

Joint U.S.-Canada Scientific Review Group Report

Morris J Wosk Centre for Dialogue
580 West Hastings Street
Vancouver, BC
February 19-22, 2013

Scientific Review Group (SRG) Members

Richard Methot, co-chair, NOAA, NMFS, NOAA
Greg Workman, co-chair, PBS, DFO
Michael Prager, NMFS, NOAA, retired
Robyn Forrest, PBS, DFO (Acting)
David Sampson, OSU, AP nominee

Joint Technical Committee (JTC) Members

Allan Hicks, NWFSC, NMFS, NOAA
Nathan Taylor, PBS, DFO
Chris Grandin, PBS, DFO
Ian Taylor, NWFSC, NMFS, NOAA
Sean Cox, SFU, AP nominee

Pacific Hake / Whiting Acoustic Survey Team

Dezhang Chu, NWFSC, NMFS, NOAA
Rebecca Thomas, NWFSC, NMFS, NOAA
Lawrence Hufnagle, NWFSC, NMFS, NOAA
Stephane Gauthier, IOS, DFO

Advisory Panel Advisors to the SRG

Mike Buston, (AP Nominee - CDN)
Rod Moore, (AP Nominee - USA)

Introduction

Under the authority of the Agreement Between The Government of The United States of America and The Government of Canada on Pacific Hake/Whiting (here after referred to as “The Treaty”), the Scientific Review Group (SRG) met in Vancouver, British Columbia, 19 to 22 February, 2013, to review a draft stock assessment document prepared by the Canada/US Joint Technical Committee (JTC). With ratification of “the Treaty” by both parties, the SRG has attempted to operate fully in accordance with “the Treaty.” The SRG based its interim terms of reference on the language of the US-Canada Pacific Whiting Agreement and on the Pacific Fishery Management Council’s Stock Assessment and Review (STAR) terms of reference. The interim terms of reference have now been approved by the Joint Management Committee (JMC) as the formal Terms of Reference for the SRG. The SRG was comprised of two US, two Canadian and two additional members designated by the JMC based on recommendations from the Advisory Panel (AP). The JMC also appointed two industry advisors to assist the SRG in its deliberations. The SRG has amended its Terms of Reference to reflect these recent appointments

The meeting convened at 9AM Tuesday, February 19, 2013, Greg Workman (meeting chair) welcomed attendees and after a round of introductions reviewed the agenda, SRG Terms of Reference and assigned reporting duties.

Summary Conclusions

1. The 2012 acoustic-trawl survey was successfully completed. The US and Canadian teams that conducted this survey are to be applauded for their effort, and we acknowledge the strain that conducting this survey in back-to-back years placed on their programs. For the first time, the survey was conducted in conjunction with the acoustic-trawl survey for sardine, and we recommend that a review of the merits of this approach be conducted.
2. The survey result was a relative biomass of 1.381 Mmt, a substantial increase from the 2011 survey biomass of 0.521 Mmt. The survey and the fishery were dominated by age 2 (63.7% survey; 34.6% fishery) and 4 (16.1% survey; 34.5%) year old fish from the 2010 and 2008 year classes, with differences due to the different selectivity of young fish to the survey vs. the fishery. The survey has verified the strength of the 2008 year class and finds that the 2010 year class seems even stronger, but it now adds uncertainty to the forecast.

3. The standard assessment model was used to analyze the time series of data extended through 2012. The median estimated female spawning biomass is 1.503 Mmt at the beginning of 2013 and is expected to be stable to increasing through 2015 due to an expected very large 2010 year class and the above average 2008 year class. This level of estimated spawning biomass has not been seen since 1993. The consistency between the 2013 assessment result and the 2012 assessment result provides strong retrospective evidence that 2009 survey and the 2011 assessment overestimated the stock abundance at that time. Unlike 2011, there is now agreement between the most recent acoustic survey and commercial fishery age composition data as well as the most recent acoustic survey biomass index. This alignment of data from separate sources engenders greater confidence in the 2013 assessment result
4. Although utilization of the total quota has been around (85%) in recent years, the retrospective estimates of fishing intensity hovered around the target from 2006 to 2011. This implies that for at least the most recent period the assessment model has overestimated the total allowable catch when applying the default harvest policy to overly optimistic estimates of stock status.
5. The model forecasts that a catch of 626,364 mt in 2013 and 715,041 mt in 2014 would have equal probabilities of producing a fishing intensity that is greater than or less than the F40% default harvest rate. The two young age groups that are supporting the fishery will be ages 3 and 5 during the 2013 fishing season, this is near their peak biomass and potential maximum contribution to lifetime yield. However, there is still considerable risk in fishing them too hard until their strength is verified, particularly the 2010 year class, which is still very young and thus not yet well characterized .
6. As noted, there is considerable uncertainty in the 2010 year class strength because it has only been observed for one year. Some recent year classes with strong occurrence at age 2 were subsequently downgraded as years of data were added (notably, the recent 2008 year class). A conservative estimate of the 2010 year class strength (using only the lower 10% of the model estimated recruitment) reduces the strength from a median estimate of 11.6 billion recruits (a near record size) to 6.9 billion recruits, which is still large and near the size of the 1970 and 1999 recruitments. If this is the true state of nature, a catch of 336,000 mt would allow the stock to maintain its biomass level, at least approximately, from 2013 to 2014. The SRG requests that the JTC document this scenario in its report.

7. The SRG and the JTC recommend a range of 336,000 – 626,000 mt as (available) harvest level in 2013. The upper end would implement the default harvest policy and would allow some continued biomass growth into 2014 if the current assessment result is accurate. The lower level would still not exceed the harvest policy even if the 2010 year class is only 51% of its current estimate. If the lower 2010 recruitment strength is realized, a 2013 harvest of 626,000 mt has a 91% chance of reducing biomass in 2014. Yield would also decline in this case.
8. The SRG finds that yield-per recruit and spawning biomass per recruit analyses are informative and should be included in the annual JTC report.
9. The JTC has made an excellent start in development of a Management Strategy Evaluation to investigate the performance of the hake/whiting assessment and management system. They have achieved a proof-of-concept, but the SRG finds that some structural concerns preclude endorsing current results as a base MSE from which conclusions could be drawn. Thus it is premature to use current results to indicate the merits of annual surveys versus biennial surveys. The next stage of MSE development will benefit from inputs from all parties.

Discussion on 2012 Coastwide Acoustic Survey Result

Dr. Rebecca Thomas presented an overview of the 2012 combined Hake/Sardine (SaKe) acoustic trawl survey. This was the first effort at a combined survey for both species; the need for a combined survey was the lack of vessel time to conduct both surveys independently. The NOAA Ship *BELL M. SHIMADA* performed daytime acoustics for hake and night time acoustics and trawling for sardine while an industry vessel the *FV FORUM STAR* volunteered as the daytime catcher vessel for hake verification tows. The CCGS *W.E. RICKER* conducted daytime acoustics and trawling in the Canadian zone for Pacific hake only. All three vessels were equipped with scientific sounders and calibrated with the same technical approach using standard targets (38.1-mm tungsten carbide with 6% cobalt, and 64-mm copper spheres). Survey activities began on June 27th aboard the *SHIMADA* and concluded on September 7th aboard the *RICKER*. One hundred and seventeen transects were completed; deviations from previous year's survey designs included sounding in shallower waters, extending track lines further offshore, re-orientating transects off Vancouver island to make them parallel to lines of latitude, some zigzag transects in the north, and dropping a total of 19 transects due to weather, mechanical breakdowns, and reorienting of transects in Canadian waters. Difficulties with the joint SaKe survey included spatial and temporal discontinuity between the *SHIMADA* and the *FORUM STAR*, sounding at a slower speed, and having to use a

fixed ping rate which occasionally resulted in false bottom readings that sometimes overlaid hake distributions.

The survey's estimate of age 2+ hake biomass was 1.381 million tonnes, of which approximately 92% (1.261×10^6 tonnes) was found in US waters. Most biomass was observed between San Francisco and Astoria with dense aggregations off Crescent Beach and from Newport to Astoria.

Age one fish were observed infrequently in trawls in the south but not seen with the acoustics thus it was not possible to determine whether or not the survey has reached the southern limit of the age one hake distribution. As is typical, larger and older fish were encountered as the survey progressed northward. Age 2 fish were distributed between 36 and 46 degree N latitude, age 3+ fish were found north of 40 degrees N and there was a noticeable shift to "older" fish at 50-51 degrees N. The 2010 year class (age 2 fish) dominated the catches. Approximately 700,000 of the estimated 1.381 million tonnes were age 2 fish from the 2010 year class.

When all target-verification trawls are grouped by nation approximately 91% of the catch by weight in US waters was hake, while in Canadian waters 41% of the catch was hake. These overall percentages may be misleading because some tows are intended as target verification and some are to measure the size composition of hake. In US waters yellowtail rockfish and Pacific herring were the main by-catch species, in Canada, Pacific Ocean perch, yellowmouth rockfish, Pacific herring and walleye pollock accounted for approximately 45% of the catch.

The FORUM STAR occasionally had difficulty keeping up with the SHIMADA and sometimes could not immediately fish the acoustic targets observed by the SHIMADA leading to questions about the reliability of target verification tows when the FORUM STAR's tows were more than a few hours later. A sample of the echograms from the FORUM STAR were compared to those of the SHIMADA when there was a lag between sounding and fishing and in all cases the lead acoustician was reasonably confident that the echograms represented the same body of fish. It was also noted that hake of age-1 and age-2 tend to be partitioned by latitude and remain partitioned into single age class aggregations, which was confirmed through fishing experiments in 2010. It appears, therefore, that there was little risk of catching the wrong species or age class even when there was a lag between sounding and fishing. The SRG asked if the presumed species or age class of an aggregation was recorded prior to fishing to assess the acoustician's ability to correctly identify aggregations. Similarly, the SRG asked whether the purpose of each fishing event was recorded (e.g., species ID or length/age of hake). Currently, neither type of recording is done consistently but the acoustic team acknowledged that both could easily be done in future surveys.

The implications of not adequately sampling the acoustic sign with verification tows were discussed and while it was acknowledged that the number of desired tows will always be limited by logistic constraints, the acousticians felt they had done enough tows to feel confident in their interpretation of images on the sounder most of the time. The SRG and Acoustic team discussed the process by which acousticians decide to conduct target verification tows. While the process is fluid, the acoustic team again reported that they felt 80-90% confident in their visual interpretation of the acoustic sign.

The effect of operational deviations from survey design was investigated by the acoustic team using simulation. The loss of transects due to weather or other complications was assessed by randomly removing up to 10 transects from the realized survey and re-computing biomass. The resultant changes were negligible, < 1% in biomass and ~0.2% change in CV. The team also investigated the implication of not sounding to the southern and western extents of the hake distributions by simulating a one to 12 nautical mile extension of transects that still had hake at their western extent. This resulted in significant increases in the total estimate of hake biomass, with extension of 12 nautical miles resulting in an increase of ~ 35%, but this is an extreme upper end to the amount of possible underestimation.

There was discussion of the implications of surveying a stock of fish that is migrating northward while the survey also tracks south to north, e.g. a Doppler effect. This could potentially result in seeing the same schools of hake more than once during the survey, which would lead to positive bias in the biomass estimate. The acoustic team responded that the survey moves fast enough for this not to be a concern. The SRG accepted this response but suggest that, if time is available, this question could be examined through a straightforward simulation to give a more objective answer.

Given the difficulties of conducting a combined Hake/Sardine survey and the compromises made to accomplish it with the limited resources available, the SRG asked if future surveys would be conducted in the same manner. The acoustic team reported they would be doing a SaKe survey again in 2013 but would be using the SHIMADA to do both the fishing and sounding thereby reducing one source of variance in the acoustic biomass estimate. The SRG sought clarification of whether a variance component due to the joint survey difficulties was included in the reported variance of the survey biomass estimate. The acoustic team confirmed that it was not, but considered it to be a small source of variance compared to the main sources of variance in the biomass estimate (i.e., from the kriging process), and the currently unmeasured variance due to the target verifications.

Discussion on Acoustic/Trawl Survey

Dr. Dezhang Chu continued with a detailed presentation of the biomass estimation process for aged 2+ Pacific hake from the 2012 Joint SaKe survey using the software EchoPro. Dr Chu first presented a schematic of the process broken down into three elements, data preparation, geostatistical analysis and abundance or biomass estimation. Data preparation involves combining biological trawl data (length, weigh, sex, age) with a target strength model and measures of echo integration data (Nautical Area Scattering Coefficient (NASC) or S_A) within a stratification scheme to generate estimates of biomass density, abundance density, or NASC at 0.5 NM intervals along transects. Geostatistics are used to assess the degree of spatial correlation within the data and account for it in the estimates of biomass and variance that are interpolated/extrapolated using kriging. The output from the kriging software is a grid of biomass density, abundance density, or NASC within the survey area mask with a resolution of 0.25 NM. The grid is then converted to numbers at age or length by sex using the proportions male or female at length or age for a given stratum, the area of the stratum and the cumulative density for the stratum. The kriging method is an improvement on methods used previously in that it accounts for spatial correlation and calculates the estimation error associated with interpolation/extrapolation. The approach does not account for other sources of uncertainty such as uncertainty in target strength, target classification, or survey extent relative to stock distribution.

Dr Chu reported an acoustic biomass estimate of 1.381 million metric tonnes with 91.3% of the biomass estimated to be in US waters; 8.7% in Canadian waters; an overall estimated CV of 4.75%; and an estimated CV in the Canadian zone of 14.27%. The CVs and reported correlation lengths in the analysis suggest hake were less patchy (i.e., had larger aggregation sizes) in 2012 than in 2011, at least in US waters.

Dr Chu undertook six different sensitivity analyses looking at the effects of: stratification scheme chosen; grid cell resolution; areal coverage; choice of kriged variables; choice of semi-variogram model; and the kriging parameter values selected. He found that none alone had a significant effect on the estimate of total biomass generated.

The SRG entered into a lengthy discussion about the target strength (TS) model chosen to convert acoustic backscatter into biomass. The acoustic team reported that they were using a published and well-established TS relationship for Pacific hake (*Merluccius productus*) (Traynor 1996, ICES JMS 53: 253-258). When asked if the relationship was known to hold true even for small fish the team reported that it did and that the relationship had been evaluated down to fish as small as 26 cm and may be extrapolated down to a few centimeters (Chu et al. 2003, ICES JMS 60:508-515). The SRG was satisfied with this response and acknowledged the need expressed by the acoustic team for dedicated research vessel time to conduct target strength studies.

The SRG asked the survey team how to interpret the semi-variogram, in particular how to interpret a normalized lag length of 0.03, the team reported that it corresponded to a distance of about 30 NM along isobaths and 6 NM across isobaths (along transects), meaning the interpolation/extrapolation algorithm used the lag calculated from the distances that were normalized by roughly 1100 NM along isobaths and 210 NM across isobaths, respectively (an aspect ratio of about 5), which reflects the anisotropic distribution of Pacific hake.

The SRG suggests that in future meetings, it may be more efficient for the details of the biomass estimation process to be presented prior to presentation of survey results, as many of the questions directed towards Dr Thomas, were referred to in the following presentation by Dr Chu. In addition, while recognising the highly technical nature of the biomass estimation process, it would be beneficial to begin with a layman's overview of the steps and assumptions involved in converting survey observations into biomass estimates.

Discussion on US Pacific Hake (aka Whiting) Fishery

US fleets had an aggregate allocation of 186,037 tonnes in 2012, divided amongst the catcher processor (~78,500 tonnes), shore based (~56,900 tonnes), and tribal (~48,500 tonnes) fleets, with 28,000 re-allocated from tribal to non-tribal late in season. The fleets caught 157,260 tonnes (~84.5% of target), the shortfall being mainly due to uncaught tribal allocation, although further discussion indicated that uncaught quota is also due to a number of other complex processes including bycatch avoidance (in the US) and carry-over amounts. At-sea catches peaked in the spring and late fall (May/June, Sept/Oct), due largely to vessels leaving to participate in the Alaskan pollock fishery in mid-summer, but also due in part to the lower bycatch rates achieved in the spring and fall. The shore-based fleet displayed the opposite pattern, with a slow start in the spring and peak landings in mid-summer. In 2012, the shoreside fishery was more protracted with catches peaking in August and remaining high through to November.

Catch from catcher-processors was largely composed of four year old fish during May through August, switching to nearly 100% two year old fish in September through October. The shore based catch was dominated by four year olds, throughout most of the season. However, there was an abrupt switch to two year olds in October. This likely reflects the dominance of the incoming 2010 year class and the growth of these fish during the fishing season.

The SRG asked about the discrepancies between the TAC values in the draft stock assessment and the presentation made at the meeting. The JTC explained this as the product of a complex allocation scheme that made it difficult to know which number to

present in the assessment document: the aggregate TAC, the TAC minus carry over or the fleet allocations. The SRG notes that the JTC's practice of including tribal catch in the at sea catch in the catch tables causes confusion when trying to document percentage utilization.

Discussion on Canadian Fishery

The Canadian fleet had an allocation in 2012 of 50,345 tonnes, plus carry-over from the previous year (15,427 mt), resulting in a Canadian TAC of ~66,000 tonnes. The 2012 Canadian catch was 46,776 tonnes. The Canadian catch was dominated by four, six and seven year old fish throughout the season with measurable numbers of the 1999 cohort, as 13 year olds, still visible in the age composition data.

The shore-side and at-sea processor fisheries continued well into December. The spatial pattern of fishing was reported to be sporadic, early in the season. Fishing improved during August, September and October when approximately 75% of the catch was taken.

The JTC presented a video of the Canadian fishery since 1996 showing the shift in fleet distribution from the lower west coast of Vancouver island in the 1990's and early 2000's to Queen Charlotte Sound in the mid 2000's and back to the west coast of Vancouver island in recent year. The industry reported a significant change in fisher behaviour over the last 2-3 years wherein larger vessels with larger nets are being used to target more diffuse aggregations of hake in deep water or near bottom. They also report that fish are now targeted at the heads of marine canyons instead of over the shelf break as they had been historically.

Discussion on Data for Assessment

The JTC presented an overview of data sources incorporated into this year's assessment, which included:

- Fishery dependent;
 - Annual catch (1966-2012);
 - Age frequency compositions (1975-2012).
- Fishery independent;
 - Acoustic survey (1995-2012);
 - Biomass index;
 - Age frequency compositions.
- Externally derived inputs;
 - Weight-at-age (1975-2012, survey & fishery combined);

- Ageing imprecision (from double & blind reads);
- Maturity schedule.

The JTC noted other potential data sources that were not include in this year's assessment:

- Fishery and survey length frequency and length at age;
- Fishery CPUE;
- Bottom trawl survey biomass indices;
- Acoustic biomass surveys prior to 1995;
- Several proposed pre-recruit indices including: juvenile hake by-catch in the Shrimp trawl fishery, the coast wide juvenile rockfish survey, and the acoustic index for age 1 hake currently being developed by the survey team;
- Several early sources of biological data;
- The CalCOFI larval fish survey hake production index;
- Recently collected maturity data;
- Other environmental or ecosystem covariates.

SRG discussion focused on clarifying why several of these data sources were rejected or included, whether or not biological sampling of the fleet was representative, why TACs in both the US and Canada were not fully subscribed, and why several possible ecosystem covariates were not included, the abundance of California sea lions for example. The SRG was satisfied that exclusion of these potential data sources was for well-founded reasons, or would require extensive work to determine how to include them.

Discussion on Stock Assessment

Allan Hicks described the 2013 stock assessment; he began with the JTC's responses to the SRG recommendations from 2012. The JTC and acoustic team were able to meet the two highest-priority recommendations: To conduct an acoustic survey in 2012 on short notice (when one had not been scheduled) and to begin development of a management strategy evaluation (MSE), a complex computer simulation that allows examining the likely effectiveness of various stock monitoring and management strategies. Because of the intense effort that went into these activities, recommendations lower on the priority list were not attempted. The SRG was satisfied with what the JTC had accomplished.

The 2013 assessment was an update of the 2012 assessment, in that no major changes were made to the assessment model or data streams. The base model includes these main features:

- 20 age classes
- Data through 2012
- A single geographic area
- No separation by sex
- Population weights at age determined directly from observations on the commercial fishery and acoustic survey trawls
- A parametric stock–recruitment relationship including process error
- Natural mortality (M) assumed constant across time, area, sex, and age, and estimated with an informative prior
- Fishery selectivity and survey selectivity were each assumed constant over time and estimated up to age 6, beyond which each was considered constant with age. Estimation did not use a parametric model.
- An additional variance component estimated for the acoustic survey, added to the statistical estimate provided by the survey team
- Several refinements to account for ageing error

Several “bridge” models demonstrated that changes from 2012 results were in accord with the data observations. In general, results from this assessment are more optimistic than from last year’s assessment, partly because the acoustic survey estimate is considerably more positive. Strong cohorts are estimated to have recruited in 2008 and 2010, which is a positive factor for fishing in the immediate future. The abundance of the 2008 cohort is approximately as estimated in 2012, and its body growth plus the incoming 2010 cohort account for the biomass increase. As well as resulting in an increase in the 2012 biomass estimate, addition of the 2012 survey data to the model also reduced the confidence intervals in estimated 2012 spawning biomass, spawning depletion and recruitment. This was noted several times during the meeting as a positive outcome of conducting the 2012 acoustic survey, especially as it confirmed the presence of a large proportion of two year old fish in the age composition data, in agreement with the commercial age composition data.

Stock biomass was estimated to have increased since the 2012 assessment. Female spawning biomass was estimated at 1.50 million mt, more than twice the projected spawning biomass in the 2012 assessment (0.64 million mt). This difference is largely driven by increases in the estimated size of the 2008 and 2010 year classes. The 2010 recruitment is estimated to be especially strong, at 13.6 billion fish. The SRG noted that the assessment has in the past overestimated the size of large recruitments, so that current estimates of stock size may be somewhat overoptimistic.

Fishing intensity in 2012 was estimated to be below the 40% SPR management benchmark. Indeed, the status was estimated to have only a 9% probability of the fishing intensity in 2012 being above its target and the spawning biomass being less than 40% of the unfished value.

The SRG was perplexed by a longstanding pattern in selectivity estimates, in which selectivity gradually increases from ages 2-5 and then jumps sharply between ages 5 and 6 (pre-SRG draft assessment Fig. 18). This is particularly pronounced in estimates of survey selectivity. In past years additional selectivity sensitivities that put the flat top transition at an older age showed a gradual increase in selectivity from age 2 to whatever age was selected. Such reduced selectivity for the survey at ages that should be fully available to the survey is counter-intuitive. Discussion ensued in which several possible causes were noted, the main ones being violation of the assumption of constant selectivity over time; the lack of strong 5-yr-old cohorts in years when biennial and triennial acoustic surveys have been done; and potential changes in M with age or time. It was noted that sensitivity runs with a parametric (smooth) selectivity curve gave unsatisfactory results.

Two sensitivity runs included in the assessment allowed selectivity to change: i) flexibly; and ii) more flexibly over time. Although increased fishery selectivity on age-2 fish has been reported, this was not seen in estimates from these sensitivity runs. While these runs did fit the data better they did not change the assessment outcome substantially in that under both scenarios the estimate of current stock status was only slightly less optimistic.

Fits to age-compositions are better than in most assessments, although the SRG noted a consistent pattern of negative residuals in the age 15+ group (both fishery and survey age compositions). This was not considered of major importance (they are based on very tiny proportions), but it will be interesting to see whether the pattern changes if the assessment model is reconfigured.

Regarding fit to the acoustic survey, the SRG noted the large negative residual in 2001 and large positive residual in 2009. Although a perfect fit is not expected, these relatively large deviations suggest that survey catchability may vary from year to year more than would be expected. The current assessment practice is to inflate the variance on all surveys to account for these large deviations. Future investigation of alternative error structures is advised.

In the discussion of estimated benchmarks and status, the SRG questioned whether the 40% SPR benchmark is a target or a limit reference point. The discussion returned to this question several times during the meeting. This uncertainty points out the need for discussion with the JMC and further development of reference points to make sure that

target and limit are distinct. If one aims for a target and is accurate, then one would expect to overshoot 50% of the time and undershoot the other 50%.

It was observed that the current assessment shows fishing above SPR 40% in some recent years. This might be misleading, as realized landings have always been below quota. The estimates reflect changes in assessment estimates over time, not poor compliance.

As noted above, the assessment estimates a 9% probability that SPR in 2012 was below 40% (i.e., F too large) and spawning biomass was below 40% of unfished. This is a relatively small probability of that undesirable state. However, the SRG noted that the corresponding phase plot (pre-SRG draft assessment Fig. 32) suggests that the system of stock monitoring and management in place was not effective at reducing F according to the 40:10 algorithm when biomass drops below the 40% level. The SRG anticipates that further MSE results should illustrate whether other harvest control rules would be more effective in this regard.

Retrospective analyses did not exhibit the systematic patterns found in some assessments. The SRG noted the value of the “squid plot,” a newly developed tool to illustrate changes in estimated recruitment deviations in subsequent assessments. Here, it revealed that recruitments initially thought to be high are generally found smaller in later assessments, as more information on the cohort becomes available (pre-SRG draft assessment Fig. 46), except for the 1999 year class that had its estimate increase over time. Still, such recruitments tend to stabilise as above average indicating that the model is able to identify above average recruitment at an early stage, but cannot characterise with high precision the magnitude of recruitment until the cohort is four or five years of age. To a lesser degree, recruitments thought to be small also revert towards the average over time, but this tendency will be strongly influenced by the level of ageing error which causes large year classes to bleed into adjacent small year classes.

The inability to estimate recruitment until a cohort has been fished for several years has been a characteristic of catch-age models since they were first developed in the 1960s. The SRG discussed the need for a fishery-independent recruitment index to help define recruitment size earlier. The SRG noted that it might be possible to expand the existing acoustic survey to include such an index, in this case an index of age-1 fish which are currently excluded from the survey. The acoustic team noted that preliminary work has been done on developing such an index but need time and resources to complete the work.

A comparison of estimates from historic assessments shows wide variation (pre-SRG draft assessment Fig. 47). The SRG noted that each of those assessments was done by qualified scientists, presumably using up-to-date methods and with great care; yet, results

vary considerably. Much of this variation is due to the different approaches used over time to assuming or estimating the catchability coefficient of the acoustic survey. At present, this is estimated in the assessment model. The SRG questioned whether it might be possible, given better target-strength research, to put a strong prior distribution on this parameter, which controls scaling of the assessment.

Projected catch in 2013 to meet the 40% SPR benchmark was 626,364 mt. (Conditional on the assumptions of the assessment model, this provides a 50% probability that $F < F_{40\%}$.) The SRG noted that this assessment includes considerable uncertainty about the strength of the 2010 year class, especially given the pattern observed of previous larger recruitments being overestimated when they are first observed.

SRG requests for additional runs

The SRG made the following nine requests of the JTC at the end of Day 2:

1. Include other catch levels in forecast decision tables
 - a. 250 (or status quo), 350, and 400 thousand mt

Rationale: Finer precision and a more complete picture of forecasts.

2. What are the consequences of setting the maximum age of estimated selectivity to ages 2, 3 and 4, for the survey and the fishery? These runs would be similar to sensitivities that have already been done (sensitivity analysis 4 in assessment document).

Rationale: To bracket the uncertainty in the scale of the estimates and create a more similar picture to the structural uncertainty presented in the 2012 assessment.

3. Conduct a model run with the assumption that the 2010 yearclass is less than predicted, and create forecast tables. Set the 2010 recruitment deviation at zero and to half of what it is estimated to be in the base model.

Rationale: Examine a worst-case scenario. What if 2010 recruitment is less than predicted, as has been seen with estimates in the recent past (i.e., 2008 year class).

4. What are the impacts of applying a smoother to the selectivity curve (survey), i.e., reduce the sudden increase from age 5 to age 6 in survey selectivity.
 - a. Would reducing the variability from age-to age also address this?

- b. Investigate a parametric form and restriction of the offset between the last two ages.

Rationale: The current selection curve for the survey is implausible because of the sudden increase (kink) at older ages (5 to 6), and the lower selectivity at young ages. Is there much difference in the assessment results when assuming a smoother selectivity curve?

5. Conduct a sensitivity run in which fishery selectivity mimics targeting of strong cohorts and a run that assumes an increase in fishery selectivity for recent years.
 - a. More specific results from the time-varying selectivity-at-age 2 and 3, specifically.
 - b. Residual plots of fits to age compositions for these runs.
 - c. Look at survey selectivity in relation to the time-varying fishery selectivity.

Rationale: To investigate the characteristics of the time-varying selectivity estimates more closely and investigate the hypothesized targeting behavior. Does time-varying fishery selectivity reduce the residual patterns seen in the age composition data?

6. How much does the survey selectivity pattern change after dropping the 2001 and 2009 survey data (biomass and age comp)?
 - a. Look at specific age composition residuals in 2001 and 2009
 - b. In the future, the JTC may want to look at the expected numbers-at-age and survey numbers-at-age to compare the fit at age. How much are large deviations in biomass associated with particular age groups?

Rationale: To understand if this is a factor in the peculiar fishery and survey selectivity pattern and if this is causing the low selectivity at age 2 in the survey.

7. Explore whether there is autocorrelation in the recruitment deviations. Produce a phase plot of recruit $[t+1]$ vs recruitment $[t]$. Compare this to the forecast recruitments (which should be close to a random uncorrelated cloud).
 - a. Does the 2011 forecast recruitment have a different distribution from an uncorrelated lognormal distribution?

Rationale: The three year forecast may be inadvertently overemphasizing the possibility of back to back large recruitment events.

8. Explore the body weight-at-age data. Is it possible to separate the fishery and survey, which use the same weight-at-age values in the base model?

Rationale: Do the differences in collection time have an effect on the results, particularly for young fish.

9. Do a yield-per-recruit analysis.

Rationale: It may be useful for managers to understand the trade-off between growth and age at first harvest, and the potential for growth overfishing on this population, especially given that the stock will be dominated by three year old fish in 2013.

The JTC addressed each of the SRG's requests by presenting the MLE results of the nine sensitivity tests. Summary plots of estimated spawning biomass and recruitment showed the changes resulting from each sensitivity test, many of which were very similar to the base case.

Request 1: The JTC addressed this request by showing additional catch levels in forecasts.

Request 2: The effect of reducing the maximum age at which selectivity was estimated in both the fishery and the survey simultaneously was to slightly reduce the estimate of final year depletion and inflate estimates of depletion early in the time series. Fixing the maximum estimated age to 2 causes some very large changes to the early biomass trajectory.

There was discussion about whether it would have been more informative to look at changes to the maximum selected age for only the survey to be consistent with the sensitivity case in the 2012 assessment. It was recognised that applying the change to both the fishery and the survey likely represented an extreme case and applying the change to only the survey would likely return an intermediate result. The JTC confirmed that fixing the maximum estimated age at 2 for the survey while using the base case maximum estimated age for the fishery resulted in final year depletion estimate between the base case and the extreme case. None of the runs resulted in a maximum likelihood estimated depletion below the 40% reference point, although the confidence interval of the final year depletion estimate increased in both directions as maximum estimated age decreased.

Request 3: In this case the SRG wanted to explore the effect of reducing the strength of the 2010 recruitment to explore the implications of overestimating the size of this year class. The JTC looked at two alternative approaches. The first was to set the recruitment

deviation in 2010 to zero (i.e., average recruitment from the stock recruit relationship). The resulting recruitment was outside the range of 2010 recruitment estimates arising from the MCMC runs, and could therefore be considered to be inconsistent with the data. The alternative suggested by the JTC was to use the MCMC results from the base model and develop a decision table based on the lower 10% (99 runs) of 2010 recruitment estimates. This was similar to the approach taken in the 2011 stock assessment, where there was very large uncertainty in the estimate of the 2008 year class that had only been seen once in the fishery and never in the survey. Looking at only results from the lowest 10% of the posterior estimates of 2010 recruitment could be interpreted as a “worst case” recruitment scenario that is still consistent with the 2012 data. The range of recruitment estimates in this 10% sample was 4.2-8.0 billion and it was noted that even at these levels 2010 recruitment is estimated to have been above average (i.e., the estimate of 2010 recruitment in the scenario where the recruitment deviation was set to zero was 2.1 billion). Estimated 2010 recruitment in the base case model was 11.6 billion. The JTC showed the decision table and reproduced pre-SRG draft Figure 34 from the stock assessment document with the probabilities from the lower 10% of the decision table. It was pointed out that, while this was a useful plot, the probabilities were not straightforwardly interpreted as they represented only 10% of the posterior sample. There was suggestion that the plot title should indicate that this is a deliberately conservative scenario and that the vertical axis be labelled “conditional probability”. Further discussion of this analysis is reported in the Catch Level Recommendation section below.

Request 4: Constraining the allowable change in selectivity from age 5 to age 6 did not remove the “kink” in the selectivity curve, instead it shifted the kink down to a lower age. A comment was made that something in the data must be responsible for this unusual pattern. Application of a parametric (logistic) selectivity curve produced results similar to the case where the maximum age was fixed at age 3.

Request 5: Plotting the selection coefficients for each age for each of the time varying selectivity scenarios showed a general trend of increased selection of young fish during the last decade but current levels similar to historic levels. The patterns also showed long-term trends towards increasing selection of 4 and 5 year old fish. It was noted that not much could be read into the details of the trends, because selectivity parameters are likely to be confounded with other parameters (e.g., natural mortality), and because time-varying selectivity added many additional parameters but did not produce much change in the model results. Constraints on selection possibly cause distortions in other aspects of the fit. Bubble plots of Pearson residuals showed little change from the base model except in a few young ages. However, given the selection curve, one should not expect much change in the residuals for older ages because they are all assumed to be fully selected.

Request 6: In this request the survey data for 2001 and 2009 were removed from the model, as these points were not well fit and would have resulted in inflation of the estimated extra survey standard deviation, which in turn could affect estimates of selectivity or mask other model uncertainties. Removing these index points did greatly reduce the estimated survey process error but did not have a great effect on the estimated spawning biomass or depletion. The conclusion from this sensitivity analysis was that the commercial age composition data were providing a consistent signal throughout the time series while the survey biomass estimates, for 2001 and 2009 at least, were inconsistent. The shape of the estimated survey selectivity curve did not change as a result of removing these data.

Request 7: This analysis did not show strong evidence of any meaningful autocorrelations in recruitment, but it did show that strong yearclasses are not followed by strong yearclasses in the following year. Whether or not this is a significant finding would require a more formal statistical test.

Request 8: Separating the survey and commercial weight at age had little effect on model results. It was pointed out that any changes in weights in the survey would be absorbed by the estimate of survey catchability. Nevertheless, the 40:10 harvest control rule is driven by changes in spawning biomass, so it seems important to use weights-at-age in the spawning biomass calculation that are as accurate as possible.

Request 9: A yield-per-recruit analysis with knife-edge selection in the fishery and pooled weight-at-age data was presented. The analysis showed that yield per recruit is likely maximised at around 3.5 – 4 years of age (under the assumption about natural mortality in the analysis). There was discussion about the potential for growth overfishing the 2010 year class, and it was noted that the 2010 year class will be three years old in the 2013 fishing season, although results would be sensitive to assumptions about natural mortality. It was suggested that the yield-per-recruit isopleths plot be included in the assessment document, as well as possibly a spawning biomass per recruit plot. The SRG suggests that the equilibrium age composition resulting from a yield-per-recruit analysis could provide some information on likely average distribution of the population in the US and Canada given the current harvest policy.

In general, the SRG found the results of the sensitivity analyses helpful in understanding the effects of selectivity and estimated size of the 2010 year class on model results.

Management Strategy Evaluation

Results from initial Management Strategy Evaluation (MSE) analyses were presented by Dr Nathan Taylor (see assessment document Appendix). A number of questions relating to the details of the operating and assessment models were clarified by the JTC. The term “management procedure” in this context is understood to include the processes of data

collection, analysis, assessment model and application of harvest control rule. Results were presented for four alternative scenarios: (i) No fishing (no assessment); (ii) Perfect information (no assessment, “correct” F applied each year; (iii) Annual survey; (iv) Biennial survey.

The SRG greatly appreciated the efforts of the JTC to produce this MSE and found the results to be an informative first step. The JTC emphasizes that the analysis was relatively simple (e.g., the operating model and assessment model were structurally identical) to assist understanding of the behaviour of an MSE and interpretation of its results. So, the results are not a test of the assessment model itself, rather it is a test principally of the harvest policy given the stock monitoring program and precision of the data it produces, the assessment model that analyzes the data and the time lags that are inherent in the system. The JTC and SRG agreed that results are too preliminary to draw meaningful conclusions about the relative value of annual vs biennial surveys for a number of reasons, including potential problems with assuming fixed selectivity and with assuming log-normal errors in survey observations, leading to some extremely large simulated stock assessment errors. An additional limitation is that many aspects of the quota-setting process are not included in the MSE, for example decisions to take less catch than the target catch set by the operating model in some years. Some additional feedback measures could be added to more realistically represent the management process (e.g., capping the catch or the percentage increase in catch). Also, the simulation model of the population in the MSE meets the assumptions of the assessment model very closely. While the current MSE takes an important early step to examine the performance of the management procedure under favorable conditions, it is not a terribly realistic or demanding test.

The SRG made a number of suggestions for extending/improving the MSE, particularly in relation to developing a more realistic operating model. In doing so, the SRG recognised that technical design of the MSE cannot be decoupled from development of objectives and identification of appropriate performance measures. There was recognition of the need to develop a process that included the JMC, JTC, AP and SRG to guide further development of the MSE framework. That said, the SRG made the following observations/recommendations, which make up a small portion of potential worthwhile developments in the MSE:

- The MSE should consider alternative, more realistic operating models.
- The assumption of log-normal errors in the survey may result in observations that are too extreme due to the long tail of the log-normal distribution. Explore alternative potential error structures (e.g., t-distribution).
- Investigate age-specific natural mortality

- Include time-varying fishery selectivity in the operating model; otherwise the estimation model will be using the fishery age composition as a very consistent sample of the population age composition, including incoming yearclasses.
- Explore capping catches and trade-offs associated with different catch caps as an alternative to the current fixed harvest-rate policy.
- Look at a constant catch harvest policy
- Examine using a different base case than the 2012 assessment for projections

A detailed discussion of survey frequency followed, with concerns raised by the acoustic team about the lack of resources to do an annual survey, as well as attendant problems with running the survey as a joint hake/sardine survey. While it is the role of the SRG and JTC to articulate research needs (i.e., the need for annual surveys), the acoustic team asked that thought be given to the trade-offs in terms of personnel and resources available to conduct annual surveys, and also the value of other analyses by the acoustic team that cannot be done in survey years (e.g., work on target strength and identification for hake and other species, and work on the age-1 index). However, the SRG emphasises that lack of funds or resources should not be used to exclude necessary research from research recommendations and that recommendations should reflect scientific need, regardless of cost. AP members emphasised the importance of the surveys and suggested that additional resources may be able to be provided (or asked for) by industry. There was a suggestion that industry may also be able to provide acoustic data from their sounders, although questions about spatial interpretation of the data might limit their usefulness.

It was noted by one industry representative that the Canadian fishery is currently in a period of change and uncertainty expressing concern that the current survey design might not be capturing changes in hake behaviour and distribution. A discussion ensued around the merits of re-designing the survey, and while it was acknowledged that there have been changes in hake distribution the consensus was that changing the survey design would in effect negate the value of all previous surveys and should not be contemplated without substantial analysis of the likely impacts on the assessment. Furthermore the survey design was thoroughly reviewed during the 2012 SRG meeting and found to be consistent with international best practices.

A potential question to be addressed by the MSE is whether there are years when survey information is particularly valuable – for example, years following an apparently large recruitment event (such as 2010 and 2012) when the estimate of the size of this year class is still not well-characterised by the model. The bridging analysis in the current year’s assessment showed the value of adding the 2012 survey information in reducing the uncertainty interval around estimates of stock biomass. It was suggested that you cannot

know the value of a survey in advance, but it does appear that the difficulty in characterising the size of incoming year classes until they have been seen three or four times, as illustrated in the “squid” plot (pre-SRG draft Figure 46), is well-understood in this assessment. Resolving some of this uncertainty for the 2010 year class was actually one of the reasons behind the request for a 2012 survey in 2011. The current situation can be compared with the 2011 assessment, where there was no survey data to confirm the presence of the large 2008 year class seen in the fishery data and correspondingly there was extremely large uncertainty in the catch advice. It was noted that some of the work on survey timing and frequency could be handed off to the member nation groups rather than tasking the JTC, given the prevalence of this question for other groundfish fisheries besides hake.

Further discussions focused on the potential for not doing an assessment every year (or only updating the assessment in some years) to allow the JTC to work on the MSE, although the SRG recognises that this may not be ideal for a volatile species such as hake. This is another potential question that could be addressed using the MSE.

SRG Research Recommendations

High Priority

- The SRG recommends continued development of the MSE and taking into account the technical recommendations documented in this SRG report. There was recognition of the need to develop a process that included the JMC, JTC, AP and SRG to guide further development of the MSE framework.
- Life-history data improvements. Current information on maturity at age is from a single study in the 1990s. A new study of maturity at age is in progress, which the SRG strongly supports. The SRG recommends regular collection and analysis of life-history data such as growth, fecundity, and maturity at age, rather than relying on static values from the literature. Possible sampling by the fishery in late fall would provide more reliable information on maturity and fecundity because the fish are coming into spawning condition. The assessment should investigate a better measure of age-specific reproductive potential, rather than the current approach that uses average weight-at-age from the surveys and fisheries.
- Acoustic research
 - For the survey, record more information on the decision process used for assigning locations for trawl sites. Some tows are taken for target verification and some are for hake size composition sampling. This extra

information is important for measuring the survey variance associated with the trawls.

- Age-1 index development. Because the current acoustic survey does not develop an index of fish below age 2, a large recruitment (when it occurs) cannot be confirmed for several years, especially given surveys only in odd years. An index of abundance of young hake could speed reaction of stock assessments to high recruitment events. Preliminary research has been done on the potential of obtaining an index from the acoustic survey. The SRG recommends that this research be carried forward.
- Target strength, target verification, measurement of diffuse hake in low backscatter regions (fuzz).

Less High Priority

- Assessment model configuration:
 - Selectivity at age: The slow increase in survey selectivity at age, then abrupt increase to age 6 is counter-intuitive. The SRG request that the JTC configure the assessment model to explore the possibility of more constant selectivity at age and declining natural mortality at age as an alternative scenario.
 - Selectivity over time: The current assessment model configuration uses constant fishery selectivity at age over time. This treats the fishery as an accurate sample of the population's age composition and contributes to underestimating the variance of assessment results. This is particularly problematic for the MSE and for management-related inferences. A renewed investigation of configurations with time-varying fishery selectivity is advised, including investigation of non-asymptotic patterns.
- A summary of annual fishery operations would assist the JTC and SRG in interpreting annual changes in the pattern, timing and characteristics of the catch and fishery, especially if this could be available as early as the late fall JTC workshop. A more extensive document describing development of this fishery since its inception would also be invaluable to analysts in interpreting historic data. The SRG requests that the JMC and AP consider the possibility of creating these summaries.
- An annual survey is still desirable, but not at the expense of acoustic research that would improve survey reliability.
- Investigate the time series of recruitment to determine if the current lognormal random assumption is supported and if there are useful correlations with other

species or oceanographic features.

- Hake may be moving north as the survey is progressing from south to north, thus causing a Doppler effect that would create a small positive bias in the survey. The potential magnitude of this effect could be modeled. However, if it is consistent from year-to-year, then the effect is already being dealt with as a component of the survey catchability (q).
- Inter-vessel calibrations. The SRG notes that calibration of acoustics gear is performed regularly on vessels conducting the survey; however, potential differences among vessels have not yet been quantified fully. We recommend periodic inter-vessel calibrations. Based on comments from experts, the SRG believes that about 10% of the survey budget might be needed for such work. This is an important aspect of quality control in this assessment.

SRG Recommendations on Harvest Advice

The SRG endorses and commends the work of the JTC and the survey team. The SRG notes that the applying the default harvest control rule with the base case model predicts a median catch of 626,000 mt. The SRG notes there is considerable uncertainty in this assessment particularly with respect to the strength of the 2010 year class.

Given this uncertainty the SRG requested the JTC produce decision tables based on model runs with 2012 estimated recruitments from only the lower 10% of the posterior distribution (see discussion above). The reasons for suggesting caution in setting the 2013 harvest include:

- The age composition of the population is very truncated (the large majority of the population in the 2013 fishery will be 3 and 5 year old fish).
- There is concern about potentially harvesting a large proportion of immature fish.
- Recent assessments are known to have overestimated the size of incoming year classes (notably the 2005, 2006 and 2008 year class).
- The base model results and decision table are conditional on structural assumptions of the model, some of which are of concern to the SRG (e.g., survey selectivity).

- The assessment indicates that fishing intensity has been above the target in recent years, probably as a result of overestimating population size in previous assessments.

The model runs using recruitment from the lower 10% of those estimated by the base case model suggests that a 2013 catch of 336,000 mt would result in stable or increasing biomass over the next two years, even under this lower-recruitment scenario. The SRG suggests that this could be used as a precautionary lower bound of the suggested catch range. The SRG also recommends the F40% catch of 626,000 mt be used as an upper limit to the range of catches to be considered by the JMC. A 2013 catch closer to the lower end of the range could extend the available catch from the 2008 and 2010 yearclasses out over more years. The SRG notes that, historically, the larger 1980, 1984, and 1999 yearclasses accounted for approximately 65% of the total historic catch and contributed significantly to landings as 6 - 10 years olds which is well past their maximum yield per recruit age of ~ 4 years. The 2008 and 2010 year classes are currently estimated to be among those top-producing year classes.