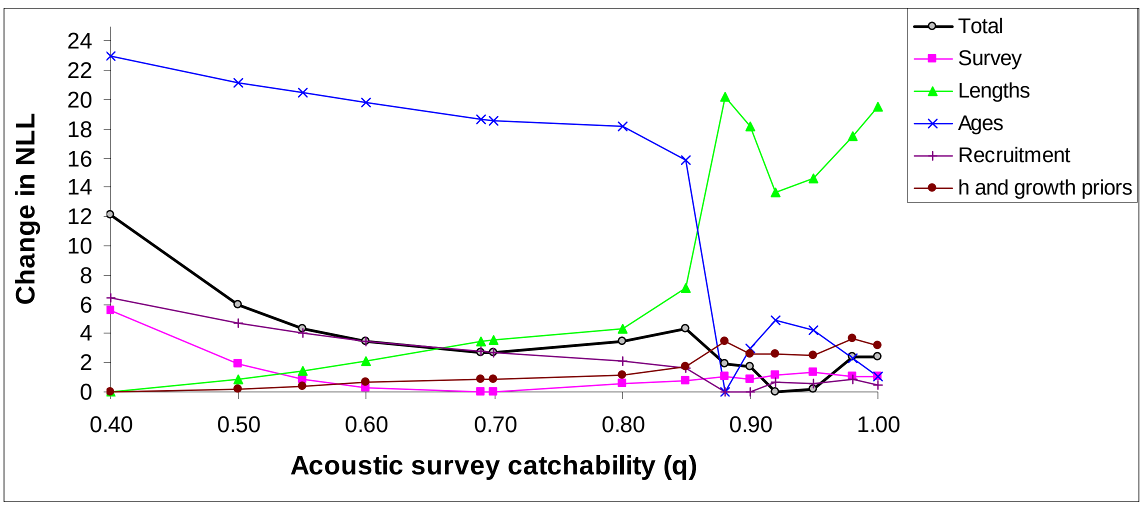

Figure 1. Likelihood contributions for the major data sources profiled across the survey q parameter, which was fixed at 0.7 in the preliminary base model.

Request #3: Provide the MCMC result (cross-correlation matrix for the old M[2] and the dome-shaped selectivity parameters) for the base model (pre-STAR base although not a fully

[1] The "survey q" is a calibration parameter that adjusts for the discrepancy between what the survey estimates for the stock biomass relative to what the model estimates that the survey should see, adjusted by the survey's selection curve and the timing of the survey.

[2] The "old M" parameter is the natural mortality coefficient for the 15+ age-group. The "base M" parameter is the natural mortality coefficient for age-groups 0 to 13, with the natural mortality coefficient for age-14 fish determined by linear interpolation between the other two values.

converged chain). This will provide more information about possible parameter confounding and the shape of the joint posterior function.

Response: Monte Carlo Markov Chain (MCMC) results from the partially-converged chain were plotted for the parameters for survey selectivity, old M, growth, stock-recruitment, and select selectivity parameters. The plots of the selectivity parameters showed much cross-correlation, but the STAT suggested that this was an expected result given that the selectivity blocking was accomplished in SS3 using offsets. There seemed little point to further pursuing this line of investigation with the pre-STAR base model.

Request #4: Provide likelihood contour plots that represent the posterior correlation around ML[3] estimates for the following three comparisons: a) survey q vs. old M, b) old M vs. old acoustic survey selection, and c) survey q vs. old selection. This will provide insight to the extent these parameters are confounded.

Response: The likelihood profiles were flat across a wide range of values for the three parameters (final survey selectivity, old M, and survey q) implying that the data provided little information to distinguish among different combinations of these parameters.

Request #5: Change the effective sample sizes with regard to Canadian and U.S. length data and the catch-at-age data. Reduce the effective sample sizes by half relative to the base model and evaluate three different selectivity scenarios ranging towards less complexity from the base model structure to a time-invariant structure. Compare the number of parameters and likelihoods for the model components for this weighting change and the base model run. The panel is seeking confirmation that the model result is not driven by weighting. In addition, the panel would like to see residual bubble plots and the estimated selectivities, the time series of female spawning biomass, and the fit to the survey.

Response: Relative to the initial STAR Base model, removing all the time blocks for the fishery selection curves coupled with a tenth of the weighting on the fishery length- and conditional age-composition data increased the overall NLL by 166 units and reduced the number of parameters by 27, with most of the change associated with the conditional age-composition data. An alternative four-year blocking structure with a tenth of the weighting on the fishery composition data did not result in significant loss of fit, with NLL increasing by about 11.4 units for a reduction of 6 parameters. The Panel concluded that the choice of blocking structure for the fishery selection curves was not unduly influenced by the weighting on the composition data and noted that the estimated acoustic survey selection curve tended to change in concert with changes in the estimated fishery selection curves. The configuration with a single set of fishery selection curves with a tenth of the weighting provided a very poor fit to the acoustic survey biomass data, especially in the earlier years of the survey. Having time-invariant selection for the fisheries clearly seemed the wrong approach.

---

[3] ML = maximum likelihood, which formally measures the best fit that the model provides to the observed data given the model form and parameter values. Note that NLL = negative log-likelihood, which is the negative value of the natural logarithm of the ML.

Request #6: The panel would like to assess the consequences of asymptotic selectivity in the survey. The panel recognizes that the selectivity pattern is uncertain; however, a fixed asymptotic acoustic survey selectivity may help to understand how this pattern affects the model results. Additionally this scenario may also reveal how acoustic selectivity is related to the time blocking assumption and the shape of the commercial selectivities estimated in the base model.

Response: Forcing asymptotic selectivity caused lack of convergence when the model was configured with recent time blocks (after 2001). However, one run using asymptotic survey selectivity with no time blocking after 2001 did converge; the results (Fig. 2) suggested that much of the large biomass estimated by the preliminary base model for the 1980s was "cryptic biomass" (from old ages that were not fully selected in the preliminary base model) and that recent time blocks are needed to account for the large catches from the 1999 year class.

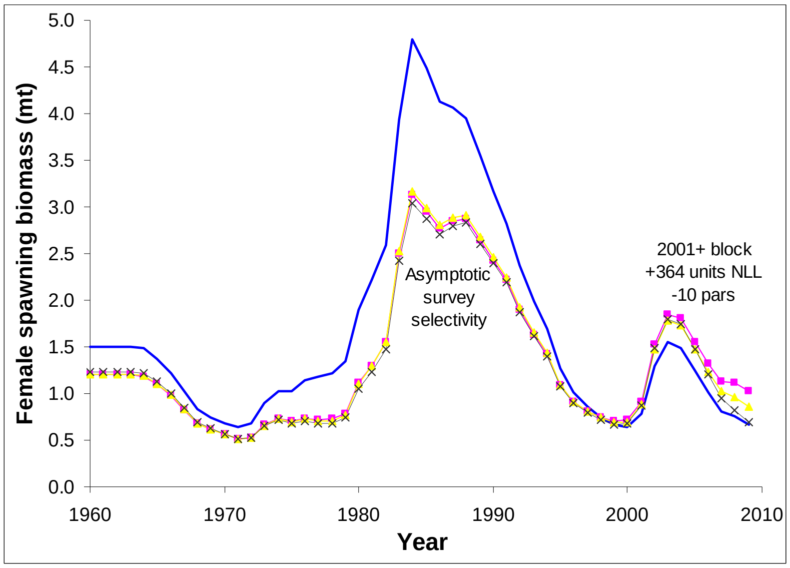

Figure 2. Comparison of spawning biomass trajectories from the preliminary model configuration with domed (solid line) versus asymptotic acoustic survey selection.

Request #7: The panel requested jitter runs to assess whether the model runs are converging appropriately using the fixed q parameter versus a freely estimated q model scenario.

Response: The STAT did not respond to this request because they were reluctant to spend time conducting jitter runs on a preliminary model. The request was subsumed by the STAR Panel's request (#13) to apply jitter runs to the final proposed base model.

Request #8: In reference to Request #2, identify components and parameters responsible for abrupt changes in likelihood with changes in the fixed q values. This will allow better understanding of the apparent trade-offs between different data sources and model configurations.

Response: As survey q was increased progressively from 0.7 to 1, estimates of female spawning biomass declined with no marked change in overall trend (Fig. 3). At the same time, the selectivity curve associated with the acoustic survey shifted towards older fish (Fig. 4). Thus the change in likelihood associated with change in survey q did not appear to reflect a marked change in the state of nature but instead appeared to be due to tension between the length and age composition data.

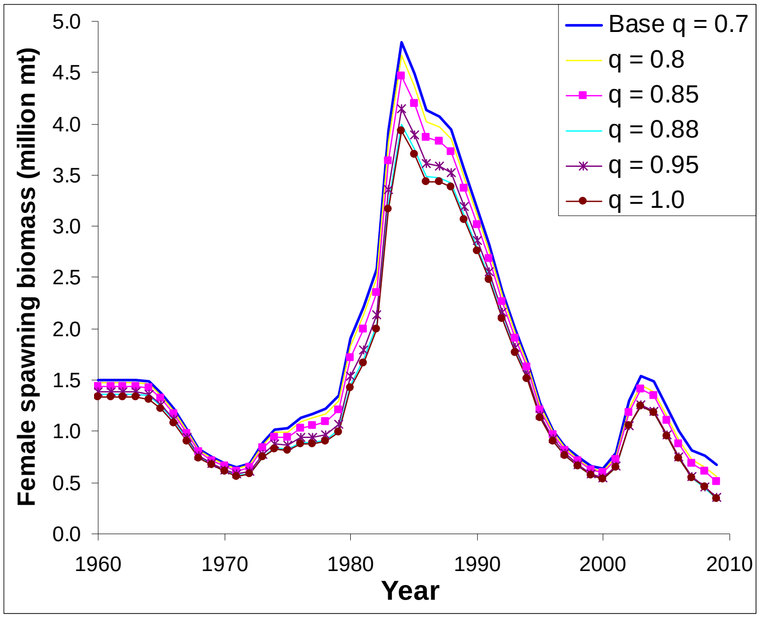

Figure 3. Estimates of female spawning biomass obtained when fitting the model using different fixed values of acoustic survey q.

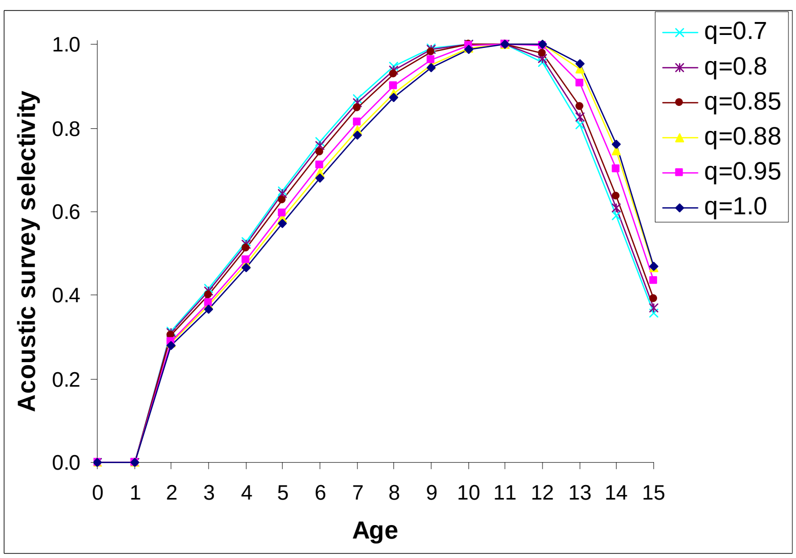

Figure 4. Estimates of the selectivity curve for the acoustic survey obtained when fitting the model using different fixed values of acoustic survey q.

Request #9: Start the 4-yr selectivity blocking structures at different years with a) no offset between U.S. and Canadian fisheries and b) a 2-yr offset between these fisheries. Collapse the blocks at the end of the time series using the base model. The current blocking selection is somewhat arbitrary and there is a need to understand how sensitive the model is to the blocking structure.

Response: The model was sensitive to the starting years used for the four-year selectivity blocking structure for the U.S. and Canadian fisheries (Fig. 5). The best fit was obtained using 1981 as the starting year for the four-year blocking structure, with no offset between the U.S. and Canadian fisheries, which was how the block boundaries were structured in the preliminary model. It appears that this structure best matched the supposed tracking by the fishery of the stronger year classes.

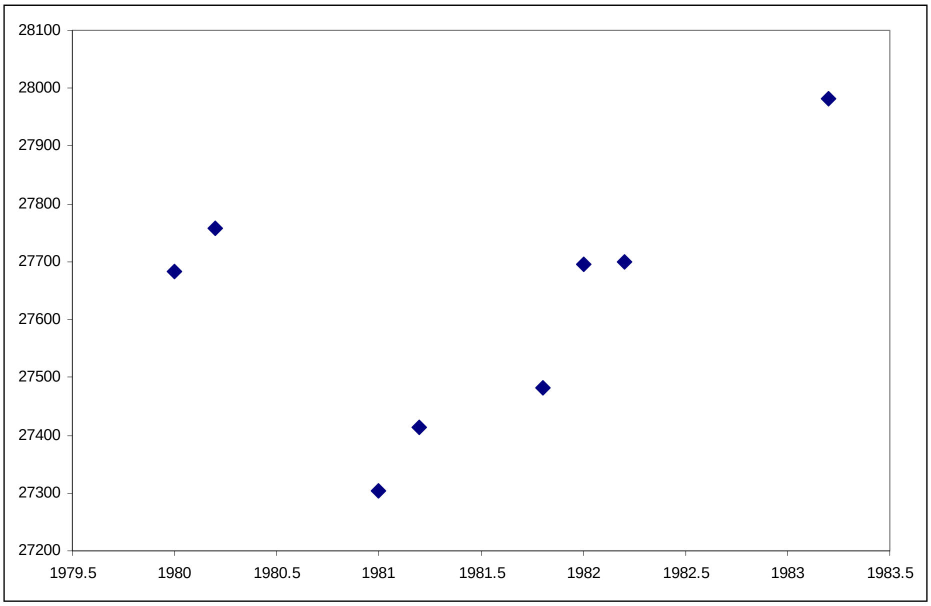

Figure 5. Comparison of the values of negative log-likelihood obtained using different start years for the blocking of the fishery selectivity curves. The first of each pair of values represents the result when the U.S. and Canadian time blocks are synchronised, while the second point reresents the results obtained when the Canadian time blocks lag the U.S. time blocks by two years. The value at 1981.8 represents the result obtained with an 8-year time block at the end.

Request #10: A retrospective analysis of the best model from Request #8 (where q is fixed and final survey selectivity is estimated) featuring runs up to 2000, 2002, 2004, 2006 and 2008 with projections based on observed catches. These runs should undertake projections using the most recent selectivities. This analysis will be used to assess potential bias of model results and projections.

Response: As years were successively stripped from the data used in the analysis, estimates of the strength of the 1999 year class in the female spawning biomass (Fig. 6) and in the estimates of recruitment (Fig. 7) increased until the data were removed for 2001, when the presence of the strong 1999 year class had not yet registered sufficiently in the age composition data from the

U.S. fishery. This suggested that the signal regarding the strength of this year class was driven primarily by the age compositions of the catches recorded by the fisheries. As the more recent years were removed, the selectivity blocking for the last years changed as the final years were collapsed into the last block. These blocking changes account for some of the differences between the curves.

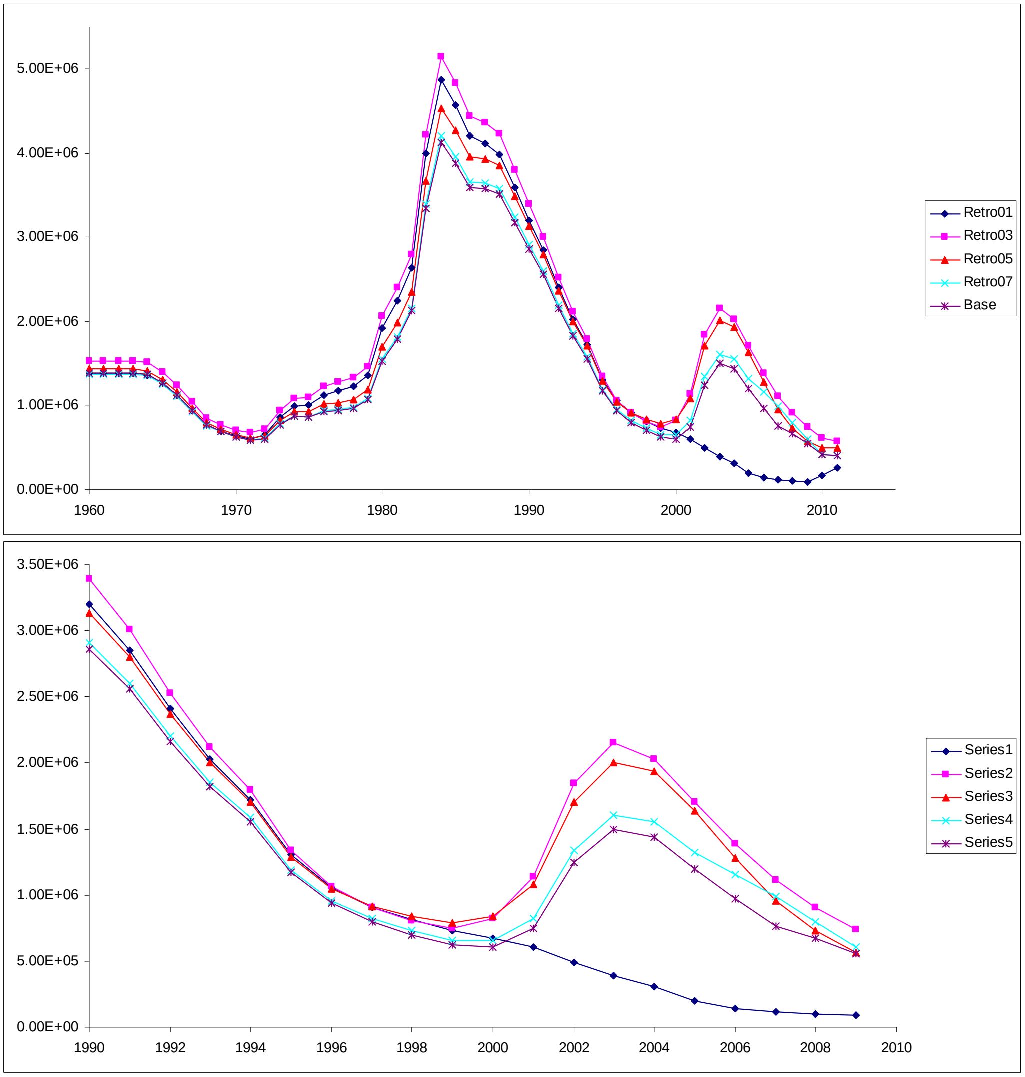

Figure 6. Trends in female spawning biomass obtained when fitting the model to data to 2001, 2003, 2005, 2007, and 2008, and forecasting using the observed values of recent catches.

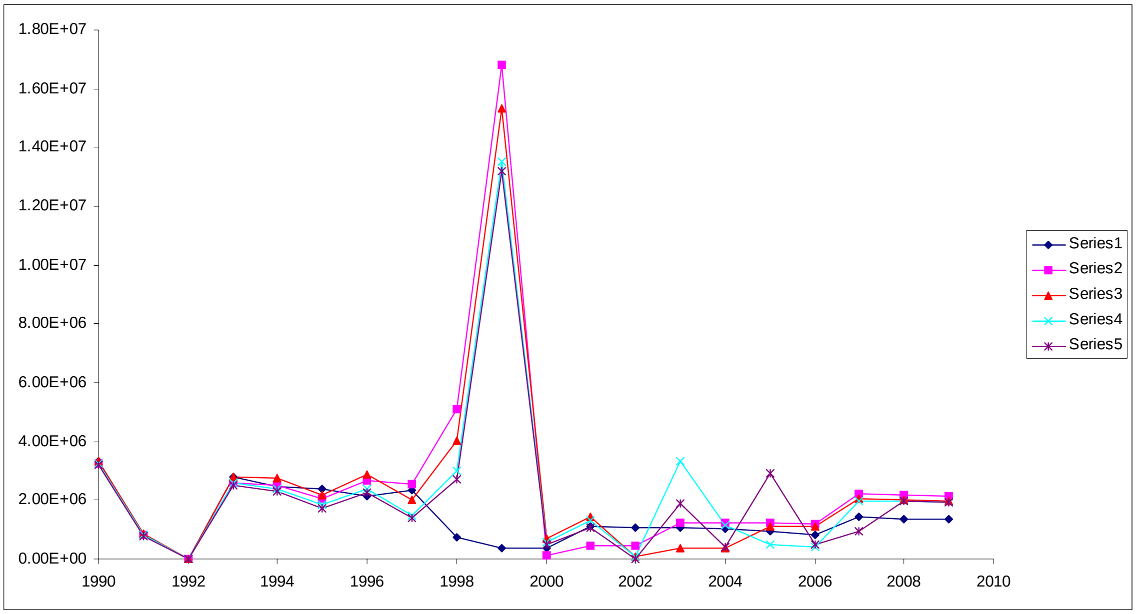

Figure 7. Trends in recruitment estimates from 1990 to 2008 obtained when fitting the model to data to 2001, 2003, 2005, 2007, and 2008, and forecasting using the observed values of recent catches. The different runs differed very little in their estimates of earlier recruitment.

Request #11: Provide runs where q is freely estimated and final survey selectivities are fixed at 0.9, 0.6, and 0.3 (assuming that q can be reliably estimated given fixed selectivity). These runs will ascertain whether the model can converge given those three fixed selectivities.

Response: Fixing the acoustic survey's selectivity coefficient for the biggest fish (final selectivity) at 0.3, 0.6, and 0.9 resulted in progressively lower estimates of female spawning biomass (Fig. 8). When the parameters for acoustic survey q and final selectivity were both freed up, the estimated q was 0.85. Convergence was obtained in all cases.

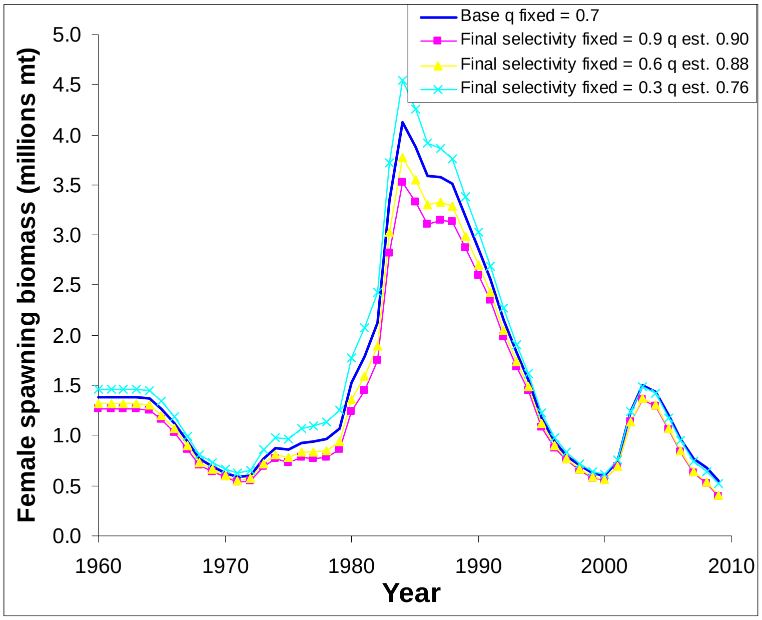

Figure 8. Estimates of female spawning biomass obtained when acoustic survey selectivity value for the biggest fish is fixed at 0.9, 0.6 and 0.3, and acoustic q is estimated.

Request #12: If there is convergence and plausible results for the Request #10 runs, provide a retrospective analysis as structured in Request #10 with the fixed selectivities under Request #11. This analysis will be used to assess potential bias of model results and projections.

Response: Truncating the series of yearly input data progressively (ending the series in years 2008, 2006, 2004, 2000) increased the estimated strength of the 1999 year class and the 2005 year class, as well as spawning biomass estimates (Fig. 9).

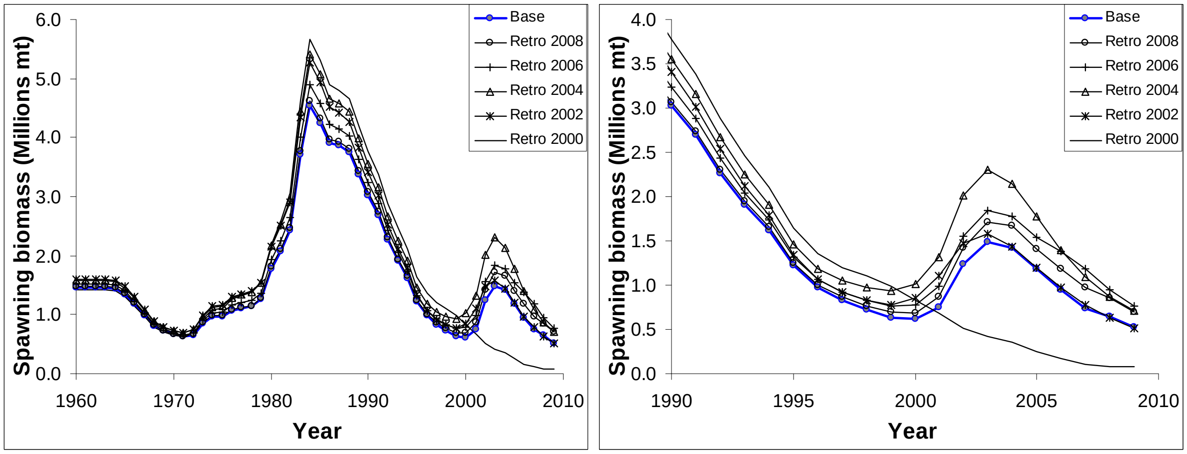

Figure 9. Changes in estimates of spawning biomass from retrospective analysis with final survey selection (on the oldest fish) fixed at 0.3.

At higher fixed values of final survey selection (0.6 and 0.9), spawning biomass estimates were progressively scaled down. In all cases, the biomass trend was maintained, except for the 2000 retrospective run, where the 1999 year class signal was lost. In additional to the requested runs, the STAT ran retrospective analyses where the parameters for acoustic survey q, the final survey selectivities, and old-M were freely estimated (a similar approach as last year's assessment). All retrospective runs showed a similar biomass trend across a wide range of acoustic survey q values.

The analyses based on data restricted to 1960-2000 demonstrate that forward projections are unreliable because the spawning biomass is highly dependent on single cohorts. Future strong cohorts cannot be reliable predicted.

The panel agreed that the best model configuration is one with a freely estimated value for acoustic survey q. The Panel recommended that analyses based on a set of fixed q's could be used to bracket uncertainty in a decision table, in the event that an MCMC run could not be completed.

Request #13: Jitter the new preferred base model with q freely estimated, final acoustic survey selectivity estimated, and a simplified blocking structure. This preferred model differs from the pre-STAR base model by allowing q to be freely estimated and simplifying the selectivity blocking structure to be four year blocks. This is the new preferred base model. The jitters will confirm convergence properties for this model configuration.

Response: A set of 200 jitter runs of the model with widely dispersed starting values resulted in 27.5% converged runs, all with near identical estimates of depletion rates and spawning biomass. This corroborates that the model is stable, and that the converged estimates represent a global solution, and not a local minimum.

Request #14: Provide sensitivity analyses using the pre-recruit survey index with the current weighting of the fishery comp data and a down-weighting of the fishery comp data. Include a run with a simplified 8-yr selectivity blocking structure and under the down-weighting scenario with the pre-recruit index included. This may reveal a general pattern of model sensitivity to such data that may be informative for recruitments. Tension in the data might be informative.

Response: The inclusion of the pre-recruit index and down-weighting the age and length data pulled down the estimated biomass associated with the strong 1999 year class. This result is probably due to the conflicting signals from the acoustic survey and fishery catch-at-age data. This alternate configuration also caused a higher estimate of recent and projected biomass because the entire series was flattened out. The panel agreed with the STAT that the inclusion of the pre-recruit survey should not occur. The pre-recruit time series is short, and the juvenile fish may experience orders of magnitude variation in survival because of changing ocean currents and natural mortality.

Request #15: Provide a sensitivity analysis of an alternative model where historical California length frequencies are dropped. This will determine the sensitivity of model results to these data.

Response: Removing the historical California length composition data had no significant or appreciable effect on the biomass estimates.

Request #16: Provide a run with asymptotic selectivities for the acoustic survey and the Canadian fishery. This will provide an alternative view where the age data are not fitted by the selectivity functions becoming more dome-shaped. Forcing asymptotic selectivity for the Canadian fishery and the survey recognizes that the Canadian fishery tends to catch larger fish reducing availability of larger fish in the U.S. fishery.

Response: Asymptotic selectivity curves for the acoustic survey and the Canadian fishery, combined with constant M =0.23 throughout the life cycle, generally forced the U.S. fishery selectivity curves to become more asymptotic. This change to the configuration also increased the estimates of acoustic survey q, reduced the spawning biomass (all above B0 in the series prior to the mid-1960s), and significantly reduced uncertainty in the recruitment estimates of the large year classes. This scenario was 184 NLL units removed from the base with a reduction of only three parameters. Acoustic survey q was estimated at 1.4 in this scenario, although full selectivity was not achieved until age 14. For the scenario when M for older fish was fixed but selectivities were not forced to be asymptotic, the result was 128 likelihood points removed from the base case with an increase of 2 parameters relative to the previous scenario A similar model with M = 0.25 throughout was109 likelihood points removed from the base, slightly better than with M = 0.23.

To better fit the data under the constraints of this scenario, the selectivity curves shifted significantly to the right, maximizing on 9-10 yr old fish in the acoustic survey, and all three selectivities were quite low at age 15+ (~0.1). For this last scenario, q was estimated at 0.34. The STAT explained that one of the reasons they originally recommended a fixed q was the confounding in the estimates of q and M.

Request #17: Provide a sensitivity analysis assuming an M of 0.25 with and without the free estimation of M for age 15+. This will explore the necessity of assuming a high M for old fish.

Response: At a fixed natural mortality rate of 0.25 with 'old M' fixed, stock trajectory shifts upwards with current spawning stock biomass estimated to be substantially greater than unfished spawning stock biomass (Fig. 10). The result was considered implausible. This run demonstrates a high degree of sensitivity in the model to relatively small changes in the value of natural mortality for ages less than 15. It follows that assuming M is fixed could lead to an important compression of uncertainty. The STAT team confirmed that a defensible prior for M existed and could be implemented in the MCMC run. However, the MCMC run completed after the STAR meeting used M fixed at 0.23^{-yr} for ages less than 14.

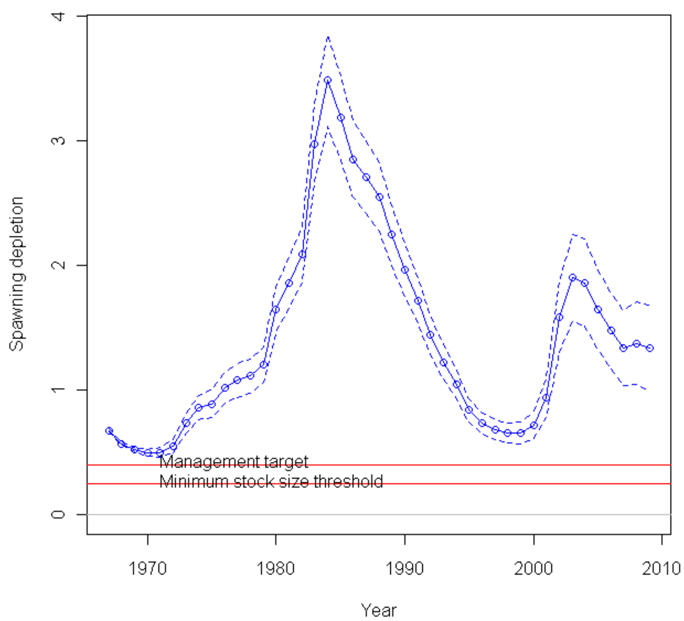

Figure 10. Predicted spawning stock biomass assuming a fixed natural mortality rate of 0.25 for all ages, but otherwise under the proposed base model configuration.

The requested comparison run with free estimation of old-M was included in the response to Request #20.

Request #18: Provide values of q that result in 12.5% and 87.5% of the female spawning biomass probability distribution. This assumes normal distribution of probabilities of the female spawning biomass around the ML estimate. This may be the basis for an alternative decision table if the MCMC does not converge.

Response: The uncertainty over female spawning stock biomass was presented by manually altering q to achieve the 12.5th and 87.5th percentiles of the proposed base model prediction. These results were included in a decision table (included in response to #21) alongside corresponding estimates of depletion over a three year projection period. The final table will include additional rows reflecting different management decisions (a range of quotas).

A different formulation of this table was considered in which depletion and spawning stock biomass are estimated with respect to changes in a key axis of uncertainty such as q. However it was concluded that it would be better to attempt to represent the uncertainty over all important axes, including the parameter controlling the final selectivity of the acoustic survey and natural mortality rate. It was acknowledged that where available the MCMC run would offer a better means of conveying uncertainty in model outputs.

Request #19: Provide a run fixing base M at 0.25 and allowing M for older fish (age 15+) to be freely estimated. This will assess whether a higher M for older fish or a higher estimate of base M is needed to explain the lack of older fish.

Response: The requested run was included in the response to Request #20.

Request #20: Profile over base M from 0.21 to 0.25 with the M for older aged fish and q freely estimated using a) dome-shaped and b) asymptotic selectivity in the acoustic survey and the Canadian fishery. This will establish model sensitivity to estimates of base M.

Response: At the lowest natural mortality rate of 0.215 the model predicts a similar trend in spawning stock biomass that is more damped, with both lower estimates of B0 and smaller recruitment events, particularly in 1980 and 1984 (Fig. 11). In scenarios where fewer individuals die, fewer individuals are recruited in the model to maintain fit to the observed data (Fig. 12). Changing M had the least impact on spawning stock biomass in the most recent years.

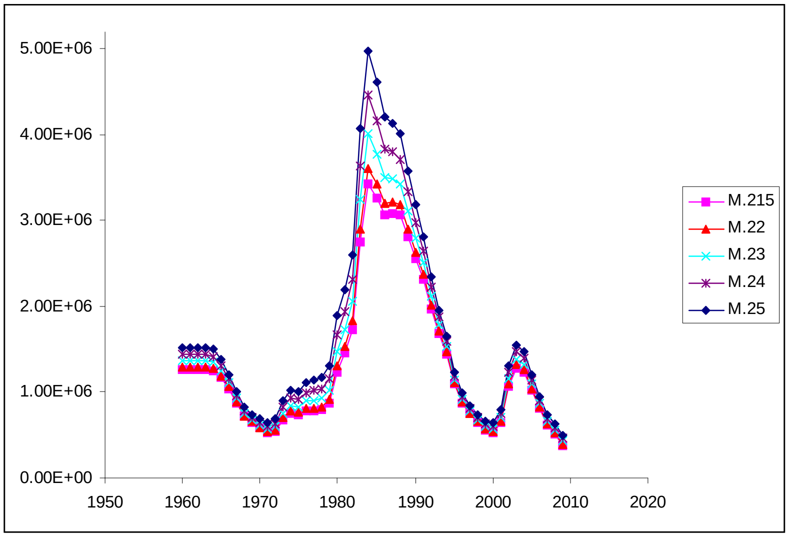

Figure 11. Predicted spawning stock biomass over a range of fixed rates of natural mortality under the proposed base model configuration.

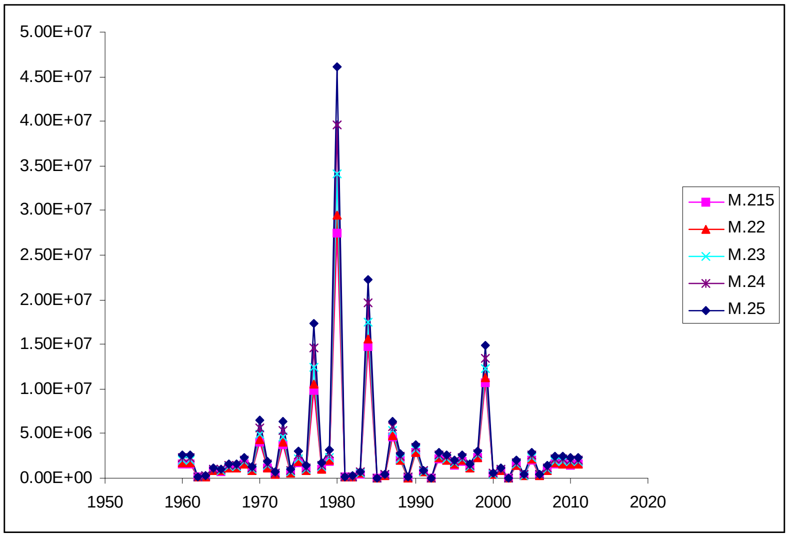

Figure 12. The predicted recruitment over a range of fixed natural mortality rates under base case model specification.

Request #21: Using the results of Request #18, develop one row of a decision table that can be used to exemplify a contingency decision table if the MCMC results do not converge.

Response: Under the proposed base model configuration, the maximum likelihood estimate of the current female spawning stock biomass is 435,000 metric tons corresponding to an estimated depletion of 32 percent (Table 1). Given the 40:10 rule, female spawning stock biomass is projected to fall to 357,000 tons by 2010, at 26 percent depletion. Subsequently, spawning stock biomass is projected to recover slightly to 363,000 tons, a depletion level of 27 percent.

In all years, there is considerable uncertainty over stock status. The most pessimistic of the two alternate models estimates current female spawning stock biomass at 294,000 tons and a considerably lower depletion level of 22%. The most optimistic alternative model predicts current female spawning stock biomass at 575,000 tons at a depletion level just above the 40 percent reference level. The projections are somewhat pessimistic in general; the most optimistic projection predicts a depletion level of 34% in 2011.

		State of nature					
		12.5th percentile of female spawning biomass from base model		Base Model		87.5th percentile of female spawning biomass from base model	
Relative probability		0.25		0.50		0.25	
Management Action							
Year	Coast-wide catch (mt)	Female spawning biomass (millions mt)	Estimated depletion	Female spawning biomass (millions mt)	Estimated depletion	Female spawning biomass (millions mt)	Estimated depletion
2009	253,582	0.294	22%	0.435	32%	0.575	41%
2010	193,109	0.230	17%	0.357	26%	0.484	35%
2011	189,054	0.251	19%	0.363	27%	0.476	34%

Table 1. Predicted female spawning stock biomass and depletion under three model scenarios including a three year forecast.

Description of base model and alternative models used to bracket uncertainty

The initial base model developed by the STAT in advance of the STAR meeting did not include all the data for the 2008 Canadian fishery. Complete data became available just days before the start of the meeting, with the consequence that the STAT did not have time to fully explore the properties of the initial base model. The primary change in the model configuration developed during the STAR meeting was the freeing up of the acoustic survey q parameter, which the STAT had fixed at a value of 0.7 in the initial base model. The final base model also had a slightly simplified blocking structure for the time-varying fishery selection curves.

The SS3 model configuration selected for the final base model had the following features.

- A single coastwide stock was assumed and there was no explicit spatial structure.
- There were separate US versus Canadian fisheries, each with its own length-composition and conditional age-at-length composition data and age-based selection curves.
- The joint US-Canada acoustic / midwater trawl survey biomass index was the only fishery independent time series and was the primary tuning index.
- Age-reading imprecision was incorporated, as it had been in the 2008 assessment, but the age-reading variability coefficients were revised from the 2008 assessment with additional data from new cross-calibration studies of age-readings. Also, the STAT adopted an adjustment for cohort-specific age-reading error, which is thought to occur when very strong cohorts are present.
- Time-varying growth parameters were estimated with two blocks for the length-at-age-12 parameter and four blocks for the growth coefficient (k). This blocking structure was slightly more complicated than the structure used in the 2008 assessment.
- Parameters of a Beverton and Holt recruitment curve were estimated using an assumed beta-prior probability distribution for the steepness parameter, as in recent past assessments, with annual recruitment deviations estimated for 1962 to 2007. The recruitment variability parameter (sigma-R) was estimated in this assessment, whereas in the 2008 assessment it was fixed at a value of 1.13.
- Acoustic survey selection was assumed to be time-invariant.
- The catchability coefficient for the acoustic survey was freely estimated.
- The selection curves for the two fisheries and the acoustic survey were estimated and not forced to be asymptotic.
- Fishery selection was time-blocked to accommodate apparent targeting of strong year-classes and structural changes in the fisheries. There was one block for the early years of the fisheries (through 1980) and then, starting with 1981, there were sets of four-year blocks for some of the fishery selection parameters for each of the two fisheries.
- The natural mortality coefficient was fixed at 0.23 yr^{-1} for ages 0 to 13, and then was allowed to ramp to higher (or lower) values for age-14 and the age-15+ group.

Alternative models used to bracket uncertainty.

The alternative models for constructing the decision table were derived from the posterior distribution of the base model rather than from alternative model formulations. However, in the event that the MCMC did not converge -- the MCMC run was completed after the close of the STAR Panel meeting -- a decision table would have been developed using the acoustic survey q parameter as the primary dimension of uncertainty (see Request #21).

Comments on the assessment

The catchability coefficient for the acoustic survey (survey q) continues to be a major source of uncertainty in the assessment of this stock. In the initial base model brought to the STAR meeting the STAT had fixed the value of survey q because they did not feel it could be well estimated. Past STAR Panels (prior to 2008) had recommended bracketing uncertainty in decision tables by using one or more fixed values of survey q. The current STAR Panelists were concerned that fixing the survey q parameter would grossly constrain the plausible set of model estimates, and the model would produce a misleading and overly optimistic view of how much is known about the status of this stock. As was done for the 2008 assessment, the survey q parameter was freely estimated so that the final model could more appropriately reflect the uncertainty associated with this crucial parameter.

Technical merits:

- Uncertainty associated with the acoustic survey q parameter was incorporated in the assessment results.
- The assessment was able to use the MCMC approach to integrate across the uncertainty associated with all the estimated parameters. For most of the Council's groundfish stocks it has not been possible to use this approach.

Technical deficiencies:

- Although the 2008 assessment developed a prior probability distribution for the natural mortality coefficient (M) (based on Hoenig's relationship between longevity and M), the STAT did not use this prior in fitting the 2009 assessment model, but instead assumed a fixed value ($M = 0.23$ /yr). The assessment model in its estimates of stock status therefore did not capture the uncertainty about M.
- There were undesirable patterns in the residuals for the length composition data, which suggest a structural misspecification in one or more aspects of the model. The STAT was of the opinion that including growth and composition data by gender in the model structure would alleviate this problem.
- The current version of the Stock Synthesis software does not allow the fishery selection curves to mimic cohort targeting, which is a phenomenon that is supposedly driving fishing behavior. This structural limitation of SS could also account for some of the odd residual patterns associated with the length composition data.

- Some of the maximum likelihood estimates showed strong parameter cross-correlation, which generally indicates over-parameterization and a ridged posterior probability distribution that is likely to inhibit convergence in MCMC routines. Subsequent to the STAR Panel meeting, when the STAT conducted the initial MCMC runs, they encountered lack of convergence due to one redundant selectivity parameter; they solved the convergence problem by fixing this parameter at a reasonable value.
- Although the acoustic survey team made significant progress at responding to some of the deficiencies in the acoustic survey data identified by the 2008 STAR Panel, there remain important gaps in the acoustic survey documentation. The acoustic survey biomass estimates are derived by matching up the acoustic back-scatter data with biological sample data (composition by species, length and sex) collected by midwater trawl tows. The process of matching these two types of data was unclear to the STAR Panel. There should be explicit criteria that determine when and where the midwater tows are taken and their duration, and explicit rules for how these biological sample data are then assigned to the various segments of the acoustic transects. The post-stratification of tows partly based on similarity in the observed length-composition of the catches is particularly problematic, and is a likely source of bias (of unknown direction) in the length composition estimates, and a downward bias in the associated variance estimates.
- Because of the very late arrival of the 2008 fishery data (especially the data from Canada but also the late-season data from the US) the STAT had very little time to assemble and analyze the data, let alone to explore model structure and uncertainty. Also, the 2008 data probably were not thoroughly error-checked because of the need for very rapid turn-around.

Explanation of areas of disagreement regarding STAR Panel recommendations

Among STAR Panel members (including GAP and GMT representatives)

There were no major disagreements.

Between the STAR Panel and STAT Team

There were no major disagreements.

Unresolved problems and major uncertainties

- The re-weighting scheme applied to the input data series in the assessment needs further investigation given the significant tension in the age- and length-composition data in the survey and fisheries. The STAR panel is concerned that the current iterative weighting scheme may be inappropriate in that poorly fitting (but unbiased) data series may simply be down-weighted out of the model. However, it is not clear how to derive a better weighting scheme that will achieve an appropriate balance between process error (e.g., time variation in selection or growth) and observation error (e.g., effective samples sizes for composition data).

- There is concern that some of the input data may be biased. The STAR Panelists suspect, in particular, that the acoustic survey age- and length-compositions may be biased because of a tendency for biological sampling to occur disproportionately on dense aggregations of fish that may not be representative. The raw acoustic survey data need to be analyzed to allow verification that an appropriate stratification was applied. Additionally, there need to be explicit rules regarding how the age and length data are collected in the survey and an explicit recounting of the rules that applied in past surveys. The methods for combining length samples into strata needs further review. A post-stratification scheme that creates more homogeneous strata by pooling tows with catches of similar length structure is not justified. This procedure could bias the estimates of length compositions applied to the acoustic survey tracks, and grossly overestimate the precision in estimated length compositions.
- There is continuing concern regarding the strong dome-shaped selectivity pattern in the acoustic survey. The mechanism responsible for this pattern remains a mystery.
- The residual pattern in the length-composition data indicates model mis-specification and potential bias in model results. Estimates of model uncertainty, either from the maximum likelihood estimates or from MCMC, do not include uncertainty or bias with regard to model structure, or uncertainty in acoustic estimates of biomass resulting from sampling variability and bias in estimated length compositions. A sex-specific model may address some of the problems evident in the length-composition residuals but would not resolve bias caused by the post-stratification scheme.
- The assessment provides very little information regarding the acoustic survey q and consequently provides highly uncertain estimates that are key to management, such as the ABC.
- The assessment provides even less information regarding M.
- The assumption that the older fish experience higher rates of natural mortality needs independent empirical verification.
- There may be a lack of correspondence between the maximum likelihood and MCMC results. The current STAR process does not allow enough time to evaluate MCMC results and determine risk-neutral recommendations.
- The static B0 construct appears to be a poor framework for managing a stock with highly dynamic recruitment, such as Pacific hake.
- Some parameters appear to be highly correlated, which will affect the efficiency of the search algorithm.

Management, data, or fishery issues raised by the GMT or GAP representatives during the STAR Panel.

The issues that were raised are adequately described elsewhere in this report.

Recommendations for future research and data collection.

The 2007 STAR Panel presented a comprehensive review of recommendations from past STAR Panels, many of which still apply but are not reiterated here. The recommendations below resulted from discussions during the 2009 STAR Panel review and subsequent email exchanges.

1) Investigate how the biological sampling in the acoustic survey occurs to determine whether these data are representative of the backscatter in the survey.
2) Investigate how the biological samples are processed and applied to the acoustic estimates, including the post-stratification of length samples.
3) The raw data in the acoustic survey, including the length samples, needs to be appropriately assembled to allow statistical analysis of these data and appropriate stratification.

The 2009 STAR Panel also considers the following recommendations, from the 2008 STAR Panel Report, to merit consideration.

4) A Management Strategy Evaluation approach is needed to evaluate the 40-10 harvest control rule when applied to a stock with dramatically episodic recruitment, such as Pacific hake stock. Related to this is the need to evaluate how well different assessment models recapture the true population dynamics. At issue is whether a simpler model such as ADAPT / VPA performs better or worse than a more complex model such as Stock Synthesis.
5) Future assessment models should explore gender- and length-based selection processes, in recognition that the genders differ in growth and that many of the more influential dynamic processes that operate in the fishery are length-based but are currently considered from an age-based perspective (for example selectivity).
6) When the raw acoustic survey data become available there should be a re-evaluation of the treatment / adjustment of pre-1995 acoustic survey data and index values. For example, the biomass index implied by the area covered by the pre-1995 surveys should be compared with the total biomass from the full area covered by the post-1995 surveys. The difference between these two indices has implications for the magnitude of the survey catchability coefficient prior to 1995.
7) There should be further exploration of geographical variations in fish densities and relationships with average age and the different fisheries, possibly by including spatial structure into future assessment models.
8) There should be exploration of possible environmental effects on recruitment and the acoustic survey.

References

Hamel, O.S., and I.J. Stewart (2009). Stock Assessment of Pacific Hake, *Merluccius productus*, (a.k.a Whiting) in U.S. and Canadian Waters in 2009.

Pacific Fishery Management Council (2008). Terms of Reference for the Groundfish Stock Assessment and Review Process for 2009-2010.