Stock Assessment Data and Modeling, Reference Point Calculation, and Migration

IPHC Scientific Review Board, Meeting 2.

Dr Sean P. Cox, Associate Professor, School of Resource and Environmental Management, Simon Fraser University, 8888 University Dr, Burnaby, B.C., Canada V5A 1S6

Dr James Ianelli, Research Scientist, National Marine Fisheries Service-NOAA, Building 4, 7600 Sand Pt Way NE, Seattle, WA, USA 98115

Dr Marc Mangel, Distinguished Research Professor and Director, Center for Stock Assessment Research, University of California, Santa Cruz, CA 95064

20 November 2013

Overview

We met from 1 pm on Tuesday 1 Oct 2013 to about 4 pm on Thursday 3 Oct 2013. The presentations by IPHC staff were thoughtful, carefully prepared, and illuminating. Similarly, the questions put to us, both in advance and during the meeting of the SRB were focused and on target.

Assessment, apportionment, harvest policy and Management Strategy Evaluation (MSE) for Pacific halibut are all conditioned on data, so it is important that we pay close attention to management, processing, refinement, and future data collection. For halibut thus far, there are few/no conflicting results between alternative data sources and types, but that may not always be true in the future. Where the data are reliable, models can help crystalize the story that the data are trying to tell and, more importantly, help to guide sensible management choices.

<u>Response to specific IPHC questions</u>

1. Are the proposed datasets and processing procedures appropriate for the 2013 halibut stock assessment?

Specific issues about data and processing procedures are list below under Question 2 to keep issues and recommended changes together. Overall, we consider that they are appropriate, but also that some improvements can be made.

2. Are there changes or improvements in data collection, processing, or treatment in the stock assessment that would improve future analyses?

We strongly recommend that the IPHC develop fishery at-sea sampling program for collecting sex composition of commercial catch. There are many volunteer or compulsory means of collecting this information.

We support the IPHC's proposed change from using O32 WPUE as the biomass index in the assessment model to Total WPUE. The values are similar historically, but the Total WPUE more fully utilizes information contained in the sub-legal halibut fishery and survey data.

We also encourage an evaluation of total number-per-unit-effort (Total NPUE) as an alternative to Total WPUE to reduce the influence of weight-at-age assumptions (see below) on age-composition data.

Data treatment in stock assessment: Removals, by-catch, and wastage

Directed commercial fishery landings are well monitored and considered accurate throughout the range of halibut in the north Pacific. Recreational fishery catch is more difficult to quantify and is also increasing. The IPHC currently assumes that halibut selectivity in recreational fisheries is the same as the survey because recreational fishery catch is not comprehensively sampled for size- or age-composition. It would be worthwhile to examine the sensitivity of stock assessment inferences to alternative assumptions about recreational fishery selectivity and also how a complete set of recreational fishery biological samples from all jurisdictions would affect the assessment. The latter study could help to determine the value of improved recreational catch sampling.

We are particularly concerned about the potential for biased estimates of by-catch in the Gulf of Alaska (GOA) fisheries. By-catch estimation in areas 3A and 3B is apparently based on low levels of observer coverage, whereas by-catch estimates from the Bering Sea (the largest halibut by-catch levels) are based on nearly 100% observer coverage. The new NMFS observer deployment plan in the GOA should help improve the estimates of by-catch in the trawl fisheries, but low and potentially inadequate observer coverage for longline vessels operating in the GOA will likely continue into the foreseeable future.

Wastage estimates for all IPHC regulatory areas are currently extrapolated from IPHC survey catch rates. Although the survey-based estimates of wastage agree quite well with direct estimates of fishery wastage available in area 2B it would be more appropriate to obtain direct estimates in the GOA by improving observer coverage rather than continuing to rely on indirect estimates.

Although the current procedures for estimating by-catch and wastage appear adequate for most areas, the resulting data do not span the entire fishery time-series, which means that missing values prior to 1991 must be imputed or assumed equal to zero in the stock assessment. One option for filling in the missing values is to extrapolate current estimates, although that entails some tenuous assumptions about the relationship between historical by-catch and abundance. An alternative approach would be to model the by-catch and wastage processes in the stock assessment as a function of groundfish effort, generate predicted by-catch and wastage values for all years, and add a likelihood component to the assessment for the years in which observed by-catch and wastage estimates exist. This latter approach is preferable because it propagates the uncertainty in historical by-catch and wastage and only conditions the model on values that are actually observed.

Data treatment in stock assessment: Empirical weight-at-age

Dramatic changes in size-at-age of halibut over the past two decades need to be accounted for in the assessment model. Empirical weight-at-age data obtained from surveys are used appropriately in the model to compute spawning and exploitable biomass. Prior to these calculations, weight-at-age observations require smoothing to remove inter-annual sampling variability. We recommend using the smoothed mean weights-at-age over time instead of the existing method, which smoothed over ages (independent of previous years or cohorts). The most important feature to maintain in the smoothing is consistency of weight-at-age over time for each cohort; that is, size-at-age for each cohort should generally be increasing albeit at different rates.

We recommend evaluating other approaches for forecasting weight-at-age in stock assessment projections that are used to populate decision tables. An example approach, used for walleye pollock assessments, is provided in the following document

http://www.afsc.noaa.gov/refm/docs/2009/EBSpollock.pdf).

We encourage IPHC research on estimating size-at-age from otolith length frequency analysis. However, it is important to address the issue of missing length data during the middle period of the fishery, when otoliths were used as a proxy. We discussed fitting otolith length frequency data directly, which may be preferred if the past transformations prove to be unreliable.

Data treatment in stock assessment: Imputing missing WPUE for the Bering Sea

Although the Bering Sea has low halibut density, the large bottom area within the 0-400 fathom zone can have an important influence on halibut biomass apportionment. However, only part of this area is covered by the annual IPHC survey. In 2006, IPHC conducted a single calibration experiment to estimate a conversion factor for translating NMFS bottom trawl surveys in selected sections of the Bering Sea into equivalent units of IPHC setline WPUE. This conversion is then applied to the entire bottom trawl survey time-series, providing WPUE for the period 1982-2013.

The original analysis of the 2006 experiment used a simple size-based calibration curve to account for different selectivity patterns between the setline gear and the bottom trawl. The IPHC's new approach compares size composition data from the setline survey in 2006 to bottom trawl survey size compositions from 2005-2007. The new method, using ratio or least-squares estimation, gives predicted cumulative length frequencies that are consistent with observed cumulative frequencies from the setline survey. Therefore, it appears to us that the new method (using either ratio or least-squares) is appropriate and an improvement over the previous approach.

We have two recommendations for the Bering Sea Index (BSI). First, sampling error from the BSI (as modified by the calibration curve) should be accounted for in Area 4 contributions to the coast-wide WPUE (or NPUE) used in the assessment. Second, the IPHC should consider replicating the calibration experiment to reduce sensitivity of the BSI to the single observation obtained in 2006. We recognize the challenges of securing vessels to complete a longline survey in this area and that the IPHC has made attempts already; therefore, it is important to consider the cost of this calibration and potential impacts on the stock assessment and biomass apportionment. As noted above for recreational catch sampling, sensitivity of stock assessment (and apportionment) inferences to deviations from the current calibration should be examined before moving forward on a new BSI field program.

Finally, we the IPHC should evaluate the potential utility of the NMFS trawl **slope** survey that occurs in even numbered years. This survey may help cover the area north of the present setline survey where halibut are apparently common. See the following website for details on these and other NMFS surveys.

http://www.afsc.noaa.gov/RACE/groundfish/survey_data/default.htm

Data treatment in stock assessment: hooking mortality and repeated capture-release

Smaller sizes of halibut imply an increasing rate of capture-release prior to fish entering the legal population and this may result in relatively high cumulative impacts of hooking mortality. There are statistical methods for estimating hooking mortality from wounding rates observed on surviving halibut. These include estimating vulnerability of aircraft (from anti-aircraft fire; Mangel and Samaniego 1984) and vulnerability of lake trout to parasitic sea lamprey, both using data obtained on survivors only.

Data treatment in stock assessment: revisions to age data

Over the past year, the IPHC updated and revised the ageing error matrix to reconcile break-and-bake and surface ageing methods. The update was similar over most halibut ages with the main effects being reduced biases in oldest ages.

Data treatment in stock assessment: other topics

- Consider retrospective analyses (using the long time-series model) that extend back to the early part of the new survey. This may shed light on how new survey data have changed stock perspectives and also evaluate the sensitivity of the model to having a reasonably precise index.

- Look at the age compositions by region of the longer time frame (perhaps simply mean length by region over time)? This will help understand the historical perspective of the regional age structures. For example, is there evidence that an area with mostly "young" halibut in recent years had a broader age range historically?

- Spatial pattern of sex ratio in the survey and catch (perhaps given a single age range) would be useful. This would help understand if some regions have different sex ratios and would pose questions about the impact of differential region-specific harvest rates.

- A model allowing the male selectivity offset parameter to vary over time might be useful (or could provide support for why it is reasonable to assume it is constant).

- An examination of total mortality (perhaps with migration rates included) should be presented (i.e., a catch-curve type analysis simply using survey data alone). This may help provide consistency with estimated movement rates by region and begin to distinguish movements (as estimated from tagging studies) from other sources of mortality.

- Finally, we encourage an examination of how alternative models for natural mortality affect stock assessment inferences. For example, the assessment could consider an allometric relationship between halibut body size and natural mortality (Lorenzen, 2000). Such effects could be examined in the context of scenarios for future management strategy evaluation research.

3. Pending a full re-analysis of the harvest policy, are there modifications that could improve the current treatment of reference points?

The IPHC harvest policy is relatively complex because it attempts to link biological reference points to an environmental time-series (e.g., the Pacific Decadal Oscillation, PDO). In particular, the average recruitment used to compute the $B_{100\%}$ reference point requires a decision about whether the PDO state is "good" or "bad" for halibut recruitment. We had extensive discussion of concerns related to these computations and linking them to the harvest policy[1]. IPHC staff presented an updated assessment approach that integrates the PDO effect into the assessment model rather than estimating the PDO effect on recruitment outputs from the model. The updated estimates of the PDO effect on recruitment is slightly lower when included in the new full time series assessment model.

[1] There is a reasonable amount of literature on this topic suggesting that such a link is unlikely to improve conservation or catch outcomes unless the environmental factor is strongly linked and is also highly predictable (see **References** section).

We agree with IPHC staff recommendations that including estimated recruitment from the new (apparently poor) PDO regime is inappropriate given the current method that is used for computing $B_{100\%}$. This would avoid a major modification to the calculation of this value.

We considered alternative methods that could be explored for deriving reference points. These include:

(1) Historical-based indicators and/or reference points derived from assessment model outputs. For example, depletion in any year could be computed relative to a base year that most stakeholders could agree represented "good" or "bad" halibut abundance or fishing conditions (or processing conditions). The same could be done for fishing mortality and SPR rates. Although the absolute value of historical-based indicators (e.g., F_{1970}) might differ for alternative assessment model assumptions, differences among models may be small when values are expressed relative to a common base year (e.g., F_{2013}/F_{1970} could be similar across alternative assessment models);

(2) Compute a "dynamic B_0", which treats unfished biomass as a stochastic rather than deterministic reference point. A dynamic B_0 would take environmental influences on recruitment into account in computing spawning biomass over the historical period in the absence of fishing. This would not require assumptions about how environment and recruitment are linked, which might help transparency;

(3) Projected replacement yield (yield in year t leading to the same spawning biomass in year t+1) is usually a simple concept to explain to stakeholders; and

(4) Compute B_{msy} or yield curves (even with steepness assumed) from the combinations of PDO and weight-at-age.

4. How should uncertainty among alternative models to treat the data and structure the stock assessment model best be incorporated into the 2013 decision table results?

We suggest working with the MSAB, beginning in 2014, to derive historical-based reference points from the long-term assessment model. This would be a useful and educational exercise as part of the MSE process. MSAB could provide area- and sector-based preferences (preliminary objectives) for years that represent favored fishery (or fishing) conditions. This way, the decision table could be used as a tool to help guide stakeholders in developing measurable objectives involving targets/limits, probabilities, and timeframes. The cumulative experience of examining decision tables should also be explicitly monitored to provide feedback on decision tendencies and the treatment of risk.

When presenting decision tables, we recommend exploring alternative definitions of probability so that a wider range of stakeholders can understand their implications. For example, probabilities could be stated in frequency format: e.g., "in 76 out of 100 situations like this, halibut biomass (or fishery CPUE) would be twice as high as it is today."[2] Recall the old Trident chewing gum slogan: "4 out of 5 dentists recommend Trident to their patients who chew gum" – would this have reached as wide an audience if it said: " 0.8 of dentists recommend Trident to their patients who chew gum"?

In general, the utility of the columns representing 20% and 30% of B_0 seem least useful to us, given that the computation of B_0 is a function of recruitment regime and the mean body mass-at-age that is assumed. It may be useful to begin familiarizing Commissioners and stakeholders with reference points that are less complex and easier to understand. For example, historical-based reference points could be used such as the current (or projected) spawning biomass relative to the lowest estimated value, or the current biomass relative to the mean (median or quantile) over some preferred period. We suggest dropping the blue line in the decision table until a formal management strategy evaluation has been conducted, because the present row simply represents a harvest policy that has been used in the past rather than one that has been fully evaluated against a suite of objectives.

Overall, we endorse the long-term assessment modeling approach because it: (1) provides longer term perspective; (2) uses much more of the existing data; (3) provides a potential link to historical tag data, which could help inform fishing mortality and selectivity patterns; (4) could be used collaboratively with stakeholders to derive historical-based reference points.

We are also very encouraged by the idea of “ensemble” modeling to more fully represent the uncertainty in stock assessment model estimates and projections. The SRB wondered how the ensemble components would be weighted in arriving at univariate indicators needed in decision tables.

5. What avenues of stock assessment model development should be pursued in the future?

As noted above, we agree with the IPHC's intent to pursue an Ensemble Modeling (EM) approach. There are two ways to involve an EM approach in halibut assessment and management. First, the EM approach can be adopted as the stock

---

2

assessment used to develop annual decision tables. As noted above though, this approach would require a method to weight EM components to produce a decision table with practical dimensions. Determining appropriate weights might be challenging because of nesting models within the ensemble. Second, an alternative is to develop a simple assessment model and test it for robustness against the suite of models in the ensemble. This is the general approach used in most applications of Management Strategy Evaluation (MSE). The simple assessment model is easier to explain and use in decision-making, while models can be added and subtracted from the ensemble without adding variability due to the "analyst effect".

A preliminary 3-model ensemble could include: (1) the long-term halibut assessment model, (2) areas as fleets, which could account for some of the demographic effects of different fishing selectivity patterns; and (3) a spatial model that uses existing tagging-derived movement estimates.

6. Comment on considerations for including migration in future stock assessment and MSE operating models.

We believe that spatial-temporal modeling of the distribution of halibut is a key area for future research. Fitting the model to data will be challenging, but not impossible. Indeed, modeling movement mechanisms is a greater challenge. A constant transition matrix could be developed from existing migration understanding derived from tagging.

Mechanisms causing halibut movement are more difficult to study, but could have important implications for determining robustness of management procedures. Density, size/age/gender, predation-risk, and environmental conditions could all drive halibut movement and distribution. The number of possible factor combinations involved in movement should serve as an early warning that spatial models might be best used in the second alternative to ensemble modeling – that is, as test beds for relatively simple, defensible assessment models and management procedures.

References

A'mar, Z. T., Punt, A. E., & Dorn, M. W. (2009). The impact of regime shifts on the performance of management strategies for the Gulf of Alaska walleye pollock (Theragra chalcogramma) fishery. *Canadian Journal of Fisheries and Aquatic Sciences*, *66*(12), 2222–2242. doi:10.1139/F09-142

Basson, M. (1999). The importance of environmental factors in the design of management procedures. *ICES Journal of Marine Science*, *56*(6), 933–942. doi:10.1006/jmsc.1999.0541

Collie, J. S., & Spencer, P. D. (1993). Management strategies for fish populations subject to long-term environmental variability and depensatory predation. In *Proceedings of the international symposium on management strategies for*

exploited fish populations (Vol. 629, p. 650). University of Alaska.
Lorenzen, K. (2000). Allometry of natural mortality as a basis for assessing optimal release size in fish-stocking programmes. *Canadian Journal of Fisheries and Aquatic Sciences 57: 2374-2381.*
Mangel, M. and F.J. Samaniego. 1984. Abraham Wald's work on aircraft survivability. Journal of the American Statistical Association 79:259-271
Satterthwaite, W.H., Kitaysky, A.S., and M. Mangel. 2012. Linking climate variability, productivity and stress to demography in a long-lived seabird. Marine Ecology Progress Series 454:221-235
Walters, C. J. (1989). Value of short-term forecasts of recruitment variation for harvest management. *Journal of the Fisheries Board of Canada*, *46*(11), 1969–1976.