CHILEAN HOKI STOCK ASSESSMENT REVIEW 2012

Ian J. Stewart[1] (Chair) and Dana H. Hanselman[2]

[1] **NOAA Fisheries, Northwest Fisheries Science Center**
Seattle, WA, USA
Ian.Stewart@noaa.gov

[2] **NOAA Fisheries, Alaska Fisheries Science Center**
Juneau, AK, USA
Dana.Hanselman@noaa.gov

Introduction

A review of the Chilean hoki (*Macroronus magellanicus*) stock assessment was conducted in Viña del Mar, Chile, 4-8 June, 2012 (Appendix 1). The review occurred by request of the Subsecretaria de Pesca (UnderSecretariat for Fisheries) to Instituto de Fomento Pesquero - IFOP (Fisheries Development Institute) for an independent review of the stock assessment, following a modeling workshop conducted in 2011 (Quinn and Cox 2011) and review of data sources conducted in 2012 (Clark 2012). The workshop was conducted under a specific Terms Of Reference (TOR) guiding the main questions to be addressed. Participants in the current review included the stock assessment team, hoki working team members and several additional attendees (Appendix 2). Prior to the review an extensive website was made available to the attendees, containing links to all important documents, model files and supporting technical publications The review followed closely to the TOR, and the preliminary agenda: generally a series of presentations by the stock assessment team, followed by questions and further discussion.

The stock assessment team did an excellent job of preparing a comprehensive set of documentation covering everything from the basic biology and life history of Chilean hoki to the history of the fishery, assessment and management actions. The review spent roughly two days critically evaluating each of the data sources and input parameters used in the stock assessment, as well as some of the basic structuring assumptions of the stock assessment. This was followed by a detailed investigation of the stock assessment model, including the formulations for the stock-recruitment function, the initial conditions, the fleet structure and selectivity approaches employed as well as other technical details. The final portion of the review consisted of a discussion of reference point calculations, harvest control approaches for hoki, model projections and uncertainty, and the management and scientific structure of the stock assessment process for this species.

This document contains a point by point summary of the primary information and discussion related to each of the TOR, followed by any recommendations. Each of these recommendations is then summarized in a table at the end of the review, categorized as short-term (to be addressed as soon as possible, if possible for the next stock assessment that will be conducted), medium-term (over the course of the next several assessments), and long-term in scope. Additionally, the recommendations are prioritized within each time-frame.

Review summary organized by the Terms of reference (TOR)

In this section we repeat the terms of reference, summarize the content of the primary documents and discussion during the review, and provide guidance on the key questions posed.

TOR 1. Structure and dynamics of the national stock

Critically analyzing the conceptual model(s) in use, the hypotheses on the structure and dynamics of the national stock, the relevance of potential connectivity with other stocks, for the purposes stock assessment and trophic interactions with other species.

Key questions:

- *How important are the migrations and connectivity with the Atlantic with respect to the endogenous?*
- *Can interactions (predation by other species) represent a relevant factor on the dynamics and fluctuations of the Chilean hoki abundance?*

Relevant documents:
CHODAR TOR1. Stock structure (Docs 18, 14; Clark 2012, Payá 2012k).
CHODAR TOR2. Monitoring system (Docs 18, 16; Céspedes M. 2012, Clark 2012).

A critical part of fisheries stock assessment is determining the spatial extent of the stock. In the case of hoki, there appears to be at least two potential layers of stock structure: between Argentinean and Chilean waters and between southern and northern waters within Chile (Doc. 14). Within Chilean waters, it is currently unclear whether there may have been some erosion of smaller sub-populations as the stock was fished down. Apparently the majority of the spawning stock occurs off southern Chile (41-47°S), but there is some evidence of spawning in the extreme south of Chilean waters (55-56°S) and in Argentinean waters. We were provided with otolith microchemistry analyses that showed high levels of mixing at an international scale, while genetic studies suggested potentially distinct populations. If the movement rates are high within the stock boundaries, mixing can be positive as it can prevent localized depletion. If movement rates are high and one area is acting as a source or sink, this can lead to strong biases in the assessment results. Low movement rates would be generally consistent with the current approach of separate stock assessments in the two countries waters.

There is an Argentinean stock assessment which has not been compared in detail to the Chilean assessment. During the review it was indicated that in recent years in Argentina a similar decreasing trend has been estimated in the hoki stock, from a maximum of about 1,000,000 t in 2000 to a current level around 700,000 t. Further, it appears that the two assessments may differ considerably in their estimates of stock productivity (particularly natural mortality). A substantial mismatch between the assumptions upon which the stock assessment and management are based and the spatial stock structure and could have short and long-term deleterious effects on both the productivity and evolution of the stock. We recommend the authors continue to research both movement rates and stock structure. Recent guidelines from the Alaska Fisheries Science Center offers advice on considering stock structure and might be useful to consult (ftp://161.55.120.150/afsc/public/Plan_Team/sept_2010/stockstructure.pdf).

We were presented with various estimates of consumption of hoki by other species. Hoki are the main prey of both southern hake and kingclip, and are also cannibalistic (23-72%). To the north of the primary spawning area, there have been increases in the occurrence of jumbo squid, as evidenced through recent fishery and survey bycatch. There has also been a large increase in biomass of squid in Peru (from 200,000 to 800,000 t (Doc 10). However, plausible values for total consumption of hoki by jumbo squid are difficult to determine. There have been projects that attempt to estimate consumption rates, but jumbo squid do not feed in captivity. The majority of predation by jumbo squid is thought to take place on hake, not hoki. An extreme example showed that jumbo squid could potentially consume 700,000 tons of hake in one month (Table 2, Doc 10). The common hake stock assessment model includes a change in natural mortality to account for increasing predation. We recognize predation and cannibalism are likely to be highly important processes that affect the annual survival of hoki. However, estimating changes in natural mortality to include these factors is not usually tenable in a situation with limited data like the hoki assessment, and should be approached cautiously. Avenues for investigation include the use of ecosystem model outputs, or time-series of abundance from other stock assessments as if they were data with which to tune temporal changes in natural mortality.

Recommendations:

Short-term:

1) Qualitatively compare the trends in biomass and recruitment between the Argentinean and Chilean assessments.

Long-term:

1) Conduct a joint stock assessment with the Argentineans.
2) Implement conventional tagging experiments, ideally with Argentina, to attempt to estimate movement rates, and a better understanding of stock structure.
3) Improve estimates of abundance of and potential consumption rates by jumbo squid and other key predators.

TOR 2. Life cycle and dynamics parameters.

Critically reviewing the hypotheses on life cycle and dynamics parameters used in the most recent stock assessments, particularly regarding properties of stock closure, hypotheses on the spawning areas and recruiting.

Key questions:

- *Is it plausible to assume a closed stock at a national level?*

- *Do historical abundance fluctuations of the resource respond to an environmental phenomenon of predator/prey interactions, or do they constitute a stochastic behavior?*
- *Could changes in abundance and reduction of the stock age structure have modified the patterns of life cycle (growth and maturity, among others)?*

Relevant documents:
CHODAR TOR 3. Life history information (Docs 18, 13, 15, 11, 10, 5, 4; Canales 2011, Clark 2012, Ojeda and Hidalgo 2012, Payá 2012h, j, l, f).

Stock structure was discussed under TOR 1. We were presented with various documents describing historical changes in predator abundance. While predator/prey interactions may be an important part of the changes in abundance of hoki, it appears more likely that fluctuations in recruitment are largely density independent (e.g. environmentally driven), and that the greatest factor in the decline of the spawning stock was high exploitation rates by the purse seine fishery in late 1990s (Doc. 4, Doc. 26).

The life history parameters of hoki appear to be reasonably and reliably estimated in recent assessment analyses (Docs. 5, 13, 15). Ages appear to be well determined (Doc. 13), as cohorts can be visually identified in the data and followed over time. IFOP age lab routinely estimates the precision of replicate observations by different readers, as a standard procedure to ensure the quality and consistency of age readings. It was presented at the Data Review meeting in January that coincidence in ageing was about 70%. However, ageing reading errors have not been included in the stock assessment models. Ageing imprecision can contribute to inaccurate estimates of cohort strengths if not accounted for in the stock assessment.

Mean weight at age appears to be highly variable with an overall reduction in average weight-at-age since 1988 (Fig. 2, Doc. 5). There appears to also be an increase in weight-at-length during the 2000s while the weight-at-age has decreased. This could be worth investigating in more detail. The assessment utilizes a reasonable approach, accounting for these changes by using observed mean weight-at-age directly in the assessment model (Doc. 28). Maturity-at-size appears to differ between histological and macroscopic observation, but appears to be declining in recent years (Doc. 4). We agree with the observation of the CHODAR review that sampling maturity only from the spawning aggregations may bias results, but would prefer that representative samples from off the spawning aggregations be collected during the early winter, rather than samples from the summer, which are very difficult to estimate maturity from in most cases (Doc. 18). Cooperation with industry may be necessary to get these samples.

The age-truncation seen in the current population is consistent with a period of high fishing mortality and a depleted stock. Some of the observed changes in life history parameters are also consistent with this historical pattern. Natural mortality is notoriously difficult to estimate, and a time-varying natural mortality is even more challenging. The authors have made a thorough investigation of the available of existing methods to estimate natural mortality with various life history parameters (Doc 15). It was suggested at the meeting to estimate differing natural mortality by age because of predation of hake and older hoki on young fish. In our experience, this can be done, but often just scales to the recruitment actually seen by the survey and fishery after the fish reach full selection. When estimated in the model, it appeared that the age data has some information regarding natural mortality, specifically, the assessment model seemed to estimate a reasonable value given the hoki life history (Doc. 28). Likelihood profiles of natural mortality would be useful to see how well-estimated natural mortality is by the model and exactly which data sources are providing this information.

Recommendations:

Medium-term:

1) Investigate the degree of ageing imprecision by performing double-reads of the same otoliths and/or exchanges with other aging laboratories.
2) Consider an age-specific (Lorenzen) natural mortality to account for predation juveniles
3) Obtain representative samples from hoki in the winter both on and off the spawning aggregations.
4) Investigate the increase in weight-at-length, but decrease in weight-at-age over the same period.

TOR3. Growth parameters, age/length keys, and matrices of catch at age/size.

Critically reviewing the growth parameters, age/length keys, and matrices of catch at age/size.

Relevant documents:
CHODAR TOR 4. Landings and discards (Docs 18, 4, 9; Clark 2012, Payá 2012f, m).
CHODAR TOR 5. Age/length composition data (Docs 18, 4, 7; Clark 2012, Payá 2012f, i).

We strongly agree with the recommendations by the CHODAR panel that "Above all else, a stock assessment requires reliable estimates of total removals." The greatest deficiency in the current stock assessment is the absence of a best estimate of total catch including both discards and underreporting. The authors showed several estimates of the potential ratios of actual catch to reported landings over time and these values were substantial (Tables 2 and 3, Doc. 9). Discussions during the review revealed both political and legal impediments to attempting to estimate discard rates, but these obstacles need to be overcome. One common misconception is that including estimates of these catches will be likely to lower the overall stock status or quotas available to industry. However, in our experience, if total catch has been underestimated over the full time series, the addition of these data may actually scale the assessment estimates of abundance upward. We suspect that the general trends estimated by the stock assessment model are reliable, but the absolute scale may be highly biased without including these removals. Based on the information provided, it appears that the biological sampling from the trawl fleet occurred prior to any catch sorting by size or grade, and the purse seine fleet had little incentive for sorting as fishmeal was the primary product of this sector. As long as this continues to be the case, the length and age data collected from the fishery should be representative of the whole catch adjusted for underreporting and discarding.

Age-length keys for processing the data in recent assessments are constructed using standard methods and sample sizes for otoliths appear generally adequate for a relatively short-lived fish. We were shown two sets of length compositions for the purse seine fishery, the second a new data set brought to the review and not provided to us (or assessment authors) in advance of the review. These data showed a longer time series of lengths collected during the jack mackerel fishery, and perhaps a greater proportion of larger/older fish that had been caught. We recommend that industry and the assessment authors collaborate on whether these data are adequate to include in the assessment.

Recommendations:

Short-term

1) Collaborate among institutions and with industry to compute the best estimates of total catch for the entire time-series including all discards and underreporting. This may require a workshop to identify key issues, data sources and participants and outline a plan under which the research will be conducted. This effort should also include recent years (2008+), which were not available for this review.
2) If there are delays in creating a working time-series of total catch estimates, the next stock assessment should perform sensitivity analyses to several plausible levels of total mortality in order to evaluate the effects on the results and uncertainty in those results.
3) Collaborate with industry to compute best estimates of the catch-at-length during the purse-seine fishery. Given the magnitude of these catches, this could be important to estimates of selectivity for this fleet and the trends in spawning stock biomass.

Medium-term

1) Conversion factors used to adjust dressed products to whole catch weights appear to be variable by source. Although this is likely a very small contributor to whole catch uncertainty relative to discarding and underreporting, an effort should be made to refine conversion factors as processing technology and market products change over time.

TOR4. Relevance and sufficiency of the data sources and indicators.

Critically analyzing relevance and sufficiency of the data sources and indicators (landings, catches, CPUE, and spawning biomass of hydroacoustic surveys) used in the stock assessment of the different fleets that have exploited this resource.

Key questions:

- *Are the following informative? i) purse seine CPUE from the Center-South zone and ii) trawling CPUE from both zones in the years previous to 1998*
- *Is the trawling fleet CPUE informative during the period in which Hoki was not a target resource?*
- *Is it necessary to differentiate the CPUE of the trawling fleets between the two Fishery Units?*
- *Can the fishing power of the trawling fleets be considered constant throughout the time series?*
- *Is it appropriate to consider the spawning biomass of the surveys as a representative abundance index of the whole national spawning stock*

Relevant documents:
CHODAR TOR 6. Abundance indices based on commercial catch rates (Docs 18, 6, 8, 17; Caballero and Payá 2012, Clark 2012, Payá 2012a, b).
CHODAR TOR 7. Abundance indices based on acoustic surveys (Docs 18, 12, 4; Clark 2012, Lillo 2012).
CHODAR TOR 8. Estimates of abundance and mortality from the summer 1972 trawl survey (Docs 18, 7; Clark 2012, Payá 2012i).
CHODAR TOR 9. Provide measures of population abundance that are appropriate for stock assessment (Doc 18; Clark 2012).

We were presented the results of extensive statistical analyses of fishery CPUE from the trawl fishery. The modeling methods applied to these data were appropriate, and consistent with those commonly used in analysis of data with many zeroes and a skewed distribution (the delta-lognormal). We generally endorse the analysis and particularly appreciate the recent move to GLMMs which generally should provide a robust estimation framework and a better estimate of the variance.

Despite the high quality of the analyses that have been conducted, fishery CPUE is rarely a quality index of abundance. This has been documented in many fisheries around the world, and is a primary factor in the development and implementation of fishery-independent research surveys wherever possible. In the case of Chilean hoki, much of the time series appears problematic, since hoki was not actually a target species of the fishery. Before 2002, hoki were mainly bycatch of the southern hake fishery. Once hoki became a targeted species, we are still concerned that hyperstability of the index may be an issue, as the fishery is focused where the fish are most aggregated during the spawning season: in a few specific canyons which are presumably the best habitat. The authors also specifically restrict the analysis to the main 14 fishing grounds (78% of historical catch) and left out the area between fishing grounds Under the "basin effect" (MacCall 1990), this approach may exclude marginal habitat, 'edge areas', where CPUE will fall faster than in the best habitat. The authors do make a comparison that shows that including all the zones has a minimal effect on the overall index in recent years, but it is difficult to determine within the scale of the comparison if this is true from 1979-1999 (Fig. 22, Doc 6). For a species that aggregates and can readily be located with echosounders, the CPUE may remain artificially high until extreme depletion has occurred. Unfortunately, there is very little other data available on relative abundance prior to 2002.

Fishing power is not considered constant through time. There are three periods of catchability considered (1970-1996, 1997-2001, and 2002-2010) which correspond to periods of strictly bycatch, fishery capitalization, and full targeted fishing. These blocks seem well defined to capture different dynamics of the fishery. However, direct proportionality of trawl fishery CPUE and population abundance during the period between 1997 and 2001 seems unlikely. Discussions during the meeting indicated a reduction in the fleet as it shifted from treating hoki as a bycatch to actively targeting the species, potentially resulting in higher CPUE.

When comparing the indices by separate fleets or separate fishery units, the results are quite different. These results may give some insight into the hyperstability of the fishery data in the spawning aggregations because the index calculated in the center-north demersal fishery area which is in feeding aggregations dropped precipitously, while the Southern demersal area CPUE remained high (Figure 51, Doc. 17). Given the increased uncertainty when these areas are separated, we think the whole area indices are more useful. We recommend using the whole area trawl CPUE index, but to continue to apply a relatively high CV.

We recommend excluding from the stock assessment model the trawl CPUE index from the period between 1997 and 2001. We recommend the continued use of the period prior to 1997, only because there is little other reliable information on abundance during this time period.

The purse seine CPUE, where hoki were also not the primary target species, is likely to be an inferior index of abundance relative to the trawl CPUE. Hoki made up little (4%) of the total annual catch, and they were caught opportunistically during the primary jack mackerel fishery. To deal with it, CPUE analyses were restricted to the hoki fishing season, that is, from September to March (next year), when 99% of hoki were caught. Most of the hoki were caught in one or two months that varied between years, but hoki were mostly caught in October and/or November. During these months about 60-70% of the whole catch of the fleet was hoki, and the rest were mostly sardines and anchovies. Hoki was an alternative when jack mackerel were not available to the fleet because they had migrated to the spawning grounds in the high sea. CPUE analyses also restricted the data to the two main coastal fishing zones (7 and 8), and excluded the zones far from the coast were jack mackerel were mostly fished (Doc. 8, model Mod 4f, Fig.1). But there is very weak information on actual effort; there is no haul level effort data (Doc. 8). Further, the index represents only juvenile hoki, which is of less value to the estimation of spawning biomass. An additional issue identified during the review was the inclusion of 1986/1987 in the GLM, but only 1988+ in the assessment model, likely due to historical precedent. Due to this aggregate of concerns, we do not recommend the continued use of this index in the assessment model.

The acoustic trawl survey index is reviewed thoroughly in TOR 7 of CHODAR 2012 (Doc. 18), and we will not repeat the technical details in that report. Its conclusion was that the acoustics group "applies appropriate instrumentation and survey analysis methods to produce valid estimates of Chilean hoki abundance and biomass estimates" and that no issues suggested that the "estimates should not be used in the Chilean hoki stock assessment..." We have some reservations about the spatial coverage of the acoustic survey, because it assumes that the majority of the spawning biomass returns to the same locations every year, and that the same proportion of the total biomass makes that return at the same time, overlapping with the timing of the survey. There was also no discussion of the 'dead-zone' near the substrate where hoki would be invisible to the acoustics, and the topic was raised that trawling is conducted during daytime, while acoustic data are collected at night. Survey fishing gear changes prior to 2005 may also have contributed to additional variability in the index. However, the concordance of the fishery CPUE and acoustic biomass in the period after 2002 is comforting, as they are examining approximately the same area. Concern was raised during the review that the fish may show peak abundance at differing times of the spawning season in different years. However, if the acoustic survey does indeed encompass the majority of the spawning biomass every year, it should be suitable as an index of total biomass after accounting for selectivity in the assessment model. If funding were available, a synoptic multispecies bottom trawl survey conducted biennially or triennially would complement the acoustic survey of the aggregations and provide benefits for multiple stock assessments.

The review included only a limited discussion of the 1972 trawl survey data, as all participants agreed with the stock assessment team's decision to remove these data from the current assessment model. We recommend no change to this approach, as the need for very strong assumptions regarding relative catchability far outweigh the benefits of a single survey observation.

Recommendations:

Short-term

1) Omit the years 1997-2001 from the trawl CPUE index time series (and GLM analysis).
2) Remove the purse seine CPUE index from the assessment model.

Long-term

1) Investigate the presence of other potential spawning aggregations. This could be achieved through extending the acoustic survey, or through industry collaboration/charter.
2) Design and conduct a multispecies synoptic trawl survey.
3) As time permits, explore further analyses of purse seine CPUE, potentially investigating the spatial-temporal interactions, the subsetting of trips targeting hoki, and changes in the fishing vessels participating.

TOR5. Relevance, sufficiency and adequacy of the approach and procedures (Doc 28, 26, 25; Quinn and Cox 2011, Payá 2012g, d).

Analyzing the relevance, sufficiency and adequacy of the approach and respective procedures used in the recent stock assessments.

The stock assessment generally utilizes modern and appropriate statistical and computational methods suitable for the relatively limited data available. In the following TORs, we recommend some specific suggestions for improvement; however they are minor in comparison to the major topics regarding improving the quality of the input data referred to in the previous TORs (i.e., including estimates of total catch).

Recent stock assessments for Chilean hoki (2010, 2011 and 2012) have estimated similar trends in stock abundance and recruitment strengths. These trends appear quite robust to alternate model configurations, stock dynamics assumptions, and alternative treatment of the data. From these analyses it is clear that the stock has declined substantially, and, due to ubiquitously below average recruitment since 1996, the current spawning biomass is estimated to be at a historical low level.

TOR6. Stock assessment model: procedure models, observation and error and uncertainty sources.

Critically reviewing the stock assessment model with respect to the procedure models, observation and mistake models used in the estimation for status determination, considering the different uncertainty sources.

6.1. Review the formulation of the stock assessment models with a focus on the most recent assessments.

Key questions

- *Is the approach adopted in the assessment adequate and sufficient to appropriately address the most relevant uncertainty sources?*
- *Is the setting up of parameters for per-fleet exploitation patterns adequate?*
- *Is the stock-recruit ratio adequately used in the assessment?*

6.2. Summary of the changes of programming language used in the assessment models.

6.3. Critically reviewing the procedure to define the assessment initial conditions. Propose modifications where appropriate.

6.4. Analyzing the model goodness to fit the data (residuals analysis) and reviewing the consistency and the main outcomes of the base model, considering the outcomes of the retrospective analyses.

6.5. Critically reviewing the treatment of the uncertainty in the model outcomes (distributions a priori and posteriori, and MC/MC chains).

6.6. Review of the base model definition procedure, the alternative cases and the sensitivity analyses conducted, considering the assessment procedure used.

6.7. Review of the significance and impact on estimations deriving from the weighting assigned to the different information sources and abundance indices used in the assessment.

Key questions

- *What procedures are most appropriate to establish the weightings more objectively?*
- *Do the relative weightings alter the biomass trend and stock status?*

For the most recent iteration of the hoki stock assessment (2011), the authors have reprogrammed their model in AD Model Builder (ADMB Project. 2012). This is a widely used platform for modern fisheries stock assessment which allows reviewers from other regions to easily review code, if necessary. ADMB provides reliable and efficient estimation of many parameters and also allows for simultaneous estimation of parameter uncertainty through asymptotic approximations and built-in Markov Chain Monte Carlo (MCMC).

The assessment follows modern and generally accepted methods to estimate the population dynamics of Chilean hoki. The assessment has been through many iterations testing alternative configurations of the model (Doc. 28). After the recent reviews (Docs. 18, 25), the assessment seems to be nearing its optimal model configuration.

The assessment has historically estimated initial conditions and recruitment deviations in a number of different ways (Doc. 28). In the most recent iteration (unlike previous models), it appears that recruitments are not bias corrected to account for the distribution created by estimating deviations from the stock-recruitment relationship in log-space. The bias correction is important when using a stock-recruitment relationship for estimating the reference point and initial conditions simultaneously. Generally, the degree of recruitment variability (σ_r) should be consistent between the input value (used to create a penalty on the recruitment deviations) and the RMSE (which should be slightly smaller than the input to account for estimation bias) over the most informed portion of the recruitment series. In addition, because observed data are never perfectly informative regarding the true degree of recruitment variability, it has been shown through simulation that estimators based on the maximum of the posterior density

(MPDs or minimization, but not full Bayesian integration) should be bias corrected in proportion to the estimated standard errors in order to avoid bias correcting deviations with expectation approaching zero (Methot and Taylor 2011). This is particularly important for early recruitment deviations and initial age structure and may be a contributor to the sensitivity observed to initial conditions in the hoki assessment. We recommend retaining the current approach of initializing the model with non-equilibrium conditions, but adding and scaling the bias correction appropriately. This may reduce the effects of selecting a particular start year, especially if the start year is selected to predate appreciable levels of fishery removals.

The assessment workshop proposed an extensive set of sensitivity analyses (Doc. 25). Model results were robust to most sensitivity runs that were conducted, especially estimates of recruitment. The exceptions were when the fishery dependent CPUE was entirely removed, natural mortality was lowered, hyperstability was accounted for, and when trawl fishery selectivity was considered static for the whole time series. With the exception of large changes to the fixed value for natural mortality, the most sensitive portion of the estimated biomass time series was the historical abundance. When fishery dependent CPUE was removed, this had a large effect on historical biomass showing that these data are providing much of the information regarding abundance in the early portion of the time-series. The alternative value for natural mortality seemed implausibly low for a short-lived species of hoki. The estimated hyperstability for the CPUE series is perhaps the most interesting alternative because, as noted in previous TORs, the CPUE models do seem to have a strong possibility of being hyperstable. The model was not particularly sensitive to the choice of the stock-recruitment relationship, in fact, assuming no stock-recruitment relationship gave the same recruitment patterns, but a difference reference point. These sensitivity analyses also tested different weightings on data sources (Scenarios 22). The model did not seem particularly sensitive to these weightings in terms of either biomass or recruitment. We suspect if the multinomial weights were given more extreme values this may not be the case, but it appears that for a reasonable range of values, the model results are robust. The assessment might also attempt to tune the weights to the standardized deviation of normalized residuals (SDNR), or the effective sample size. The SDNR is closely related to the root mean squared error (RMSE) or effective sample size; values of SDNR of approximately 1 indicate that the model is fitting a data component as well as would be expected for a given specified input variance. The normalized residuals for a given year *i* of the abundance index can be computed as

$$\delta_i = \frac{\ln(I_i) - \ln(\hat{I}_i)}{\sigma_i}$$

where σ_i is the input sampling standard deviation of the estimated abundance index, *I*. For age or length composition data assumed to follow a multinomial distribution, the normalized residuals for age/length group *a* in year *i* can be computed as

$$\delta_{i,a} = \frac{(y_{i,a} - \hat{y}_{i,a})}{\sqrt{\hat{y}_{i,a}(1-\hat{y}_{i,a})/n_i}}$$

where y and $\hat{y}$ are the observed and estimated proportion, respectively, and n is the input assumed sample size for the multinomial distribution. The effective sample size can be computed for the age and length compositions modeled with a multinomial distribution, and for a given year i was computed as

$$E_i = \frac{\sum_a \hat{y}_a * (1 - \hat{y}_a)}{\sum_a (\hat{y}_a - y_a)^2}.$$

An effective sample size that is nearly equal to the input sample size can be interpreted as having a model fit that is consistent with the ability to fit the data, and therefore the degree of uncertainty can be considered generally reasonable. Patrick Cordue, Mark Maunder, Chris Francis and others have suggested versions of these techniques at reviews we have been involved in. The model can be reweighted by setting an objective function penalty to reduce the deviations of average SDNR of a data component from one or the deviations from mean input to mean output effective sample size. We recommend that this be done via comparison to the mean input values across years for a data source, not for individual years. In addition, the actual sample sizes for many of the compositional data are quite heterogeneous across years. We recommend that these interannual differences in sample sizes be accounted for explicitly in the data weighting, which also allows years of small sample sizes to be used, rather than choosing an arbitrary minimum threshold for inclusion in the assessment. Subsequent plotting compositional fits with the sample size (or weight) in each graph can be helpful to interpret and importance of lack of fit. Similarly, heterogeneity among years in the variance of the survey index of abundance should be included, with the mean scaled to the relative desired precision. This can be informed via the evaluation of residual patterns and the RMSE of the fit to the index.

The purse seine CPUE residuals were strongly patterned with periods of negative, positive, and back to negative. Trawl fleet CPUE and acoustic biomass residuals were much better (Figs. 8 and 9, Doc. 26). Age residuals in the acoustic survey were slightly non-Gaussian, but not troubling. However, the residual patterns for the trawl age compositions were quite strong after the selectivity change in 1991 to asymptotic selectivity. The model fit very poorly to ages greater than 10, and particularly the plus group, in most years from 1991 through 2000 (Fig. 14, Doc. 26). The choice of 1991 appears to be an arbitrary one; the fleet changed to targeting hoki under an ITQ-program in 2002, and this year would appear to be a more reasonable place to change selectivity. We recommend starting the new selectivity block in 2002.

There was a considerable discussion regarding the rationale for allowing a change in the selectivity of the acoustic survey and trawl fleets in 2008. Concern was raised that the change in age-structure evident in the data might be a result of additional depletion rather than a change in the availability of the larger/older hoki in the stock. However, evidence of a change in the maturity schedule, combined with the fact that both the trawl fleet and survey trawling seem to have observed the same phenomena indicate that these two sources of data originating from the spawning grounds may realistically have different selectivity patterns in the current period. We therefore did not recommend a change to this approach.

Retrospective analysis (within-model) was conducted where one year of data, starting with the terminal year, is sequentially removed. In recent years, there was virtually no retrospective pattern. Before 2008, the models retrospectively appear to overestimate the biomass. This may be due to the selectivity change that occurred in 2008 for the survey, and may suggest that the new selectivity block should begin earlier than 2008. We recommend that retrospective analysis be continued as a primary model diagnostic. That is, if a strong retrospective pattern exists, parameters such as natural mortality, catchability and selectivity may be misspecified. We do not recommend a retrospective "correction" because it may correct the pattern, but could just as easily be further from the truth, rather than closer. A good reference on retrospective patterns is from the NMFS retrospective working group:
(http://www.nefsc.noaa.gov/nefsc/publications/crd/crd0901/crd0901.pdf).

The assessment uses prior distributions to constrain some parameters, and has used both MCMC methods and the Hessian standard deviations to convey uncertainty. We appreciate the assessments attempts to estimate risk of various management scenarios. However, risk is underestimated, because uncertainty in these cases is underestimated due to at least two major factors. The first is that the model has several important but fixed parameters, such as natural mortality. The second is that the uncertainty reported is only related to uncertainty within the preferred model, and not alternative structural assumptions or configurations of the underlying dynamics. Some methods to further look into assessment uncertainty are to free up as many parameters as possible, while still achieving convergence, and examine the posterior distributions from those model runs. Likelihood profiles are an effective technique for investigating the information content of each data component, as well as a means for expressing uncertainty in fixed quantities that is difficult to achieve via asymptotic or Bayesian methods. We recommend routine use of likelihood profiles, particularly for parameters such as natural mortality, survey catchability, steepness, and SB_0.

Recommendations:

Short-term

1) Review the current model code and include appropriate bias-correction terms.
2) Investigate the estimated level of recruitment variability and iteratively ensure that it is consistent with the variability used to determine reference points and initial conditions.
3) Attempt to fit asymptotic selectivity from 2002 on instead of from 1991 on for the trawl fishery.
4) Ensure that the Francis method for tuning compositional weights has converged by conducting more iterations.
5) Include interannual variability in sample sizes in the weights on data sources.

6) Correct the minor inconsistency in the stock-recruitment plotting code causing mislabeling of some recruitment years.

Medium-term

1) Explore other methods of conveying the uncertainty in the model outputs, such as alternate states of nature, likelihood profiles, and examination of posterior distributions when freeing up more key parameters.
2) Attempt to provide consistency in the tuning and weighting methods employed among stock assessments for groups of similar species. This would be helpful to avoid inconsistent and/or subjective decisions by different assessment authors.
3) Attempt to tune input to output error by using standardized deviation of normalized residuals and/or effective sample sizes.

TOR 7. Providing technical assistance to improve the robustness of the assessment against currently-known uncertainty factors (Doc 26, 25; Quinn and Cox 2011, Payá 2012d).

Providing technical assistance to improve the robustness of the assessment against currently-known uncertainty factors. Characterizing the uncertainty sources in the assessment, according to importance for the reliability of the estimates reached.

Key questions:

- *Are the criteria used to choose the assessment starting year adequate?*
- *Are the criteria used to estimate the assessment initial conditions appropriate?*
- *Is the spawning biomass index weighting derived from the surveys adequate (with respect to the other abundance indices)?*
- *Was the stock in an equilibrium condition and in a stationary status at the beginning?*
- *Can the survey catchability be estimated?*

The starting year of the hoki assessment has gone through considerable modification over time (Doc. 28). In Alaska, the assessment model starting year is usually chosen based on when reasonably reliable catch records exist that show that the population was being exploited. On the west coast of the U.S., models are started much earlier (as early as the beginning of the 1900s) in order to include all significant removals from the population due to fishing. Due to scarce of absent data for many of these early years, those models generally estimate recruitment levels largely dictated by the stock-recruitment relationship until data becomes available to directly influence population estimates. The latest hoki assessment model begins with non-equilibrium conditions which we strongly support, because a fish stock, particularly one with high recruitment variability is never expected to show an equilibrium age distribution at any stock size and only varies around a central tendency corresponding to carrying capacity, or SB_0. We recommend that the hoki assessment uses a starting year that is consistent with the known beginning of substantial catches, and continues to allow for non-equilibrium conditions (allowing the initial age composition to be estimated, rather than static).

For the 2010 and 2011 stock assessments, the weighting of the acoustic survey was high (CV of 0.15) relative to any of the fishery CPUE data (CVs of 0.3-0.4). While this may not reflect an estimate of the actual sampling error of the survey, it does seem appropriate to give the survey additional weight. It represents the only fishery-independent index of abundance, making it reasonable to assume that it is the data source most likely to be proportional to stock abundance. Catchability of this survey can be estimated, but the assessment team thought that the very high catchability that resulted from previous attempts was unrealistic, and so have constrained its estimation with a tight prior (CV=5%). The acoustic time series is still relatively short, and without considerable contrast in abundance when compared to earlier years when the stock was declining rapidly. A likelihood profile on catchability would be a useful exercise to evaluate which data sets are informing the estimate and how strong the improvement in model fit may be to freely estimating the expected value. We recommend continued periodic testing of whether catchability, natural mortality, and steepness are estimable as the time-series' of catch, fishery data and the acoustic survey index grow longer.

The assessment team should consider creating decision tables for management that can be focused on a single 'axis of uncertainty' which should be identified by the stock assessment team. For hoki, survey catchability, recruitment variability, natural mortality, historical catch are likely candidates as the dominant source of uncertainty. On the west coast of the U.S., a 'base-case' is selected as the most probable model and two alternate cases are presented that 'bookend' the likely range of uncertainty and are expected to be about half as probable as the base-case. We recommend routine use of likelihood profiles and decision tables to illustrate components of the total uncertainty not currently included in the assessment due to fixed parameters. An example of a useful set of scenarios for hoki would be to show a set of alternative model results using the full set of recruitment estimates as a base-case, a set using only the recent low estimates, and a set using only the historic high estimates. The U.S. west coast stock assessment TOR might be a useful reference for a general discussion regarding determining and reporting the primary axis of uncertainty (http://www.pcouncil.org/wp-content/uploads/GF_Stock-Assessment_ToR_2011-12.pdf).

Recommendations:

Short-term

1) Evaluate, after addressing the recommendations regarding bias-correction and initial conditions above, whether starting the model prior to the first year of appreciable catches would reduce the sensitivity to the specific choice of year.
2) Perform periodic testing in future assessments of whether catchability, natural mortality, and steepness are estimable. Routinely use likelihood profiles to better understand which data sources are providing information on these parameters.

3) Consider the use of simple decision tables to convey a broader spectrum of uncertainty to fisheries managers.

TOR 8. Critically analyzing the BRP and reviewing the criteria and procedures used to determine the stock status (Doc 28, 27, 22; Payá 2012g, c, e).

Critically analyzing the BRP and reviewing the criteria and procedures used to determine the stock status.

8.1. Reviewing the BRP used for diagnostic purposes (precautionary and limit): Are they appropriate? Is it plausible to estimate species-specific BRPs?

8.2. Critically reviewing the stock status indicators and comment on their robustness against current levels and uncertainty sources.

Key questions

- *Is the status estimation reliable (considering the uncertainty)?*
- *Is the potential egg ratio an adequate index to represent the reproductive value/contribution of the spawning stock?*
- *Is the stock reduction index a good indicator given the estimation uncertainty of the assessment initial condition?*
- *Given the precedent that the age at first maturity has been reduced in a year in females, is the spawning biomass estimation procedure a robust indicator throughout the assessment time series?*

The current hoki assessment uses reference points based on *MSY*-proxies of $SB_{40\%}$ and $F_{40\%}$ calculated from SB_0 (the central tendency of the stock size implied by the dynamics in the absence of fishing mortality; not necessarily a condition that has actually occurred in the early or any portion of the time-series). The biomass limit reference point $SB_{20\%}$ also appears reasonable for a species with the life history of hoki, however it is evident that with considerable density-independent recruitment variability occurring, the stock may decline to or below this threshold even in the absence of appreciable fishing mortality. Although there are a range of theoretically-based alternative proxies (e.g., $F_{35\%}$, $SB_{50\%}$) and methods for calculating relative stock size (e.g., relative to a specific time-period's recruitment, relative to a reference year, relative to an estimate of what stock size might have occurred in a specific year in the absence of fishing [dynamic SB_0]), those selected for hoki appear to consistent with a large body of current fisheries research as well as with management in many areas around the world.

The assessment team has been using an eggs-per-recruit approach in place of the more standard spawning biomass-per-recruit analysis. There has been criticism of this approach due to additional uncertainty related to the fecundity information. During the review it was noted that the fecundity curve compared to the weight-at-age curve appears to accentuate the younger fish over the older fish, which is counter to most literature that has shown that egg-quality is usually higher in older fish. We recommend changing from an egg-output-based calculation to spawning biomass-based calculation in the short-term, due to the availability of only a single fecundity vector and variable growth over time. Spawning biomass appears to be a more robust quantity; at least until a more thorough investigation into time-varying fecundity (which should be done by age rather than converted from length) can be performed. Fecundity data exist from 2001-2011, with an average sample size of about 100. Depending on the outcome of analyzing these data, it might be reasonable in the future to allow the maturity schedule to be updated and reverting to egg-output-based reference points.

The phase plots of fishing mortality and spawning biomass that the authors have generated are very informative and represent a nice innovation beyond those used in many other stock assessments. We recommend that management work with a group of assessment authors (not hoki-specific) to develop a general control rule for "data-medium" species. The figure below shows an example of the Alaska control rule for "data-medium" species. Essentially, $SB_{40\%}$ is the target, and when spawning biomass declines below the target, there is an automatic feedback (or "ramp") that lowers the allowable fishing mortality. This linear ramp means that as biomass decreases, catch decreases more sharply. The west cost of the U.S. uses a similar system, but the ramp controls catch, not fishing mortality. A control rule in place would prevent arbitrary recommendations by assessment teams deciding how to lower fishing mortality when a stock is declining, particularly if it was applied consistently and across stocks. Another system could be utilized that is similar to what the hoki assessment team is currently doing. That is, make a rule (we call it "P^*" in the U.S.) that sets the catch at the level that there is less than XX% probability of being below a set limit reference point next year (e.g., $SB_{20\%}$). Clearly, as the stock gets closer to the LRP, catch would decrease faster than biomass. An important aspect to successfully developing harvest control rules is to consider them across a broad group of similar stocks, and either have laws that enforce them, or cooperation with industry to follow them. The Scientific Committee needs to work with the stock assessment teams, but ultimately they need to set targets and limits, not individual stock assessment authors.

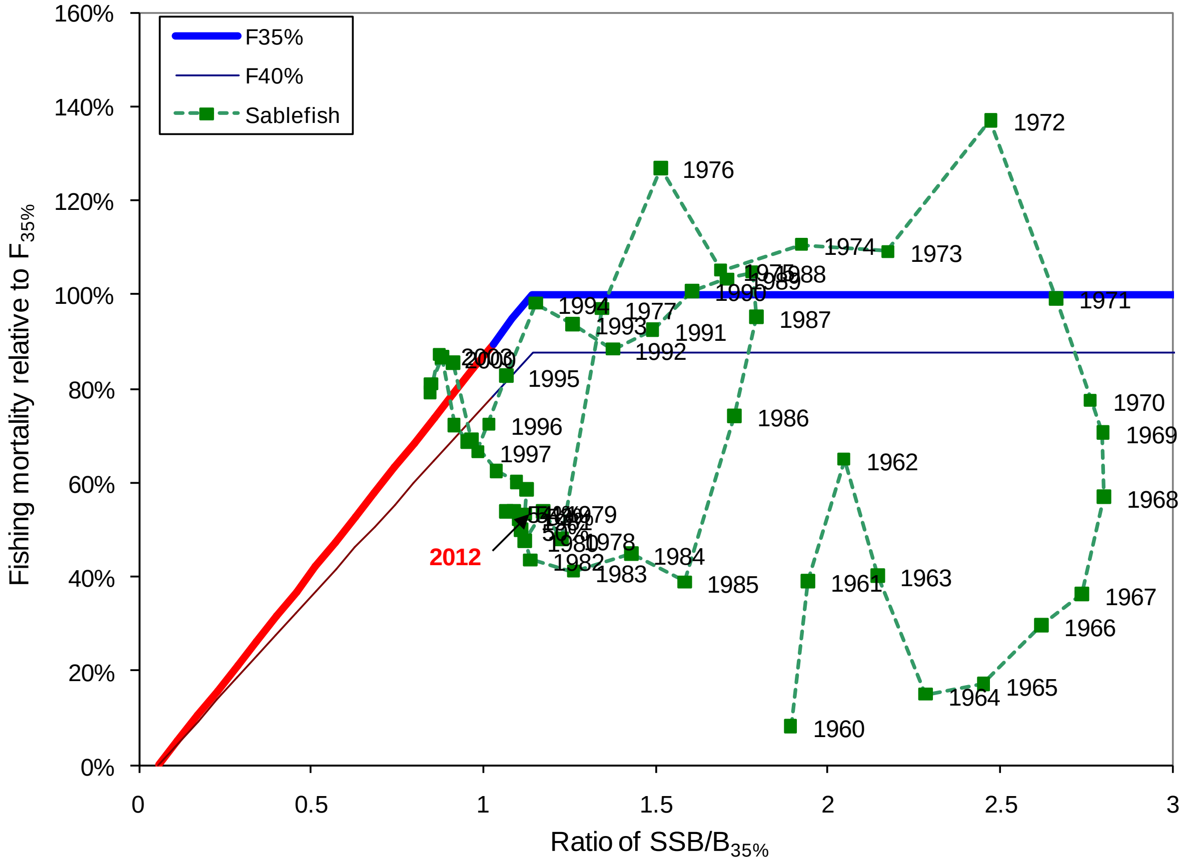

Figure. Phase-plane diagram of time series of sablefish estimated spawning biomass relative to the unfished level and fishing mortality relative to F_{OFL} ($F_{35\%}$).

Recommendations:

Short-term

1) Change from an egg-output-based calculation of reference points to a spawning biomass-based calculation.
2) Develop an explicit harvest control rule, working with the scientific committee, that is mutually agreeable and explicit in the identification of reference point targets and limits for both biomass and fishing mortality rates.
3) In the harvest control rule recommended above, specifically consider the use of a ‘ramp-down’ policy, allowing for a disproportionate reduction in fishing mortality rate when the stock is below target biomass levels that will generally allow for stock recovery rather than just stabilization.

Medium-term

1) Depending on the results from analysis of recent fecundity data, consider allowing the maturity schedule to be updated for recent years, and reverting to egg-output-based reference points.

TOR 9. Short and medium-term improvements to address the uncertainty of the stock assessment (Doc 28, 27, 25; Quinn and Cox 2011, Payá 2012g, c).

Providing technical assistance to introduce short and medium-term improvements to address the uncertainty of the stock assessment.

9.1. Recommending short and medium-term improvements for the stock assessment.

9.2. Make suggestions to establish a standard for the Chilean Hoki stock assessment.

As discussed above, the use of fixed values used for natural mortality, steepness, and an informative prior distribution for survey catchability, results in a gross underestimation of the uncertainty in the current Chilean hoki stock assessment. Also noted above, providing decision tables that include probability distributions of alternate states of nature (such as "low productivity", "base case", and "high productivity" for future scenarios), while still maintaining that the "base case" as the highest probability outcome could be valuable in capturing more of the uncertainty in assessment results.

TOR10. Key aspects to be improved in the short and medium terms (Doc 27, 25; Quinn and Cox 2011, Payá 2012c).

Identifying the key aspects to be improved in the short and medium term for the stock assessment, and advising research actions aimed to increase the knowledge on the resource and reducing uncertainty of the stock assessment.

10.1. Recommending research lines which improve the knowledge on structure, dynamics, distribution and connectivity of possible hoki stock units.

These recommendations have been included in the other TORs as applicable and are summarized in a table at the end of this document.

TOR 11. Reviewing the analysis procedure for the recommendation of Biologically Acceptable Catches (Doc 28, 27, 26, 22; Payá 2012g, c, d, e).

Reviewing the analysis procedure for the recommendation of Biologically Acceptable Catches of Chilean hoki (based on NRC 1998)

Key questions:

- *Is the BAC measurement procedure adequate?*
- *Is the way in which uncertainty is incorporated in the BAC analysis adequate?*
- *Is the way in which the risk assessment is conducted adequate?*

In most commercial fisheries it is generally considered a sign of a depleted stock when the annual catches are far below the quotas that have been set. This appears to have been the case for hoki over most of the last decade. Discussions during the review revealed that this was due to a shift in the spatial and fleet distribution away from the northern area and the purse seine fleet. However, it is clear that, although catches have not greatly exceeded the stock assessment recommendations, future shifts in allocation and/or spatial distribution of the fleets could lead to this result. It is important in the short-term for management to identify a harvest control rule (described in TOR 8) under which to operate, in order to ensure the best scientific advice available is used to set the annual BAC. Additionally, it appears that the authors have continued to present their best estimates of what BAC should be, but scientific advice has not been followed by management when quotas are set. According to Figure 1 (Doc. 4), the quota has been the same since the 2005 stock assessment, despite different recommendations since. The Undersecretary has consistently proposed quotas that are double or triple the recommended BAC from the stock assessment, and the National Management Council has accepted the Undersecretary’s recommendations. In 2012, the Undersecretary attempted to recommend a lower quota (still significantly larger than the assessment BAC), but this was rejected by the Management Council so the quota was automatically reduced to 80% of the previous year. Therefore, there is a strong disconnect between the scientific analysis and the actual quotas that are being set at the end of the process.

The assessment presents a perfectly adequate method of showing the risk by quantifying the probability of being below reference points. But as discussed in previous TORs, the risk is underestimated because uncertainty is underestimated. While a risk tolerance of a probability of 10% would be considered conservative, but the true probability is likely much higher. We recommend attempting to more fully capture the uncertainty of the current level of spawning biomass compared to reference points to better assess risk.

Short-term

1) Develop an explicit harvest control rule, working with the scientific committee, that is mutually agreeable and explicit in the identification of reference point targets and limits for both biomass and fishing mortality rates.
2) When a stock assessment is rejected, management needs to reduce quotas. Rejection of the assessment recommendations implies higher uncertainty, and thus, higher risk.

Long-term

1) Evaluate the potential for benefits of a marine protected area on part of the essential spawning stock areas.

TOR 12. Critically reviewing the stock dynamics simulation models (Doc 28, 27, 26, 22; Payá 2012g, c, d, e).

Critically reviewing the stock dynamics simulation models

Key questions:

- *Is the approach adopted by IFOP (non-participatory) adequate?*
- *What information should be considered to establish the initial condition of the projection*
- *Is it necessary to consider a stock-recruit ratio?*
- *On what basis or information should parameters be set?*
- *Are current recruitment projections consistent?*
- *Is it adequate the way in which stock estimation/assessment uncertainty is incorporated in the projection/simulation analysis?*

One deficiency of the projections in the 2011 stock assessment was detected that was not found in 2010 stock assessment. The projections in Doc. 26 (e.g., Figs 34-35) present only one iteration of random recruitment draws. This results in the 'bumpy' future projections of biomass. It would be preferable to plot the median/mean over many iterations of recruitment variability along with the 95% quantiles or perform a full Bayesian integration (via MCMC) and present the credible intervals. The way it is presented now, using the Hessian standard deviation from just one iteration, underestimates the uncertainty and gives an unrealistic expectation of the stock quickly rebounding with high probability (although this is not necessarily the case from one realization). The initial conditions should always be the age structure in the last year. We recommend focusing on short-term projections: 1 year and 5 years would be reasonable to avoid a strong influence of unobserved recruitments drawn from the S-R curve. It is not necessary to use a stock–recruitment relationship for the projections, although this could be presented as one of several alternatives. The authors also might consider drawing recruitments from just the mean and standard deviation of past recruitment (parametric) or resampling the prior recruitment estimates (nonparametric). If the assessment authors anticipate that allocation or gear changes are a strong possibility in the future, it is important to simulate projections where the relative distribution of catches among fleets or gears change over time.

Simulation studies could be conducted to 1) validate the performance assessment model code and determine the estimability of parameters (the bootstrap approach), 2) Test how robust the assessment is to alternate assumptions and structural dynamics (different models/hypotheses for generating the data) and 3) full exploration of the performance of the assessment and management/harvest control rule (Management Strategy Evaluation; MSE). These approaches are generally applicable to all stock assessments, and represent a considerable amount of additional research that must usually be conducted outside the normal stock assessment preparation schedule. For Chilean hoki, type 1 simulation would be useful to validate the current model software, as it is a custom application that has not been tested with a similar set of data in the manner that many more generalized stock assessment software applications have been tested (e.g., Multifan, Stock Synthesis, AMAK, Coleraine and others). This would be useful in the medium term. Type 2 simulation experiments should be undertaken when specific hypotheses are of concern (e.g., changes in natural mortality over time) but the data are insufficient to resolve their validity. This type of experiment consists of generating data sets under the alternative hypotheses and determining when the current model fails to perform adequately. It does not appear that there are well-defined alternative hypotheses ready to be tested in this manner for hoki, but this could be investigated as the few remaining technical issues are resolved and after a type 1 analysis has been conducted. Full management strategy evaluation is currently a goal for many fisheries, but might be premature for Chilean hoki until the harvest control rules, reference points and management system have been fully specified, but may be useful to examine potential control rules being formulated.

Recommendations

Short-term
1) Refine projections to include either full MCMC projections with credible intervals or quantiles of many recruitment draws.
2) Conduct a type 1 simulation (does simulated input data give the right answer) to validate new custom model
3) Consider using resampled or parametrically generated recruitments instead of a S-R relationship

Long-term
1) Conduct management strategy evaluations to examine effects of various catch, allocation, productivity and environmental scenarios on future stock health.

Conclusions

The assessment team has gone through considerable effort to make the most of a “medium-data” situation. In both the 2010 and 2011 assessment the model is robust in defining a distinct downward trend and that the stock is likely below the target spawning biomass ($B_{40\%}$) and possibly below the proposed limit spawning biomass

($B_{20\%}$). We have compiled many recommendations, but none of them are likely to significantly change this general result. Collaborative research needs to be conducted with industry and management to compute the best possible estimates of total catch for all stocks where discards and underreporting are a problem. We believe that another important task to be completed is to establish better protocols for interaction between the Scientific Committee, industry and the assessment team. This will be critical in establishing appropriate reference points and a control rule to regulate catches when the stock is in decline. This recommendation is not hoki-specific in that standardizing the reference point calculations and harvest control rule should be done for all "data-medium" stocks to avoid entangling the general discussions with specific catch quotas. In some type of working group or workshop forum, a TOR or set of guidelines should be created for stock assessment authors including responses to previous reviews, reference point calculations, BAC methods and other issues determined to be important by the scientific committee. If the Scientific Committee does not already, they should publish minutes of their meetings and the assessment authors should explicitly respond to committee recommendations or concerns. The Scientific Committee could request specific model hypotheses to be evaluated and considered for potential inclusion in the annual stock assessment, but should limit the number of hypotheses tested annually and provide adequate time to complete the requested analyses. There is not currently an explicit process for including these recommendations and ensuring the confidence of all parties. This can lead to a loss of confidence in the scientific process. Additionally, parties should mutually agree upon formats for figures (particularly the phase plot) and tables (such as a decision table, or a table of catch options and probabilities of various biological outcomes [e.g. the probability of stock decline]), such that managers can accurately and easily interpret the results among assessments.

The assessment team has only a short time-period between the most recent acoustic survey data becoming available and the annual stock assessment being produced. Although this is common for many stock assessments, it underscores the need for prioritization of research in between stock assessments and a focus on the most important issues, perhaps as outlined in the last several independent reviews. For future reviews, we would recommend that only one year's assessment be specifically reviewed because it created confusion in this review concerning which model we were evaluating.

Recommendations by Priority

Short-term:

1) Qualitatively compare the trends in biomass and recruitment between the Argentinean and Chilean assessments.
2) Collaborate among institutions and with industry to compute the best estimates of total catch for the entire time-series including all discards and underreporting. This may require a workshop to identify key issues, data sources and participants and outline a plan under which the research will be conducted. This effort should also include recent years (2008+), which were not available for this review.
3) If there are delays in creating a working time-series of total catch estimates, the next stock assessment should perform sensitivity analyses to several plausible levels of total mortality in order to evaluate the effects on the results and uncertainty in those results.
4) Collaborate with industry to compute best estimates of the catch-at-length during the purse-seine fishery. Given the magnitude of these catches, this could be important to estimates of selectivity for this fleet and the trends in spawning stock biomass.
5) Omit the years 1997-2001 from the trawl CPUE index time series (and GLM analysis).
6) Remove the purse seine CPUE index from the assessment model.
7) Review the current model code and include appropriate bias-correction terms.
8) Investigate the estimated level of recruitment variability and iteratively ensure that it is consistent with the variability used to determine reference points and initial conditions.
9) Attempt to fit asymptotic selectivity from 2002 on instead of from 1991 on for the trawl fishery.
10) Ensure that the Francis method for tuning compositional weights has converged by conducting more iterations.
11) Include interannual variability in sample sizes in the weights on data sources.
12) Correct the minor inconsistency in the stock-recruitment plotting code causing mislabeling of some recruitment years.
13) Evaluate, after addressing the recommendations regarding bias-correction and initial conditions above, whether starting the model prior to the first year of appreciable catches would reduce the sensitivity to the specific choice of year.
14) Perform periodic testing in future assessments of whether catchability, natural mortality, and steepness are estimable. Routinely use likelihood profiles to better understand which data sources are providing information on these parameters.
15) Consider the use of simple decision tables to convey a broader spectrum of uncertainty to fisheries managers.

16) Change from an egg-output-based calculation of reference points to a spawning biomass-based calculation.
17) Develop an explicit harvest control rule, working with the scientific committee, that is mutually agreeable and explicit in the identification of reference point targets and limits for both biomass and fishing mortality rates.
18) In the harvest control rule recommended above, specifically consider the use of a 'ramp-down' policy, allowing for a disproportionate reduction in fishing mortality rate when the stock is below target biomass levels that will generally allow for stock recovery rather than just stabilization.
19) Develop an explicit harvest control rule, working with the scientific committee, that is mutually agreeable and explicit in the identification of reference point targets and limits for both biomass and fishing mortality rates.
20) When a stock assessment is rejected, management needs to reduce quotas. Rejection of the assessment recommendations implies higher uncertainty, and thus, higher risk.
21) Refine projections to include either full MCMC projections with credible intervals or quantiles of many recruitment draws.
22) Conduct a type 1 simulation (does simulated input data give the right answer) to validate new custom model.
23) Consider using resampled or parametrically generated recruitments instead of a S-R relationship.

Medium-term:

1) Investigate the degree of ageing imprecision by performing double-reads of the same otoliths and/or exchanges with other aging laboratories.
2) Consider an age-specific (Lorenzen) natural mortality to account for predation juveniles.
3) Obtain representative samples from hoki in the winter both on and off the spawning aggregations.
4) Investigate the increase in weight-at-length, but decrease in weight-at-age over the same period.
5) Conversion factors used to adjust dressed products to whole catch weights appear to be variable by source. Although this is likely a very small contributor to whole catch uncertainty relative to discarding and underreporting, an effort should be made to refine conversion factors as processing technology and market products change over time.
6) Explore other methods of conveying the uncertainty in the model outputs, such as alternate states of nature, likelihood profiles, and examination of posterior distributions when freeing up more key parameters.
7) Attempt to provide consistency in the tuning and weighting methods employed among stock assessments for groups of similar species. This would be helpful to avoid inconsistent and/or subjective decisions by different assessment authors.
8) Attempt to tune input to output error by using standardized deviation of normalized residuals and/or effective sample sizes.

9) Depending on the results from analysis of recent fecundity data, consider allowing the maturity schedule to be updated for recent years, and reverting to egg-output-based reference points.

Long-term:

1) Conduct a joint stock assessment with the Argentineans.
2) Implement conventional tagging experiments, ideally with Argentina, to attempt to estimate movement rates, and a better understanding of stock structure.
3) Improve estimates of abundance of and potential consumption rates by jumbo squid and other key predators.
4) Investigate the presence of other potential spawning aggregations. This could be achieved through extending the acoustic survey, or through industry collaboration/charter.
5) Design and conduct a multispecies synoptic trawl survey.
6) Evaluate the potential for benefits of a marine protected area on part of the essential spawning stock areas.
7) Conduct management strategy evaluations to examine effects of various catch, allocation, productivity and environmental scenarios on future stock health.
8) As time permits, explore further analyses of purse seine CPUE, potentially investigating the spatial-temporal interactions, the subsetting of trips targeting hoki, and changes in the fishing vessels participating.

References

ADMB Project. 2012. AD Model Builder: automatic differentiation model builder. Developed by David Fournier and freely available from admb-project.org.

Caballero, L. and I. Payá. 2012. Abundance index based on purse seine CPUE. Chilean Hoki Stock Assessment Review 2012. Doc. 8. 9 p.

Canales, C. 2011. Natural mortality estimates for Chilean hoki (*Macroronus magellanicus*). Chilean Hoki Stock Assessment Review 2012. Doc 15. 5 p.

Céspedes M., R. 2012. Review of monitoring system for Chilean hoki. Chilean Hoki Stock Assessment Review 2012. Doc 16. 29 p.

Clark, W. G. 2012. Review of data and external parameter values used in the assessment of the Chilean stock of hoki (merluza de cola). Chilean Hoki Stock Assessment Review 2012. Doc 18. 23 p.

Lillo, S. 2012. Hydroacoustic survey methods to estimate biomass of Chilean hoki. Chilean Hoki Stock Assessment Review 2012. Doc 12. 14 p.

MacCall, A. D. 1990. Dynamic geography of marine fish populations. Washington Sea Grant Program, Seattle.

Methot, R. D. and I. G. Taylor. 2011. Adjusting for bias due to variability in estimated recruitments in fishery assessment models. Canadian Journal of Fisheries and Aquatic Sciences **68**:1744-1760.

Ojeda, V. and H. Hidalgo. 2012. Chilean hoki age. Chilean Hoki Stock Assessment Review 2012. Doc 4. 41 p.

Payá, I. 2012a. Abundance index based on catch rates of the trawl fishery from 1979 to 2009. Chilean Hoki Stock Assessment Review 2012. Doc. 6. 39 p.

Payá, I. 2012b. Abundance index based on catch rates of the trawl fishery from 1979 to 2010. Chilean Hoki Stock Assessment Review 2012. Doc 17. 74 p.

Payá, I. 2012c. Biological reference points (BRP) for Chilean hoki (*Macruronus magellanicus*): comparing BRP on the basis of egg production per recruit and spawning biomass per recruit. Chilean Hoki Stock Assessment Review 2012. Doc 27. 5 p.

Payá, I. 2012d. Chilean hoki (*Macroronus Magellaniccus*) stock assessment in 2011. Chilean Hoki Stock Assessment Review 2012. Doc 26. 66 p.

Payá, I. 2012e. Chilean hoki stock assessment results, 2010. Part 2., Chilean Hoki Stock Assessment Review 2012. Doc 22. 17 p.

Payá, I. 2012f. Data and parameter review for Chilean hoki stock assessment. Chilean Hoki Stock Assessment Review 2012. Doc. 4. 24 p.

Payá, I. 2012g. The evolution of Chilean hoki (*Macrorunos magellanicus*) stock assessments conducted by the Instituto de Fomento Pesquero since 1999. Chilean Hoki Stock Assessment Review 2012. Doc 28. 24 p.

Payá, I. 2012h. Fish predation on Chilean hoki. Chilean Hoki Stock Assessment Review 2012. Doc 11. 4 p.

Payá, I. 2012i. Is the swept area biomass 1972 a virgin biomass at equilibrium? , Chilean Hoki Stock Assessment Review 2012. Doc. 7. 8 p.

Payá, I. 2012j. Review of Humbold squid in Chilean waters and its probable consume of hake. Chilean Hoki Stock Assessment Review 2012. Doc 10. 16 p.

Payá, I. 2012k. Stock Units in South American hoki, *Macruronus magellanicus*. Chilean Hoki Stock Assessment Review 2012. Doc 14. 9 p.

Payá, I. 2012l. Testing the hypothesis of stability of weight at age by year in commercial catches. Chilean Hoki Stock Assessment Review 2012. Doc. 5. 11 p.

Payá, I. 2012m. Whole catch uncertainty factors. Chilean Hoki Stock Assessment Review 2012. Doc. 9. 6 p.

Quinn, T. J., II and S. P. Cox. 2011. Workshop Report. Chilean hoki stock assessment workshop 2011. Chilean Hoki Stock Assessment Review 2012. Doc 25. 74 p.

Appendix 1: Agenda

CHILEAN HOKI STOCK ASSESSMENT REVIEW 2012 AGENDA

Monday 4/6
8:30 – 9:30

- Meeting of experts with IFOP steering committee at the hotel

10:00-11:30

- Meeting of experts with fishery managers at the Under-Secretary of fisheries in Valparaiso Port.
- Return to the Hotel in Viña del Mar

12:00-13:00

- Welcome and administrative stuff
    - Review of TOR (Doc 1) and Agenda (Doc 2).
    - General Introduction (CHODAR2012)(Doc 3). Presentation. I. Payá.

13:00-14:30 Lunch
14:30-16:00

- TOR 1. Structure and dynamics of the national stock.
    - CHODAR TOR1. Stock structure (Docs 18, 14).
    - CHODAR TOR2. Monitoring system (Docs 18, 16).

- TOR 2. Life cycle and dynamics parameters.
    - CHODAR TOR 3. Life history information (Docs 18, 13, 15, 11, 10, 5, 4).

- TOR3. Growth parameters, age/length keys, and matrices of catch at age/size.
    - CHODAR TOR 4. Landings and discards (Docs 18, 4, 9,).
    - CHODAR TOR 5. Age/length composition data (Docs 18, 4, 7).

16:00-16:30 Coffee break
16:30-17:30

- TOR4. Relevance and sufficiency of the data sources and indicators.
    - CHODAR TOR 6. Abundance indices based on commercial catch rates (Docs 18, 6, 8, 17).
    - CHODAR TOR 7. Abundance indices based on acoustic surveys (Docs 18, 12, 4).
    - CHODAR TOR 8. Estimates of abundance and mortality from the summer 1972 trawl survey (Docs 18, 7).
    - CHODAR TOR 9. Provide measures of population abundance that are appropriate for stock assessment (Doc 18).

17:30 End

Tuesday 5/6
9:00-11:00

- TOR5. Relevance, sufficiency and adequacy of the approach and procedures (Doc 25).

- TOR6. Stock assessment model: procedure models, observation and error and uncertainty sources.
  - *6.1. Review the formulation of the stock assessment models with a focus on the most recent assessments.* (Docs 25, 20 & 19)

11:00-11:30 Coffee break
11:30-13:00

6.1. Continuing.

13:00-14:30 Lunch
14:30-16:00

6.1. Continuing.

16:00-16:30 Coffee break
16:30-17:30

6.2. Summary of the changes of programming language used in the assessment models.

17:30 End

Wednesday 6/6
9:00-11:00

6.3. Critically reviewing the procedure to define the assessment initial conditions. Propose modifications where appropriate (Docs 25, 18, 7).

11:00-11:30 Coffee break
11:30-13:00

6.4. Analyzing the model goodness to fit the data (residuals analysis) and reviewing the consistency and the main outcomes of the base model, considering the outcomes of the retrospective analyses (Doc 25 & 21).

13:00-14:30 Lunch.

14:30-16:00

6.5. Critically reviewing the treatment of the uncertainty in the model outcomes (distributions a priori and posteriori, and MC/MC chains)(Docs. 25, 23 , 22).

6.6. Review of the base model definition procedure, the alternative cases and the sensitivity analyses conducted, considering the assessment procedure used (Doc 25 & 21).

16:00-16:15 Coffee break
16:30-17:30

6.7. Review of the significance and impact on estimations deriving from the weighting assigned to the different information sources and abundance indices used in the assessment (Doc 25 & 21).

17:30 End

Thursday 7/6
9:00-11:00

- TOR 7. Providing technical assistance to improve the robustness of the assessment against currently-known uncertainty factors (Doc 25).

11:00-11:30 Coffee break
11:30-13:00

- TOR 8. Critically analyzing the BRP and reviewing the criteria and procedures used to determine the stock status.

Lunch: 13:00-14:30
14:30-16:00

- TOR 8. Continuing.

16:00-16:15 Coffee break
16:15-17:30

- TOR 9. Short and medium-term improvements to address the uncertainty of the stock assessment (Doc 25).
- TOR10. Key aspects to be improved in the short and medium terms (Doc 25).

17:30 End

Friday 8/6
9:00-11:00

- TOR 11. Reviewing the analysis procedure for the recommendation of Biologically Acceptable Catches

11:00-11:30 Coffee break
11:30-13:00

- TOR 11. Continuing.

Lunch: 13:00-14:30
14:30-16:00

- TOR 12. Critically reviewing the stock dynamics simulation models

16:00-16:15 Coffee break
16:15-17:30

- TOR 12. Continuing.
- Final Conclusions and recommendations

17:30 End

Appendix 2: List of Review Participants

Name	Background	Institution
Chairman/reporter		
Dr. Ian Stewart	Lead assessment author for Pacific hake	NWFSC, NMFS, NOAA
Dr. Dana Hanselman	Lead assessment author for Sablefish in Alaska.	AFSC, NMFS, NOAA
Stock Assessment Team		
Ignacio Payá	Stock Assessment Models	IFOP
Cristian Canales	Stock Assessment Models	IFOP
Renzo Tascheri	CPUE analysis	IFOP
Hoki Working Team Members		
Sergio Lillo	Acoustic Surveys	IFOP
Renato Céspdes	Fishery Monitoring	IFOP
Patrício Galvez	Fishery Monitoring	IFOP
Dario Rivas	Fishery manager	UnderSecretariat for Fisheries
Ruben Alarcón	CPUE analysis	CEPES, fishery industry
Alejandro Zuleta	Stock Assessment	CEPES, fishery industry
Observers		
Analía Giussi	Stock Assessment	INIDEP, Argentina
Fernando Balbontín	Reproductive Biology	U. Valparaíso
Carlos Moreno	Ecology	U. Austral
Juan C. Quiroz	Stock Assessment	IFOP
Elson Leal	Stock Assessment	IFOP
Luis Adasme	Technician	IFOP
Doris Bucarey	Stock Assessment	IFOP