Report of the 2008 U.S. / Canada Pacific Hake (Whiting) Stock Assessment Review (STAR) Panel.

Review Panelists:

David Sampson,
Panel chair and representative for the Scientific and Statistical Committee of the Pacific Fishery Management Council (PFMC)

Malcolm Haddon,
Center for Independent Experts

Noel Cadigan,
Fisheries and Oceans, Canada

Dan Waldeck,
Representative for the PFMC Groundfish Advisory Panel

John Wallace,
Representative for the PFMC Groundfish Management Team

Overview

During 11-14 February 2008, a joint Canada-U.S. Pacific Hake / Whiting Stock Assessment Review (STAR) Panel met in Seattle, Washington, to review three stock assessment documents, by Helser *et al* (2008), Sinclair & Grandin (2008), and Martell (2008). The Panel operated under the U.S. Pacific Fishery Management Council's Terms of Reference for STAR Panels (SSC 2006), but as in previous years, the Panel attempted to adhere to the spirit of the Canada-U.S. Treaty on Pacific Hake / Whiting. As was the case in 2004, 2005, 2006, and 2007 the Panel included a member from Canada and the stock assessment team also included Canadian participants (see *List of Participants*). The revised stock assessments and the STAR Panel review will be forwarded to the Pacific Fishery Management Council and its advisory groups, and to Canadian Department of Fisheries and Oceans (DFO) managers and the PSARC (Pacific Scientific Advice Review Committee) Groundfish Sub-committee.

All members of the stock assessment team (STAT) – Drs. Thomas Helser, Ian Stewart, Owen Hamel, Alan Sinclair, Chris Grandin, and Steve Martell – attended and actively participated in the meeting. Public comment was entertained throughout the four-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, and this was sufficient time to adequately prepare for the review of the three assessments.

The Panel convened at 08:30 on Monday February 11th. Stacey Miller (US National Marine Fisheries Service, NMFS) welcomed the group and then Dr. Elizabeth Clarke (NMFS) briefly reviewed the status of the Pacific Hake / Whiting treaty. The treaty now needs to be ratified by the Canadian parliament and until that occurs the necessary committees cannot be formed. Nevertheless the STAR panel review could continue and would attempt to meet the needs of both parties to the treaty. David Sampson (STAR Panel Chair) then opened the meeting with an overview of the review process including the terms of reference, Panel membership, expected products, and a timeline for completion of the Panel's report. A preliminary meeting between the assessment team groups had occurred and they had all used the same available data in their assessments, although each emphasized different aspects and aggregated the data to different degrees. Tom Helser provided the STAR Panel with a detailed description of the available data inputs. Rebecca Thomas (NMFS) provided a detailed overview of the acoustic survey work. Chris Grandin described the fishery distribution changes in Canadian waters during 2006 and 2007. In addition, John Horne (University of Washington) gave a presentation on a revision of target strengths used in the acoustic survey for Pacific Hake / Whiting. Then the following three stock assessments were presented. Tom Helser (NMFS) presented the Stock Synthesis II (SS2) catch-at-age model (Helser *et al*, 2008), Alan Sinclair (DFO) presented an ADAPT / VPA model (Sinclair & Grandin, 2008), and Steve Martell (University of British Columbia) presented an assessment model that directly estimated parameters of management interest (named TINSS; Martell, 2008).

Based on discussion of the stock assessment documents and related presentations, the Panel requested 24 clarifications, some of which included additional model runs, to help identify the base case, the full range of uncertainty in the stock assessment, and the similarities and differences between the three assessment models. This large number of requests reflected the complexity of reviewing three distinct assessment models contributed from two nations. This iterative process of making additional model runs and discussing the results continued through the end of the day on February 13th. The Panel spent the morning of February 14th reviewing an outline structure of its report; the meeting was adjourned at 14:00. A draft Panel report was distributed by email to all Panel participants for serial development. A draft final Panel report was completed on February 22nd so that it could be included in the "Briefing Book" for PFMC’s March meeting.

After careful review of results and diagnostics from the three assessment models (SS2, ADAPT / VPA, and TINSS), the Panel recommended acceptance of a particular scenario from the SS2 model as the base case. This scenario, developed during the review period, estimated the most important parameters more freely and reflected a broad but realistic range of uncertainty in the relative depletion level and productivity of the stock. The base model was developed with careful consideration of knowledge and uncertainty about Pacifc hake stock dynamics, and fisheries and survey’s for this stock. Although all three models had the same data streams available for use, the models differed in the amount of data used, the degree of data aggregation, and assumptions on the magnitude of observation error relative to process error. The basic data sets consisted of the following: total catches from the US and Canadian fisheries between 1966 – 2007; length compositions from the US fishery (1975-2007) and the Canadian fishery (1988-2007); conditional age-at-length compositions from the US fishery (1975-2007) and the Canadian fishery (1988-2007); standard age composition data (derived from age-length keys) from the US fishery (1973-1974) and the Canadian fishery (1977-1987); biomass indices, length compositional 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); plus biological data relating to pre-recruit abundance, growth, maturity at age and length, and natural mortality.

The SS2 catch-at-age model involved the least degree of data aggregation. Both the ADAPT and TINSS models combined the US and Canadian age compositions and assigned equal weight over all years. A major structural difference between the models was the pattern ascribed to the selectivity for the surveys and commercial fisheries. In the SS2 model, evidence from the US fishery and acoustic survey age-compositions favored "domed selectivity", in which the oldest age-classes were less apparent than intermediate age-classes. Sinclair & Grandin (2008) concluded from a catch curve analysis, an analysis of the ratio of catch-at-age from the fishery and from the survey, and from the VPA that fishery and survey selectivity were asymptotic, meaning that the oldest age-classes were as apparent as intermediate age-classes. The TINSS model assumed that fishery and survey selection were asymptotic and an analysis showed that the estimated steepness parameter (h) was not overly sensitive to the assumption.

Responses to the STAR Panel's requests for alternative model runs indicated that all three models provided similar predictions about the resource biomass trajectory when the model assumptions were made to be the same or similar. At the same time the three models made similar predictions for the parameters of management interest (ABC, $F_{40/MSY}$, and depletion), but with differing ranges of uncertainty. The SS2 model that was originally brought to the STAR Panel bracketed uncertainty by using alternative models corresponding to a low and high acoustic survey selectivity at the final-age, but freely estimated the survey catchability. The final SS2 model agreed upon by the STAR and STAT involved more freely estimating the acoustic survey selectivity parameters, as well as acoustic survey catchability and the natural mortality coefficient for ages 14 and 15+. This had the effect of increasing the breath of uncertainty around key management parameters, such that it encompassed the uncertainty expressed by the alternative ADAPT / VPA and TINSS models. From the base case the estimated 2007 spawning stock biomass (SSB) is just below the target level of 0.40 SSB_0. A comparison of model outputs is provided in the table below.

Character	SS2 Base Case	VPA	TINSS
Model Platform	Stock Synthesis 2.0n	ADAPT	AD-Model Builder
Ageing error matrix	Yes	No	No
Selectivity pattern	Domed	Asymptotic	Asymptotic
Fishery composition observation error	Yes	No	Yes but < SS2
Survey composition observation error	Yes	Yes	Yes but < SS2
Length-compositions	Yes	No	No
2007 Depletion	0.379 (0.22 – 0.54)	0.280	0.519 (0.334 – 0.796)
2008 Catch '000s t	527 (141 – 942)	346 (40 - 520)	446 (182 – 864)

Table notes: The SS2 estimates for 2007 Depletion and 2008 Catch are the maximum likelihood estimates with approximate 95% confidence limits. The corresponding VPA estimates are from Run 1A and the 2008 catch range values are the catches from this run that will exceed the target exploitation rate with 20% and 80% probability. The corresponding TINSS estimates are from the marginal posterior distributions, with the ranges showing the 95% confidence limits.

There was debate over what would constitute a safe level of catch. Pacific hake / whiting exhibit highly variable episodic recruitment and the fishery during the last 40 years has been driven largely by three large year classes (1980, 1984, and 1999). Questions were raised over whether the Council's 40/10 harvest control rule, by itself, would be sufficient to maintain the stock above the B_{25} level that triggers rebuilding. It was pointed out that: (1) the fishery currently depends on the 1999 cohort, which is declining in abundance and biomass, (2) fishing mortality is increasing and in recent years has been higher than most previous years, and (3) recent catches have been relatively high. These risk factors concerning the fishery are increasing and should be a cause for concern. It is unknown exactly how much risk is involved with the use of the current assessments and harvest control rule with a species such as Pacific hake / whiting. There was general consensus among the STAR panel and STAT that there would be great value in developing and conducting a detailed Management Strategy Evaluation to determine the most robust combination of data collection, applied stock assessment, and harvest control rule that should be applied to achieve sustainable use of the Pacific hake / whiting resource.

In the meantime, the Panel concurred that the stock assessment is suitable for use by the Council and Council advisory bodies for ABC and optimal yield (OY) determination, and for stock projections. However, the risk factors listed above, when coupled with the observation that SSB has been in decline since 2003 (and is now predicted to be below SSB_{40}) while ABC has increased substantially over the same period, strongly suggests there may be cause for concern if managers elected to take the full ABC.

The STAR Panel commends the STAT for the quality of the documentation provided for review and their cooperation in performing additional analyses requested during the meeting.

Analyses requested by the STAR Panel

Monday Questions for the Stock Synthesis Analysts

1. *A major axis of uncertainty is the survey and commercial fishery selectivity. A domed selectivity provides a better fit. Can the specific data (i.e. age+year+fishery), or components where the fit is improved be identified. Rationale: if the improvement in fit is specific to just a small part of the data, as opposed to broadly based, then the improvement in fit may be for the wrong reason.*

Response: The STAT team produced Figure 1 showing the change in negative log-likelihood for the SS2 models with the survey selex[1] = 0.7 and selex = 0.5. The total difference in log-likelihoods between these models was 300 units, indicating that the selex = 0.5 assumption resulted in a substantially better fit overall. A negative in Figure 1 indicates that the selex = 0.7 assumption (less domed) fit the data more poorly than the selex = 0.5 assumption. About 50% of the improvement in fit from the more domed selex=0.5 model was associated with US age-composition data in 1990-1992 and the survey age composition data in 1997; however, 50% of the improvement was broadly distributed. The conclusion from this analysis was that the improvement in fit was not an artifact caused by some other type of model misspecification or unusual data.

[1] "Selex" is the name of the parameter that controls the selection coefficient on the 15+ age-class. A selex value of 1.0 is equivalent to asymptotic selection, where the oldest ages are fully selected.

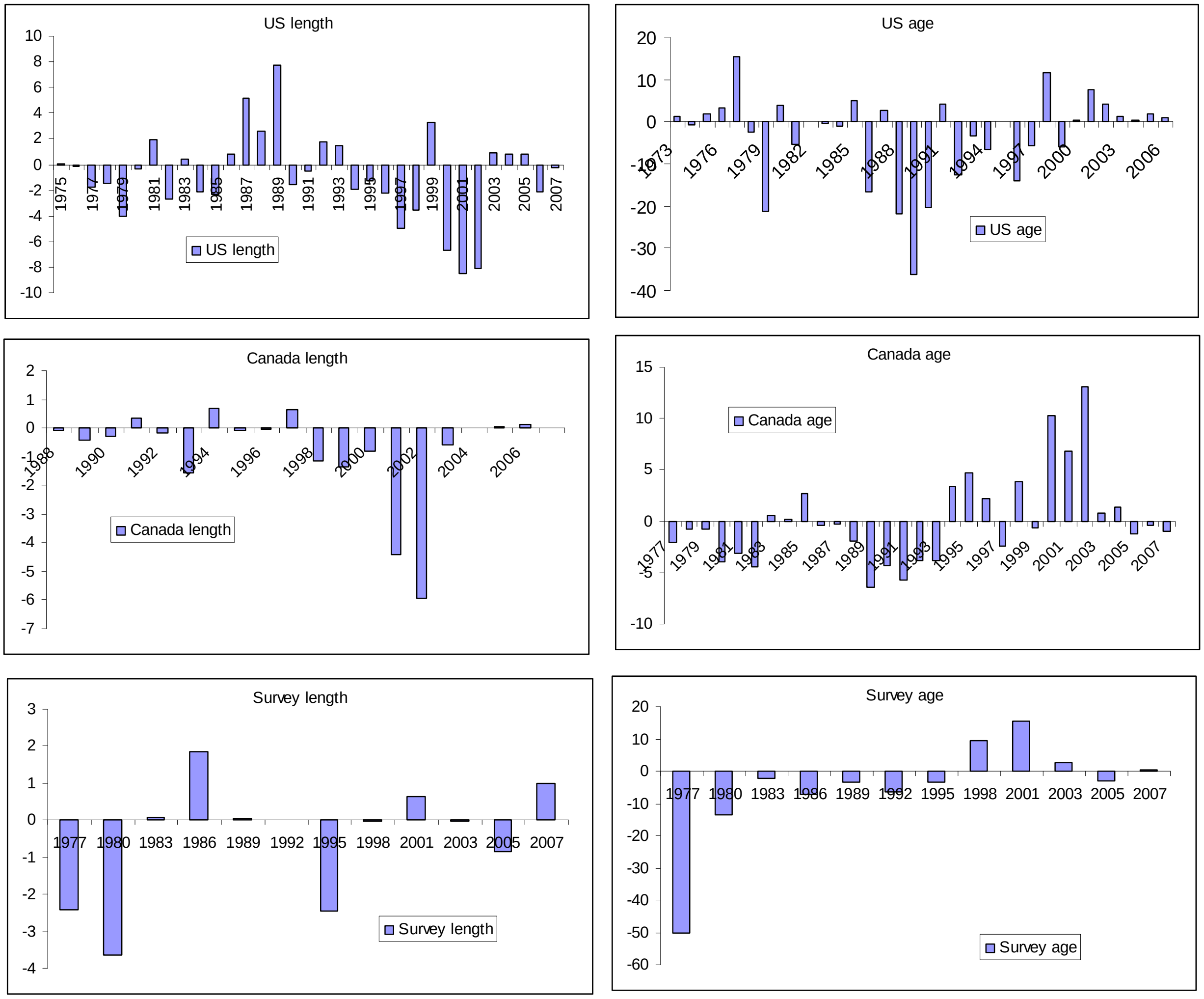

Figure 1. Change in negative log-likelihood values by data component between the SS2 models with selex = 0.7 and selex = 0.5.

2. *There is an inconsistency between the Canadian and US fishery age compositions. Are there specific data elements that are responsible for this inconsistency? Rationale: The end-result will be very dependent on the relative weighting applied to the two data sources.*

Response: Evidence for dome-selectivity was broadly distributed throughout the US age composition data (e.g., Figure 1, does not indicate major lack of fit due to the degree of domed-ness in selection), and prior to 1995 in the Canadian age composition data. However, in the period 1995-2003 the Canadian age composition data suggests that fishery had a less domed selectivity. Hence, there is stronger evidenced for domed-selectivity in the US age composition data.

3. *Tabulate discards in non-directed fisheries. Rationale: Demonstrate that the discards are trivial.*

Response: The hake discards in the non-hake fisheries reported by the NMFS Northwest Fisheries Science Center observer program were 822 mt in 2005, and 941 mt in 2006. The amounts are trivial compared to the directed fishery.

4. *Bailey et al. (1982) suggested that the reported foreign catches during1966-1976 were underestimated. Can the potential magnitude be quantified? Rationale: Unaccounted catches could influence assessment results.*

Response: The magnitude of under-reported catch in 1966-1976 was quantified, and an adjusted US catch was derived (see Bailey's US Catch in Figure 2). This was a provisional analysis, and the STAT reported that they would like to explore this as part of future research.

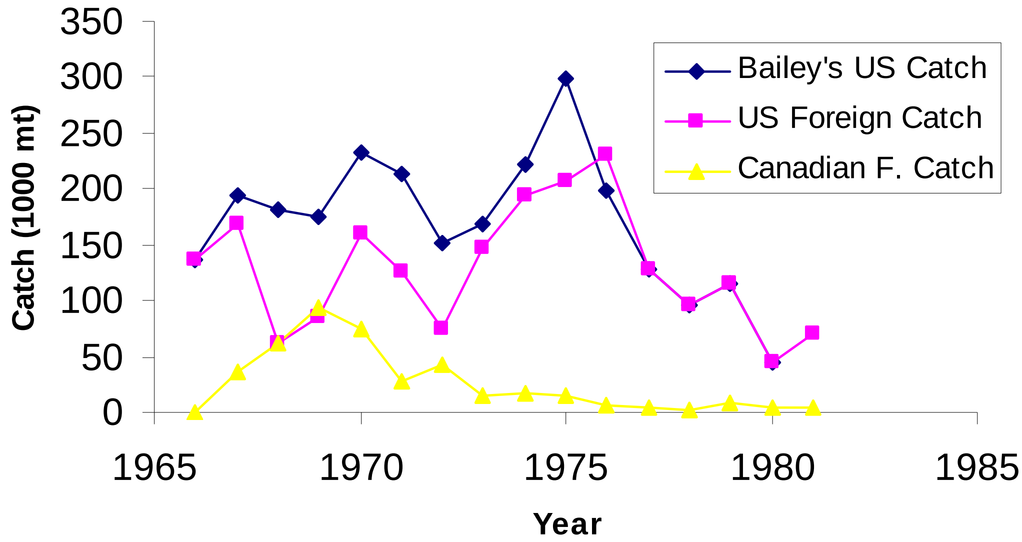

Figure 2. Foreign catch from US waters in 1966-1976, adjusted for mis-reporting. Reported catch (US Foreign Catch) and Canadian catch are included for reference.

5. *Tabulate the timing of the acoustic surveys. Rationale: Demonstrate that there have been no significant seasonal changes, which could affect catchability.*

Response: The timing of the surveys is shown in Figure 3. The STAT felt that the duration and changes in the timing of the survey would not have an important effect on the survey catchability.

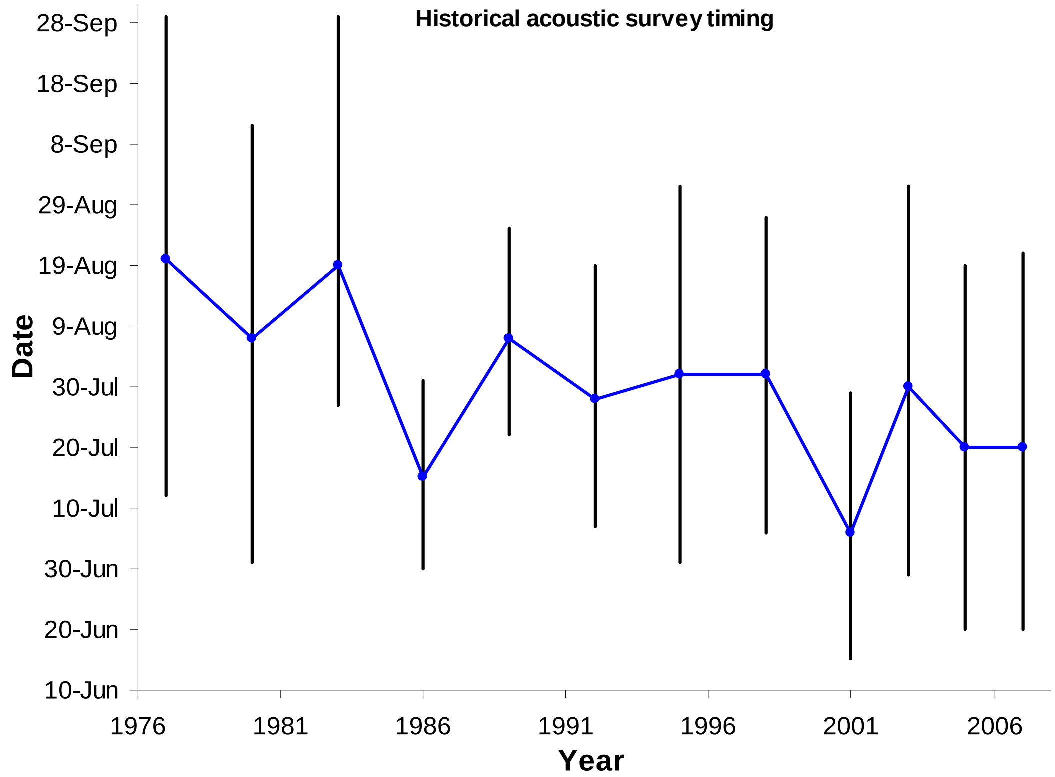

Figure 3. The start and end dates of the acoustic surveys, shown as vertical lines. The blue line connects the annual mid-point dates.

6. *Provide evidence for no sex-differences in growth and/or spatial distribution. Rationale: If these differences do occur then they have implications for future model development.*

Response: The STAT presented estimates of growth rates by sex (see Figure 4) and estimates of the proportion of females (Figures 5a,b). The evidence suggests that a model structured as length- and gender-based could produce considerable improvements in fits to the data. This was clearly not possible to do within the time frame of the meeting; however, it is a recommended area for future research.

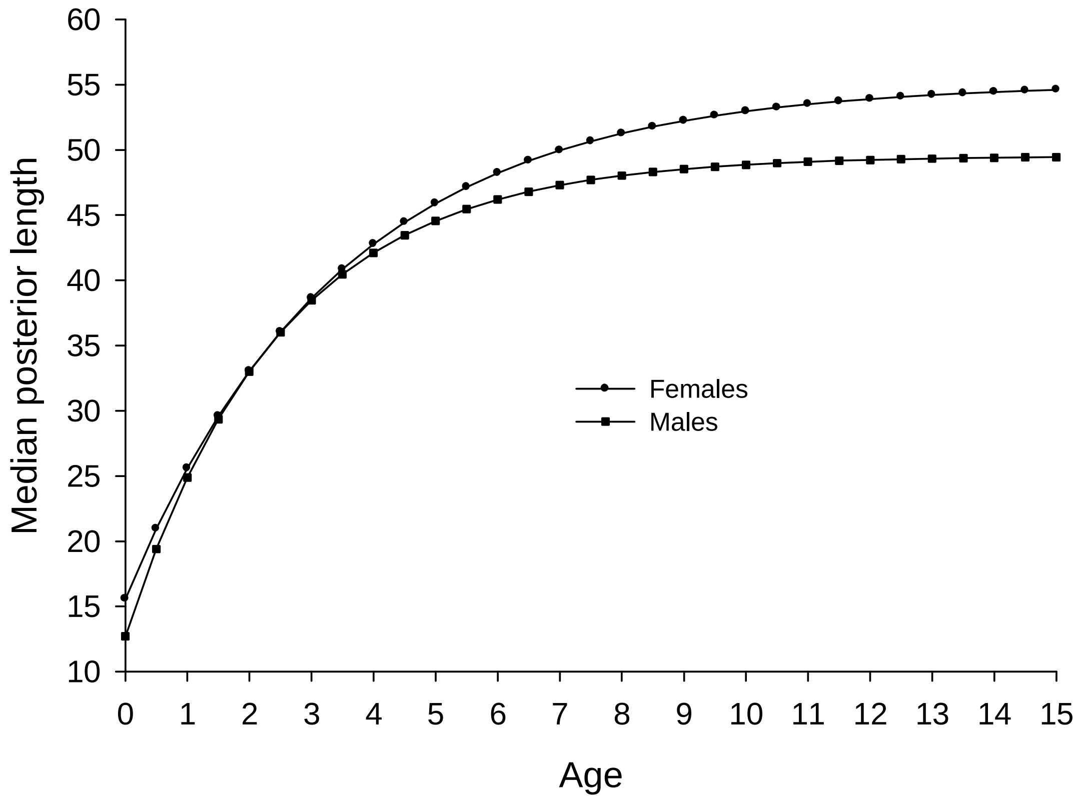

Figure 4. Mean growth curves by sex estimated over numerous cohorts from the 1975-2000 cohorts.

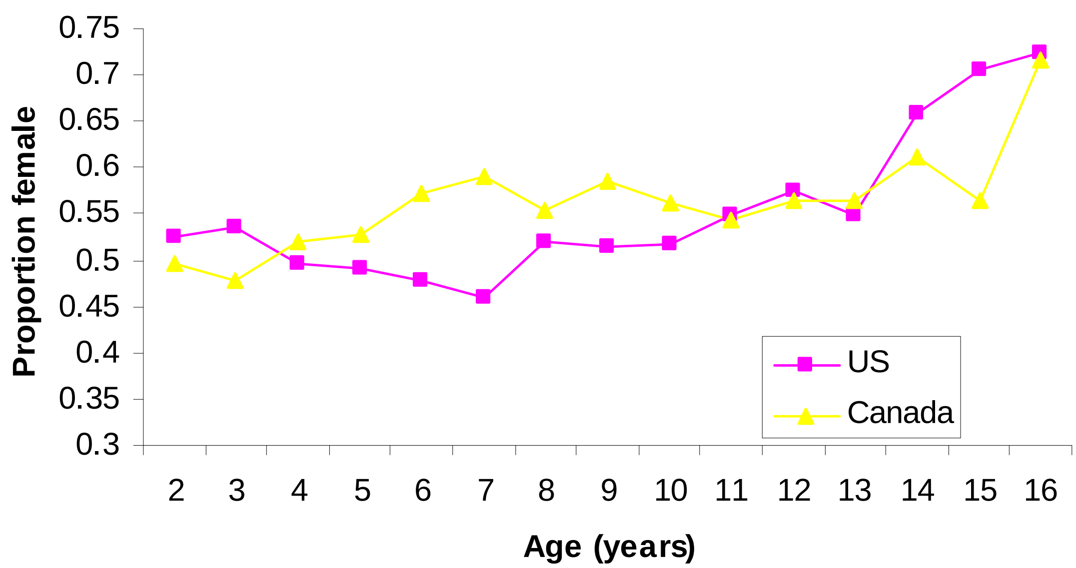

Figure 5a. Proportion female by age from commercial fishery samples during 1991-2006.

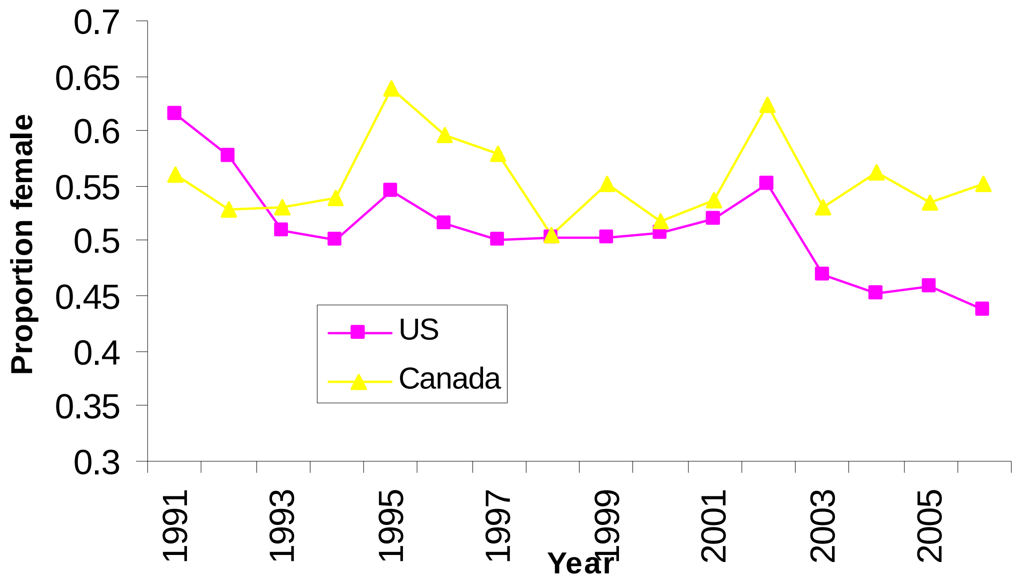

Figure 5b. Proportion female (all ages) from commercial fishery samples during 1991-2006.

7. *Provide rationale for age-based selection in the fishery and survey, as opposed to length-based selection. Rationale: real processes affecting catchability would more likely be length-based rather than age-based.*

Response: The STAT reported that this was a useful area for future research. The dominant source of variation in fishery selectivity and survey catchability may actually be length rather than age. However, in many fisheries models selectivity and catchability are commonly modeled as a function of age, and a motivation in designing the SS2 model was to keep it as standard as possible, while at the same time using the observed data more directly for estimation.

Tuesday Morning Questions for ADAPT / VPA and TINSS Analysts

1. *Compare predicted weight-at-age with empirical observations of weight-at-age, by year or cohort. Rationale: Confirm validity of the assumptions about length-weight relationships.*

Response: Text in the document describing the TINSS model implied that it had used empirical estimates of weight-at-age from field samples. In fact, in both the ADAPT / VPA and the TINSS models had used the same data on weight-at-age when referring to biomass estimates. These data were derived from the empirical data on length-at-age using a time-invariant weight-at-length relationship.

2. *Provide a plot of annual fishery selectivity. Rationale: To examine the assumption of annually constant selectivity.*

Response: The VPA provides estimates of fishery selectivity by age for each year of the analysis, and a major contrast between the SS2 versus the VPA and TINSS models was whether selectivity was domed or asymptotic. Estimates of average selectivity through time from Run 1A (Figure 6 and 7), averaged using the same time-blocks as the SS2 model, indicate some variation between each 7-year block, especially the 1984 - 1992 block compared to the other blocks. All the curves were asymptotic and of similar shape.

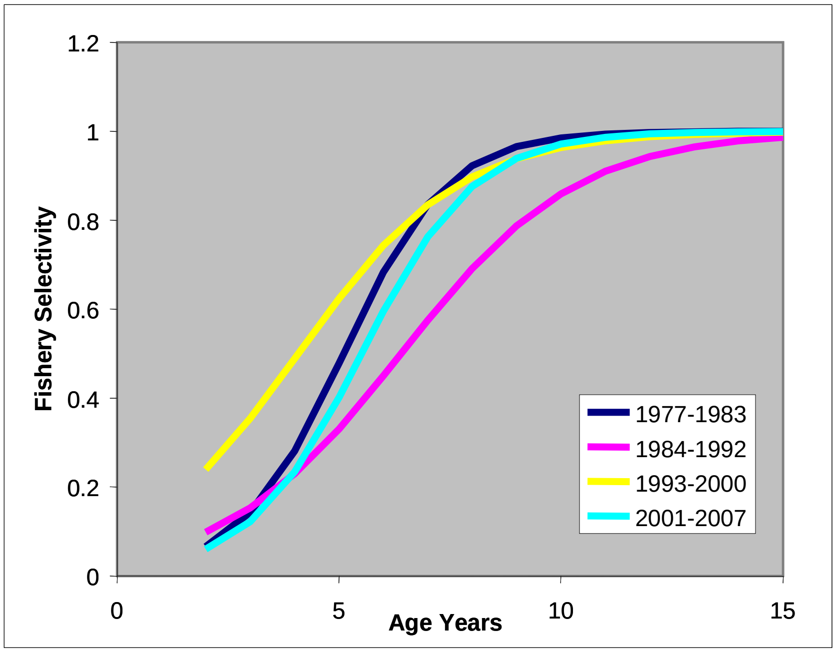

Figure 6. The predicted selectivity by age from the VPA analysis (selectivity Run 1A; Sinclair & Grandin, 2008) for four-year blocks.

The fishing mortality rates on the oldest age-classes indicate increased mortality rates in the most recent years.

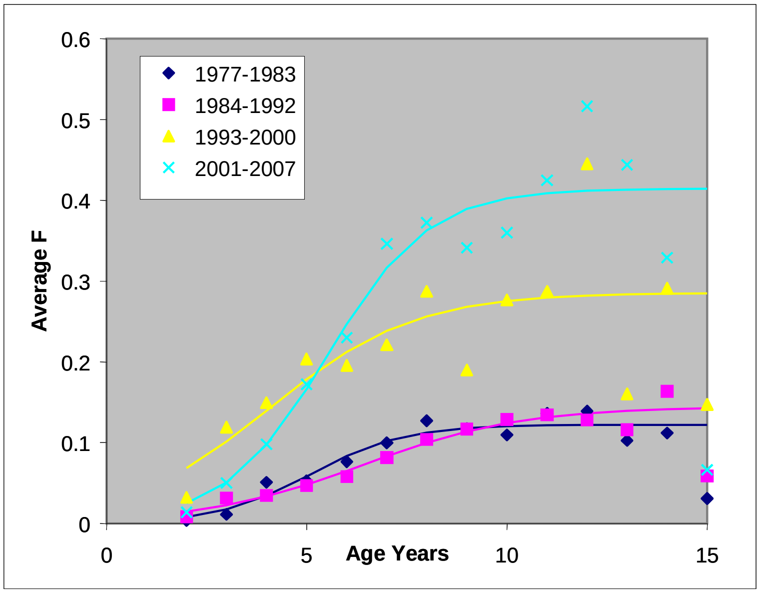

Figure 7. Fishing mortality relative to the selectivity in each group of years (*cf* Fig. 6).

3. *Provide plots of the VPA survey catchability. Rationale: To examine the assumption of annually constant and asymptotic selectivity in other models.*

Response: The survey catchability values as implied by the different VPA runs indicate asymptotic patterns (Figure 8). Also, the analyses of residuals of catch-at-age in the VPA assessment report (Sinclair & Grandin, 2008, reproduced below as Fig. 9) indicated that total mortality and survey catchability was relatively constant over ages 7-14 years.

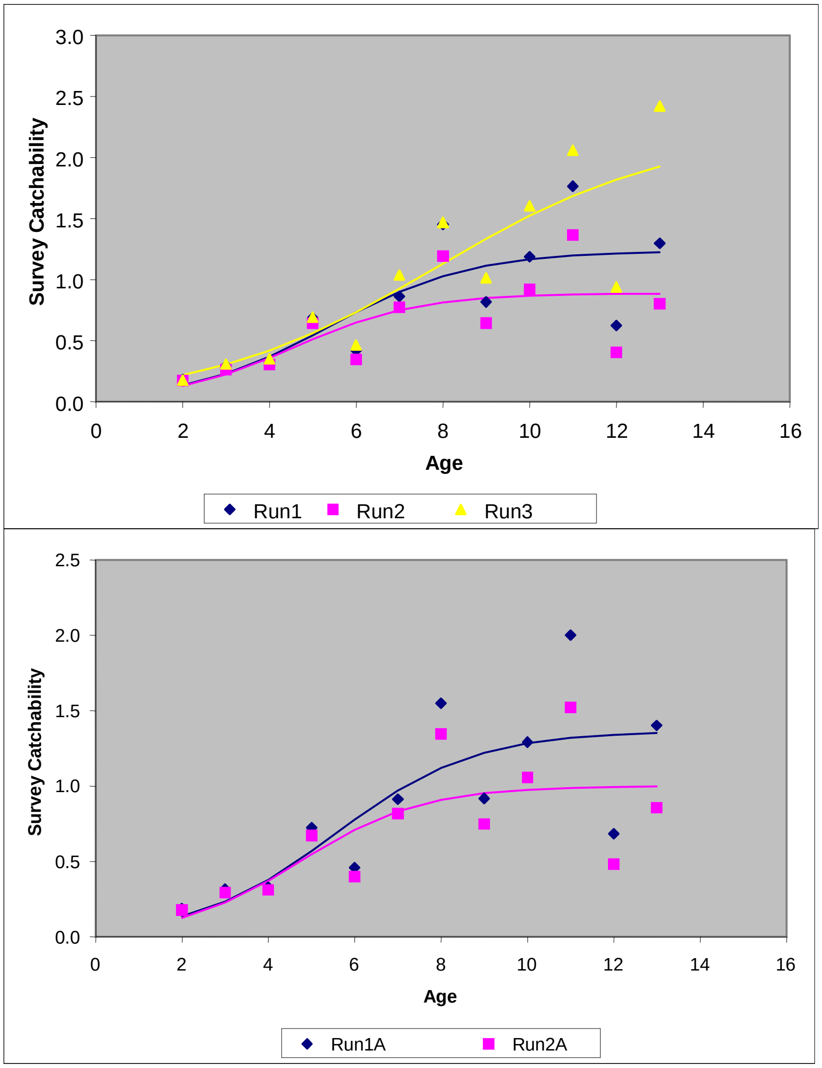

Figure 8. The survey catchability-at-age under the VPA analyses with the ADAPT runs 1, 1A, 2, 2A, and 3 (Run 1 F_T = weighted average 7+ fish, Run 2 F_T = Wt Av 4+ fish, and Run 3 F_T = Wt Av 12+ fish. Run 1A F_T = Wt Av 7+ fish and 10 more year parameters, Run 2A F_T = Wt Av 4+ fish and 10 more year parameters).

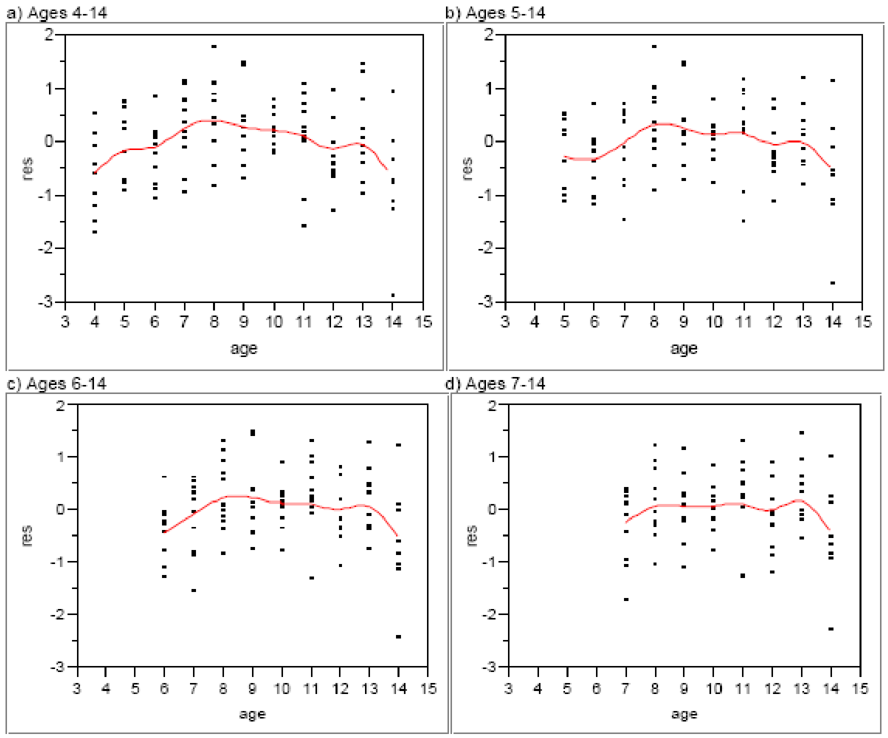

Figure 9. Figure 18 in Sinclair and Grandin (2008) showing residual patterns with respect to age from preliminary GLM analyses of total mortality of Pacific hake based on the results of the acoustic survey.

4. *Fmsy prior sensitivity. Shift the prior plus/minus 20%. Rationale: How sensitive is the management advice (e.g. Table 2 and 5) to the prior.*

Response: The posterior probability on the F_{MSY} is effectively coincident with the prior, indicating that the data are not informative for the target fishing mortality rate. Because other parameters are correlated with F_{MSY} the influence of the original prior was explored. In particular, the sensitivity of parameters of management interest were considered. While the changes in F_{MSY} have direct influences on steepness and ABC, the *MSY* appears to be relatively insensitive to the prior on F_{MSY} (Figure 10). Similarly, the predicted depletion level, estimates of *M* (natural mortality), and unfished spawning biomass were insensitive to the influence of F_{MSY} (Figure 11). The $\mathrm{F}_{\mathrm{MSY}}$ management target does not appear to be unduly influenced by the prior probability for $\mathrm{F}_{\mathrm{MSY}}$.

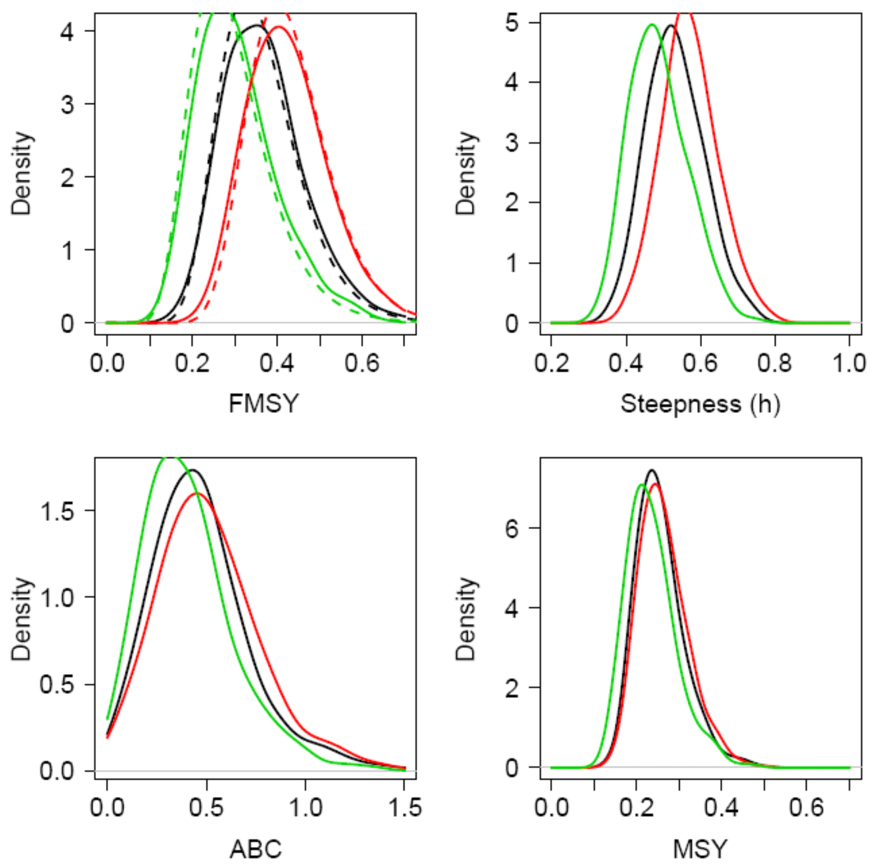

Figure 10. Shifted plots of the prior and posterior for F_{MSY} (solid line =prior, dotted line = posterior), with its implications for steepness, the ABC and the MSY. The insensitivity of MSY to the prior imposed on F_{MSY} is apparent.

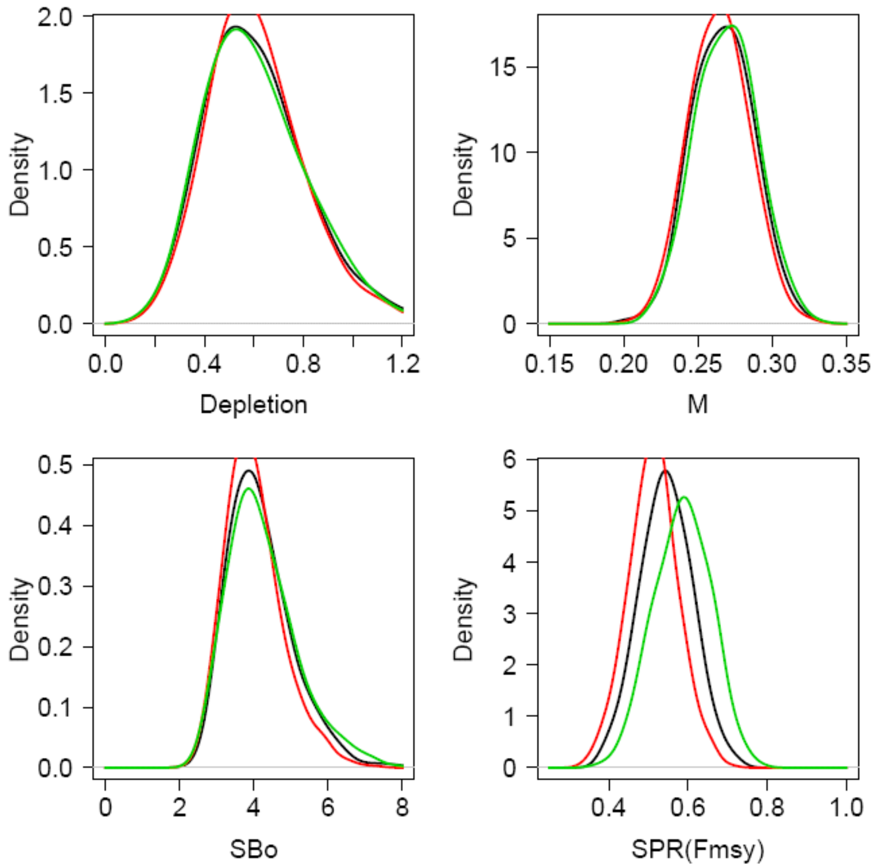

Figure 11. The insensitivity of stock depletion levels, natural mortality (*M*), and initial spawning biomass relative to shifts in the prior on F_{MSY} (Fig. 10).

<u>Tuesday Afternoon for the Stock Synthesis Analysts.</u>

1. *What is the impact on values in Table f. in Helser et al. when natural mortality is estimated, with a reasonable prior. Rationale: Fixing natural mortality, and profiling only over selex, may over-state uncertainty in depletion, etc. because of confounding in the effects of selex and natural mortality on population outcomes.*

Response: The top panel of Figure 12 illustrates data from Table 13b in the original SS2 assessment document, which subsumes the original Table f, while the bottom panel indicates how a less informative prior on *M* (natural mortality) alters the profile over the survey selectivity parameter for the oldest fish. The net effect was to compress the lower limits upwards. This question led to the Wednesday afternoon Request 4 (below).

Derived Parameter	Base model Final selex=0.5 MLE	Asymptotic 95% CI		Alt. Low Final selex=0.7 MLE	Asymptotic 95% CI		Alt. High Final selex=0.3 MLE	Asymptotic 95% CI	
2007 Depletion	0.437	0.293	0.581	0.291	0.212	0.370	0.570	0.418	0.723
2008 Depletion	0.429	0.254	0.604	0.292	0.156	0.428	0.597	0.413	0.782
MSY	346,130	247,101	445,159	219,270	153,310	285,230	467,030	320,273	613,787
B_{MSY}	637,580	359,397	915,763	434,510	248,255	620,765	917,560	504,980	1,330,140
SPR_{MSY}	0.234	0.107	0.360	0.248	0.104	0.393	0.247	0.108	0.385
2008 Catch	401,720	190,765	612,675	111,090	22,335	199,845	750,820	411,034	1,090,606
Rzero (billions)	1.210	1.010	1.410	0.787	0.700	0.874	1.674	1.376	1.971
Bzero (millions, mt)	1.836	1.531	2.141	1.193	1.060	1.326	2.538	2.086	2.989

Derived Parameter	Base model Final selex=0.5 MLE	Asymptotic 95% CI		Alt. Low Final selex=0.7 MLE	Asymptotic 95% CI		Alt. High Final selex=0.3 MLE	Asymptotic 95% CI	
2007 Depletion	0.472	0.324	0.620	0.307	0.213	0.400	0.568	0.417	0.720
2008 Depletion	0.485	0.302	0.668	0.271	0.147	0.395	0.603	0.417	0.789
MSY	406,060	275,863	536,257	284,320	189,227	379,413	476,520	321,950	631,090
B_{MSY}	742,810	400,535	1,085,085	516,020	281,878	750,162	932,550	510,464	1,354,636
SPR_{MSY}	0.242	0.106	0.378	0.239	0.104	0.374	0.248	0.110	0.386
2008 Catch	532,400	251,160	813,640	180,080	28,264	331,896	770,080	414,399	1,125,761
Rzero (billions)	1.503	1.170	1.835	1.043	0.788	1.297	1.728	1.362	2.095
Bzero (millions, mt)	2.086	1.692	2.480	1.461	1.188	1.734	2.567	2.088	3.047

Figure 12. The impact on parameters of management interest of estimating natural mortality using a broader prior than originally used in the SS2 modelling. The top panel is the original outputs while the lower panel illustrates the effect of the estimation of *M*.

2. *Explore estimating the initial age-composition in 1966. Rationale: The steady-state assumptions may have implications on model results.*

The SS2 model assumes that the population has an equilibrium age structure in 1966, but the age compositions from the earliest samples indicate that equilibrium was unlikely. This also is expected from the very high variation in recruitment leading to episodic recruitment. In fact, the use of bounded recruitment residuals (forcing a sum to zero) limited the number of years which could include recruitment deviations. 1963 was the earliest year in which recruitment deviations could be successfully imputed (Fig. 13). The additional early recruitment deviations had a relatively minor effect on the subsequent recruitment deviations (Fig. 14) and the spawning stock biomass trajectory (Fig. 15).

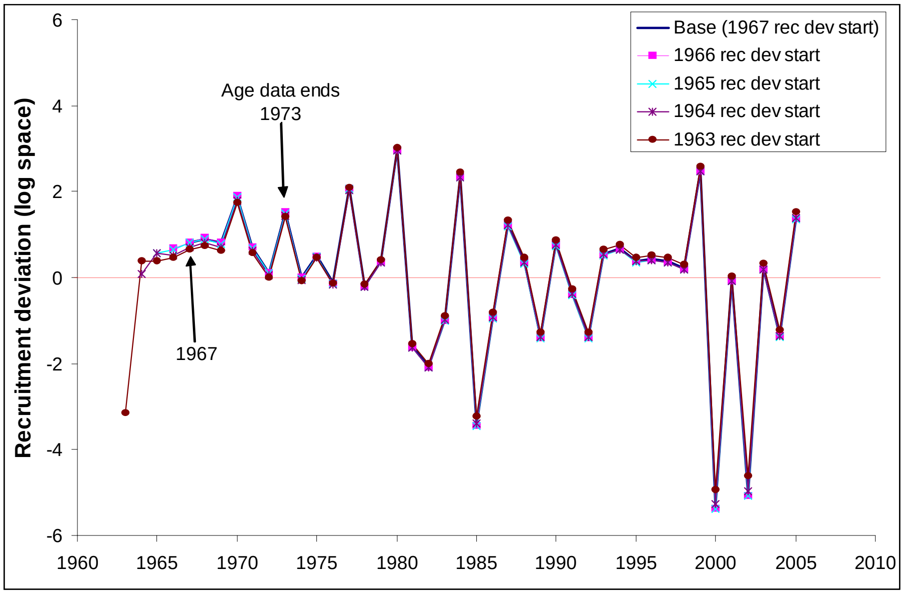

Figure 13. The imputation of recruitment deviations to the years prior to available data in an attempt to duplicate the non-equilibrium conditions expected at the start of the fishery.

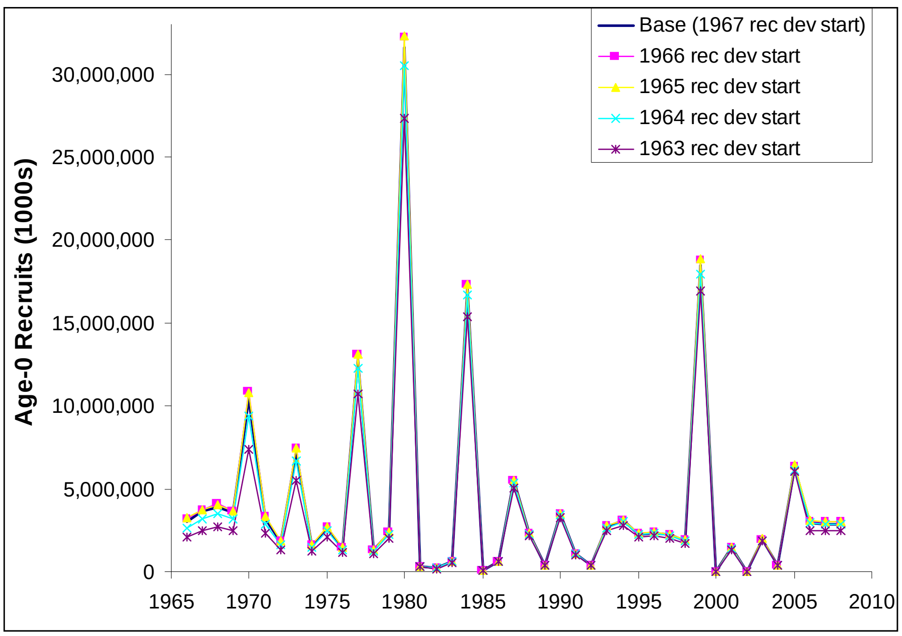

Figure 14 The impact on the predicted sequence of recruitment deviations of extending the time series of recruitment deviations back before the available data (leading to a non-equilibrium age structure in 1966 – the assumed start of the fishery).

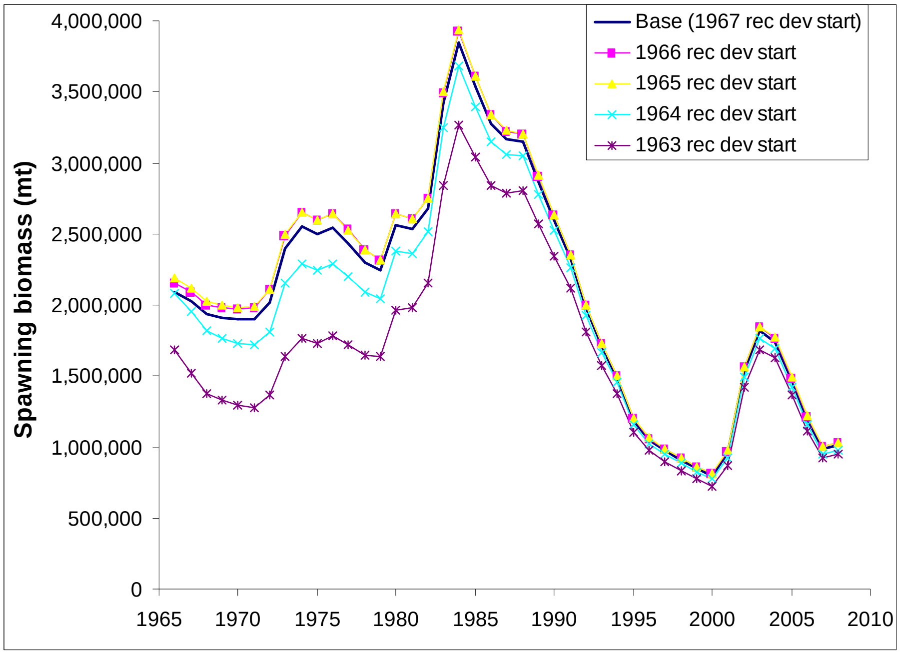

Figure 15. The impact on the predicted time series of spawning stock biomass of extending the start of recruitment deviations at the beginning of the time series.

<u>Wednesday Morning for All Analysts</u>

1. *Compute landings divided by age 2+ beginning of year biomass. Rationale: Want a consistent measure of harvest across models.*

The requested estimates of catch divide by age-2+ biomass from the three models (SS2 final base, ADAPT / VPA Run 1A, and TINSS) are shown in the middle panel of Fig. 16. The SS2 and TINSS estimates are very similar. The ADAPT / VPA estimates are generally elevated above the estimates of the other two models, which is consistent with the lower biomass estimated by the ADAPT / VPA model.

2. *Provide comparison of SSB and age-0 recruitment. Rationale: These will illustrate similarities and differences between models.*

The requested estimates of Age-0 recruitment and SSB from the three models (SS2 final base, ADAPT / VPA Run 1A, and TINSS) are shown in the top and bottom panels of Fig. 16. All three models agree on which year-classes are dominant, but the models differ in their estimates of absolute year-class strength. The SS2 and TINSS models have similar spawning biomass trajectories in the early part of the time series but diverge in recent years. The ADAPT / VPA model estimates of spawning biomass are consistently smaller than the estimates from the other two models and are considerably different for the early part of the time-series.

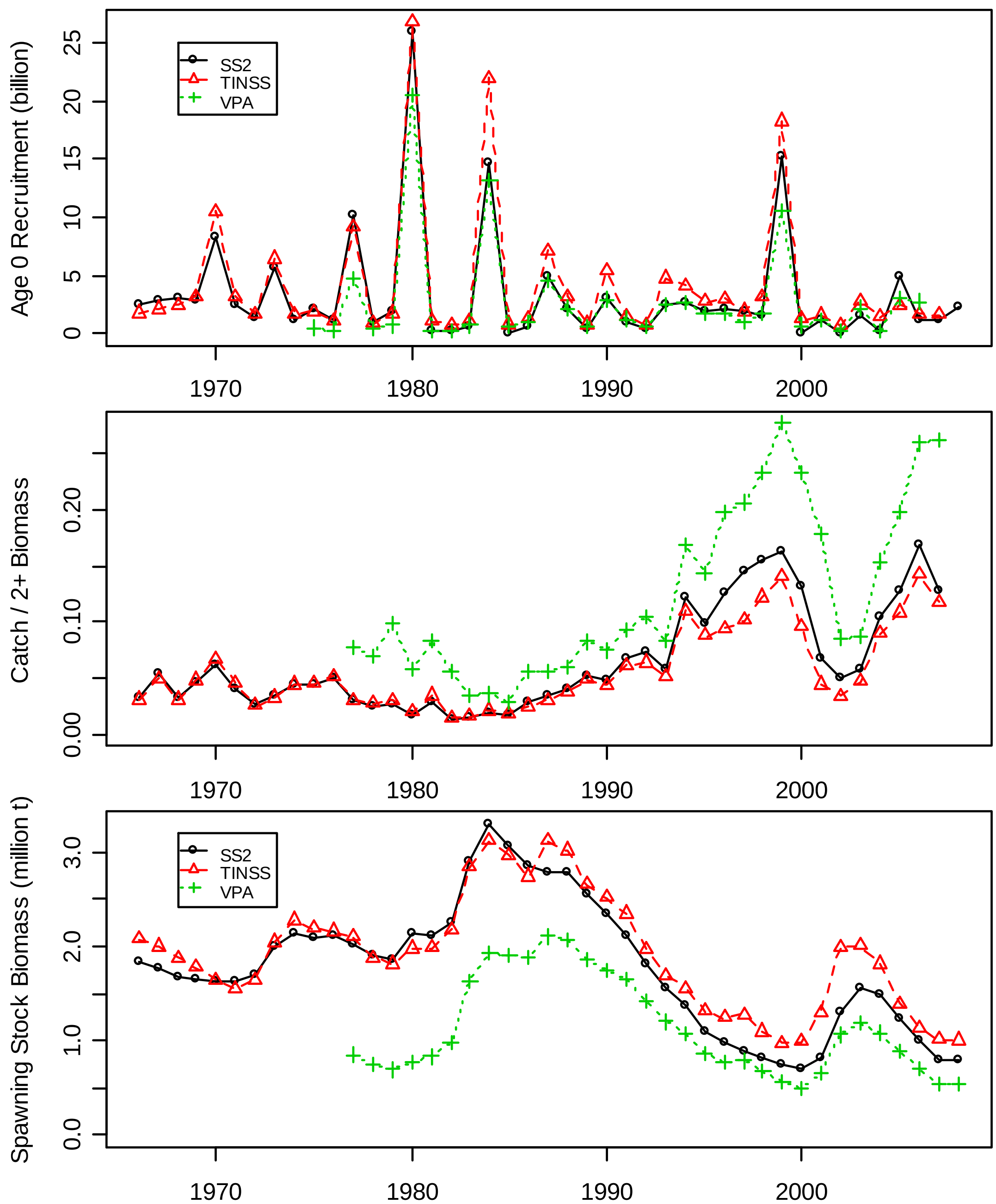

Figure 16. Comparison of estimates of Age-0 recruitment, harvest rate (catch / Age-2+ biomass), and spawning stock biomass from the SS2 base-case model, the ADAPT / VPA Run1A model, and the TINSS model.

3. *Provide one-year (2008) catch forecasts based on a B_o calculation using the earliest growth and the 40:10 rule, linear in catch. Use Fmsy and F40% where possible. Rationale: These will illustrate similarities and differences between models.*

The STAT provided the requested information, which is summarized below.

	SS2	ADAPT / VPA	TINSS
40-10 Catch in 2008	527,180	346,000	325,000

4. *Provide a comparison across models of retrospective patterns. Rationale: These comparisons will illustrate how the models respond to changes in assessment data.*

The retrospective analyses illustrated the similarities between the models. The general trend in the spawning stock biomass trajectory was approximately repeated for all models. The importance of the survey data is apparent in the shifts in the trajectory that occur following the removal of years of survey data (Fig. 17 to 19).

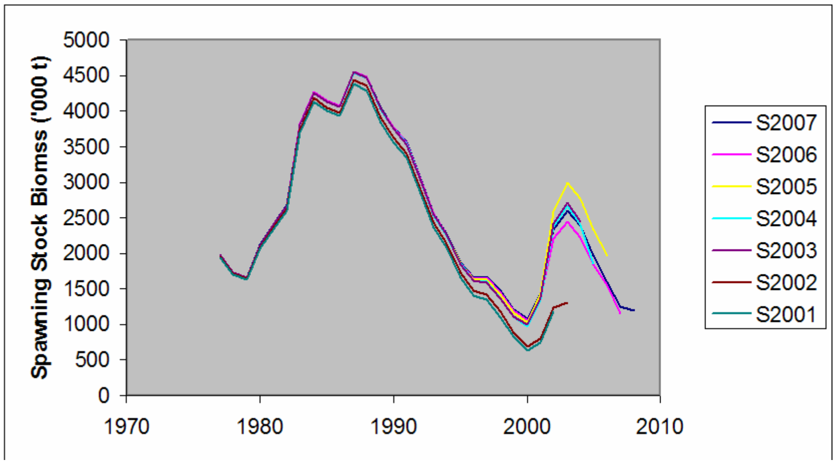

Figure 17. The retrospective analysis of Spawning Stock Biomass from the ADAPT / VPA analysis.

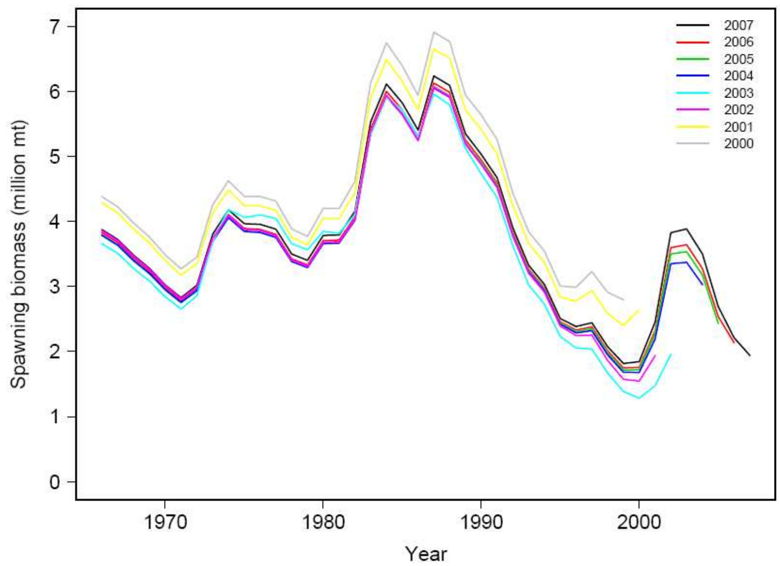

Figure 18. The retrospective analysis on Spawning Stock Biomass from the TINSS modelling.

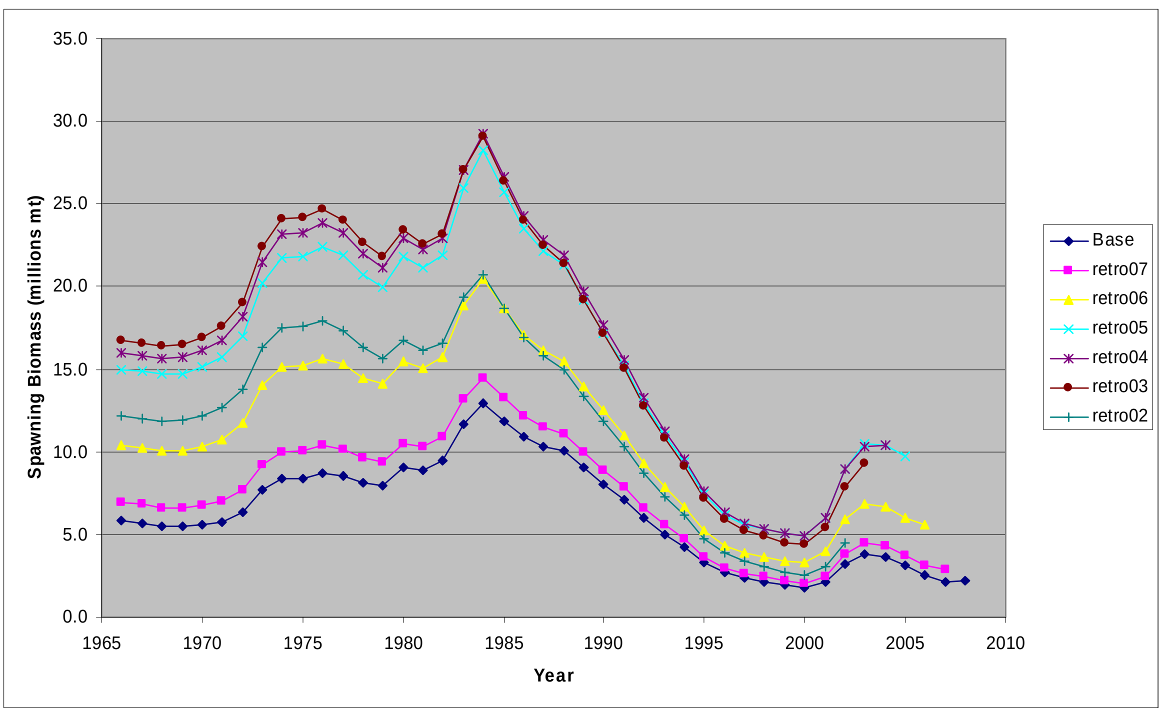

Figure 19. The retrospective analysis on Spawning Stock Biomass from the SS2 base-case model.

5. *With respect to Tues Pm request 1, try an age-dependent M. Fix young M at 0.23 and estimate old M. Rationale: The current specification for the SS2 decision table may over-state uncertainty. The new specification may fix this problem.*

The M for young fish was fixed up to age-13 and then allowed to change. Relative to this same summary information from the original assessment model (top panel of Fig. 12), the change in model specification resulted in the desired contraction in the range of values encompassed by the low and high alternatives

Derived Parameter	Base model Final selex=0.5 MLE	Asymptotic 95% CI		Alt. Low Final selex=0.7 MLE	Asymptotic 95% CI		Alt. High Final selex=0.3 MLE	Asymptotic 95% CI	
2007 Depletion	0.353	0.240	0.466	0.324	0.225	0.423	0.386	0.254	0.519
2008 Depletion	0.357	0.217	0.497	0.322	0.197	0.447	0.398	0.237	0.559
MSY	452,320	237,151	667,489	423,950	248,467	599,433	499,660	238,568	760,752
B_{MSY}	1,191,500	629,294	1,753,706	1,045,200	561,394	1,529,006	1,350,100	704,280	1,995,920
SPR_{MSY}	0.332	0.114	0.550	0.317	0.116	0.517	0.337	0.115	0.559
2008 Catch	463,510	154,144	772,876	370,290	127,132	613,448	591,290	170,008	1,012,572
Rzero (billions)	1.858	1.532	2.185	1.682	1.430	1.933	2.083	1.612	2.553
Bzero (millions, mt)	2.631	2.171	3.092	2.379	2.024	2.734	2.958	2.293	3.623

Figure 20. The effect of adding an age-dependent M to the base model configuration brought to the STAR Review..

6. *In the VPA, compute a "domed-run", with F at age 14 equal to one-half the average F at ages 7-12. Rationale: Explore the reasons for differences between ADAPT and SS2 SSB estimates, which we think is due to domed-selection.*

The effect of using an imposed dome-shaped selectivity on the VPA was to increase the apparent spawning stock biomass (Fig. 21) in such a manner as to make the VPA output much more similar to the spawning stock biomass trajectories from the SS2 and TINSS models (Fig. 22). However, the mean square residual for the asymptotic (flat) selectivity was 0.664 while it was 0.857 for the dome-shaped selectivity, indicating that the quality of the model fit declined when selectivity was dome-shaped.

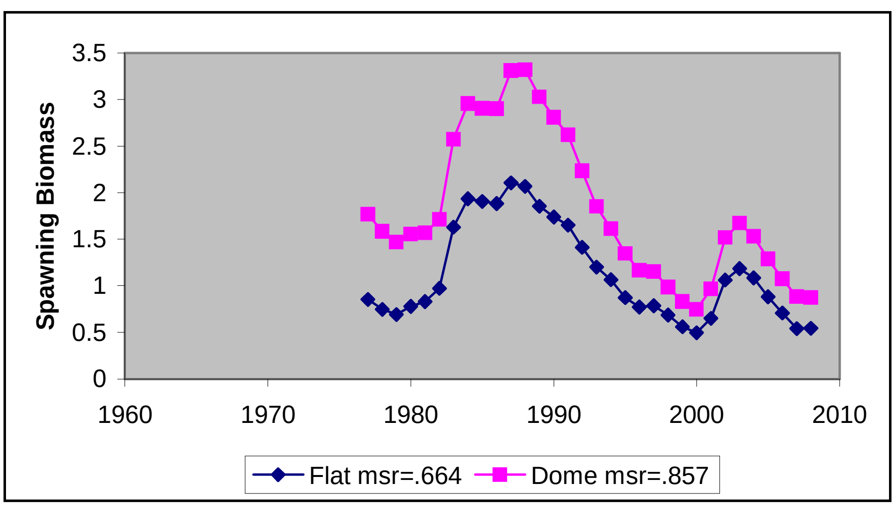

Figure 21. The effect of the spawning stock biomass trajectory of forcing the VPA to use a dome-shaped selectivity curve.

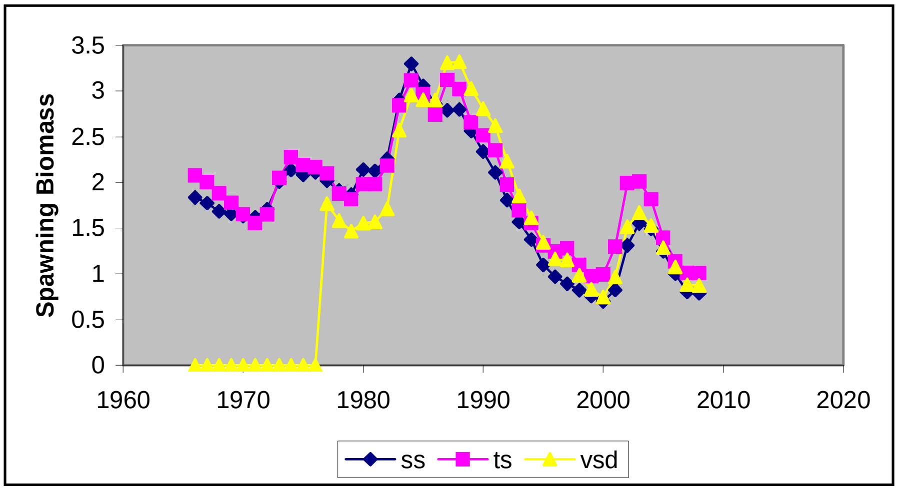

Figure 22. The effect of the spawning stock biomass trajectory of forcing the VPA to use a dome shaped selectivity curve. The VPA (vsd) is compared with the SS2 (ss) base-case and TINSS (ts) models.

<u>Wednesday. Afternoon.</u>

1. *With respect to Tues. PM, request 2, plot confidence limits and point estimates for SSB and depletion in 2008 from different recruitment deviation starting points. Rationale: Estimating the initial age distribution may affect uncertainty in the final results.*

The STAT produced a plot (Fig. 23) showing spawning biomass estimates and confidence limits for different recruitment deviation starting years. The plot indicated that uncertainty in the estimates of final biomass was not strongly affected by the assumption of an equilibrium age distribution in 1967. The STAT did not produce a similar plot for estimated depletion in 2008.

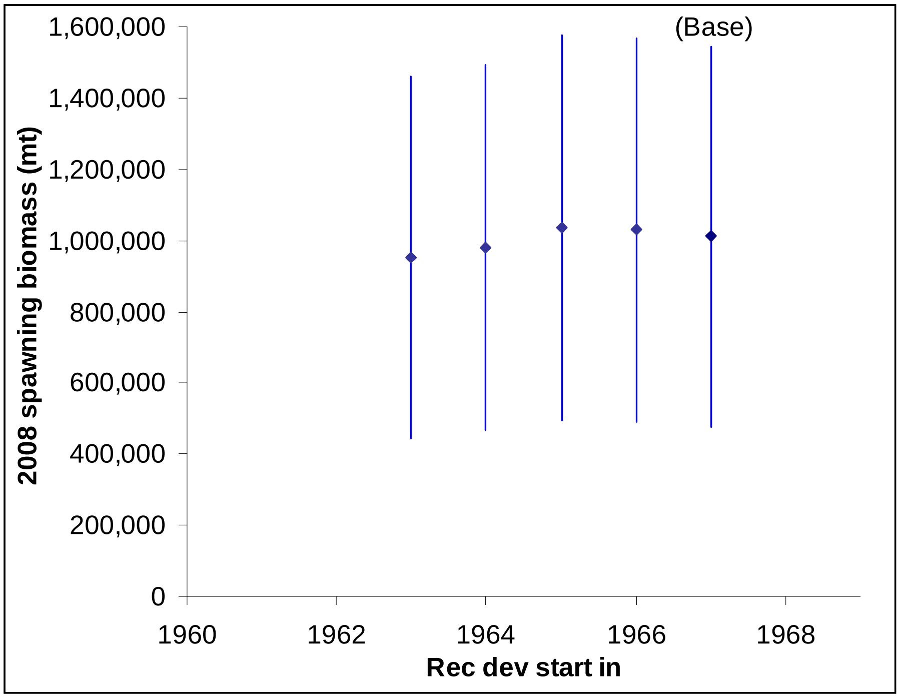

Figure 23. The impact on the estimated spawning stock depletion level in 2008 of extending the start of the recruitment deviations back to 1963.

2. *With respect to request 5, Wed. AM, do a run with final selection (selex) estimated. Rationale: If there is sufficient information to do this estimation, then this would provide a more objective basis for assigning probabilities to the SS2 model states of nature in the decision table.*

The overall effect of estimating the final selectivity parameter (selex), along with survey catchability and the natural mortality coefficient for the oldest age-class, was to broaden the uncertainty around the estimated 2008 catch (Fig. 24). Generally, the uncertainty in this final model encompassed the uncertainty expressed in the other SS2 model scenarios and in the ADAPT / VPA and TINSS models. Subject to some additional diagnostic tests, the Panel and STAT were of the opinion that the run "final selex est" would be suitable for use as a base model.

Derived Parameter	Base model Final selex=0.5 MLE	Asymptotic 95% CI		Asymptotic selex MLE	Asymptotic 95% CI		Final selex est MLE	Asymptotic 95% CI	
2007 Depletion	0.353	0.240	0.466	0.265	0.193	0.337	0.362	0.236	0.489
2008 Depletion	0.357	0.217	0.497	0.248	0.151	0.345	0.372	0.217	0.527
MSY	452,320	237,151	667,489	383,790	263,961	503,619	466,270	212,391	720,149
B_{MSY}	1,191,500	629,294	1,753,706	796,640	428,101	1,165,179	1,343,800	712,602	1,974,998
SPR_{MSY}	0.332	0.114	0.550	0.277	0.108	0.445	0.352	0.121	0.582
2008 Catch	463,510	154,144	772,876	216,180	65,131	367,229	527,180	141,707	912,653
Rzero (billions)	1.858	1.532	2.185	1.403	1.254	1.552	1.728	1.362	2.095
Bzero (millions, mt)	2.631	2.171	3.092	1.987	1.776	2.198	2.567	2.088	3.047

Figure 24. The effect of altering the assumptions in the SS2 modelling with respect to selectivity (asymptotic versus estimated final selection) given estimation of the survey catchability and natural mortality for the oldest age classes.

3. *For the SS2 base model (to be decided), provide evidence of global convergence. Rationale: to confirm convergence.*

The STAT conducted a series of runs with the proposed SS2 base model in which the initial parameter values were perturbed by random "jitter". Many of the runs failed to converge. Most of those that seemed to have converged did so with the same value of log-likelihood and M for the oldest age-class as the proposed base model (Fig. 25). None of the jittered runs produced a smaller negative log-likelihood value, which suggests that the proposed base model had fully converged to the global maximum likelihood estimates.

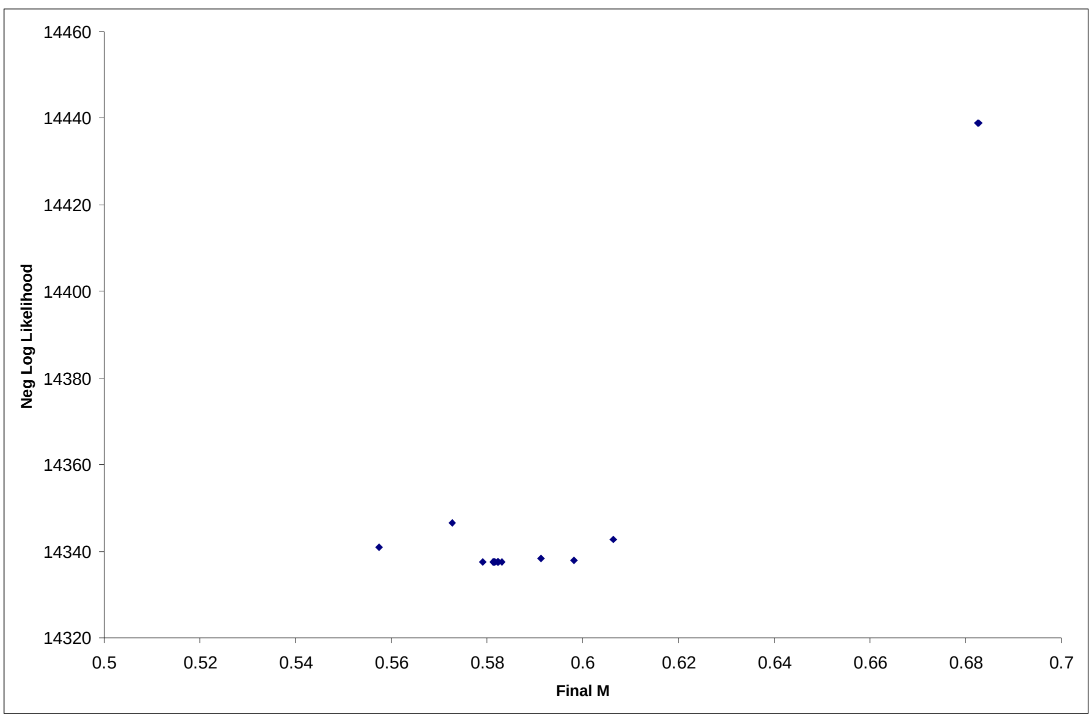

Figure 25. Demonstration of global convergence of the SS2 base model using randomly perturbed initial parameter values.

4. *Identify the change in fit in specific data (i.e. age+year+fishery) components, between the "final-selex est" model and the initial base model with M=0.23 and selex=0.5. The fit to the acoustic survey index appears to be worse in the final-selex est model compared to the base model. If time permits, compare final-selex est with the M-estimated (selex=0.5) model. Rationale: Better fits to a single or only a few components is less convincing from a robustness perspective than improvements in fits that are broadly distributed across most data.*

The adopted SS2 base case, where the Age-15+ natural mortality, survey *q*, and selectivity are estimated, improves the fit over the selex = 0.5 model by about 258 negative log-likelihood units (a highly significant change, Fig. 26). Most of that change is a result of changes in the fit to the age composition data. In particular the fit is especially improved with the US fishery age composition data and the acoustic trawl survey age composition data. However, for reasons that are not presently clear, the age composition data for the Canadian fishery declined in their quality of fit. While it is the case that these data tend to be in opposition to each other, it is not clear why this change in the fitting strategy should adversely influence the fit to the Canadian fishery age composition data.

	selex = .5	free M selex	difference
LIKELIHOOD	14595.4	14337.6	-257.8
indices	-6.86409	-2.60188	4.26221
length_comps	1883.36	1892.21	8.85
age_comps	12661.7	12400.8	-260.9
Recruitment	55.5339	43.4585	-12.0754
us lgt	1241.1	1244.6	3.5
can lgt	533.138	530.324	-2.814
surv lgt	109.117	117.288	8.171
us age	8218.97	8070.83	-148.14
can age	2757.38	2800.46	43.08
surv age	1685.38	1529.5	-155.88
survey	-6.86409	-2.60188	4.26221

Figure 26. Changes in negative log-likelihood resulting from model configuration changes from the preliminary base case SS2 model in the original assessment document.

5. *For the final-selex est SS2 model, provide estimates of cryptic biomass. Rationale: We want to establish how much of the older spawning biomass is unobserved by the survey.*

Cryptic biomass is predicted to make up a variable amount of the stock at different times in the history of the fisheries (Fig. 27). Once the 1980 and 1984 year classes began to join the exploited stock, the cryptic biomass attained levels of more than 500,000 tonnes. Currently the proportion of cryptic biomass is at a low level, being less than 5%, but this is expected to increase because the 1999 year-class is just entering the cryptic phase (age-9+).

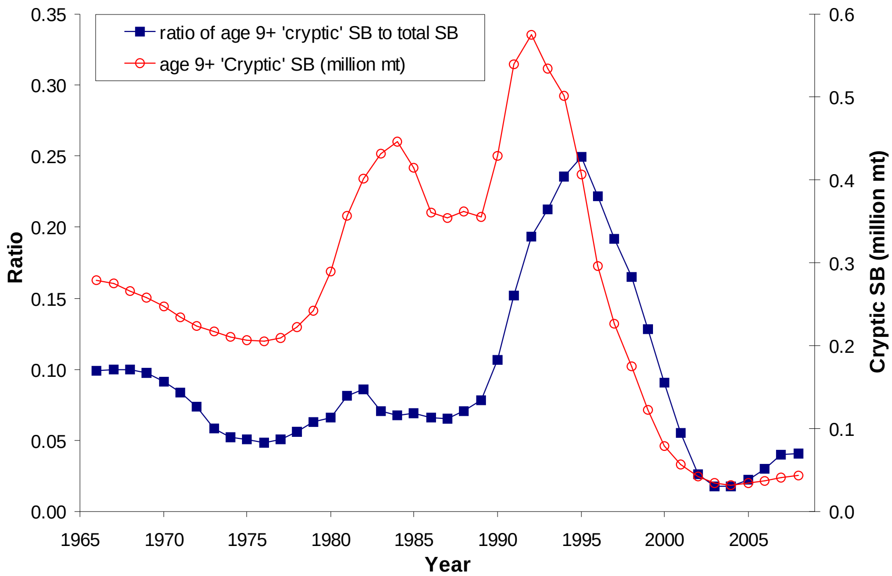

Figure 27. The ratio of cryptic biomass (aged 9+) and total biomass expressed as a ratio and as an absolute measure of cryptic biomass. The impact of the 1980 and 1984 year-classes is apparent while the effect of the 1999 year-class has yet to appear.

Description of base model

Three distinct stock assessment models were brought to the STAR Panel meeting and were carefully reviewed by the Panel. While all three models worked from the same basic set of data, they used different approaches for aggregating the data and made different structural assumptions to model the data. The STAR Panel chose to use the SS2 modeling platform for the base model and decision table because the SS2 model made the most comprehensive use of the available data and provided a more flexible tool for evaluating different plausible sets of assumptions regarding underlying uncertainties in the data (e.g., relative error among different data sources, imprecision in age-readings) and in the model structure (e.g., domed versus asymptotic selection, time-varying selection, age-dependent natural mortality). Further, results from the SS2 model configuration chosen for the base model encompassed the range of results produced by the other model platforms (ADAPT / VPA and TINSS). Requested model runs for the ADAPT / VPA and TINSS models demonstrated that these models were able to produce spawning stock biomass trajectories that were very similar to those produced by the SS2 model.

The SS2 model configuration selected for the 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 primary tuning index.

- Age-reading imprecision was incorporated, but there were insufficient data to estimate ageing bias.
- Time-varying growth parameters were estimated.
- A Beverton and Holt recruitment curve was estimated using an assumed beta-prior probability distribution for the steepness parameter and a variability parameter (sigma-R) value of 1.13, with annual recruitment deviations estimated for 1967 to 2005.
- Fishery selection was time-blocked to accommodate apparent targeting of strong year-classes and structural changes in the fisheries (four independent blocks for each of the two fisheries).
- 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.
- The natural mortality coefficient was fixed at 0.23^{-yr} for ages 0 to 13, and then was allowed to ramp to higher (or lower) values for age-14 and the age-15+ group.

<u>Alternative models used to bracket uncertainty.</u>

The alternative models for constructing the decision table were derived from the posterior distribution of the base model rather than from alternative model formulations. As previously noted, however, numerous other model configurations were explored during the STAR Panel review, including formulations based on the ADAPT / VPA and TINSS models. The approximate confidence intervals surrounding the SS2 base model estimates generally encompassed the range of values estimated by other reasonable model forms and configurations.

Technical merits / deficiencies in the assessment

In past assessments for this stock the catchability coefficient for the acoustic survey (survey-Q) was the major dimension of uncertainty. Past STAR Panels have recommended bracketing uncertainty in decision tables by using one or more fixed values of survey-Q. Discussion during the current STAR Panel review focused primarily on the issue of the form of the selection curves: domed versus asymptotic. The ADAPT / VPA and TINSS models assumed that selection curves for the two fisheries and the survey should all be asymptotic. The SS2 model, in contrast, used a formulation for selection that allowed the data to indicate its preference for domed versus asymptotic selection; that is, SS2 estimated the amount of dome.

The SS2 base model and the ADAPT / VPA and TINSS models made the strong but unverified assumption that the weight-at-length (or age) relationship and the maturity-at-length (or age) relationship have been time-invariant, despite radical changes in stock biomass and cohort strength that could affect these key biological components.

<u>The Stock Synthesis model</u>

- The SS2 model as formulated in the current assessment allowed the STAT to conduct a very full exploration of how key parameters (natural mortality, survey catchability, shape of the selection curves) influenced goodness-of-fit to the data.

- Despite the very flexible modeling structure used, the various likelihood profiles indicated clear tension between the US versus Canadian age-composition data. The reason for this tension is unclear but probably indicates one or more structural problems with the current model formulation. Possible issues include accounting for spatially related stock dynamics, the need to distinguish the genders, and having the selection processes be explicitly length based.
- The STAT explored the effects of assuming an initial equilibrium age-composition and showed that the assumption had little impact on the uncertainty of the estimates of biomass or depletion levels, but this result was very counter-intuitive.

<u>The ADAPT / VPA model</u>

- The ADAPT / VPA model, relative to the other two models, provided the most flexible approach to modeling fishery selection. It did not assume any particular form for selection except at the oldest true age (14) in the model. However, the model was based on the assumption that acoustic survey catchability at ages 13 and 14 were equal.
- The model did not estimate fishing mortality values for the age-15+ fish. As a consequence the issue of reduced selection for the terminal age-class was not investigated.
- Results from a VPA are subject to error due to selection of the so-called terminal fishing mortality coefficients. The influence of this error dissipates as the estimates of stock size propagate to younger ages, but a high cumulative fishing mortality is required to produce rapid dissipation. Because relatively low fishing mortality rates have been applied to the Pacific hake stock, especially prior to 1993, it seems likely that the estimates of abundance and biomass may still be tainted by error from the terminal fishing mortality values.

<u>The TINSS model</u>

- The approach of formulating the model in terms of the management variables MSY and F(MSY) seems very sensible and preferable to having these variables be derived from other less meaningful parameters (e.g., steepness).
- The model provided a simple representation of the dynamic processes that was uncluttered by nuisance parameters.
- The model results presented to the STAR Panel did not provide much evidence that the model's simple structure provided an adequate representation of the available data. For example, residual plots from the model fits to the age-composition data showed evidence of systematic lack of fit to the youngest and oldest ages, consistent with the notion that the fit could be improved by allowing domed selection, but the magnitude of the improvement was not evaluated.

<u>Recommendations for remedies</u>

- The importance of possible structural problems in the SS2 model could be explored by constructing more complex models that incorporate processes based on length, gender, and space. However, overly complex models may not produce reliable results on which to base management decisions.

- The VPA approach is appealing because of its simplicity and transparency, and it provides a useful contrast to integrated analysis approaches such as SS2 and TINSS. Use of alternative VPA derivatives, such as XSA or other approaches, might provide a useful contrast to the ADAPT approach.
- The TINSS model could usefully be expanded to include other processes affecting the dynamics of the stock (e.g., time-varying selection) and the available data (e.g., ageing error). It would useful to include measures such as AIC for formally evaluating model parsimony.
- A full Management Strategy Evaluation would permit the formal evaluation of the relative value of each modeling approach (e.g., SS2, VPA, TINSS) for the production of management advice. The Management Strategy Evaluation approach is internationally accepted as the best way of evaluating the performance of stock assessment methods and their interplay with management decisions.

Areas of disagreement regarding STAR Panel recommendations

Among STAR Panel members

There were none

Concerns raised by GAP, GMT, and DFO advisors

There were none

Between the STAR Panel and STAT Team

The analysts responsible for preparing the ADAPT / VPA model disagreed with the STAR Panel's recommendation to use the SS2 model for developing a base model and decision table. Their minority report is included as an appendix.

Unresolved problems and major uncertainties.

Data problems and uncertainties

- Although the SS2 model included age-reading imprecision, the age-composition data are assumed to be unbiased, but the validity of this assumption has not been evaluated.
- There continues to be considerable uncertainty regarding the acoustic target strength of Pacific hake. This uncertainty may be consistent with the variability in survey-Q implied by the three models, but this consistency should be established to verify that the models have correctly incorporated the uncertainty associated with the acoustic survey.
- It was disconcerting to learn that the acoustic survey biomass estimates are based on very sparse sampling to establish the species, size and age composition of the acoustic signs. While it is accepted that this is typical of acoustic surveys, it would have been reassuring to have been shown some evidence that a single short tow from a long acoustic transect provides a reliable and unbiased estimate of the species, size, and age composition of identified fish aggregations.

Modeling problems and uncertainties

- The SS2 and TINSS models both estimated the acoustic survey-Q to be less than 1, but the ADAPT / VPA model estimated the survey-Q to be greater than 1 for some older ages. The mechanisms that account for the discrepancies of survey-Q from 1 need to be understood.
- It is unclear what mechanisms are responsible for the apparent domed selection in the fisheries and survey that is implied by the SS2 model.
- Spatial changes in fishery operations have the potential to cause high inter-annual variation in fishery selection. The SS2 model uses four time-blocks to accommodate changes in fishery selection but this may be too rigid a structure. The consequences of imposing an overly rigid selection structure are unknown.
- The issue of an appropriate objective method for iteratively re-weighting observed data remains unresolved. The approach taken to develop the SS2 base model seems reasonable, but we have no basis for presuming that the approach produced a correct balance of the uncertainties among the different data sources.

Management, data, or fishery issues raised by the GMT, GAP, or DFO advisors.

Discussions during the STAR Panel review identified several important risk factors that, in the interest of being precautionary, should be taken into consideration when setting catch quotas for 2008. For several years the fishery has been very dependent on the exceptionally strong 1999 year-class; this year-class is now diminishing in biomass. None of the more recent year-classes show evidence of being as strong as the 1999 year-class. Successful recruitment in the future depends on leaving the stock with adequate spawning biomass. Despite catches being constant or even declining, fishing mortality in recent years has been increasing and is now estimated to be at higher levels than it was during most of the history of the fisheries.

The standard decision table developed for the Council does not fully address the Canadian Request for Catch Advice which asked how the expected trajectory of stock biomass would be affected by a range of annual catch quotas. Consequently, the Panel asked the STAT to develop a risk plot with the SS2 base model showing the effect of different levels of catch (Fig. 28, below; Fig. 58 in the SS2 assessment document).

The Canadian Request for Catch Advice also asked for an analysis of appropriate biological reference points for the stock. Specific analyses to address this request were not examined by the STAR Panel.

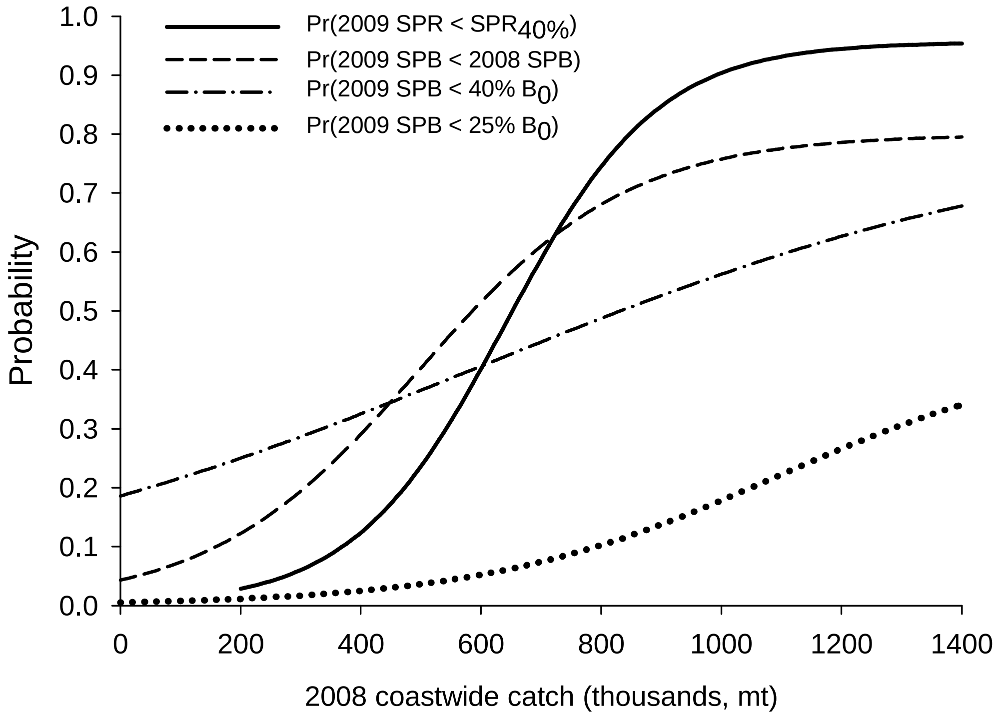

Figure 58. Risk profiles showing probability of the 2009 SPR rate being less than target SPR40% and 2009 spawning biomass being less than 25% Bzero for a suite of different coastwide catches in 2008.

Prioritized recommendations for future research and data collection

The Panel notes that the 2007 STAR Panel presented a comprehensive review of recommendations from past STAR Panels. Many of these recommendations still apply, but they are not reiterated here. The recommendations below resulted from discussions during the 2008 STAR Panel review and subsequent email exchanges.

1. The Panel recommends that a Management Strategy Evaluation approach be used to evaluate whether the current 40-10 harvest control rule is sufficient to produce the management advice necessary to ensure the sustainable use of the Pacific hake stock with its dramatically episodic recruitment. The 40-10 rule assumes that simply reducing catches in a linear fashion as stock biomass declines will be sufficient to guide the fishery back towards the target spawning biomass level. However, with the fishery being dependent upon a single declining cohort just reducing the catch may achieve the status quo but it rebuilding will not occur without new recruitment.

2. Related to Recommendation 1, the operating model developed for the Management Strategy Evaluation should evaluate how well the different assessment models recapture 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 SS2.

3. Female Pacific hake grow differently than male Pacific hake and 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). The Panel recommends

that future assessment models explore the need for including both gender- and length-based selection into the dynamics.

4. The inclusion of ageing error was found to be influential on the model fit in the SS2 model. However, issues with ageing still remain. Further ageing error analyses are required, especially focused on estimating any bias in the ageing. It will be important to conduct a cross-validation of ageing error from the different laboratories conducting the ageing. It is especially important to include otoliths that were read by AFSC staff.
5. In light of current acoustic survey information, re-evaluate treatment / adjustment of pre-1995 acoustic survey data and index values. For example, compare the biomass index implied by the area covered by the pre-1995 surveys 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.
6. 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.
7. There should be exploration of possible environmental effects on recruitment and the acoustic survey.
8. There should be further investigation and resolution of possible under-reporting of foreign catch.

List of STAR Panelists

- David Sampson, Panel Chair and Scientific and Statistical Committee representative
- Malcolm Haddon, Panel Reviewer from the Center for Independent Experts
- Noel Cadigan, Panel Review from Fisheries and Oceans Canada
- Jeff Fargo, Advisor from Fisheries and Oceans Canada
- Dan Waldeck, Advisor from the Groundfish Advisory Panel
- John Wallace, Advisor from the Groundfish Management Team

List of STAT Members

- Tom Helser, NWFSC / NMFS, lead author of the SS2 assessment
- Ian Stewart, NWFSC / NMFS, co-author of the SS2 assessment
- Owen Hamel, NWFSC / NMFS, co-author of the SS2 assessment
- Alan Sinclair, Fisheries and Oceans Canada, lead author of the ADAPT / VPA assessment
- Chris Grandin, Fisheries and Oceans Canada, co-author of the ADAPT / VPA assessment
- Steve Martell, University of British Columbia, author of the TINSS assessment

References

Helser, T.E., Stewart, I.J., and O.S. Hamel (2008) Stock assessment of Pacific Hake (Whiting) in U.S. and Canadian waters in 2008. 129 p.

Martell, S. (2008) Assessment and management advice for Pacific hake in U.S. and Canadian waters in 2008. 47 p.

Sinclair, A.F. and C.J. Grandin (2008) Canadian fishery distribution, Index analysis, and Virtual Population Analysis of Pacific Hake, 2008. 59p.